Stochastic Optimization by Simulation: Convergence Proofs for the $GI/G/1$ Queue in Steady-state

Pierre L'Ecuyer • Peter W. Glynn
Département d'I.R.O., Université de Montréal, C.P. 6128, Montréal, Quebec, Canada, H3C 3J7
Operations Research Department, Stanford University, Stanford, California 94305

Approaches like finite differences with common random numbers, infinitesimal perturbation analysis, and the likelihood ratio method have drawn a great deal of attention recently as ways of estimating the gradient of a performance measure with respect to continuous parameters in a dynamic stochastic system. In this paper, we study the use of such estimators in stochastic approximation algorithms, to perform so-called "single-run optimizations" of steady-state systems. Under mild conditions, for an objective function that involves the mean system time in a $GI/G/1$ queue, we prove that many variants of these algorithms converge to the minimizer. In most cases, however, the simulation length must be increased from iteration to iteration; otherwise the algorithm may converge to the wrong value. One exception is a particular implementation of infinitesimal perturbation analysis, for which the single-run optimization converges to the optimum even with a fixed (and small) number of ends of service per iteration. As a by-product of our convergence proofs, we obtain some properties of the derivative estimators that could be of independent interest. Our analysis exploits the regenerative structure of the system, but our derivative estimation and optimization algorithms do not always take advantage of that regenerative structure. In a companion paper, we report numerical experiments with an $M/M/1$ queue, which illustrate the basic convergence properties and possible pitfalls of the various techniques.
(*Discrete Event Systems; Stochastic Approximation; Gradient Estimation; Optimization; Steady-state*)

1. Introduction

Simulation has traditionally been used to *evaluate* the performance of complex systems, especially when analytic formulae are not available. Using it to perform *optimization* is much more challenging. Consider a (stochastic) simulation model parameterized by a vector θ of continuous parameters, and suppose one seeks to minimize the expected value $\alpha(\theta)$ of some objective function. In principle, if $\alpha(\theta)$ is well behaved, one could estimate its derivative (or gradient) by simulation and use adapted versions of classical nonlinear programming algorithms. Recently, the question of how to estimate the gradient of a performance measure (defined as a mathematical expectation) with respect to continuous parameters, by simulation, has attracted a great deal of attention. (See, e.g., Glasserman 1991, Glynn 1990, L'Ecuyer 1990, Rubinstein and Shapiro 1993, and Suri 1989.) For steady-state simulations, a single-run iterative optimization scheme based on stochastic approximation (SA) has been suggested (Meketon 1987, Pflug 1990, Suri and Leung 1989). At each iteration, this scheme uses an estimate of the gradient of α to modify the current parameter value. These methods could enlarge substantially the class of stochastic optimization problems that can be solved in practice.

0025-1909/94/4011/1562$01.25

In this paper, we investigate the combination of SA with different derivative estimation techniques (DETs) in the context of a single queue. The general theory of SA has been studied extensively (see Kushner and Clark 1978, Metivier and Priouret 1984, and many references cited there), but not so much their combination with various DETs for discrete-event systems in the steady-state context, as we do here. Preliminary empirical experiments have been undertaken by Suri and Leung (1989) for a $M/M/1$ queue. These authors looked at two SA methods, which they presented as heuristics. One was based on infinitesimal perturbation analysis (IPA), while the other was an adaptation of the Kiefer-Wolfowitz (KW) algorithm, which uses finite differences (FD) to estimate the derivative. They did not prove convergence. We examine in this paper many DETs, including some based on FD, with and without common random numbers, IPA, and variants of the likelihood ratio (LR) or score function (SF) method. These techniques can be combined with SA in different ways. For example, at iteration n of SA, one can use either a (deterministic) truncated horizon t_n, or a fixed number t_n of regenerative cycles. Also, t_n can increase with n or remain constant. We prove a.s. (almost sure) convergence to the optimizer for many SA/DET variants. Within each class of DET (FD, LR, IPA), there are variants for which we require $t_n \to \infty$ as $n \to \infty$, and others for which there is no constraint on t_n (e.g., it can be constant). For the latter, the DETs are all regenerative estimators, with one exception. That exception is IPA with the same number of customers at each SA iteration, for which we prove weak convergence.

Chong and Ramadge (1992a, 1993) also analyzed (in parallel to us) IPA-based SA algorithms to optimize a single queue and proved a.s. convergence to the optimum, using different proof techniques than ours and different assumptions. In their first paper, they studied the case of an $M/G/1$ queue, while in their second, they considered a $GI/G/1$ queue and an SA algorithm which updates the parameter after an arbitrary number of customers. That includes in particular the case of one customer per SA iteration. In Chong and Ramadge (1992b), they extended their analysis and convergence proofs to more general regenerative systems. Fu (1990) previously analyzed a different variant of SA-IPA algorithm, for which he proved a.s. convergence. His algorithm exploited the regenerative structure of the system and the special form of the objective function (1) (see §2.2). His result corresponds to our Proposition 6(c). Wardi (1988) also suggested and analyzed a different variant of SA, combined with IPA, for which he showed a nonstandard kind of convergence which he called convergence in zero upper density. In all those papers, only IPA was considered.

A different approach for stochastic optimization, called the *stochastic counterpart* method, is proposed and thoroughly analyzed in Rubinstein and Shapiro (1993). The basic idea is to estimate the whole objective function as a function of θ in a parameterized form, using a likelihood ratio technique, and then to optimize that sample function by a standard (deterministic) optimization method. In this paper, we do not consider that approach.

In §2, we consider a $GI/G/1$ queue for which the decision variable is a parameter of the service time distribution. The aim is to minimize a function of the average system time per customer. We feel that many important questions that would arise in more general models, when SA is used to optimize infinite-horizon (steady-state) simulations, are well illustrated by this simple example. We recall the classical SA algorithm and give (in Appendix I) sufficient conditions for its convergence to the optimum. Section 3 reviews different ways of estimating the derivative (DETs). For a variety of SA-DET combinations, we prove convergence to the optimum under specific conditions (see Propositions 3–7). In the conclusion, we discuss briefly how all this can be extended to more general systems and mention prospects for further research. A companion paper (L'Ecuyer, Giroux, and Glynn 1994) reports numerical investigations and discusses the question of convergence rates, for which further analysis would be needed.

All our proofs are relegated to Appendix II. Since θ changes constantly between the iterations of SA, some convergence properties of the derivative estimators (e.g., bounded variance and convergence in expectation to the steady-state derivative) must be shown to hold *uniformly* in θ. As a by-product of our proofs, we obtain original results concerning $GI/G/1$ queues that could be of independent interest. For instance, it follows from the renewal-reward theorem (Wolff 1989) that for a stable queue, the average sojourn time of the first t customers in the queue converges in expectation, as

$t \to \infty$, to the infinite-horizon average sojourn time per customer. We prove, under appropriate conditions, that this convergence is uniform over θ and s, where s is the initial *state* (taken over some compact set), which corresponds to the waiting time of the first customer, and θ lies in a compact set in which the system is (uniformly) stable. We also derive a similar uniform convergence result for the derivative of the expected average sojourn time with respect to θ and a few additional characterizations of this expectation.

2. Example: A *GI/G/1* Queue

2.1. The Basic Model

Consider a $GI/G/1$ queue (Asmussen 1987, Wolff 1989) with interarrival and service-time distributions A and B_θ respectively, both with finite expectations and variances. The latter depends on a parameter $\theta \in \bar{\Theta} = [\bar{\theta}_{\min}, \bar{\theta}_{\max}] \subset \mathbb{R}$. We assume that for each $\theta \in \bar{\Theta}$, the system is stable. Let $w(\theta)$ be the average sojourn time in the system per customer, in steady-state, at parameter level θ. The objective function is:

$$\alpha(\theta) = w(\theta) + C(\theta) \tag{1}$$

where $C : \bar{\Theta} \mapsto \mathbb{R}$ is continuously differentiable and analytically available. We want to minimize $\alpha(\theta)$ over $\Theta = [\theta_{\min}, \theta_{\max}]$, where $\bar{\theta}_{\min} < \theta_{\min} < \theta_{\max} < \bar{\theta}_{\max}$. Let θ^* be the optimum. We define $\bar{\Theta}$ and Θ this way to be able to do FD derivative estimation at any point of Θ (see §3.1). This is also useful for LR and IPA. Let $\bar{\Theta}^0$ be an open interval that contains $\bar{\Theta}$.

A $GI/G/1$ queue can be described as a discrete-time Markov chain as follows. For $i \geq 1$, let W_i, ζ_i, and $W_i^* = W_i + \zeta_i$ be the *waiting* time, *service* time, and *sojourn* time for the ith customer, and ν_i be the time between arrivals of the ith and $(i+1)$th customer. For our purposes, W_i will be the state of the Markov chain at step i. The state space is $S = [0, \infty)$ and $W_1 = s$ is the initial state. $W_1 = 0$ corresponds to an initially empty system. For $i \geq 1$, one has

$$W_i^* := W_i + \zeta_i \quad \text{and} \quad W_{i+1} := (W_i^* - \nu_i)^+ \tag{2}$$

where x^+ means $\max(x, 0)$. Since $C(\theta)$ can be evaluated directly, we will estimate only the derivative of $w(\theta)$ and then add $C'(\theta)$ separately. Here and throughout the paper, the "prime" denotes the derivative with respect to θ.

We can view the Markov chain $\{W_i, i = 1, 2, \ldots\}$ as being defined over the probability space $(\Omega, \Sigma, P_{\theta,s})$, where $\{P_{\theta,s}, \theta \in \Theta, s \in S\}$ is a family of probability measures defined over (Ω, Σ). The sample point $\omega \in \Omega$ represents the "randomness" that drives the system, and $P_{\theta,s}$ depends (in general) on θ and s (where $W_1 = s \in S$ is deterministic). Let $E_{\theta,s}$ denote the corresponding mathematical expectation. When the quantities involved do not depend on s, we sometimes denote $E_{\theta,s}$ by E_θ to simplify the notation. For $t \geq 1$, let

$$h_t(\theta, s, \omega) = \sum_{i=1}^{t} W_i^*, \tag{3}$$

$$w_t(\theta, s) = \int_\Omega h_t(\theta, s, \omega) dP_{\theta,s}(\omega). \tag{4}$$

Here, $h_t(\theta, s, \omega)$ represents the *total* sojourn time in the system for the first t customers, and $w_t(\theta, s)$ its expectation. Let $\mathcal{F}_t$ be the σ-field generated by $(\zeta_1, \nu_1, \ldots, \zeta_t, \nu_t)$. Then, $h_t(\theta, s, \omega)$ is $\mathcal{F}_t$-measurable. Also, if $s = 0$ and if τ denotes the number of customers in the first busy cycle, then $\tau + 1$ is a stopping time with respect to $\{\mathcal{F}_t, t \geq 1\}$. Let $\bar{S} = [0, c]$ be the set of *admissible* initial states, where c is a (perhaps large) constant. It is well known from renewal theory that for each fixed $\theta \in \bar{\Theta}$ and $s \in \bar{S}$, $\lim_{t\to\infty} w_t(\theta, s)/t = w(\theta)$.

2.2. Variants of the Optimality Equation

If α is convex and θ^* lies inside Θ, then the minimization problem is equivalent to finding a root of

$$\alpha'(\theta) = w'(\theta) + C'(\theta). \tag{5}$$

Even if θ^* is on the boundary of Θ, the minimization problem can be solved by a descent method which, at each step, computes $\alpha'(\theta)$ at the current point θ and moves opposite to its sign. Here, we will use a *stochastic* descent method (see §2.3), which at each iteration moves in the direction of an *estimate* of $\alpha'(\theta)$.

Alternative formulae for the direction of descent can be derived using a regenerative approach as follows. Let $s = 0$ and let τ be the number of customers in the first regenerative cycle (busy period). From elementary renewal theory one has $w(\theta) = u(\theta)/l(\theta)$ where

$$u(\theta) = E_{\theta,0}\left[\sum_{i=1}^{\tau} W_i^*\right], \quad l(\theta) = E_{\theta,0}[\tau].$$

If $w'(\theta)$ exists, then, from standard calculus, one has

$$w'(\theta) = \frac{u'(\theta)l(\theta) - l'(\theta)u(\theta)}{l^2(\theta)} = \frac{u'(\theta) - w(\theta)l'(\theta)}{l(\theta)}. \quad (6)$$

One can combine estimators for each of the four quantities on the right-hand side of (6) to obtain an estimator for $w'(\theta)$. Alternatively, finding a root of (5) is the same as finding a root of

$$u'(\theta) - w(\theta)l'(\theta) + l(\theta)C'(\theta) \quad (7)$$

or of

$$u'(\theta)l(\theta) - l'(\theta)u(\theta) + l^2(\theta)C'(\theta). \quad (8)$$

So, instead of using an estimate of (5) in the descent method, one use an estimate of (7) or (8). That was first suggested by Fu (1990) for (7) and Glynn (1986) for (8). The interest of (7–8) is that *unbiased* estimators of them can be obtained based on a few regenerative cycles, which is not the case for (5). For example, an unbiased estimator of (8) is easily built from an unbiased estimator of $(l(\theta), l'(\theta), u'(\theta))$ and an *independent* unbiased estimator of $(l(\theta), u(\theta))$. Such estimators can be constructed via the LR method, based on two regenerative cycles (see §3.3). Similarly, an unbiased estimator of (7) can be constructed via IPA, based on one regenerative cycle, often in spite of the fact that the estimators of $u'(\theta)$ and $l'(\theta)$ are individually biased (see §3.4 and Heidelberger et al. 1988).

Equations (5) and (7–8) are specific to the form of our objective function (1). For a more general case, let $\alpha(\theta) = \varphi(\theta, w(\theta))$, where φ is convex and continuously differentiable. Then, as in Chong and Ramadge (1992b),

$$\alpha'(\theta) = \frac{\partial\varphi}{\partial w}(\theta, w(\theta))w'(\theta) + \frac{\partial\varphi}{\partial\theta}(\theta, w(\theta)),$$

where $\partial\varphi/\partial\theta$ and $\partial\varphi/\partial w$ denote the partial derivatives of φ w.r.t. its first and second component, respectively. With appropriate conditions on φ, our development for the DETs which aim at estimating $\alpha'(\theta)$ would go through for this more general case. Generalization to vectors of parameters is straightforward. Of course, more complicated nonconvex functions, e.g., with multiple local minima, are more difficult to deal with, as in the deterministic case.

Equations (7–8) are more dependent on the form of α than (5). For another illustration, let $\alpha(\theta) = D(\theta)w^2(\theta) + C(\theta)$, where D and C are known differentiable functions. Here,

$$\alpha'(\theta) = D'(\theta)w^2(\theta) + 2D(\theta)w(\theta)w'(\theta) + C'(\theta).$$

But in a descent algorithm, one can use instead an unbiased estimator of

$$l^3(\theta)\alpha'(\theta) = D'(\theta)u^2(\theta)l(\theta) + 2D(\theta)u(\theta)[l(\theta)u'(\theta) - u(\theta)l'(\theta)] + l^3(\theta)C'(\theta),$$

which can be obtained via LR, based on three (independent) regenerative cycles.

2.3. The Stochastic Approximation Scheme

We consider a stochastic approximation (SA) algorithm of the form

$$\theta_{n+1} := \pi_\Theta(\theta_n - \gamma_n Y_n), \quad (9)$$

for $n \geq 1$, where θ_n is the parameter value at the beginning of iteration n ($\theta_1 \in \Theta$ is fixed, or random with known distribution), Y_n is an estimator of either (5), (7), or (8), obtained at iteration n, $\{\gamma_n, n \geq 1\}$ is a (deterministic) positive sequence decreasing to 0 such that $\sum_{n=1}^{\infty} \gamma_n = \infty$, and π_Θ denotes the projection on the set Θ (i.e., $\pi_\Theta(\theta)$ is the point of Θ closest to θ). To obtain Y_n, in each case, we compute directly $C'(\theta_n)$, and estimate only the remaining terms, by simulating the system for one or more "subrun(s)" of finite duration. Specific estimators are discussed in §3.

Let $s_n \in \bar{S}$ denote the *state* of the system at the beginning of iteration n. For all the estimators that we consider, the distribution of (Y_n, s_{n+1}), conditional on (θ_n, s_n), is completely specified by n and P_{θ_n, s_n}, and is independent of the past iterations. In other words, $\{(Y_n, \theta_{n+1}, s_{n+1}), n \geq 0\}$ is a (nonhomogeneous) Markov chain (Y_0 is a dummy value). Denote by $E_{n-1}(\cdot)$ the conditional expectation $E(\cdot \mid \theta_n, s_n)$, i.e., the expectation conditional on what is known at the beginning of iteration n. Suppose that each Y_n is viewed as an estimator of (5) and is integrable. Then, $E_{n-1}(Y_n)$ exists and we can write

$$Y_n = \alpha'(\theta_n) + \beta_n + \epsilon_n \quad \text{where} \quad (10)$$

$$\beta_n = E_{n-1}[Y_n] - \alpha'(\theta_n) \quad (11)$$

represents the (conditional) bias on Y_n given (θ_n, s_n), while ϵ_n is a random sequence, with $E_{n-1}(\epsilon_n) = 0$ and $E_{n-1}(\epsilon_n^2) = \mathrm{var}(Y_n \mid \theta_n, s_n)$. If Y_n is an estimator of (7) or (8) instead, then replace $\alpha'(\theta_n)$ in (10–11) by $l(\theta_n)\alpha'(\theta_n)$ or $l^2(\theta_n)\alpha'(\theta_n)$, respectively.

2.4. Convergence to the Optimum

Proposition 1, proved in Appendix I, gives (simplified) sufficient conditions for the convergence of (9) to the optimum. It treats the case where the (conditional) bias β_n goes to zero and the variance of Y_n does not increase too fast with n. For some of the regenerative methods, one has $\beta_n = 0$ for each n. Otherwise, when the DET uses the same number of customers at all iterations, β_n typically *does not* go to zero. But sometimes, $E_0(\beta_n) \to 0$ and the algorithm might still converge to the optimum. This is treated by Theorem 1 of Appendix I, which ensures weak convergence under appropriate conditions.

PROPOSITION 1. *Suppose that α is differentiable and convex or strictly unimodal over Θ. If $\lim_{n\to\infty} \beta_n = 0$ a.s. and $\sum_{n=1}^{\infty} E_0(\epsilon_n^2)\gamma_n^2 < \infty$ a.s., then $\lim_{n\to\infty} \theta_n = \theta^*$ a.s.*

For convenience in the following sections, in the case where Y_n is a truncated-horizon estimator of (5), we will decompose β_n as $\beta_n = \beta_n^F + \beta_n^R$, where β_n^F is the bias component due to the fact that we simulate over a finite horizon, and β_n^R represents the possibility that Y_n may itself be a biased estimator of the derivative of the finite-horizon expected cost. Typically, with FD, $\beta_n^R \neq 0$. If we use a deterministic horizon t_n at iteration n, then

$$\beta_n^F = w'_{t_n}(\theta_n, s_n)/t_n - w'(\theta_n).$$

To make sure that the latter converges to zero a.s., we will show, under appropriate conditions, that $w'_t(\theta, s)/t - w'(\theta)$ converges to zero *uniformly* in (θ, s) as t goes to infinity.

2.5. Continuous Differentiability and Uniform Convergence

We want sufficient conditions under which α is convex or strictly unimodal, w and each $w_t(\cdot, s)$ are differentiable, and the following uniform convergence results hold:

$$\lim_{t\to\infty} \sup_{\theta\in\bar{\Theta}, s\in\bar{S}} |w_t(\theta, s)/t - w(\theta)| = 0 \quad \text{and} \tag{12}$$

$$\lim_{t\to\infty} \sup_{\theta\in\bar{\Theta}, s\in\bar{S}} |w'_t(\theta, s)/t - w'(\theta)| = 0. \tag{13}$$

In Proposition 11, we establish (12–13) under Assumption A below. We also prove, under Assumptions A–C, that $w_t(\theta, s)/t$ is convex and continuously differentiable in θ for each s and t, and that α is also convex and continuously differentiable. Note that these properties can be expected to hold only when appropriate regularity conditions are imposed on the service time distribution B_θ. On the other hand, the properties that are exploited here are merely *sufficient* for the validity of SA, not necessary. Assumptions A and B are used for IPA and LR derivative estimation, respectively (they are typical IPA and LR assumptions), while C is used to ensure the convexity of α. For example, an exponential service time distribution with mean θ verifies all these assumptions; see L'Ecuyer, Giroux, and Glynn (1994) for the details. Define $U_i \overset{\text{def}}{=} B_\theta(\zeta_i)$. Then, U_i is a $U(0, 1)$ random variable and

$$\zeta_i = B_\theta^{-1}(U_i) \overset{\text{def}}{=} \min\{\zeta \mid B_\theta(\zeta) \le U_i\}.$$

Define also

$$Z_i = \frac{\partial}{\partial\theta} B_\theta^{-1}(U_i).$$

ASSUMPTION A. (i) *There is a distribution $\tilde{B}$ such that $\sup_{\theta\in\bar{\Theta}^0} B_\theta^{-1}(u) \le \tilde{B}^{-1}(u)$ for each u in $(0, 1)$. The queue remains stable when the service times are generated according to $\tilde{B}$. Also, $\tilde{E}[\zeta^8] \le K_\zeta$, where $1 \le K_\zeta < \infty$ and $\tilde{E}$ is the expectation that corresponds to $\tilde{B}$.*

(ii) *For each $u \in (0, 1)$, $B_\theta^{-1}(u)$ is differentiable in θ. There exists a measurable $\Gamma : (0, 1) \mapsto \mathbb{R}$ such that*

$$\sup_{\theta\in\bar{\Theta}^0} \left| \frac{\partial}{\partial\theta} B_\theta^{-1}(u) \right| \le \Gamma(u)$$

for each u and

$$1 \le K_\Gamma \overset{\text{def}}{=} \int_0^1 (\Gamma(u))^8 du < \infty.$$

ASSUMPTION B. (i) *Assumption A* (i) *holds, and the moment generating function associated with $\tilde{B}$ is finite in some neighborhood of zero.*

(ii) *Let B_θ have a density b_θ, whose support $\{\zeta \ge 0 \mid b_\theta(\zeta) > 0\}$ is independent of θ for $\theta \in \bar{\Theta}^0$.*

(iii) *Everywhere in $\bar{\Theta}^0$, $b_\theta(\zeta)$ is continuously differentiable with respect to θ, for each $\zeta \ge 0$.*

(iv) *For each $\theta_0 \in \bar{\Theta}$ and $K > 1$, there exists $\Upsilon = (\theta_0 - \epsilon_0, \theta_0 + \epsilon_0) \subset \bar{\Theta}^0$ and $\bar{\theta} \in \bar{\Theta}^0$ such that*

$$\sup_{\theta \in \Upsilon}\left(\frac{b_\theta(\zeta)}{b_{\bar{\theta}}(\zeta)}\right) \le K \quad \text{and} \quad E_{\bar{\theta}}\left[\left(\sup_{\theta \in \Upsilon} \frac{\partial}{\partial \theta} \ln b_\theta(\zeta)\right)^8\right] < \infty.$$

ASSUMPTION C. *C is convex and continuously differentiable and for each u, $B_\theta^{-1}(u)$ is convex in θ.*

3. Ways of Estimating the Derivative

One crucial ingredient for the SA algorithm considered here is an efficient derivative estimation technique (DET). In this section, we survey some possibilities and state convergence results regarding their combination with SA. All the propositions are proved in Appendix II.

3.1. Finite Differences (FD)

This method is described, for instance, in Glynn (1989) and Kushner and Clark (1978), without the projection operator. When used in conjunction with FD, the SA algorithm (9) is known as the Kiefer-Wolfowitz (KW) algorithm. Here, we describe and use *central* (or *symmetric*) FD. For other variants, like *forward* FD, see the above references. When θ is a d-dimensional vector, the latter uses only $d + 1$ instead of $2d$ subruns per iteration. However, its asymptotic convergence rate is not as good (Glynn 1989). Spall (1992) analyzes a different FD method for SA, called "simultaneous perturbation," and shows that it could be significantly more efficient than FD when d is large.

Take a deterministic positive sequence $\{c_n, n \ge 1\}$ that converges to 0. At iteration n, simulate from some initial state $s_n^- \in \bar{S}$ at parameter value

$$\theta_n^- = \pi_{\bar{\Theta}}(\theta_n - c_n) = \max(\theta_n - c_n, \bar{\theta}_{\min})$$

for t_n customers. Simulate also (independently) from state $s_n^+ \in \bar{S}$ at parameter value

$$\theta_n^+ = \pi_{\bar{\Theta}}(\theta_n + c_n) = \min(\theta_n + c_n, \bar{\theta}_{\max})$$

for t_n customers. Let ω_n^- and ω_n^+ denote the respective sample points. A FD estimator of (5) is

$$Y_n = C'(\theta_n) + \frac{h_{t_n}(\theta_n^+, s_n^+, \omega_n^+) - h_{t_n}(\theta_n^-, s_n^-, \omega_n^-)}{(\theta_n^+ - \theta_n^-)t_n}. \quad (14)$$

The conditional bias

$$\beta_n^R = E_{n-1}[Y_n - C_n'(\theta) - w_{t_n}'(\theta_n, s_n)/t_n]$$

can itself be decomposed as $\beta_n^R = \beta_n^D + \beta_n^I$, where

$$\beta_n^D = \frac{w_{t_n}(\theta_n^+, s_n) - w_{t_n}(\theta_n^-, s_n)}{(\theta_n^+ - \theta_n^-)t_n} - \frac{w_{t_n}'(\theta_n, s_n)}{t_n} \quad \text{and}$$

$$\beta_n^I = \frac{w_{t_n}(\theta_n^+, s_n^+) - w_{t_n}(\theta_n^+, s_n) + w_{t_n}(\theta_n^-, s_n) - w_{t_n}(\theta_n^-, s_n^-)}{(\theta_n^+ - \theta_n^-)t_n} \quad (15)$$

represent respectively the bias due to finite differences and the bias due to the possibly different initial states.

PROPOSITION 2. *Let Assumptions A–B hold, $t_n \to \infty$ and $c_n \to 0$. Then $\lim_{n\to\infty} \beta_n^D \overset{\text{a.s.}}{=} 0$.*

The term β_n^I can be eliminated by picking $s_n^- = s_n^+ = s_n$. Otherwise, if $|s_n^+ - s_n^-|$ is bounded, the numerator in (15) should be in $O(1/t_n)$ (asymptotically). In that case, to get $\beta_n^I \to 0$, take $1/(t_n c_n) \to 0$. Even when $\beta_n^I = 0$, taking t_n constant may lead to problems, because β_n^F is usually *not* zero. As $c_n \to 0$, when ω_n^- and ω_n^+ are distinct ("independent"), the variance of Y_n usually increases to infinity. However, we have:

PROPOSITION 3. *Suppose one uses SA with the estimator (14). Let Assumption A–C hold, $t_n \to \infty$, $c_n \to 0$, and $\sum_{n=1}^{\infty} t_n^{-1} c_n^{-2} \gamma_n^2 < \infty$. Assume that $\beta_n^I \to 0$ a.s. as $n \to \infty$ (this can be achieved trivially by taking $s_n^- = s_n^+ = s_n$). Then, $\theta_n \to \theta^*$ a.s.*

Note that in the proofs of Propositions 2 and 3, Assumption A is used to prove (13), while C is used to prove the convexity of $w(\cdot)$, and B is used only to prove the continuous differentiability. These remarks also apply to Proposition 6.

A different approach is to estimate (8) instead of (5) using finite differences. A *forward* FD approximation of (8) at $\theta = \theta_n$, adapted from Glynn (1986), is

$$\frac{u(\theta_n^+) - u(\theta_n)}{\theta_n^+ - \theta_n} l(\theta_n) - \frac{l(\theta_n^+) - l(\theta_n)}{\theta_n^+ - \theta_n} u(\theta_n) + l(\theta_n^+)l(\theta_n)C'(\theta_n) \quad (16)$$

$$= \frac{u(\theta_n^+)l(\theta_n) - l(\theta_n^+)u(\theta_n)}{\theta_n^+ - \theta_n} + l(\theta_n^+)l(\theta_n)C'(\theta_n). \quad (17)$$

To estimate (17), simulate for $2t_n$ independent regenerative cycles using parameter value θ_n [θ_n^+] for the odd

[even] numbered cycles. Let τ_j denote the number of customers during the jth cycle and h_j denote the total system time for those τ_j customers. Then, an unbiased estimator of (17) is

$$Y_n = \frac{1}{t_n} \sum_{j=1}^{t_n} \left(\frac{h_{2j}\tau_{2j-1} - h_{2j-1}\tau_{2j}}{\theta_n^+ - \theta_n} + \tau_{2j}\tau_{2j-1}C'(\theta_n) \right). \quad (18)$$

Here, $t_n \to \infty$ is not required. For instance, one can use $t_n = 1$ for all n.

PROPOSITION 4. *Let Assumptions A–C hold, $\sum_{n=1}^{\infty} t_n^{-1} c_n^{-2} \gamma_n^2 < \infty$, and suppose one uses SA with the estimator (18). Then, $\theta_n \to \theta^*$ a.s.*

3.2. Finite Differences with Common Random Numbers (FDC)

One way to reduce the variance in (14) is to use common random numbers across the subruns at each iteration, start all the subruns from the same state: $s_n^- = s_n^+ = s_n$, and synchronize. More specifically, one views ω as representing a sequence of $U(0, 1)$ variates, so that all the dependency on (θ, s) appears in $h_t(\theta, s, \cdot)$. Take $\omega_n^+ = \omega_n^- = \omega_n$. Since the subruns are aimed at comparing very similar systems, $h_{t_n}(\theta_n^+, s_n, \omega_n)$ and $h_{t_n}(\theta_n^-, s_n, \omega_n)$ should be highly correlated, especially when c_n is small, so that considerable variance reductions should be obtained. Conditions that *guarantee* variance reductions are given in Glasserman and Yao (1992). Proposition 2 still applies. However, taking $t_n \to \infty$ is essential. In L'Ecuyer, Giroux, and Glynn (1994), we discuss implementation issues related to FDC and show that if t_n is kept constant, SA with FDC typically converges to the wrong value.

3.3. A Likelihood Ratio (LR) Approach

The LR approach (Glynn 1990, L'Ecuyer 1990, Reiman and Weiss 1989, and Rubinstein 1989) can be used as follows to estimate $w_t'(\theta, s)$. Let us view ω as representing the sequence of interarrival and service times for the first t customers, that is, $\omega = (\zeta_1, \nu_1, \ldots, \zeta_t, \nu_t)$. For any $s \in S$, to differentiate the expectation (4) with respect to θ, take a fixed $\theta_0 \in \Theta$ and rewrite:

$$w_t(\theta, s) = \int_\Omega h_t(\theta, s, \omega) L_t(\theta_0, \theta, s, \omega) dP_{\theta_0,s}(\omega) \quad (19)$$

where

$$L_t(\theta_0, \theta, s, \omega) = \prod_{i=1}^{t} \frac{b_\theta(\zeta_i)}{b_{\theta_0}(\zeta_i)}$$

is a likelihood ratio. Under appropriate *regularity conditions* (see L'Ecuyer 1994), one can differentiate w_t by differentiating inside the integral:

$$w_t'(\theta, s) = \int_\Omega \psi_t(\theta, s, \omega) dP_{\theta_0,s}(\omega) \quad \text{where}$$

$$\begin{aligned}\psi_t(\theta, s, \omega) &= h_t(\theta, s, \omega) L_t'(\theta_0, \theta, s, \omega) \\ &= h_t(\theta, s, \omega) L_t(\theta_0, \theta, s, \omega) S_t(\theta, s, \omega) \quad (20)\end{aligned}$$

is the LR estimator,

$$S_t(\theta, s, \omega) = \frac{L_t'(\theta_0, \theta, s, \omega)}{L_t(\theta_0, \theta, s, \omega)} = \sum_{i=1}^{t} d_i \quad (21)$$

is called the *score function*, and

$$d_i = \frac{\partial}{\partial\theta} \ln b_\theta(\zeta_i).$$

Only *one* simulation experiment (using $P_{\theta_0,s}$) is required to estimate the derivative. From Proposition 14, under Assumption B, (20) is an unbiased estimator of $w_t'(\theta, s)$ for θ in some neighborhood of θ_0. After adding the derivative of the deterministic part and taking $\theta_0 = \theta_n$, the LR derivative estimator at iteration n of SA becomes

$$\begin{aligned}Y_n &= C'(\theta_n) + \psi_{t_n}(\theta_n, s_n, \omega_n)/t_n \\ &= C'(\theta_n) + h_{t_n}(\theta_n, s_n, \omega_n) S_{t_n}(\theta_n, s_n, \omega_n)/t_n. \quad (22)\end{aligned}$$

Note that the variance of $S_t(\theta, s, \omega)$ (and of (20) at $\theta = \theta_0$) increases with t. This is a significant drawback and must be taken into account when making the tradeoff between bias and variance. Here, $\beta_n^R = 0$ and $\beta_n^F \to 0$ as $t_n \to 0$. But the variance on Y_n then goes to infinity also. One remedy, as in FD, is to increase t_n more slowly. We show in Proposition 17 that under Assumption B, the variance of Y_n does not increase faster than linearly in t_n. The conditions of Proposition 1 can then be verified with $\gamma_n = \gamma_0 n^{-1}$ and $t_n = t_a + t_b n^p$ for $0 < p < 1$, where t_a and $t_b > 0$ are two constants. In the finite-horizon case, SA with LR converges at a rate of $t^{-1/2}$ (Glynn 1989) in terms of the total simulation length t. But when the variance increases with t_n and t_n increases with n, this is no longer true.

One can circumvent the bias/variance problem of LR by exploiting the regenerative structure (Glynn 1986 and 1990, Reiman and Weiss 1989). One approach is to combine estimators for each of the four quantities on the right-hand side of (6) to construct an estimator for $w'(\theta)$. One can estimate $l(\theta)$ and $u(\theta)$ as usual, and $u'(\theta)$ and $l'(\theta)$ by:

$$\psi_u(\theta, \omega) = \left(\sum_{i=1}^{\tau} W_i^*\right) S_\tau(\theta, 0, \omega) \quad \text{and} \tag{23}$$

$$\psi_l(\theta, \omega) = \tau S_\tau(\theta, 0, \omega), \tag{24}$$

respectively. From Proposition 15, under Assumption B, these estimators are unbiased. Suppose one simulates for r regenerative cycles. Let τ_j be the number of departures during the jth cycle, h_j the total system time for those τ_j customers in cycle j, and S_j the score function associated with that cycle. Then, an estimator of $w'(\theta)$ is given by

$$\psi_w(r, \theta, \omega) = \frac{\sum_{j=1}^r \tau_j \sum_{j=1}^r h_j S_j - \sum_{j=1}^r h_j \sum_{j=1}^r \tau_j S_j}{(\sum_{j=1}^r \tau_j)^2}. \tag{25}$$

This estimator is biased for finite r. However, we show in Proposition 18 that under Assumption B, as $r \to \infty$, (25) has bounded variance and converges in expectation to $w'(\theta)$, uniformly with respect to θ. The corresponding estimator of $\alpha'(\theta_n)$ for iteration n, based on t_n regenerative cycles, is:

$$Y_n = C'(\theta_n) + \psi_w(t_n, \theta_n, \omega_n). \tag{26}$$

Now, instead of trying to estimate $\alpha'(\theta_n)$ at each iteration, one can estimate (8). Since (23–24) provide unbiased estimators of $u'(\theta)$ and $l'(\theta)$, an unbiased estimator of (8) can be obtained from two independent regenerative cycles as described in the text that follows (8). One can also use more than two cycles and average out. Further, estimators of all quantities can be computed from each cycle and combined in a splitting scheme. Take $2t_n$ cycles at iteration n, and let τ_j, h_j, and S_j be defined as for (25). Then, an unbiased estimator of (8) at $\theta = \theta_n$ is

$$Y_n = \frac{1}{t_n} \sum_{j=1}^{t_n} \left(\frac{1}{2} (h_{2j} S_{2j} \tau_{2j-1} + h_{2j-1} S_{2j-1} \tau_{2j} - h_{2j-1} S_{2j} \tau_{2j} - h_{2j} S_{2j-1} \tau_{2j-1}) + \tau_{2j} \tau_{2j-1} C'(\theta_n) \right). \tag{27}$$

Since this estimator is unbiased, t_n can be taken constant in n (e.g. $t_n = 1$ for all n).

The estimators (22), (26), and (27) can be integrated into the SA algorithm. The following proposition tells us about the a.s. convergence of such a scheme.

PROPOSITION 5. *Let Assumptions A–C hold and* $\sum_{n=1}^{\infty} \gamma_n^2 < \infty$.

(a) *Suppose one uses SA with the LR estimator* (22). *If* $s_n \in \bar{S}$ *for all* n, $t_n \to \infty$, *and* $\sum_{n=1}^{\infty} t_n \gamma_n^2 < \infty$, *then* $\theta_n \to \theta^*$ *a.s.*

(b) *If one uses SA with the regenerative LR estimator* (26) *and* $t_n \to \infty$, *then* $\theta_n \to \theta^*$ *a.s.*

(c) *If one uses SA with the estimator* (27), *then* $\theta_n \to \theta^*$ *a.s.*

3.4. Infinitesimal Perturbation Analysis (IPA)

The basic idea of IPA, applied to our context, is to estimate $w_t'(\theta, s)$ by the *sample derivative*

$$h_t'(\theta, s, \omega) = \sum_{i=1}^{t} \delta_i, \tag{28}$$

where ω is interpreted as the sequence $U_1, U_2, \ldots,$ defined before Assumption A,

$$\delta_i = \sum_{j=v_i}^{i} \frac{\partial B_\theta^{-1}(U_j)}{\partial \theta} = \sum_{j=v_i}^{i} Z_j, \tag{29}$$

and v_i is the first customer with index ≥ 1 in the busy period to which customer i belongs. That is, $v_i = \max\{j \mid 1 \leq j \leq i \text{ and } W_j = 0\}$ if that set is nonempty, $v_i = 1$ otherwise. For further details and justifications, see Glasserman (1991) or Suri (1989). Then,

$$Y_n = C'(\theta_n) + h_t'(\theta_n, s_n, \omega)/t_n, \tag{30}$$

which can be computed easily during the simulation. The sum (29) is called the *IPA accumulator*. Observe that imposing $v_i \geq 1$ means that we consider only the service time perturbations of the customers who left during the current SA iteration. In other words, (28) assumes that the IPA accumulator is reset to zero between iterations. The initial state s_n of iteration n can be either 0 for all n (always restart from an empty system), or the value of $(W_t^* - \nu_t)^+$ from the previous iteration (for $n > 1$).

We can consider another variant of IPA in which the IPA accumulator is *not* reset to zero between iterations. In that case, both s_n and the initial value a_n of the IPA accumulator are taken from the previous iteration. The value of a_n is the value of δ_t from the previous iteration if $s_n > 0$, and is 0 otherwise. It must be considered as part of the "state." For this IPA variant, (28) must be modified to:

$$h_t'(\theta, s, a, \omega) = ak_t^* + \sum_{i=1}^{t} \delta_i, \tag{31}$$

where a is the initial value of the IPA accumulator and

$$k_t^* = \min(t, \min\{i \geq 0 \mid W_{i+1} = 0\}) \tag{32}$$

represents the number of customers in the current iteration who are in the same busy period as the last customer of the previous iteration, when $W_1 = s > 0$.

As for LR, we can also construct regenerative IPA estimators. With $s = 0$, the value of (28) for the first cycle is

$$h_\tau'(\theta, 0, \omega) = \sum_{i=1}^{\tau} \sum_{j=1}^{i} Z_j. \tag{33}$$

With r cycles, let τ_j and h_j' denote the respective values of τ and $h_\tau'(\theta, 0, \omega)$ for the jth cycle. An estimator of $w'(\theta)$ is then

$$\frac{\sum_{j=1}^{r} h_j'}{\sum_{j=1}^{r} \tau_j}. \tag{34}$$

At iteration n, take $r = t_n$ regenerative cycles and let $T_n = \sum_{j=1}^{t_n} \tau_j$. This yields

$$Y_n = C'(\theta_n) + \frac{\sum_{j=1}^{t_n} h_j'}{\sum_{j=1}^{t_n} \tau_j} = C'(\theta_n) + \frac{1}{T_n} \sum_{i=1}^{T_n} \delta_i. \tag{35}$$

Unfortunately, for finite t_n, this estimator is biased for $\alpha'(\theta_n)$. To better exploit the regenerative structure, one can estimate (7) instead of (5), using IPA. This was suggested in Fu (1990). From Proposition 11 and (6),

$$E_\theta[h_\tau'(\theta, 0, \omega)] = l(\theta)w'(\theta) = u'(\theta) - w(\theta)l'(\theta),$$

so that IPA provides an unbiased estimator of (7) based on a *single* regenerative cycle. Using t_n cycles at iteration n and averaging out yields the following estimator:

$$Y_n = \frac{1}{t_n} \sum_{j=1}^{t_n} (h_j' + \tau_j C'(\theta_n)). \tag{36}$$

As pointed out by Fu (1990), proving a.s. convergence of SA to the optimizer is relatively easy with (36) because it is unbiased for (7) at $\theta = \theta_n$ for any t_n. With (30) and (35), it is more difficult.

Heidelberger et al. (1988) argue that (28) divided by t is a consistent estimator of $w'(\theta)$ for a rather general class of $GI/G/1$ queues and give a proof for the $M/G/1$ case. To prove convergence of SA using Proposition 1, what we need is not convergence of (28) divided by t to $w'(\theta)$ a.s. (as $t \to \infty$), but convergence in expectation, uniformly over θ. In fact, *both* kinds of convergence, as well as variance boundedness, follow from Propositions 9–11. Proposition 9 also shows that the IPA estimator (28) is unbiased for $w_t'(\theta, s)$ under Assumption A. This leads to:

PROPOSITION 6. *Let Assumptions A and C hold, and* $\sum_{n=1}^{\infty} \gamma_n^2 < \infty$.

(a) *Suppose one uses SA with the IPA estimator* (30). *If* $s_n \in \bar{S}$ *and* $a_n = 0$ *for all* n, *and* $t_n \to \infty$, *then* $\theta_n \to \theta^*$ *a.s.*

(b) *If one uses SA with the estimator* (35), *with* $t_n \to \infty$, *then* $\theta_n \to \theta^*$ *a.s.*

(c) *If one uses SA with the estimator* (36), *then* $\theta_n \to \theta^*$ *a.s.*

If a_n is not reset to 0 between iterations, proving Proposition 6(a) appears more difficult, but we believe that the result still holds. Proposition 6(c) corresponds to the result of Fu (1990).

For this $GI/G/1$ example, IPA has the stronger property that even when using a truncated horizon t_n that is constant with n, if the IPA accumulator a_n is *not* reset between iterations and under mild additional assumptions, SA converges to the optimizer. On the other hand, if a_n is reset to zero at the beginning of each iteration, then we have the same problem as with FDC. By keeping the value of a_n across iterations, the estimator takes into account the perturbations on the service times of the customers who left during previous iterations. It is true that the structure of the busy periods, and (in general) the individual terms of the sum (28), could depend on θ, which changes between iterations. But as θ_n converges to some value, that change becomes negligible under appropriate continuity assumptions. (In the present $GI/G/1$ context, the Z_j's are in fact *independent* of θ, but not the v_i's.) With this intuitive reasoning, we should expect SA-IPA to converge to θ^* for whatever t_n. Proposition 7 states that this is effectively true under a few (sufficient) conditions. Here, we cannot use Proposition 1 because $\beta_n \not\to 0$. Instead, we give a *weak convergence* proof by verifying the assumptions of the main theorem of Kushner and Shwartz (1984). Using different proof techniques, Chong and Ramadge (1992a, 1993) have shown a.s. convergence (which is stronger) for SA-IPA with constant truncated horizon. On the other hand, with the regenerative IPA estimator (35), SA does not converge to θ^* in general if $t_n \not\to \infty$.

PROPOSITION 7. *Consider the SA algorithm with IPA, under Assumptions A–C, with $\{\gamma_n, n \geq 0\}$ satisfying* W4 *of Appendix I, and constant truncated horizon $t_n = t$. Let the interarrival time distribution have a bounded density. Suppose that Z_i can be expressed as $Z_i = \varphi(\theta, \zeta_i)$, where $\varphi : \bar{\Theta} \times \mathbb{R}^+ \to \mathbb{R}$ is a function such that $\varphi(\cdot, \zeta)$ is continuous for each $\zeta \geq 0$ (and is not expressed as a function of U_i). Suppose also that the IPA accumulator is not reset to 0 between iterations. Then, θ_n converges in probability to the optimum θ^*.*

4. Conclusion

Through a simple example, we have seen how a derivative estimation technique, such as FD, IPA, or LR, can be incorporated into a SA algorithm to get a provably convergent stochastic optimization method. In the companion paper (L'Ecuyer, Giroux, and Glynn 1994), we report numerical investigations and point out some dangers associated with different kinds of bias. The performance of these algorithms when there are many parameters to optimize, the incorporation of proper variance reduction techniques, the study of convergence rates, and comparisons between SA and the stochastic counterpart approach (Rubinstein and Shapiro 1993) are other interesting subjects for further investigation. In principle, IPA and LR can be used to estimate higher-order derivatives, but the variance is likely to be high. Is it too high to permit the implementation of good second-order algorithms based on these estimates? Again, further investigation is needed.

The convergence results of §3 can be extended to more general models than the $GI/G/1$ queue. Consider for example a general discrete-time Markov chain model parameterized by θ. Let $w_t(\theta, s)/t$ be the expected average cost per step for the first t steps, if the initial state is s (we have removed $C(\theta)$). Suppose that (12–13) hold (which implies that the derivative exists), that $w(\theta)$ is strictly unimodal, and that an unbiased LR or IPA derivative estimator for $w_t'(\theta, s)$ is available. If the variance of the LR estimator is in $O(t)$, then Proposition 5(a) applies, while if the variance of the IPA estimator is in $O(1/t)$, then Proposition 6(a) applies. Further, if the system is regenerative, and if unbiased LR estimators are available for $l'(\theta)$ and $u'(\theta)$, then one can construct estimators for $w'(\theta_n)$ and $l^2(\theta_n)w'(\theta_n)$ as in (26) and (27). If those estimators have bounded variances and converge in expectation uniformly in θ, as $t_n \to \infty$, then Proposition 5(b–c) applies. If a FD or FDC estimator is used and if $w(\cdot)$ and $w_t(\cdot, s)$ are continuously differentiable (for each s), then Proposition 3 applies. All this generalizes straightforwardly to the case where θ is a vector of parameters. Derivatives are then replaced by gradients.[1]

[1] The work of the first author was supported by NSERC-Canada grant no. OGP0110050 and FCAR-Québec grant no. EQ2831. The second author's work was supported by the IBM Corporation under SUR-SST Contract 12480042 and by the U.S. Army Research Office under Contract DAAL-03-91-G-0101. Part of the work was done while the first author was visiting the Operations Research Department at Stanford University in the fall of 1988. We also wish to thank S. Asmussen, N. Giroux, Y. T. Leung, G. Perron, and R. Suri for helpful hints and suggestions.

Appendix I. Sufficient Convergence Conditions

In this appendix, we prove Proposition 1 and give a second set of sufficient conditions, which imply *weak* convergence of the SA algorithm (9) to θ^*.

PROOF OF PROPOSITION 1. For each n, the sequence $\{\sum_{i=n}^{j} \gamma_i \epsilon_i, j \geq 1\}$ is a martingale. For each $\epsilon > 0$, from Doob's inequality, we have

$$P\left(\sup_{j\geq n} \left\| \sum_{i=n}^{j} \gamma_i \epsilon_i \right\| \geq \epsilon\right) \leq \frac{K}{\epsilon^2} \sum_{i=n}^{\infty} \gamma_i^2 E_0[\epsilon_i^2]$$

for some constant K. This upper bound goes to zero as $n \to \infty$. Hence, we obtain condition A2.2.4″ of Kushner and Clark (1978), and the result then follows from their Theorem 5.3.1.

Often, $\beta \not\to 0$, but $E_0[\beta_n] \to 0$ as $n \to \infty$, and the algorithm converges as well to the optimum. This is addressed by the following (weaker) result, which follows from the results of Kushner and Shwartz (1984) and by adapting the proof of the second part of Theorem 4.2.1 in Kushner and Clark (1978) (note that in the last paragraph of the latter proof, the max should be replaced by a min). We now restate the assumptions of Kushner and Shwartz (1984), with slight adaptations. See the latter reference for further details.

W1. Denote $\xi_n = (Y_n, s_{n+1}) \in \mathbb{R} \times S$. Assume that $P_{\theta,s}$ is weakly continuous in (θ, s), in the sense that $P_{\theta,s} \Rightarrow P_{\theta_0,s_0}$ when $(\theta, s) \to (\theta_0, s_0)$, and that $E[Y_{n+c} | (\theta_n, s_n) = (\theta, s)]$ is continuous in (θ, s) for some integer $c \geq 0$. Assume that for each fixed $\theta \in \Theta$, i.e., if $\gamma_n = 0$ for all

n, $\{\xi_n, n \geq 1\}$ is a Markov process with unique invariant measure P^θ and corresponding mathematical expectation E^θ. Denote $v(\theta) = E^\theta(Y_n)$. Let $\{P^\theta, \theta \in \Theta\}$ and $\{\xi_n, n \geq 1\}$ be tight (the latter uniformly over θ and s; see Kushner and Shwartz 1984).

W2. For each compact $C \subset \mathbb{R} \times S$, there is an integer $n_C < \infty$ such that for each $T > 0$, the set of probability measures $\{P[(\theta_{n+j}, \xi_{n+j-1}) \in \cdot \mid \theta_n = \theta, \xi_n = \xi], \theta \in \Theta, \xi \in C, n \geq 1, j \geq n_C, \sum_{i=n+1}^{n+j} \gamma_i \leq T, C$ compact subset of $S\}$ is tight.

W3. For some constant $\kappa > 0$, $\sup_{n\geq 1} E_0[|Y_n|^{1+\kappa}] < \infty$.

W4. $\gamma_n > 0$ for all n, $\lim_{n\to\infty} \gamma_n = 0$, $\sum_{n=1}^{\infty} \gamma_n = \infty$ and $\sum_{n=1}^{\infty} |\gamma_{n+1} - \gamma_n| < \infty$.

W5. The function v is nondecreasing in Θ and has a unique root at $\theta^* \in \Theta$.

THEOREM 1. *Under W1 – W5, $\theta_n \to \theta^*$ in probability, i.e., for each $\epsilon > 0$, $\lim_{n\to\infty} P(\|\theta_n - \theta^*\| \geq \epsilon) = 0$. Also, $v(\theta)$ is continuous in θ.*

Appendix II. Convergence Proofs

In this appendix we prove that under our assumptions LR and IPA provide unbiased estimators for $w_t'(\theta, s)$. We obtain variance bounds for these derivative estimators and for their regenerative counterparts, which are asymptotically unbiased and converge in quadratic mean, uniformly in θ. We also show that $w_t(\cdot, s)$ and $w(\cdot)$ are continuously differentiable and that (12)–(13) hold. We then prove Propositions 3 to 7.

PROPOSITION 8. *Let $s \in \bar{S}$ be the initial state and τ be the number of customers served before the system empties out for the first time. Let $h_t'(\theta, s, \omega)$ be defined as in (28). Under A (i), there exists finite constants K_ζ, K_τ, and K_h, all ≥ 1, such that*

$$\sup_{\theta\in\bar{\Theta}^0} E_\theta[\zeta^8] \leq K_\zeta; \tag{37}$$

$$\sup_{s\in\bar{S},\theta\in\bar{\Theta}^0} E_{\theta,s}[\tau^8] \leq K_\tau; \tag{38}$$

$$\sup_{s\in\bar{S},\theta\in\bar{\Theta}^0} E_{\theta,s}[(h_\tau(\theta, s, \omega))^4] \leq K_h. \tag{39}$$

Under A (i–ii), there also exists a finite constant $K_h' \geq 1$ such that

$$\sup_{s\in\bar{S},\theta\in\bar{\Theta}^0} E_{\theta,s}[(h_\tau'(\theta, s, \omega))^2] \leq K_h'. \tag{40}$$

PROOF. From A (i), $E_\theta[\zeta^8] \leq \tilde{E}[\zeta^8] \stackrel{\text{def}}{=} K_\zeta < \infty$, which gives (37). From Theorem III.3.1 (i) in Gut (1988, p. 78), one has

$$E[\tau^8 \mid s = c] \stackrel{\text{def}}{=} K_\tau < \infty.$$

Now, let $\tilde{\zeta}_i = \tilde{B}^{-1}(U_i)$ and $\tilde{\tau}$ be the respective values of ζ_i and τ obtained if B_θ is replaced by $\tilde{B}$ while $\omega = (U_1, U_2, \cdots)$ remains the same. One has $\tilde{\zeta}_i \geq \zeta_i$. But increasing any service time or increasing s cannot decrease the number of customers in the first busy cycle. Therefore, τ is stochastically dominated by $\tilde{\tau}$, which is itself stochastically nondecreasing in s. From basic stochastic ordering principles (Wolff 1989), this implies (38).

For each i, one has $W_i \leq s + \sum_{j=1}^{i-1} \zeta_j$, so that for each $t \geq 1$,

$$h_t(\theta, s, \omega) = \sum_{i=1}^{t} (W_i + \zeta_i) \leq t\left(s + \sum_{i=1}^{t} \zeta_i\right).$$

Recall that if k is any (integer) power of two, then

$$(x + y)^k \leq 2^{k-1}(x^k + y^k)$$

(easy to check by induction on k). Therefore,

$$(h_t(\theta, s, \omega))^k \leq t^k\left(s + \sum_{i=1}^{t} \zeta_i\right)^k \leq 2^{k-1}t^k\left(s^k + \left(\sum_{i=1}^{t} \zeta_i\right)^k\right). \tag{41}$$

This holds in particular for $t = \tau$ and $k = 4$. Then, from Theorem I.5.2 in Gut (1988, page 22) (used in the third inequality), there is a constant $K_1 < \infty$, independent of θ and s, such that

$$\begin{aligned}
&E_{\theta,s}[(h_\tau(\theta, s, \omega))^4] \\
&\leq 8s^4 E_{\theta,s}[\tau^4] + 8E_{\theta,s}\left[\left(\tau \sum_{i=1}^{\tau} \zeta_i\right)^4\right] \\
&\leq 8s^4 E_{\theta,s}[\tau^4] + 8E_{\theta,s}[\tau^8] + 8E_{\theta,s}\left[\left(\sum_{i=1}^{\tau} \zeta_i\right)^8\right] \\
&\leq 8s^4 E_{\theta,s}[\tau^4] + 8E_{\theta,s}[\tau^8] + 8K_1(E_{\theta,s}[\tau^8]E_{\theta,s}[\zeta^8])^{1/2} \\
&\leq 8(c^4K_\tau + K_\tau + K_1K_\tau K_\zeta) \stackrel{\text{def}}{=} K_h < \infty.
\end{aligned}$$

From (28), for any $k \geq 1$, one has

$$(h_t'(\theta, s, \omega))^k \leq \left(t \sum_{i=1}^{t} Z_i\right)^k. \tag{42}$$

Again, since this holds for $t = \tau$ and $k = 2$, and using Theorem I.5.2 of Gut (1988), there is a constant $K_2 < \infty$ such that

$$\begin{aligned}
E_{\theta,s}[(h_\tau'(\theta, s, \omega))^2] &\leq E_{\theta,s}\left[\left(\tau \sum_{i=1}^{\tau} Z_i\right)^2\right] \\
&\leq E_{\theta,s}[\tau^4] + K_2(E_{\theta,s}[\tau^4]E_\theta[Z_1^4])^{1/2} \\
&\leq K_\tau + K_2K_\tau K_\Gamma \stackrel{\text{def}}{=} K_h' < \infty. \quad \square
\end{aligned}$$

PROPOSITION 9. *Under Assumption A, for each s, $w_t(\cdot, s)$ is differentiable, and for each $(\theta, s) \in \bar{\Theta} \times \bar{S}$, (28) is an unbiased estimator of $w_t'(\theta, s)$.*

PROOF. For fixed ω, $h_t(\theta, s, \omega)$ is continuous in each ζ_j, and therefore continuous in θ from A (ii). It is also differentiable in θ everywhere except when two events (arrival or departure) occur simultaneously, which happens at most for a denumerable number of values of θ for almost any (fixed) ω. Also, for any fixed θ, this happens with probability zero. Using (42) and (37), one obtains, for all $\theta \in \bar{\Theta}^0$,

$$E_{\theta,s}[(h_t'(\theta, s, \omega))^2] \leq t^2 E_\theta\left[\left(\sum_{j=1}^{t} Z_j\right)^2\right] \leq t^4 E_\theta[Z_1^2] \leq t^4 K_\Gamma.$$

Since every $\theta \in \bar{\Theta}$ has a neighborhood contained in $\bar{\Theta}^0$, the conclusion then follows from Theorem 1 in L'Ecuyer (1990). $\square$

We now show that the mean-square error of $h_t(\theta, s, \omega)/t$ as an estimator of $w(\theta)$, as well as the mean-square error of $h_t'(\theta, s, \omega)/t$ as an estimator of $w'(\theta)$, are in the order of $1/t$, uniformly in $(\theta, s) \in \bar{\Theta} \times \bar{S}$. As a consequence, the variance and squared bias of these estimators also decrease uniformly, linearly in t. The uniform convergence properties (12)–(13) will follow from that.

PROPOSITION 10. *Under Assumption A (i), there is a finite constant K_e such that for all $t \geq 1$,*

$$\sup_{s \in \bar{S}, \theta \in \bar{\Theta}} E_{\theta,s}\left[\left(\frac{h_t(\theta, s, \omega)}{t} - w(\theta)\right)^2\right] \leq \frac{K_e}{t}. \tag{43}$$

Under Assumption A, $w(\theta)$ is differentiable everywhere in $\bar{\Theta}$, and there is a finite constant K'_e such that for all $t \geq 1$,

$$\sup_{s \in \bar{S}, \theta \in \bar{\Theta}} E_{\theta,s}\left[\left(\frac{h'_t(\theta, s\ \omega)}{t} - w'(\theta)\right)^2\right] \leq \frac{K'_e}{t}. \tag{44}$$

PROOF. We adopt the convention that the jth busy cycle ends when the system empties out for the jth time. When $s \neq 0$, the first busy cycle does not obey the same probability law than the others, but all the busy cycles are nevertheless independent. For $j \geq 1$, let τ_j be the number of customers in the jth busy cycle, h_j the total sojourn time of those τ_j customers, and $A_j = h_j - w(\theta)\tau_j$. For $j \geq 2$, one has $E_\theta[A_j] = 0$. From Proposition 8, $w^2(\theta) \leq E_\theta[h_2^2] \leq K_h$ and

$$E_{\theta,s}[A_j^2] \leq E_{\theta,s}[h_j^2 + K_h\tau_j^2] \leq K_h(1 + K_\tau)$$

for all $j \geq 1$. Let

$$M(t) = \sup\left\{i \geq 0 \,\middle|\, \sum_{j=1}^{i} \tau_j \leq t\right\} \quad \text{and} \quad \Lambda(t) = \sum_{j=1}^{M(t)+1} \tau_j.$$

Since the A_j's are independent and have zero expectation (for the fourth inequality), applying Wald's equation (for the fifth inequality), and observing that $M(t) \leq t$ (for the sixth inequality), one obtains:

$$\begin{aligned}
&E_{\theta,s}\left[\left(\frac{h_t(\theta, s, \omega)}{t} - w(\theta)\right)^2\right] \\
&= E_{\theta,s}\left[\left(\frac{1}{t}\sum_{i=1}^{t}(W_i^* - w(\theta))\right)^2\right] \\
&\leq \frac{1}{t^2} E_{\theta,s}\left[\left(\left(\sum_{j=1}^{M(t)+1} A_j\right) - \sum_{i=t+1}^{\Lambda(t)}(W_i^* - w(\theta))\right)^2\right] \\
&\leq \frac{2}{t^2} E_{\theta,s}\left[\left(\sum_{j=1}^{M(t)+1} A_j\right)^2 + (h_{M(t)+1})^2 + (w(\theta)\tau_{M(t)+1})^2\right] \\
&\leq \frac{2}{t^2} E_{\theta,s}\left[\sum_{i,j=1}^{M(t)+1} A_iA_j + \sum_{j=1}^{M(t)+1}(h_j^2 + w^2(\theta)\tau_j^2)\right] \\
&\leq \frac{2}{t^2} E_{\theta,s}\left[\sum_{j=1}^{M(t)+1}(A_j^2 + h_j^2 + w^2(\theta)\tau_j^2)\right] \\
&\leq \frac{2}{t^2}[E_{\theta,s}[A_1^2 + h_1^2 + w^2(\theta)\tau_1^2] \\
&\quad + E_{\theta,0}[M(t)]E_{\theta,0}[A_2^2 + h_2^2 + w^2(\theta)\tau_2^2]] \\
&\leq \frac{2}{t^2}(1 + t)(2K_h(1 + K_\tau)) \leq \frac{8K_h(1 + K_\tau)}{t},
\end{aligned}$$

which proves (43), with $K_e = 8K_h(1 + K_\tau)$.

To prove (44), let h'_j denote the sum of the δ_i's associated with the τ_j customers of the jth busy cycle, $\tilde{w}'(\theta) = E_\theta[h'_2]/E_\theta[\tau_2]$, and $A'_j = h'_j - \tilde{w}'(\theta)\tau_j$. One has $E_\theta[A'_j] = 0$ for $j \geq 2$ and, from Proposition 8,

$$(\tilde{w}'(\theta))^2 \leq E_\theta[(h'_2)^2] \leq K'_h$$

(since $\tau \geq 1$) and

$$E_{\theta,s}[(A'_j)^2] \leq E_{\theta,s}[(h'_j)^2 + K'_h\tau_j^2] \leq K'_h(1 + K_\tau)$$

for all $j \geq 1$. Then, from the same reasoning as above, with W_i^* replaced by δ_i,

$$E_{\theta,s}\left[\left(\frac{h'_t(\theta, s, \omega)}{t} - \tilde{w}'(\theta)\right)^2\right] \leq \frac{8K'_h(1 + K_\tau)}{t} \stackrel{\text{def}}{=} \frac{K'_e}{t}.$$

It remains to show that $\tilde{w}'(\theta) = w'(\theta)$. Using Proposition 9 for the first equality and the expected-value version of the renewal-reward Theorem (Wolff 1989) for the second one, one has

$$\lim_{t\to\infty} \frac{w'_t(\theta, s)}{t} = \lim_{t\to\infty} \frac{E_\theta[h'_t(\theta, s, \omega)]}{t} = \tilde{w}'(\theta).$$

Furthermore,

$$\begin{aligned}
\left(\frac{w'_t(\theta, s)}{t}\right)^2 &= \left(\frac{E_{\theta,s}[h'_t(\theta, s, \omega)]}{t}\right)^2 \leq E_{\theta,s}\left[\left(\frac{h'_t(\theta, s, \omega)}{t}\right)^2\right] \\
&\leq \frac{K'_e}{t} + (\tilde{w}'(\theta))^2 \leq \frac{K'_e}{t} + K'_h.
\end{aligned}$$

Then, since $\bar{\Theta}$ is compact, $w'_t(\cdot, s)/t$ is integrable over $\bar{\Theta}$. From the Fundamental Theorem of Calculus (e.g., Theorem 8.21 in Rudin 1974), one has

$$\frac{w_t(\theta, s)}{t} = \frac{w_t(\bar{\theta}_{\min}, s)}{t} + \int_{\bar{\theta}_{\min}}^{\theta} \frac{w'_t(\xi, s)}{t}\, d\xi.$$

Taking the limit as $t \to \infty$, using the dominated convergence theorem to interchange the limit with the integral, and then differentiating yields

$$w'(\theta) = \lim_{t\to\infty} w'_t(\theta, s)/t = \tilde{w}'(\theta). \qquad \square$$

PROPOSITION 11. *Under Assumption A, (12–13) hold.*

PROOF. Equation (12) follows easily from (43) and the fact that

$$\begin{aligned}
\left(\frac{w_t(\theta, s)}{t} - w(\theta)\right)^2 &= \left[E_{\theta,s}\left(\frac{h_t(\theta, s, \omega)}{t} - w(\theta)\right)\right]^2 \\
&\leq E_{\theta,s}\left[\left(\frac{h_t(\theta, s, \omega)}{t} - w(\theta)\right)^2\right] \leq \frac{K_e}{t}.
\end{aligned}$$

The proof of (13) is similar. $\square$

PROPOSITION 12. *For a given regenerative cycle, let τ be the number of customers in that cycle and $h'_\tau(\theta, 0, \omega)$ as in (33). Under Assumption A, one has*

$$w'(\theta) = \frac{E_\theta[h'_\tau(\theta, 0, \omega)]}{E_\theta[\tau]}. \tag{45}$$

PROOF. This has been shown in the proof of Proposition 10. $\square$

PROPOSITION 13. *Under Assumption A, as $r \to \infty$, (34) has bounded variance and converges in quadratic mean to $w'(\theta)$, uniformly with respect to θ, for $\theta \in \bar{\Theta}$.*

PROOF. Define $A_j' = h_j' - w'(\theta)\tau_j$. As seen in the proof of Proposition 10, $E_\theta[A_j'] = 0$, $E_\theta[(A_j')^2] \le K_h'(1 + K_\tau)$, and $E_\theta[A_j'A_i'] = 0$ for $j \ne i$. Then,

$$E_\theta\left[\frac{\sum_{j=1}^r h_j'}{\sum_{j=1}^r \tau_j} - w'(\theta)\right]^2 = E_\theta\left[\frac{\sum_{j=1}^r (h_j' - w'(\theta)\tau_j)}{\sum_{j=1}^r \tau_j}\right]^2$$

$$\le E_\theta\left[\left(\frac{1}{r}\sum_{j=1}^r A_j'\right)^2\right] \le \frac{1}{r} E_\theta[(A_j')^2] \le \frac{K_h'(1+K_\tau)}{r}.$$

As r goes to infinity, this converges to zero uniformly in θ. This also provides a uniform upper bound on the variance of (34). □

PROPOSITION 14. *Consider the truncated horizon LR derivative estimator (20), under Assumption B. Then, for each $\theta_0 \in \bar{\Theta}$, there is a neighborhood Υ of θ_0 such that for all $\theta \in \Upsilon$ and $s \in S$, (20) is an unbiased estimator of $w_t'(\theta, s)$ with finite variance. Further, each $w_t(\cdot, s)$ is continuously differentiable over $\bar{\Theta}$.*

PROOF. Here, for fixed ω, $h_t(\theta, s, \omega)$ does not depend on θ. Since each $b_\theta(\zeta_i)$ is assumed differentiable in θ, A2 (a) in L'Ecuyer (1994) is satisfied. Let $K > 1$, $\epsilon_0 > 0$, and $\bar{\theta}$ satisfy B (iv). Then, using (41),

$$E_{\bar\theta}\left[\sup_{|\theta-\theta_0|<\epsilon_0}\left(\frac{\frac{\partial}{\partial\theta}b_\theta(\zeta)}{b_{\bar\theta}(\zeta)}\right)^2 + \sup_{|\theta-\theta_0|<\epsilon_0}\left(\frac{b_\theta(\zeta)}{b_{\bar\theta}(\zeta)}\right)^2 + (h_t(\theta,s,\omega))^2\right]$$

$$\le E_{\bar\theta}\left[\sup_{|\theta-\theta_0|<\epsilon_0}\left(\frac{b_\theta(\zeta)}{b_{\bar\theta}(\zeta)}\frac{\frac{\partial}{\partial\theta}b_\theta(\zeta)}{b_\theta(\zeta)}\right)^2\right] + K^2 + 2t^2s^2 + 2t^2E_{\bar\theta}\left[\left(\sum_{i=1}^t \zeta_i\right)^2\right]$$

$$\le E_{\bar\theta}\left[K^2\sup_{|\theta-\theta_0|<\epsilon_0}\left(\frac{\partial}{\partial\theta}\ln b_\theta(\zeta)\right)^2\right] + K^2 + 2t^2s^2 + 2t^4E_{\bar\theta}(\zeta^2) < \infty.$$

The conclusion then follows from Proposition 2 of L'Ecuyer (1994) (with $q = b_{\bar\theta}$). □

PROPOSITION 15. *Under Assumptions B (i–iii), $u(\theta)$, $l(\theta)$, and $w(\theta)$ are finite and continuously differentiable in θ, for $\theta \in \bar{\Theta}$. Also, $\psi_u(\theta, \omega)$ and $\psi_l(\theta, \omega)$, defined in (23) and (24), are unbiased estimators of $u'(\theta)$ and $l'(\theta)$, respectively, for θ in a small enough neighborhood of θ_0.*

PROOF. We first prove the second part of the proposition, and for that we will use Proposition 3 of L'Ecuyer (1994). For fixed ω, τ and $\sum_{i=1}^\tau W_i^*$ do not depend on θ. Therefore, τ, $\sum_{i=1}^\tau W_i^*$, and each $b_\theta(\zeta_i)$ are differentiable in θ everywhere in $\bar{\Theta}^0$. This implies A2 (a) in L'Ecuyer (1994) with t replaced by τ. From B (i), there is an $\hat{s} > 0$ such that for all $s \le \hat{s}$, $\tilde{E}[e^{s\zeta}] < \infty$. Then, from Theorem III.3.2 in Gut (1988, p. 81) there is an $\epsilon_1 > 0$ such that $\tilde{E}[e^{\epsilon_1\tau}] < \infty$. Let $0 < K \le e^{\epsilon_1/8}$, ϵ_0, and $\bar{\theta}$ satisfy B (iii). One has

$$\prod_{i=1}^\tau \sup_{|\theta-\theta_0|<\epsilon_0}\left(\frac{b_\theta(\zeta)}{b_{\bar\theta}(\zeta)}\right)^8 \le K^{8\tau} \le e^{\epsilon_1\tau},$$

and by a similar stochastic ordering argument as in the proof of (38),

$$E_{\bar\theta}[K^{8\tau}] \le E_{\bar\theta}[e^{\epsilon_1\tau}] \le \tilde{E}[e^{\epsilon_1\tau}] < \infty.$$

From Wald's equation and B (iv),

$$E_{\bar\theta}\left[\sum_{j=1}^\tau \sup_{|\theta-\theta_0|<\epsilon_0}\left(\frac{\frac{\partial}{\partial\theta}b_\theta(\zeta_j)}{b_{\bar\theta}(\zeta_j)}\right)^4\right] = E_{\bar\theta}[\tau]E_{\bar\theta}\left[\sup_{|\theta-\theta_0|<\epsilon_0}\left(\frac{\frac{\partial}{\partial\theta}b_\theta(\zeta)}{b_{\bar\theta}(\zeta)}\right)^4\right]$$

$$\le E_{\bar\theta}[\tau]E_{\bar\theta}\left[K^4\sup_{|\theta-\theta_0|<\epsilon_0}\left(\frac{\partial}{\partial\theta}\ln b_\theta(\zeta)\right)^4\right] < \infty.$$

Then, from (39), all the requirements of A3 in L'Ecuyer (1994) are satisfied, with $h(\theta, \omega)$ there replaced by either τ or $\sum_{i=1}^\tau W_i^*$ (which here do not depend on θ for ω fixed), and $\Gamma_{1i}(\zeta) = K^8$. This holds in a neighborhood of θ_0 for each $\theta_0 \in \bar{\Theta}$. This implies the result, except for the continuous differentiability of w, which follows from (6) and the continuous differentiability of u and l. □

PROPOSITION 16. *Under Assumptions B, $\sup_{\theta\in\bar\Theta} E_\theta[d_i^8] < \infty$.*

PROOF. Let $K > 1$. From B (iv), for each $\theta_0 \in \bar{\Theta}$, there is an open interval $\Upsilon(\theta_0) = (\theta_0 - \epsilon_0, \theta_0 + \epsilon_0)$, a $\bar{\theta} \in \bar{\Theta}$, and a constant $\tilde{K}(\theta_0) < \infty$ such that

$$E_{\bar\theta}\left[\sup_{\theta\in\Upsilon(\theta_0)}\left(\frac{\partial}{\partial\theta}\ln b_\theta(\zeta)\right)^8\right] \le \tilde{K}(\theta_0).$$

It follows that

$$\sup_{\theta\in\Upsilon(\theta_0)} E_\theta[d_i^8] = \sup_{\theta\in\Upsilon(\theta_0)} E_{\bar\theta}\left[\left(\frac{\partial}{\partial\theta}\ln b_\theta(\zeta)\right)^8\frac{b_\theta(\zeta)}{b_{\bar\theta}(\zeta)}\right]$$

$$\le KE_{\bar\theta}\left[\sup_{\theta\in\Upsilon(\theta_0)}\left(\frac{\partial}{\partial\theta}\ln b_\theta(\zeta)\right)^8\right] \le K\tilde{K}(\theta_0).$$

Now, $\{\Upsilon(\theta_0), \theta_0 \in \bar{\Theta}\}$ is a family of open sets that covers $\bar{\Theta}$. Since $\bar{\Theta}$ is compact, there is a finite subset of that family, say $\{\Upsilon(\theta^{(1)}), \ldots, \Upsilon(\theta^{(N)})\}$, that covers $\bar{\Theta}$, and one has

$$\sup_{\theta\in\bar\Theta} E_\theta[d_i^8] \le \max_{1\le i\le N} K\tilde{K}(\theta^{(i)}) < \infty. \qquad \square$$

PROPOSITION 17. *Consider the LR estimator (20). Under Assumption B,*

$$\sup_{\theta\in\bar\Theta, s\le c, t\ge 1} E_{\theta,s}\left[\frac{\psi_t^2(\theta,s,\omega)}{t^3}\right] < \infty.$$

PROOF. Since $E_\theta[\zeta_i^4] < \infty$ (Proposition 8), from §VIII.2 of Asmussen (1987), since $E_{\theta,s}[(W_i^*)^4] \le E_{\theta,c}[(W_i^*)^4]$, and from Proposition

16, there exists a constant $K_d < \infty$ such that

$$\sup_{\theta \in \bar{\Theta}, s \le c, i \ge 1} E_{\theta,s}[(W_i^*)^4 + d_i^4] \le K_d.$$

Recall that $E[d_j] = 0$ and that the d_j's are independent. Then,

$$E_{\theta,s}[\psi_t^2(\theta, s, \omega)] = E_{\theta,s}\left[\left(\sum_{i=1}^{t} W_i^*\right)^2 \left(\sum_{j=1}^{t} d_j\right)^2\right]$$

$$\le \left(t^4 \sup_{1 \le i \le t} E_{\theta,s}[(W_i^*)^4] \sum_{i=1}^{t} \sum_{j=1}^{t} E_{\theta,s}[d_i^2 d_j^2]\right)^{1/2}$$

$$\le [t^4 K_d t^2 K_d]^{1/2} = t^3 K_d. \qquad \square$$

PROPOSITION 18. *Suppose that Assumption B holds. Then, as $r \to \infty$, the regenerative LR estimator (25) has bounded variance and converges in quadratic mean to $w'(\theta)$, uniformly with respect to θ in $\bar{\Theta}$.*

PROOF. From Proposition 8, $E_\theta[\tau^8] \le K_\tau$ and $E_\theta[(h_\tau(\theta, 0, \omega))^4] \le K_h$. From Theorem I.5.2 in Gut (1988), there is a constant K_s independent of θ, such that $1 \le K_s < \infty$ and

$$E_\theta[(S_\tau(\theta, 0, \omega))^8] = E_\theta\left[\left(\sum_{j=1}^{\tau} d_j\right)^8\right] \le K_s.$$

Let $K = \max(K_\tau, K_h, K_s)$. Define $A_{1j} = h_j S_j - w(\theta)\tau_j S_j - w'(\theta)\tau_j$ and $A_{2j} = h_j - w(\theta)\tau_j$. Note that $E_\theta[A_{1j}] = E_\theta[A_{2j}] = 0$, since $w(\theta) = E_\theta[h_j]/E_\theta[\tau_j]$ and $w'(\theta) = (E_\theta[h_j S_j] - w(\theta)E_\theta[\tau_j S_j])/E_\theta[\tau_j]$ (from Proposition 15). Also, since $E_\theta[\tau_j] \ge 1$ (used in the first two lines), one has

$$w(\theta) \le E_\theta[h_j] \le K^{1/4},$$

$$w'(\theta) \le E_\theta[h_j S_j] - w(\theta)E_\theta[\tau_j S_j]$$

$$\le (E_\theta[h_j^4])^{1/4}(E_\theta[S_j^8])^{1/8} + K^{1/4}(E_\theta[\tau_j^8]E_\theta[S_j^8])^{1/8}$$

$$\le K^{3/8} + K^{1/2} \le 2K^{1/2},$$

$$E_\theta[A_{1j}^2] \le 2E_\theta[h_j^2 S_j^2] + 4w^2(\theta)E_\theta[\tau_j^2 S_j^2] + 4(w'(\theta))^2 E_\theta[\tau_j^2]$$

$$\le 2(E_\theta[h_j^4])^{1/2}(E_\theta[S_j^8])^{1/4}$$

$$+ 4K^{1/2}(E_\theta[\tau_j^8]E_\theta[S_j^8])^{1/4} + 16K(E_\theta[\tau_j^8])^{1/4}$$

$$\le 2K^{3/4} + 4K + 16K^{5/4} \le 22K^{5/4},$$

$$E_\theta[A_{2j}^4] = E_\theta[(h_j - w(\theta)\tau_j)^4]$$

$$\le 8E_\theta[h_j^4] + 8(w(\theta))^4 E_\theta[\tau_j^4]$$

$$\le 8K + 8K \cdot K^{1/2} \le 16K^{3/2},$$

$$E_\theta\left[\left(\frac{1}{r}\sum_{i=1}^{r} A_{2i}\right)^4\right] = E_\theta\left[\frac{1}{r^4}\sum_{i=1}^{r}\sum_{j=1}^{r} A_{2i}^2 A_{2j}^2\right) \le \frac{1}{r^2} E_\theta[A_{2j}^4] \le 16K^{3/2}/r^2.$$

Keeping in mind that $\tau_j \ge 1$ and $E_\theta[A_{1j}] = 0$ for each j, one obtains

$$E_\theta[\psi_w(r, \theta, \omega) - w'(\theta)]^2$$

$$\le E_\theta\Bigg[\frac{1}{r}\left(\sum_{j=1}^{r} h_j S_j - w(\theta)\sum_{j=1}^{r} \tau_j S_j - w'(\theta)\sum_{j=1}^{r} \tau_j\right.$$

$$\left. + \left(\sum_{j=1}^{r} \tau_j\right)^{-1}\left(w(\theta)\sum_{j=1}^{r}\tau_j - \sum_{j=1}^{r} h_j\right)\sum_{j=1}^{r}\tau_j S_j\right)\Bigg]^2$$

$$\le 2E_\theta\left[\left(\frac{1}{r}\sum_{j=1}^{r} A_{1j}\right)^2\right] + 2E_\theta\left[\left(\frac{1}{r}\sum_{j=1}^{r} A_{2j}\right)^2\left(\frac{1}{r}\sum_{j=1}^{r}\tau_j S_j\right)^2\right]$$

$$\le \frac{2}{r}E_\theta[A_{1j}^2] + 2\left(\frac{16K^{3/2}}{r^2}E_\theta[\tau_j^4 S_j^4]\right)^{1/2}$$

$$\le \frac{2}{r}(22K^{5/4} + 16K^{3/4}K^{1/2}) \le 76K^{5/4}/r.$$

As $r \to \infty$, this converges to zero uniformly in θ. These inequalities also provide a uniform upper bound of $76K^{5/4}/r$ on the variance of (25). $\square$

PROPOSITION 19. *Under Assumptions A (i) and C, for each $t \ge 1$ and $s \in \bar{S}$, $w(\theta)$ and $w_t(\theta, s)$ are convex in θ over $\bar{\Theta}$.*

PROOF. Since $\zeta_i = B_\theta^{-1}(U_i)$ and from (2), each W_i and W_i^* are convex in θ (for fixed U_i's). Therefore, for each (s, t), $w_t(\theta, s)$ is convex in θ. This implies that $w(\theta) = \lim_{t\to\infty} w_t(\theta, s)/t$ is also convex in θ. From Assumption C, it follows that $\alpha(\theta)$ is convex. $\square$

PROPOSITION 20. *Suppose that Assumptions A–C hold, that the system was originally started from state $s = 0$, and that the service time of the jth customer overall has distribution B_{θ_j} with $\theta_j \in \bar{\Theta}$ (the θ_j's can be different and might even be the values taken by correlated random variables, provided that these values are in $\bar{\Theta}$). Let v_i be defined as in (28) and K_h' be as in Proposition 8. Then,*

$$\sup_{k \ge 0, t \ge 1} E\left[\left(\frac{1}{t}\sum_{i=k+1}^{k+t}\sum_{j=v_i}^{i} Z_j\right)^2\right] \le K_h'. \tag{46}$$

Here, E denotes the expectation associated with the above sequence of θ_j's and we assume that it is well defined. (Note that here, we do not assume that $W_{k+1} \in \bar{S}$.)

PROOF. Suppose first that all the service times follow the distribution $\tilde{B}$. Then, the queue is stable (Asmussen 1987, Chapter VIII). Let τ be the number of customers in a regenerative (busy) cycle, let $\tilde{\delta}_i = \sum_{j=v_i}^{i} \Gamma(U_j)$, and let us view for the moment $\tilde{\delta}_i^2$ as a "cost" associated with customer i. The expected "cost" over a regenerative cycle is then, using the same argument as in the proof of (40) and assuming that $s = 0$,

$$\tilde{E}\left[\sum_{i=1}^{\tau} \tilde{\delta}_i^2\right] \le \tilde{E}\left[\tau\left(\sum_{i=1}^{\tau} \Gamma(U_j)\right)^2\right] \le K_h'.$$

From the renewal-reward theorem (Wolff 1989), one then has

$$\lim_{t\to\infty} \frac{1}{t}\sum_{i=1}^{t} \tilde{E}[\tilde{\delta}_i^2] = \tilde{E}\left[\sum_{i=1}^{\tau} \tilde{\delta}_i^2\right] \Big/ \tilde{E}[\tau] \stackrel{\text{def}}{=} \tilde{K} \le K_h'. \tag{47}$$

We will now show that $\tilde{\delta}_i$ is stochastically nondecreasing in i. For fixed values of ζ_i, ν_i, and v_i, v_{i+1} is a nonincreasing function of W_i, and it is easily seen that $(W_{i+1}, \tilde{\delta}_{i+1})$ is a nondecreasing function of $(W_i, \tilde{\delta}_i)$ for any value of U_{i+1}. Since $(W_1, \tilde{\delta}_1) = (0, \Gamma(U_1))$ while $(W_2, \tilde{\delta}_2) \geq (0, \Gamma(U_2))$, it follows that $(W_1, \tilde{\delta}_1)$ is stochastically dominated by $(W_2, \tilde{\delta}_2)$ and, by induction on i, that $(W_i, \tilde{\delta}_i)$ is stochastically nondecreasing in i. Then $\tilde{E}[\tilde{\delta}_i^2]$ is nondecreasing in i, and from (47) it follows that $\tilde{E}[\tilde{\delta}_i^2] \leq \lim_{i\to\infty} \tilde{E}[\tilde{\delta}_i^2] = \tilde{K} \leq K_h'$.

We will complete the proof using stochastic ordering arguments similar to those used in the proof of Proposition 8. For fixed U_j, replacing B_θ by $\tilde{B}$ for customer j increases ζ_j and does not affect the other service and interarrival times. Clearly, increasing a service time can never split a busy period, i.e., can never increase any v_i. Therefore, τ and each δ_i, generated under B_θ or under the assumptions of the Proposition, are stochastically dominated by τ and $\tilde{\delta}_i$ generated under $\tilde{B}$. This implies that $E[\delta_i^2] \leq \tilde{E}[\tilde{\delta}_i^2] \leq K_h'$, where E is the same as in (46). The expectation in (46), which is the second moment of the average of $\delta_{k+1}, \ldots, \delta_{k+t}$, is then bounded by K_h'. □

PROOF OF PROPOSITION 2. From Propositions 14 and 15, $w_t(\cdot, s)/t$ and $w(\cdot)$ are continuously differentiable in $\bar{\Theta}$ for each $s \in \bar{S}$ and $t \geq 0$. Further, the continuity of $w'(\theta)$ with respect to θ is uniform in θ over $\bar{\Theta}$, because the latter set is compact. Also, from Proposition 11, (13) holds. From Taylor's theorem, one has

$$\beta_n^D = (w'_{t_n}(\xi_n, s_n) - w'_{t_n}(\theta_n, s_n))/t_n$$

for $\theta_n^- \leq \xi_n \leq \theta_n^+$. Note that as $n \to \infty$, one has $\theta_n^+ - \theta_n^- \to 0$ and therefore $\xi_n - \theta_n \to 0$ a.s. Then,

$$\beta_n^D = [w'_{t_n}(\xi_n, s_n)/t_n - w'(\xi_n)] + [w'(\xi_n) - w'(\theta_n)] + [w'(\theta_n) - w'_{t_n}(\theta_n, s_n)/t_n],$$

and each bracketed term converges to zero a.s., uniformly in (θ_n, s_n), from (13) and from the uniform continuity of w'. □

PROOF OF PROPOSITION 3. From Proposition 10, the mean-square error of $h_t(\theta, s, \omega)/t$ is in $O(1/t)$, uniformly in (θ, s), so that $E_{n-1}[\epsilon_n^2]$, which is the conditional variance of Y_n, is in $O(t_n^{-1}c_n^{-2})$ and $\sum_{n=1}^{\infty} E_{n-1}[\epsilon_n^2]\gamma_n^2 < \infty$ a.s. From Proposition 11, $\lim_{n\to\infty} \beta_n^F \overset{\text{a.s.}}{=} 0$, and the result then follows from Proposition 1. □

PROOF OF PROPOSITION 4. From Proposition 10, the τ_j's and h_j's have uniformly bounded second moments. Therefore, the conditional variance of Y_n is in $O(t_n^{-1}c_n^{-2})$ and, $\sum_{n=1}^{\infty} E_{n-1}[\epsilon_n^2]\gamma_n^2 < \infty$. From Proposition 15 and since $\bar{\Theta}$ is compact, $u(\cdot)$ and $l(\cdot)$ are continuously differentiable, uniformly over $\bar{\Theta}$. It is then easy to see, using (16) and Taylor's theorem, that $\beta_n \to 0$ a.s., where β_n here is the difference between (16) and (8) evaluated at $\theta = \theta_n$. The result then follows from Proposition 1. □

PROOF OF PROPOSITION 5. From Proposition 14, $\psi_{t_n}(\theta_n, s_n, \omega_n)$ is an unbiased estimator of $w'_{t_n}(\theta_n, s_n)$, so that $\beta_n^R = 0$. From Proposition 11, we know that $\beta_n = \beta_n^F \to 0$ a.s. when $t_n \to 0$. From Proposition 17, there exists a constant $K_d < \infty$ such that $E_{n-1}[\epsilon_n^2] \leq K_d t_n$ for all n. Therefore,

$$\sum_{n=1}^{\infty} E_{n-1}[\epsilon_n^2]\gamma_n^2 \leq \sum_{n=1}^{\infty} K_d t_n \gamma_n^2 < \infty.$$

The first result then follows from Proposition 1.

For (b), Proposition 18 says that as $r \to \infty$, (25) has bounded variance and converges in quadratic mean to $w'(\theta)$, uniformly in θ. This implies uniform convergence in expectation. Then, $\lim_{n\to\infty} \beta_n = 0$ a.s., the variance of Y_n in (26) is uniformly bounded, and Proposition 1 applies.

For (c), one has $\beta_n = 0$. From Proposition 8 and the proof of Proposition 18, τ_j, h_j, $\tau_j S_j$, and $h_j S_j$ have bounded second moments for each j, uniformly in θ_n. Therefore, the variance of (27) is bounded uniformly in θ_n, and the result follows again from Proposition 1. □

PROOF OF PROPOSITION 6. From Proposition 9, $h'_{t_n}(\theta_n, s_n, \omega_n)$ is an unbiased estimator of $w'_{t_n}(\theta_n, s_n)$, so that $\beta_n^R = 0$. From Proposition 11, we know that $\beta_n = \beta_n^F \to 0$ a.s. when $t_n \to 0$. From Proposition 10, the variance of $h'_{t_n}(\theta_n, s_n, \omega_n)/t_n$ is bounded uniformly in θ_n and t_n. The first result then follows from Proposition 1.

For (b), Proposition 13 says that as $r \to \infty$, (34) has bounded variance and converges in quadratic mean to $w'(\theta)$, uniformly in θ. This implies uniform convergence in expectation. Then, $\lim_{n\to\infty} \beta_n = 0$ a.s., the variance of Y_n in (35) is uniformly bounded, and Proposition 1 applies.

For (c), $\beta_n = 0$ for each n. From Proposition 8, τ_j and h_j' have bounded second moments. So, the variance of (27) is bounded uniformly in θ_n, and the result follows again from Proposition 1. □

PROOF OF PROPOSITION 7. We will verify W1 to W5 of Appendix I, and the result will follow from Theorem 1. For this proof, we will redefine differently the state of the Markov chain. Remove the restriction $s_n \leq c$ and redefine the state at iteration n of SA as $s_n = (x_n, a_n)$, where x_n is the *sojourn* time of the last customer of iteration $n - 1$ ($x_1 = 0$), and a_n is the value of the IPA accumulator at the beginning of iteration n. Here, we assume that the arrival time of the first customer of an iteration is "unknown" (not part of the state) at the beginning of the iteration. We do that in order to facilitate the verification of the continuity conditions required in W1. Let $s = (x, a)$ be the system state at the beginning of an iteration, k_t^* be defined as in (32),

$$\psi^* = ak_t^* + \sum_{i=1}^{t} \sum_{j=v_i}^{i} Z_j \quad \text{and}$$

$$\xi = \left(\psi^*, W_t^*, I(k_t^* = t)a + \sum_{j=v_t}^{t} Z_j\right).$$

Here, ψ^* is the value of the IPA estimator (31), while the other two components of ξ give the initial state for the next iteration. At iteration n,

$$(\theta, x, a) = (\theta_n, x_n, a_n) \quad \text{and} \quad \xi = \xi_n = (\psi_n^*, x_{n+1}, a_{n+1}).$$

Since t_n is fixed at t, $P_{\theta,x,a}(\xi_n \in \cdot)$ does not depend on n.

To prove the weak continuity, let $g : \mathbb{R}^3 \to \mathbb{R}$ be continuous and bounded in absolute value by a constant K_g. We need to show that

$E_{\theta,x,a}[g(\xi)]$ is continuous in (θ, x, a). One difficulty is that for fixed U_j's, the components of ξ are discontinuous in θ. To prove the continuity of the expectation, we will use a likelihood ratio approach. Let $\theta_0 \in \Theta$, $K > 1$, ϵ_0, and $\bar{\theta}$ be as in B (iv). Let $x_0 \geq 0$ and $a_0 \geq 0$. Let us view ω as $(\nu_0, \zeta_1, \ldots, \nu_{t-1}, \zeta_t)$ and assume that ω is generated under $P_{\bar{\theta}}$. Now, for $|\theta - \theta_0| < \epsilon_0$, $x \geq 0$ and $a \geq 0$, define

$$\begin{aligned} A(\theta, x, a, \omega) = g\Bigg(& ak_t^* + \sum_{i=1}^{t} \sum_{j=v_i}^{i} \varphi(\theta, \zeta_j),\ W_t^*, \\ & I(k_t^* = t)a + \sum_{j=v_t}^{t} \varphi(\theta, \zeta_j)\Bigg) \prod_{i=1}^{t} \frac{b_\theta(\zeta_i)}{b_{\bar{\theta}}(\zeta_i)} \\ - g\Bigg(& a_0 k_{t,0}^* + \sum_{i=1}^{t} \sum_{j=v_{i,0}}^{i} \varphi(\theta_0, \zeta_j),\ W_{t,0}^*, \\ & I(k_{t,0}^* = t)a_0 + \sum_{j=v_{t,0}}^{t} \varphi(\theta_0, \zeta_j)\Bigg) \prod_{i=1}^{t} \frac{b_{\theta_0}(\zeta_i)}{b_{\bar{\theta}}(\zeta_i)}, \end{aligned}$$

where $k_{t,0}^*$, $v_{i,0}$, and $W_{t,0}^*$ are the respective values of k_t^*, v_i, and W_t^* when x is replaced by x_0, while $\omega = (\nu_0, \zeta_1, \ldots, \nu_{t-1}, \zeta_t)$ remains the same. Let $I(x, x_0) = 1$ if $v_{i,0} \neq v_i$ for at least one i, and $I(x, x_0) = 0$ otherwise. Note that $I(x, x_0) = 0$ implies that $k_{t,0}^* = k_t^*$. Also,

$$\begin{aligned} & P_{\bar{\theta}}(I(x, x_0) = 1) \\ & \leq \sum_{i=1}^{t} P_{\bar{\theta}}\left[\left|\nu_0 - x_0 + \sum_{j=1}^{i-1} (\nu_j - \zeta_j)\right| < |x - x_0|\right] \\ & = \sum_{i=1}^{t} E_{\bar{\theta}}\left[P_{\bar{\theta}}\left[|\nu_0 - x_0 + z| < |x - x_0| \,\middle|\, \sum_{j=1}^{i-1} (\nu_j - \zeta_j) = z\right]\right] \\ & \leq 2tK_\nu |x - x_0|, \end{aligned}$$

where $E_{\bar{\theta}}$ integrates over the values of z, and K_ν is a bound on the density of the interarrival time ν_0. Conditional on $I(x, x_0) = 0$, $A(\theta, x, a, \omega)$ is continuous in (θ, x, a), because g is continuous, W_t^* is continuous in x and does not depend on (θ, a), $b_\theta(\zeta)\varphi(\theta, \zeta)$ are continuous in θ for each ζ, $k_{t,0}^* = k_t^*$, and $v_{i,0} = v_i$ for each i. Further, $|A(\theta, x, a, \omega)|$ is bounded by $2K_g K^t$ and is zero when $(\theta, x, a) = (\theta_0, x_0, a_0)$. Therefore,

$$\begin{aligned} & \lim_{(\theta,x,a)\to(\theta_0,x_0,a_0)} |E_{\theta,x,a}[g(\xi)] - E_{\theta_0,x_0,a_0}[g(\xi)]| \\ & = \lim_{(\theta,x,a)\to(\theta_0,x_0,a_0)} |E_{\bar{\theta}}[A(\theta, x, a, \omega)]| \\ & \leq \lim_{(\theta,x,a)\to(\theta_0,x_0,a_0)} |E_{\bar{\theta}}[A(\theta, x, a, \omega)(1 - I(x, x_0))] \\ & \qquad + E_{\bar{\theta}}[2K_g K^t I(x, x_0)]| \\ & \leq E_{\bar{\theta}}[\lim_{(\theta,x,a)\to(\theta_0,x_0,a_0)} |A(\theta, x, a, \omega)(1 - I(x, x_0))|] \\ & \quad + \lim_{(\theta,x,a)\to(\theta_0,x_0,a_0)} 2K_g K^t 2tK_\nu |x - x_0| \\ & = 0, \end{aligned}$$

where the dominated convergence theorem has been used to pass the limit inside the expectation. This proves the required weak continuity. This also implies (as a special case) that $E_{\theta,x,a}[\psi^*]$ is continuous in (θ, x, a), which verifies the second requirement of W1, with $c = 0$.

For fixed $\theta \in \Theta$, since the system is stable, $\{\xi_n, n \geq 1\}$ is regenerative and is a Markov chain with some steady-state distribution P^θ (see Asmussen 1987, chapter VIII). Regeneration occurs whenever an iteration starts with an empty system. From Proposition 20, $\sup_{n\geq 1} E_0[(\psi_n^*/t_n)^2] \leq K_h'$ and $\sup_{n\geq 1} E_0[a_n^2] \leq K_h'$. This yields W3. By similar arguments as in the proof of Proposition 8, one can show that $\sup_{n\geq 1} E_0[x_n^2] \leq K_h$. Take $K = \max(K_h, K_h')$. For any $\epsilon > 0$, one has

$$K \geq E_0[(\psi_n^*/t_n)^2] \geq (3K/\epsilon)P[(\psi_n^*/t_n)^2 > 3K/\epsilon],$$

so that

$$\sup_{n\geq 1} P[(\psi_n^*/t_n)^2 > 3K/\epsilon] \leq \epsilon/3.$$

Similarly,

$$\sup_{n\geq 1} P[x_n^2 > 3K/\epsilon] \leq \epsilon/3 \quad \text{and}$$

$$\sup_{n\geq 1} P[a_n^2 > 3K/\epsilon] \leq \epsilon/3.$$

Then,

$$\sup_{n\geq 1} P[\max((\psi_n^*/t_n)^2, x_{n+1}^2, a_{n+1}^2) \leq 3K/\epsilon] \geq 1 - \epsilon.$$

This reasoning also holds for θ varying in any manner inside $\bar{\Theta}$. This implies the tightness properties required in W1.

For W2, let C be a compact subset of $\mathbb{R} \times S$, $c < \infty$ such that $C \subseteq [0, c]^3$, and let $\xi_n \in C$. Let i denote the ith customer overall and $nt + 1 + \tau_n^*$ be the index of the first nonwaiting customer from the beginning of iteration $n + 1$. One has $\tau_n^* = 0$ if iteration $n + 1$ starts with a new busy cycle, and otherwise τ_n^* is the number of customers, from the beginning of iteration $n + 1$, who are in the same busy cycle as the last customer of iteration n. From the same argument as in the proof of Proposition 8, there exists $K_\tau(c) < \infty$ such that $\tilde{E}_{s=c}[(\tau_1^*)^2] \leq K_\tau(c)$. Then, from straightforward stochastic ordering,

$$E[(\tau_n^*)^2 | \xi_n] \leq \tilde{E}_{s=c}[(\tau_1^*)^2] \leq K_\tau(c).$$

This implies that for all $\epsilon > 0$,

$$P[\tau_n^* \geq K_\tau(c)/\epsilon | \xi_n] \leq \epsilon.$$

Let $\epsilon > 0$, $n^*(c) = \lceil K_\tau(c)/\epsilon \rceil$, $\tilde{c} = (3K/\epsilon)^{1/2}$, and $\tilde{C} = [0, \tilde{c}]^3$. Let $n_c = 1 + \lceil n^*(c)/t \rceil$ and $i \geq n_c$. For each $0 \leq j \leq n_c$, from the same argument as we used above to prove W1, one has

$$P[\xi_{n+i} \in \tilde{C} | \tau_n^* = j] \geq 1 - \epsilon.$$

Then,

$$\begin{aligned} P[\xi_{n+i} \in \tilde{C} | \xi_n] & \geq \sum_{j=0}^{n^*(c)} P[\xi_{n+i} \in \tilde{C}, \tau_n^* = j | \xi_n] \\ & = \sum_{j=0}^{n^*(c)} P[\tau_n^* = j | \xi_n] P[\xi_{n+i} \in \tilde{C} | \tau_n^* = j] \\ & \geq (1 - \epsilon)P[\tau_n^* \leq n^*(c) | \xi_n] \geq (1 - \epsilon)^2. \end{aligned}$$

Here, P denotes the probability law associated with the Markov chain $\{\xi_n, n \geq 1\}$ when θ varies according to the algorithm and n_c can be interpreted as a time that we give to the system to stabilize. Roughly, if c is larger, the initial state could be larger (e.g., large initial queue size), and we will take a larger n_c. This implies W2.

When θ is fixed, from Proposition 9, δ_i is an unbiased estimator of the derivative of the expected system time of the ith customer (overall). Then, ψ_n^* is unbiased for the gradient of the expected total system time of customers $nt, \ldots, (n+1)t - 1$. When $n \to \infty$, from (13), the expectation of $\psi_n^*/t_n + C'(\theta)$ thus converges to $\alpha'(\theta)$. Therefore, $v(\theta) = \alpha'(\theta)$ and W5 follows. □

References

Asmussen, S., *Applied Probability and Queues*, Wiley, 1987.

Chong, E. K. P. and P. J. Ramadge, "Convergence of Recursive Optimization Algorithms Using Infinitesimal Perturbation Analysis Estimates," *Discrete Event Dynamic Systems: Theory and Applications*, 1 (1992a), 339–372.

—— and ——, "Stochastic Optimization of Regenerative Systems using Infinitesimal Perturbation Analysis," Submitted for publication, 1992b.

—— and ——, "Optimization of Queues using an IPA Based Stochastic Algorithm with General Update Times," *SIAM J. Control and Optimization*, 31 (1993), 698–732.

Fu, M. C., "Convergence of a Stochastic Approximation Algorithm for the GI/G/1 Queue Using Infinitesimal Perturbation Analysis," *J. Optimization Theory and Applications*, 65, 1 (1990), 149–160.

Glasserman, P., *Gradient Estimation via Perturbation Analysis*, Kluwer Academic, 1991.

—— and D. D. Yao, "Some Guidelines and Guarantees for Common Random Numbers," *Management Sci.*, 38, 6 (1992), 884–908.

Glynn, P. W., "Stochastic Approximation for Monte-Carlo Optimization," *Proceedings of the Winter Simulation Conference* 1986, IEEE Press (1986), 356–364.

——, "Optimization of Stochastic Systems Via Simulation," *Proceedings of the Winter Simulation Conference* 1989, IEEE Press (1989), 90–105.

——, "Likelihood Ratio Gradient Estimation for Stochastic Systems," *Comm. ACM*, 33, 10 (1990), 75–84.

Gut, A., *Stopped Random Walks*, Springer-Verlag, 1988.

Heidelberger, P., X.-R. Cao, M. A. Zazanis, and R. Suri, "Convergence Properties of Infinitesimal Perturbation Analysis Estimates," *Management Sci.*, 34, 11 (1988), 1281–1302.

Kushner, H. J. and D. S. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer-Verlag, Applied Math. Sciences, 26, 1978.

—— and A. Shwartz, "An Invariant Measure Approach to the Convergence of Stochastic Approximations with State Dependent Noise," *SIAM J. Control and Optimization*, 22, 1 (1984), 13–24.

L'Ecuyer, P., "A Unified View of the IPA, SF, and LR Gradient Estimation Techniques," *Management Sci.*, 36, 11 (1990), 1364–1383.

——, "On the Interchange of Derivative and Expectation for Likelihood Ratio Derivative Estimators," to appear in *Management Sci.*

——, N. Giroux, and P. W. Glynn, "Stochastic Optimization by Simulation: Numerical Experiments with the M/M/1 Queue in Steady-state," *Management Sci.*, 40 (1994), 1245–1261.

Meketon, M. S., "Optimization in Simulation: a Survey of Recent Results," *Proceedings of the Winter Simulation Conference* 1987, IEEE Press (1987), 58–67.

Métivier, M. and P. Priouret, "Application of a Kushner and Clark Lemma to General Classes of Stochastic Algorithms," *IEEE Transactions on Information Theory*, IT-30, 2 (1984), 140–151.

Pflug, G. Ch., "On-line Optimization of Simulated Markovian Processes," *Math. Oper. Res.*, 15, 3 (1990), 381–395.

Reiman, M. I., and A. Weiss, "Sensitivity Analysis for Simulation via Likelihood Ratios," *Oper. Res.*, 37, 5 (1989), 830–844.

Rubinstein, R. Y., *Monte-Carlo Optimization, Simulation and Sensitivity of Queueing Networks*, Wiley, 1986.

——, "Sensitivity Analysis and Performance Extrapolation for Computer Simulation Models," *Oper. Res.*, 37, 1 (1989), 72–81.

—— and A. Shapiro, *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method*, Wiley, 1993.

Rudin, W., *Real and Complex Analysis* (2nd Ed.), McGraw-Hill, 1974.

Spall, J. C., "Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation," *IEEE Transactions on Automatic Control*, AC-27, 3 (1992), 332–341.

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.

—— and Y. T. Leung, "Single Run Optimization of Discrete Event Simulations—An Empirical Study Using the M/M/1 Queue," *IIE Transactions*, 21, 1 (1989), 35–49.

Wardi, Y., "Simulation-Based Stochastic Algorithm for Optimizing GI/G/1 Queues," Manuscript, Department of Industrial Engineering, Ben Gurion University, 1988.

Wolff, R., *Stochastic Modeling and the Theory of Queues*, Prentice-Hall, 1989.

Accepted by James R. Wilson, former Departmental Editor; received December 6, 1991. This paper has been with the authors $7\frac{1}{2}$ months for 2 revisions.