Proceedings of the 2007 Winter Simulation Conference
S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, eds.

NON-LINEAR CONTROL VARIATES FOR REGENERATIVE STEADY-STATE SIMULATION

Sujin Kim

School of Industrial Engineering
Purdue University, 315 N. Grant St
West Lafayette, IN 47907-2023, U.S.A.

Shane G. Henderson

School of Operations Research and Information Engineering
230 Rhodes Hall, Cornell University
Ithaca, N.Y. 14853, U.S.A.

ABSTRACT

We assume the existence of a parameterized family of control variates that could be used in a regenerative steady-state simulation. We show how such controls can be generated in the Markov-process setting, discuss the optimization problem of searching for a good choice of parameterization, and develop a strong law and central limit theorem for the resulting estimator.

1 INTRODUCTION

In this paper we describe a variance reduction scheme, based on the method of control variates, for estimating a steady-state mean. We introduce the method by first recalling the key ideas behind the method of *linear* control variates for *finite-horizon* simulations.

The variance-reduction method of linear control variates for finite-horizon simulations (see, e.g., Law 2007, Chapter 11) works as follows. We wish to compute EX, where X is observable from some finite-horizon simulation. The simulation also produces a random (column) vector C with mean 0. We can therefore estimate EX by a sample mean of random variables of the form $X + \theta^T C$. (Here, θ^T denotes the transpose of the column vector of constants θ.) The parameter vector θ is chosen to minimize the variance of the estimator. This is straightforward, since the variance is a quadratic function of θ, and the optimizing choice of θ is easily estimated from the simulated data.

The situation is more complicated, even in the finite-horizon setting, when one estimates EX by $X + C(\theta)$, where $C(\cdot)$ is some complicated nonlinear function of the parameter vector θ and random variables that has the property that for all choices of θ, $EC(\theta) = 0$. (Such functions arise naturally when one is simulating Markov chains; see Kim and Henderson 2007, Henderson and Glynn 2002. Also, it is worth noting that this setup is fundamentally different from control variates that are given as a nonlinear function of means, as studied in Glynn and Whitt 1989.) Kim and Henderson (2007) studied this setting in detail. The main issue is that the variance function is not as tractable as a quadratic, even when one sets up the parameterization in such a way as to yield a differentiable variance function, so it is unclear how to select the parameter θ. One would like to select a θ that minimizes the variance, but the main goal is the estimation of EX, and these 2 goals need to be balanced in some appropriate way.

Kim and Henderson (2007) described 2 main approaches to this finite-horizon problem. The first approach is a stochastic-approximation scheme whereby an estimate of the variance-minimizing θ is updated with each simulation iteration. The second approach involves the use of the "sample-average approximation" framework where a choice of θ, $\hat{\theta}$ say, is made and fixed in a first stage, and then the estimation of EX proceeds in a second stage by averaging conditionally independent replicates of $X + C(\hat{\theta})$. The stochastic-approximation approach has low computational requirements in each iteration, but requires many iterations to converge, and its success depends strongly on identifying good values of certain tuning parameters, which is difficult. The sample-average approximation approach requires more computation, but its performance is much more robust.

Given the success of the sample-average approximation approach in this early work, and given the high variances associated with certain steady-state estimation problems such as those that arise in simulations of queueing systems in heavy traffic (Whitt 1989, Asmussen 1992), we would like to extend these ideas to steady-state performance measures. This paper represents a first attempt at such an extension.

We extend the finite-horizon case explored in Kim and Henderson (2007) to the regenerative steady-state case. We focus on regenerative steady-state simulation primarily because the regenerative case enjoys many of the advantages of the finite-horizon case. However, the analysis is still technically challenging.

We recognize that, at least currently, the restriction to the regenerative setting represents a significant restriction in practice. Ultimately we would like to extend this analysis

to other variance estimators that are more easily implemented and widely used in practice. These include batch means methods (adjusted to provide consistent estimates of the variance constant), and spectral methods. The primary technical challenge in those contexts appears to be establishing a uniform (in the parameters) convergence theory for variance estimators and their derivatives.

In related work, Newton (1994) described a number of variance reduction techniques for simulated diffusions. One of the techniques proposed there, but not explored in any depth, is essentially the stochastic approximation method for finite-horizon simulations studied in Kim and Henderson (2007). Henderson, Meyn, and Tadić (2003) developed a stochastic-approximation scheme for minimizing an approximation of the steady-state variance constant. They did not assume a regenerative structure. Tadić and Meyn (2004) have also made some initial steps on this problem, again using a stochastic-approximation scheme. These papers all use martingales as control variables; see Henderson and Glynn (2002) for a detailed treatment of how to get control variates from martingales. There are a number of papers that involve adaptive importance sampling. See Juneja and Shahabuddin (2006) for an entry point to that literature.

The remainder of this paper is organized as follows. In Section 2 we define the steady-state estimation problem, recall the main theory of regenerative simulation and give an expression for the appropriate variance constant as a function of the parameters. Then, in Section 3 we impose conditions on the parameterization that ensure differentiability of the variance function. This then sets the stage for laying out our estimation methodology, which includes an optimization step, in Section 4. Section 5 discusses how to obtain a parameterized source of control variates when simulating a general-state-space Markov chain, and Section 6 gives numerical results for a simple example. Section 7 contains some concluding remarks.

2 THE VARIANCE FUNCTION

Let $X = (X_k : k \geq 0)$ be a discrete-time stochastic process on the state space S. Let $f : S \to \mathbb{R}$ be a cost function on the state space. Suppose that there exists a constant α such that

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \to \alpha \tag{1}$$

as $n \to \infty$ almost surely (a.s.). We call α the steady-state mean, and our goal is to estimate it. A natural estimator of α is the time average

$$\alpha_n = \frac{1}{n}\sum_{k=0}^{n-1} f(X_k).$$

Suppose further that for each $\theta = (\theta(1), \ldots, \theta(p)) \in \Theta \subseteq \mathbb{R}^p$, $h(\cdot;\theta) : S \to \mathbb{R}$ is a real-valued non-linear function of $x \in S$ with steady-state mean 0, i.e.,

$$\frac{1}{n}\sum_{k=0}^{n-1} h(X_k;\theta) \to 0,$$

as $n \to \infty$ a.s. (We will see below that such functions can be constructed in interesting application settings.) Then we obtain a family of consistent estimators of α, $(\alpha_n(\theta) : \theta \in \Theta)$, where

$$\alpha_n(\theta) = \frac{1}{n}\sum_{k=0}^{n-1} [f(X_k) + h(X_k;\theta)].$$

Here $n^{-1}\sum_{k=0}^{n-1} h(X_k;\theta)$ serves as a control variate. But how should we select amongst this family of control variates?

A natural approach is to select the estimator with minimal "steady-state variance." To define this notion precisely, we impose a very general regenerative structure. Suppose that X is a 1-dependent regenerative process with regeneration times $T(0) = 0 < T(1) < \cdots$. Denote the cycle lengths by $\tau_i = T(i) - T(i-1)$, $i \geq 1$. Define the cost accumulated over the ith cycle by

$$F_i = \sum_{k=T(i-1)}^{T(i)-1} f(X_k), \quad i \geq 1.$$

Under appropriate moment conditions, we can rigorously define, and obtain expressions for, the steady state mean and variance for the above estimators. We first state a familiar result for α_n.

Theorem 1. *Suppose that* $E(|F_1| + \tau_1) < \infty$.

(i) *The strong law*

$$\alpha_n \to \alpha = \frac{EF_1}{E\tau_1} \tag{2}$$

holds as $n \to \infty$ *a.s.*

(ii) *Define* $Z_k = F_k - \alpha\tau_k$ *for* $k \geq 1$, *and let* $\Rightarrow$ *denote convergence in distribution. If, in addition to the conditions already stated,* $EZ_1^2 < \infty$, *then the central limit theorem (CLT)*

$$n^{1/2}(\alpha_n - \alpha) \Rightarrow \sigma N(0,1) \tag{3}$$

holds as $n \to \infty$, *in which case*

$$\sigma^2 = \frac{EZ_1^2 + 2EZ_1Z_2}{E\tau_1}.$$

The proof is nearly identical to the classically regenerative case given in Glynn and Iglehart (1993), and so we omit it here. The constant σ^2 is known as the *time-average variance constant*, and Theorem 1 immediately suggests that we can estimate it using the estimator

$$\sigma_n^2 = \frac{1}{n}\sum_{k=1}^{l(n)-1}[Z_k^2(n) + 2Z_k(n)Z_{k+1}(n)],$$

where $Z_k(n) = F_k - \alpha_n\tau_k$ and $l(n) = \sup\{k \geq 0 : T(k) \leq n\}$ is the number of completed regenerative cycles by time n. The next result affirms that σ_n^2 is a consistent estimator for σ^2 under appropriate conditions.

Theorem 2. *Assume that the conditions in Part (ii) of Theorem 1 hold. Then*

$$\sigma_n^2 \Rightarrow \sigma^2$$

as $n \to \infty$. Furthermore, if $E(F_1^2 + \tau_1^2) < \infty$, then $\hat{\sigma}_n^2$ is strongly consistent, i.e., $\sigma_n^2 \to \sigma^2$ as $n \to \infty$ a.s.

The proof follows as in Glynn and Iglehart (1993), and so is omitted.

We can apply Theorems 1 and 2 to the family of estimators $(\alpha_n(\theta) : \theta \in \Theta)$ as follows. For any fixed $\theta \in \Theta$, define

$$\begin{aligned} H_k(\theta) &= \sum_{i=T(k-1)}^{T(k)-1} h(X_i;\theta), \\ Z_k(\theta) &= F_k + H_k(\theta) - \alpha\tau_k, \text{ and} \\ \sigma^2(\theta) &= \frac{EZ_1^2(\theta) + 2EZ_1(\theta)Z_2(\theta)}{E\tau_1}. \end{aligned}$$

If the corresponding moment conditions hold, then for each $\theta \in \Theta$, $\alpha_n(\theta) \to \alpha$ a.s., and $n^{1/2}(\alpha_n(\theta) - \alpha) \Rightarrow \sigma(\theta)N(0,1)$ as $n \to \infty$. Furthermore, the variance estimator

$$\sigma_n^2(\theta) = \frac{1}{n}\sum_{k=1}^{l(n)-1}[Z_k^2(n;\theta) + 2Z_k(n;\theta)Z_{k+1}(n;\theta)] \qquad (4)$$

is consistent for $\sigma^2(\theta)$ as $n \to \infty$, where $Z_k(n;\theta) = F_k + H_k(\theta) - \alpha_n(\theta)\tau_k$.

The variance constants $\sigma^2(\theta)$ determine the rate of convergence of the estimators $\alpha_n(\theta)$, as quantified by the CLT just mentioned. Therefore, a reasonable goal is to select the parameter value, θ^* say, that minimizes the time-average variance constant $\sigma^2(\cdot)$. We are therefore interested in the (simulation) optimization problem

$$\mathscr{P}: \quad \min_{\theta\in\Theta} \sigma^2(\theta).$$

This optimization problem is easier to solve if the function $\sigma^2(\cdot)$ possesses structure. For example, if $h(x;\theta)$ is linear in θ, then $\mathscr{P}$ reduces to the setting of internal linear control variates in the regenerative setting, the theory of which is essentially a straight-forward extension of linear control variates for finite-horizon simulation; see Iglehart and Lewis (1979). If $h(x;\theta)$ is nonlinear in θ, then the problem is not as straightforward.

3 DERIVATIVES AND ESTIMATION

In this section we impose conditions that ensure that the variance function $\sigma^2(\cdot)$ is differentiable, and provide an estimator of its gradient, thereby aiding a numerical search for parameter values that yield low variance.

Assumption A1 The parameter set Θ is compact, and for all $x \in S$ the function $h(x,\cdot)$ is $\mathscr{C}^1$ on $\mathscr{U}$, where $\mathscr{U}$ is a bounded open set containing Θ. Moreover, $EH_1(\theta) = 0$ for any $\theta \in \mathscr{U}$.

Assumption A2 Define $W_k(\theta) = \sum_{i=T(k-1)}^{T(k)-1} |h(X_i;\theta)|$. The moment conditions $E(\tau_1^2 + F_1^2) < \infty$ and $EW_1^2(\theta_0) < \infty$ for some fixed $\theta_0 \in \mathscr{U}$ hold.

Assumption A3 For all $x \in S$, $h(x;\cdot)$ is Lipschitz on $\mathscr{U}$, i.e., $\exists c(x) > 0$ such that for all $\theta_1, \theta_2 \in \mathscr{U}$,

$$|h(x;\theta_1) - h(x;\theta_2)| \leq c(x)\|\theta_1 - \theta_2\|,$$

where $\|\cdot\|$ is a metric on $\mathbb{R}^p$. Therefore,

$$\sup_{\theta\in\mathscr{U}} \left|\frac{\partial h(x;\theta)}{\partial\theta(j)}\right| \leq c(x)$$

for all $x \in S$ and $j = 1,\ldots,p$. For $k \geq 1$, define $C_k = \sum_{i=T(k-1)}^{T(k)-1} c(X_i)$ to be the sum of the Lipschitz constants over the jth regenerative cycle, and assume that $EC_1^2 < \infty$.

Remark 1. *Suppose that $U(\theta)$ is a random variable for each $\theta \in \mathscr{U}$. We say that $U(\cdot)$ is* dominated *by an integrable random variable $\tilde{U}$ if $|U(\theta)| \leq \tilde{U}$ a.s. for every $\theta \in \mathscr{U}$, and $E\tilde{U} < \infty$. Under **A1-A3**, $H_1(\cdot)^2$ is dominated by an integrable random variable, and hence so is $Z_1^2(\cdot)$. To see why, note that for any $\theta \in \mathscr{U}$,*

$$\begin{aligned} H_1^2(\theta) &= [H_1(\theta_0) + (H_1(\theta) - H_1(\theta_0))]^2 \\ &\leq 2H_1^2(\theta_0) + 2(H_1(\theta) - H_1(\theta_0))^2 \\ &\leq 2W_1^2(\theta_0) + 2C_1^2\|\theta - \theta_0\|^2. \end{aligned}$$

But $\mathscr{U}$ is bounded, and hence $\|\theta - \theta_0\|^2$ is bounded.

Our next result gives conditions under which the variance function is differentiable, and gives an expression for

its gradient that also suggests an estimator for the gradient of the variance function.

Proposition 3. *Assume that **A1-A3** hold. Then $\sigma^2(\cdot)$ is $\mathscr{C}^1$ on $\mathscr{U}$ and*

$$\nabla_\theta \sigma^2(\theta) = \frac{E\nabla_\theta Z_1^2(\theta) + 2E\nabla_\theta(Z_1(\theta)Z_2(\theta))}{E\tau_1}.$$

Proof. It suffices to show that $EZ_1(\cdot)$ and $EZ_1(\cdot)Z_2(\cdot)$ are $\mathscr{C}^1$ on $\mathscr{U}$ and the gradient and expectation can be exchanged. We apply Proposition 1 in L'Ecuyer (1995) to $Z_1(\theta)$ and $Z_1(\theta)Z_2(\theta)$ component by component. Consider the jth component, for some $j \in \{1,\ldots,p\}$. The only condition that requires explicit verification is that $\partial Z_1(\theta)/\partial\theta(j)$ and $\partial Z_1(\theta)Z_2(\theta)/\partial\theta(j)$ are dominated by an integrable function of X. With probability 1,

$$\begin{aligned}\frac{\partial Z_k(\theta)}{\partial\theta(j)} &= -\frac{\partial}{\partial\theta(j)}\left(\sum_{i=T(k-1)}^{T(k)-1} h(X_i;\theta)\right)\\ &= -\sum_{i=T(k-1)}^{T(k)-1}\frac{\partial h(X_i;\theta)}{\partial\theta(j)}\end{aligned}$$

for $k \geq 1$. Hence,

$$\begin{aligned}\left|\frac{\partial Z_1^2(\theta)}{\partial\theta(j)}\right| &= 2\left|\frac{\partial}{\partial\theta(j)}\left(\sum_{i=0}^{T(1)-1} h(X_i;\theta)\right)Z_1(\theta)\right|\\ &\leq 2C_1|Z_1(\theta)|, \qquad (5)\end{aligned}$$

and

$$\begin{aligned}\left|\frac{\partial(Z_1(\theta)Z_2(\theta))}{\partial\theta(j)}\right| &= \left|\frac{\partial}{\partial\theta(j)}\left(\sum_{i=0}^{T(1)-1} h(X_i;\theta)\right)Z_2(\theta)\right|\\ &+ \left|\frac{\partial}{\partial\theta(j)}\left(\sum_{i=T(1)}^{T(2)-1} h(X_i;\theta)\right)Z_1(\theta)\right|\\ &\leq C_1|Z_2(\theta)| + C_2|Z_1(\theta)|. \qquad (6)\end{aligned}$$

By **A3** and Remark 1, the right-hand sides of equations (5) and (6) are dominated by an integrable function. □

4 OPTIMIZATION AND ESTIMATION

Solving the problem $\mathscr{P}$ above is really an intermediate step on the way to estimating the steady-state mean α, which is our true goal. Accordingly, we propose a two-phase approach. In the first phase we generate and fix a sample path $\tilde{X}_0,\tilde{X}_1,\ldots,\tilde{X}_m$, and then identify θ_m as the solution of the optimization problem

$$\mathscr{P}_m: \quad \min_{\theta\in\Theta}\sigma_m^2(\theta),$$

where $\sigma_m^2(\theta)$ is the estimator (4) of $\sigma^2(\theta)$ constructed from $\tilde{X}_0,\tilde{X}_1,\ldots,\tilde{X}_m$. (We refer to $\mathscr{P}_m$ as the sample-average approximation (SAA) problem corresponding to $\mathscr{P}$.) We then use the solution θ_m to $\mathscr{P}_m$ in a second phase, where α is estimated using

$$\alpha_n(\theta_m) = \frac{1}{n}\sum_{i=0}^{n-1}[f(X_i) + h(X_i;\theta_m)],$$

and the sample path $X_0,X_1,\ldots,X_{n-1}$ is independent of $\tilde{X}_0,\tilde{X}_1,\ldots,\tilde{X}_m$.

In this section we look at the asymptotics of this estimator as m and n become large. We begin with a uniform law of large numbers that we will then use to prove a strong law for $\alpha_n(\theta_m)$ as $n\to\infty$.

Proposition 4. *Suppose that $\{U_k(\theta): k\geq 1\}$ is a κ-dependent stationary sequence of random variables for any $\theta\in\Theta$, where Θ is a compact parameter set. Let $\{l(n): n\geq 1\}$ be a family of random indices such that $l(n)/n\to\lambda$ a.s as $n\to\infty$ for some constant $\lambda\in(0,\infty)$. Suppose that $U_i(\cdot)$ is continuous on Θ w.p.1 and dominated by an integrable random variable for all $i\geq 1$. Then*

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{i=1}^{l(n)}U_i(\theta) - \lambda EU_1(\theta)\right| \to 0 \qquad (7)$$

as $n\to\infty$ a.s.

Sketch of Proof. First write (7) as

$$\frac{l(n)}{n}\sup_{\theta\in\Theta}\left|\frac{1}{l(n)}\sum_{i=1}^{l(n)}[U_i(\theta) - EU_1(\theta)]\right|$$

plus a remainder term that converges to 0 a.s., uniformly in θ. Then we extend Proposition 7 in Shapiro (2004) for i.i.d. random variables to the κ-dependent case to obtain that

$$\sup_{\theta\in\Theta}\left|\frac{1}{m}\sum_{k=1}^{m}U_k(\theta) - EU_1(\theta)\right| \to 0$$

almost surely, as $m\to\infty$. The result then follows since $l(n)\to\infty$ as $n\to\infty$ a.s. ■

We can now state a version of the strong law and central limit theorem for $\alpha_n(\hat{\theta})$, where $\hat{\theta}$ is random.

Theorem 5. *Suppose that **A1-A3** hold, and that $\hat{\theta}$ is independent of the path $X_0,\ldots,X_{n-1}$ used to compute $\alpha_n(\hat{\theta})$ for every n. Then $\alpha_n(\hat{\theta})\to\alpha$ as $n\to\infty$ a.s., and*

$$\sqrt{n}(\alpha_n(\hat{\theta}) - \alpha) \Rightarrow \sigma(\hat{\theta})N(0,1)$$

as $n \to \infty$, where $N(0,1)$ is a standard normal random variable that is independent of $\hat{\theta}$.

Sketch of Proof. For the strong law note that

$$\begin{aligned} |\alpha_n(\hat{\theta}) - \alpha| &\leq |\alpha_n - \alpha| + \left| \frac{1}{n} \sum_{i=0}^{n-1} h(X_i, \hat{\theta}) \right| \\ &\leq |\alpha_n - \alpha| + \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{i=0}^{n-1} h(X_i, \theta) \right|. \quad (8) \end{aligned}$$

The first term in (8) converges to 0 a.s. as $n \to \infty$ by (2). For the second term, we write it as the sum of 2 terms, the first of which is a remainder term that converges to 0 as $n \to \infty$ a.s., uniformly in θ. The second term is of the form given in Proposition 4 and therefore also converges to 0 as $n \to \infty$ a.s., uniformly in θ. The CLT can be proved by conditioning on $\hat{\theta}$, invoking the regenerative CLT, and then unconditioning, and is very similar to the proof of a similar result in Kim and Henderson (2007). ■

Theorem 5 identifies the asymptotic behavior of the estimator $\alpha_n(\theta_m)$ when m is held fixed and $n \to \infty$. However, as the computational budget increases, one would probably devote an increasing amount of effort to estimating the true minimizer of Problem $\mathscr{P}$. Suppose that $m = m(n)$ is a function of n such that $m(n) \to \infty$ as $n \to \infty$. Our next result shows that if $\theta_{m(n)}$ converges in probability to a (potentially random) parameter setting θ^* as $n \to \infty$, then $\alpha_n(\theta_{m(n)})$ has the same SLLN and CLT behavior as $\alpha_n(\theta^*)$.

Theorem 6. *Suppose that $\theta_{m(n)} \to \theta^*$ as $n \to \infty$ in probability, for some random variable θ^*. Suppose further that **A1 - A3** hold and that $\theta_{m(n)}$ is independent of the path $X_0, \ldots, X_{n-1}$ used to compute $\alpha_n(\theta_{m(n)})$ for every n. Then $\alpha_n(\theta_{m(n)}) \to \alpha$ as $n \to \infty$ a.s., and*

$$\sqrt{n}(\alpha_n(\theta_{m(n)}) - \alpha) \Rightarrow \sigma(\theta^*) N(0,1)$$

as $n \to \infty$, where $N(0,1)$ is independent of θ^*.

Sketch of Proof. The strong law can be proved exactly as in Theorem 5. For the central limit theorem, note that

$$\begin{aligned} &\sqrt{n}(\alpha_n(\theta_{m(n)}) - \alpha) \\ &= \sqrt{n}(\alpha_n(\theta^*) - \alpha) + \sqrt{n}(\alpha_n(\theta_{m(n)}) - \alpha_n(\theta^*)) \\ &= D_{1,n} + D_{2,n}, \text{ say.} \end{aligned}$$

Now, θ^* is independent of the samples used to compute α_n for every n, so Theorem 5 establishes that $D_{1,n} \Rightarrow \sigma(\theta^*) N(0,1)$ as $n \to \infty$. Thus, it suffices to show that

$$D_{2,n} = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} [h(X_i, \theta_{m(n)}) - h(X_i, \theta^*)] \Rightarrow 0$$

as $n \to \infty$. One can show that this difference is essentially a sum of regenerative cycle quantities and a remainder term that converges to 0 in probability. The analysis of the sum borrows techniques from Janson (1983) and Henderson and Glynn (2001), and is omitted. ■

Theorem 6 shows that it is a good idea to try to select θ_m to solve the Problem $\mathscr{P}_m$. The best that we can hope for from a computational point of view though, is that θ_m is a first-order critical point of Problem $\mathscr{P}_m$. Hence, it is important that the set of first-order critical points of Problem $\mathscr{P}_m$ converge to those of Problem $\mathscr{P}$ as $m \to \infty$. Theorem 7 below shows that if we can uniformly approximate the gradient of a function, then the first-order critical points will converge. This is only a slight modification of a result due to Bastin, Cirillo, and Toint (2007) and so the proof is omitted. For a point x and set A define the distance $d(x,A) = \inf\{\|x - y\| : y \in A\}$.

Theorem 7. *Suppose that*

- *(i)* *Θ is convex and compact,*
- *(ii)* *$g : \Theta \to \mathbb{R}$ is $\mathscr{C}^1$ on an open set containing Θ,*
- *(ii)* *the random functions $g_m(\cdot) : \Theta \to \mathbb{R}$ are $\mathscr{C}^1$ on an open set containing Θ w.p.1, for all $m \geq 1$, and*
- *(iv)* *$\sup_{\theta \in \Theta} ||\nabla_\theta g_m(\theta) - \nabla_\theta g(\theta)|| \to 0$ a.s. as $n \to \infty$.*

Let θ_m be a first-order critical point of $g_m(\cdot)$ on Θ and $S(g(\cdot), \Theta)$ be the set of first order critical points of $g(\cdot)$ on Θ. Then $d(\theta_m, S(g, \Theta)) \to 0$ as $m \to \infty$ a.s.

Now, the estimate σ_m^2 of the variance function is essentially a nonlinear function of means. This observation can be exploited, together with Theorem 7 above, to show that the distance between θ_m and the set of first-order critical points of $\sigma^2(\cdot)$ converges to 0 as m grows. The proof is somewhat mechanical, and omitted.

Corollary 8. *Suppose that **A1-A3** hold and Θ is convex. Then $d(\theta_m, S(\sigma^2, \Theta)) \to 0$ as $m \to \infty$ a.s.*

Corollary 8 shows that our methodology is sound, in the sense that numerically minimizing an *estimate* of the variance function, will give parameter settings that are close to local minimizers of the *true* variance function, so long as the estimate of the variance function is reasonably accurate, i.e., so long as m is large enough. Since the variance function is continuous, this then implies that we get a close-to-locally-optimal variance.

5 A SOURCE OF CONTROLS

The previous sections have explored the asymptotic theory of a parameterized control variate. But where can one obtain such controls? In this section we explain how to obtain a source of parameterized controls when simulating a Markov chain on a general state space, essentially reviewing ideas from Henderson and Glynn (2002) and Kim and Henderson (2007). This is a very general context that includes the special case of Markov chains on a countable state space. It also includes a (very) large class of discrete-event systems, via the Markov chain that arises by observing the system and event clocks just after each event occurs. (See Henderson and Glynn 2001 for more details on this well-known point.)

We now assume that $X = (X_k : k \geq 0)$ is a Markov chain on the general state space S. To ensure that the steady-state mean α defined in (1) is well defined, we assume that X is positive Harris recurrent with stationary distribution π, and that $\pi|f| = E_\pi|f(X_0)| < \infty$, where E_π denotes expectation when X_0 is distributed according to π. The positive recurrence assumption ensures that the chain has a unique steady-state distribution, and the law of large numbers then holds under moment conditions (Meyn and Tweedie 1993). The moment condition $\pi|f| < \infty$ ensures that the law of large numbers holds for α_n, so that $\alpha_n \to \alpha$ as $n \to \infty$ a.s.

We can now develop the controls. Let $u : S \to \mathbb{R}$ be a real-valued function on the state space of the chain, and let $Pu(x) = E[u(X_1)|X_0 = x]$ be the expected value of $u(X_1)$ when we start in state x. (This notation suggests that in the finite state space case, Pu is a matrix-vector product, which is correct.) Now, for $i \geq 1$, consider

$$u(X_i) - Pu(X_{i-1}) = u(X_i) - E[u(X_i)|X_{i-1}],$$

which has mean 0 under any initial distribution for which the appropriate expectations exist. A sum of such terms also has mean 0, suggesting that such a sum can serve as a control variate. In particular, define

$$\begin{aligned} M_n &= \sum_{i=1}^{n} [u(X_i) - Pu(X_{i-1})] \\ &= u(X_n) - u(X_0) - \sum_{i=0}^{n-1} [Pu(X_i) - u(X_i)] \end{aligned}$$

where the second line is obtained by rearranging terms. Let $h(\cdot) = Pu(\cdot) - u(\cdot)$, so that $M_n = u(X_n) - u(X_0) - \sum_{i=0}^{n-1} h(X_i)$. If we divide by n, then the terms $u(X_n)/n$ and $u(X_0)/n$ are asymptotically negligible, so that one might expect

$$C_n \stackrel{\triangle}{=} \frac{1}{n} \sum_{i=0}^{n-1} h(X_i)$$

to be a good candidate for a control variate. Indeed, if $\pi|u| < \infty$, so that u is π-integrable, then

$$\begin{aligned} \pi|h| &= \pi|Pu - u| \\ &\leq \pi(P|u|) + \pi|u| \\ &= (\pi P)|u| + \pi|u| \\ &= \pi|u| + \pi|u| \\ &< \infty, \end{aligned}$$

where we have used Fubini's theorem and the fact that $\pi P = \pi$. Hence h is π-integrable, so the strong law applies and then

$$\begin{aligned} C_n &\to \pi h \\ &= \pi(Pu - u) \\ &= (\pi P)u - \pi u \\ &= \pi u - \pi u \\ &= 0 \end{aligned}$$

as $n \to \infty$ a.s. So C_n could serve as a control variate, so long as u is π-integrable.

Now, if u is parameterized by θ, say $u(\cdot;\theta)$, and we let $h(\cdot;\theta) = Pu(\cdot;\theta) - u(\cdot;\theta)$ then, provided that $u(\cdot;\theta)$ is π-integrable, we can use

$$C_n(\theta) = \frac{1}{n} \sum_{i=0}^{n-1} h(X_i;\theta)$$

as a control variate in conjunction with α_n, giving an estimator

$$\alpha_n(\theta) = \frac{1}{n} \sum_{i=0}^{n-1} [f(X_i) + h(X_i;\theta)].$$

But how should one choose the parameterized set of functions $u(\cdot;\theta)$? The key issues are that we need to be able to compute the one-step expectations $Pu(\cdot;\theta)$ efficiently (so that we can compute $h(\cdot;\theta)$ efficiently), and that we get good variance reductions, in the sense that there should be some θ for which

$$h(\cdot;\theta) = Pu(\cdot;\theta) - u(\cdot;\theta) \approx -f(\cdot) + \alpha \qquad (9)$$

since, if this approximation holds as an equality, the estimator $\alpha_n(\theta)$ has zero variance. The equality version of (9) is known as Poisson's equation, and it plays a key role in Markov chain theory. So essentially, we want a "good" approximation to the solution to Poisson's equation within our parameterized class of functions $(h(\cdot;\theta) : \theta \in \Theta)$.

6 AN EXAMPLE

In this section we examine the performance of the adaptive control variate method discussed in Section 4 based on a stationary first-order autoregressive process. Let $\{\varepsilon_k : k = 1,2,\ldots\}$ be a sequence of independent normally distributed random variables with mean 0 and variance η^2. A first-order autoregressive (AR(1)) process satisfies the following difference equation:

$$X_i = \mu + \phi X_{i-1} + \varepsilon_i \text{ for } i \geq 1,$$

where μ and ϕ are any constants. To ensure that X has a stationary version, we assume that $|\phi| < 1$. Then the process $X = (X_k : k \geq 0)$ is a positive Harris recurrent Markov chain with stationary distribution $\pi \sim N(\mu/(1-\phi), \eta^2/(1-\phi^2))$.

To identify the regeneration times, we use the "splitting Markov chain method" (Athreya and Ney 1978, Nummelin 1978). At each transition, an auxiliary coin-flip is generated, and a regeneration occurs if the coin-flip results in success. For a discussion of the method in a simulation setting, see Glynn and Iglehart (1993) and Henderson and Glynn (2001). To construct parameterized controls, we use the idea presented in Section 5. Let $u(\cdot,\theta), \theta \in \Theta$ be given and $h(\cdot;\theta) = Pu(\cdot;\theta) - u(\cdot;\theta)$. One can show that A1-A3 and the conditions in Theorem 7 are satisfied in the following examples, so that all of our previous results apply.

For the simulation experiment, we take $\mu = 0$, $\eta = 1$ and $X_0 = 1$. We wish to estimate the steady-state mean with cost functions $f(x) = x$ and x^2. Note that the true steady state means are zero and $\eta^2/(1-\phi)^2$, respectively. We take

$$u(x;\theta) = xe^{\theta(1)(x-\theta(2))^2+\theta(3)},$$

so that

$$\begin{aligned} \theta &= (\theta(1),\theta(2),\theta(3)) \in \Theta, \text{ and} \\ \Theta &= \{x \in \mathbb{R}^3 : a(j) \leq x(j) \leq b(j), j = 1,2,3\}. \end{aligned}$$

In the second case, we take

$$u(x;\theta) = \theta(1)x^2 + xe^{\theta(2)(x-\theta(3))^2+\theta(4)},$$

where

$$\begin{aligned} \theta &= (\theta(1),\ldots,\theta(4)) \in \Theta, \text{ and} \\ \Theta &= \{x \in \mathbb{R}^4 : a(j) \leq x(j) \leq b(j), j = 1,\ldots,4\}. \end{aligned}$$

(It turns out that in both cases, the control variate that leads to zero variance is contained within the set of parameterized controls.)

We examine the performance of the adaptive estimators relative to the standard Monte Carlo technique. In the adaptive method, we take $m = 100$ samples and obtain θ_m by applying a quasi-Newton method with a linesearch using the gradients discussed in Section 4 to solve the sample average approximation problem. As an estimator of α, we use the time average $\alpha_n(\theta_m)$ over $n = 10,000$ replicates, where θ_m is viewed as fixed, in the sense of Theorem 5. To ensure a fair comparison, we allocate equal amounts of CPU time to both methods.

Table 1: Estimated variance reduction ratio with cost function $f(x) = x$.

ϕ	α_{Naive}	σ^2_{Naive}	α_{ACV}	σ^2_{ACV}	$\sigma^2_{Naive}/\sigma^2_{ACV}$
0.2	-2.5E-4	1.6	-1.5E-6	6.8E-8	2.3E+7
0.4	7.9E-3	2.8	-3.1E-7	2.6E-9	1.1E+9
0.6	6.7E-3	6.3	-1.0E-4	8.2E-4	7.6E+3
0.8	0.037	24	3.4E-6	2.5E-7	9.5E+7
0.9	6.4E-3	101	-3.1E-8	8.3E-9	1.2E+10
0.99	2.3	2.0E+4	-0.073	1.1E+3	18

Table 2: Estimated variance reduction ratio with cost function $f(x) = x^2$.

ϕ	α_{Naive}	σ^2_{Naive}	α_{ACV}	σ^2_{ACV}	$\sigma^2_{Naive}/\sigma^2_{ACV}$
0.2	1.04	2.4	1.04	3.2E-16	7.4E+15
0.4	1.20	3.8	1.19	2.2E-16	1.7E+16
0.6	1.55	10	1.56	3.4E-11	2.9E+11
0.8	2.78	82	2.78	4.3E-15	1.9E+16
0.9	5.28	616	5.26	1.6	397
0.99	53.7	4.0E+5	50.2	10	3.90E+4

Tables 1 and 2 show the simulation results for varying values of ϕ. We use the terms "Naive" and "ACV" to represent the estimators obtained through naïve Monte Carlo estimation and the adaptive control variate method, respectively. The values α_{Naive} and α_{ACV} denote, respectively, the estimated steady-state means obtained from the naïve and adaptive estimators. Similarly, σ^2_{Naive} and σ^2_{ACV} are the estimated time average variance constants. The sixth columns in both tables show that our adaptive estimators significantly outperform the naïve estimators, yielding smaller variances for every value of ϕ.

We see that the naïve variance σ^2_{Naive} steadily increases with ϕ. In fact, we expect this to happen, because larger values of ϕ increase both the autocorrelation in, and scale of, the process X. However, the adaptive variance σ^2_{ACV} does not behave in the same way. This is because the performance of our adaptive estimator depends on our parameterization for $u(\cdot,\theta)$ as well as on our solution for the optimization problem $\mathscr{P}_m$. In particular, our technique for solving $\mathscr{P}_m$ is only guaranteed to find a locally optimal point. As a result, the upward trend in the variance is not clear from σ^2_{ACV}.

7 CONCLUSIONS

We have extended adaptive control variates to the regenerative steady-state simulation context. Limited numerical experiments show that the method is usable and can significantly outperform standard Monte Carlo. The performance depends on the parameterization used for the control variate as well as the quality of the optimization algorithm used.

In terms of future research, there is the matter of how to divide the total computational budget between estimating the optimal parameter value θ^* and estimating the steady-state mean α. This allocation is expressed in the choice of sample sizes m and n. Kim and Henderson (2007) argue that, in finite-horizon simulation, the optimal choice of m in the optimization stage is of the order of the square root of the total budget. We expect a similar result to hold in the regenerative setting. Also, we used a regenerative estimator of the time-average variance constant. It can be difficult to identify the regeneration times of the underlying process, so we plan to consider other variance estimators based on methods related to batch means.

ACKNOWLEDGMENTS

This research was partially supported by NSF Grant DMI 0400287.

REFERENCES

Asmussen, S. 1992. Queueing simulation in heavy traffic. *Mathematics of Operations Research* 17:84–111.

Athreya, K. B., and P. Ney. 1978. A new approach to the limit theory of recurrent Markov chains. *Transactions of the American Mathematical Society* 245:493–501.

Bastin, F., C. Cirillo, and P. L. Toint. 2007. Convergence theory for nonconvex stochastic programming with an application to mixed logit. *Mathematical Programming B* 108:207–234.

Glynn, P. W., and D. L. Iglehart. 1993. Conditions for the applicability of the regenerative method. *Management Science* 39:1108–1111.

Glynn, P. W., and W. Whitt. 1989. Indirect estimation via $l = \lambda w$. *Operations Research* 37 (1): 82–103.

Henderson, S. G., and P. W. Glynn. 2001. Regenerative steady-state simulation of discrete event systems. *ACM Transactions on Modeling and Computer Simulation* 11:313–345.

Henderson, S. G., and P. W. Glynn. 2002. Approximating martingales for variance reduction in Markov process simulation. *Mathematics of Operations Research* 27:253–271.

Henderson, S. G., S. P. Meyn, and V. Tadić. 2003. Performance evaluation and policy selection in multiclass networks. *Discrete Event Dynamic Systems* 13:149–189. Special issue on learning and optimization methods.

Iglehart, D. L., and P. A. W. Lewis. 1979. Regenerative simulation with internal controls. *Journal of Association for Computing Machinery* 26:271–282.

Janson, S. 1983. Renewal theory for m-dependent variables. *Annals of Probability* 11:558–568.

Juneja, S., and P. Shahabuddin. 2006. Rare event simulation techniques: An introduction and recent advances. In *Simulation*, ed. S. G. Henderson and B. L. Nelson, Handbooks in Operations Research and Management Science, 291–350. Amsterdam: Elsevier.

Kim, S., and S. G. Henderson. 2007. Adaptive control variates for finite-horizon simulation. *Mathematics of Operations Research*. To appear.

Law, A. M. 2007. *Simulation modeling & analysis*. 4th ed. Boston: McGraw Hill.

L'Ecuyer, P. 1995. On the interchange of derivative and expectation for likelihood ratio derivative estimators. *Management Science* 41:738–748.

Meyn, S. P., and R. L. Tweedie. 1993. *Markov chains and stochastic stability*. London: Springer-Verlag.

Newton, N. J. 1994. Variance reduction for simulated diffusions. *SIAM Journal on Applied Mathematics* 54 (6): 1780–1805.

Nummelin, E. 1978. A splitting technique for Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete* 43 (4): 309–318.

Shapiro, A. 2004. Monte Carlo sampling methods. In *Stochastic Programming*, ed. A. Ruszczynski and A. Shapiro, Handbooks in Operations Research and Management Science. Elsevier.

Tadić, V. B., and S. P. Meyn. 2004. Adaptive Monte Carlo algorithms using control variates. Unpublished manuscript.

Whitt, W. 1989. Planning queueing simulations. *Management Science* 35:1341–1366.

AUTHOR BIOGRAPHIES

SUJIN KIM received her Ph.D. degree in Operations Research from Cornell University. She is currently Visiting Assistant Professor in the Department of Industrial Engineering at Purdue University. Her research interests include simulation methodology, simulation optimization, and applications in electric power and health service systems. Her e-mail address is ⟨sujin@purdue.edu⟩.

SHANE G. HENDERSON is an associate professor in the School of Operations Research and Information Engineering at Cornell University. He is the simulation area editor at *Operations Research*, and an associate editor for the *ACM Transactions on Modeling and Computer Simulation* and *Operations Research Letters*. He co-edited the

handbook *Simulation* as part of Elsevier's series of Handbooks in Operations Research and Management Science, and also co-edited the Proceedings of the 2007 Winter Simulation Conference. He likes cats but is allergic to them. His research interests include discrete-event simulation and simulation optimization, and he has worked for some time with emergency services. His web page can be found via ⟨http://www.orie.cornell.edu⟩.