European Journal of Operational Research 36 (1988) 360–370
North-Holland

Theory and Methodology

Antithetic-variate splitting for steady-sate simulations

Barry L. NELSON
Department of Industrial and Systems Engineering, The Ohio State University, Columbus, OH 43210, USA

Abstract: Obtaining precise estimates of parameters of infinite-horizon or steady-state simulations can be expensive because of the need to discard initial outputs to mitigate the effects of initial conditions. We consider splitting independent replications at the point of output truncation into dependent replications to reduce point estimator variance and/or simulation cost.

Keywords: Sampling statistics, simulation, stochastic processes, time series

1. Introduction

Variance reduction techniques (VRTs) are used to reduce the population variances of point estimators based on simulations of stochastic processes; for a survey of VRTs, see Nelson (1987a). Most VRTs are designed for terminating (sometimes called 'transient' or 'finite-horizon') processes. An example of a terminating process is a store that is open for a fixed number of hours each day. The natural design for such experiments is to sample many independent and identically distributed (i.i.d.) replications of the process. On the other hand, when simulating the large class of steady-state (sometimes called 'infinite-horizon') processes, it may be necessary to sample very long, and thus expensive, replications to overcome the effects of initial conditions. For example, to determine the long-run performance characteristics of a continuously operating production system, the simulation might be initialized with the system uncharacteristically empty and idle.

Variance reduction tailored to steady-state simulation experiments is largely unexplored. An exception is Kelton (1986), who examined the following problem: Consider performing a simulation experiment to estimate $\theta \equiv \lim_{i\to\infty} E[Y_{ij}]$, which is the same for all j, from the simulation output process $\{Y_{ij};\ i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, k\}$, where different replications (j) are independent, but outputs within a replication (i) may be neither independent nor identically distributed, and $n = km$ is fixed. Kelton quantified the effect of different sample allocation strategies (choices of k) on the variance of the overall sample mean

$$\bar{Y} \equiv k^{-1} \sum_{j=1}^{k} m^{-1} \sum_{i=1}^{m} Y_{ij} = k^{-1} \sum_{j=1}^{k} \bar{Y}_j \tag{1}$$

where $\bar{Y}_j$ is the sample mean of the j-th replication. When the outputs within a replication are positively correlated, Kelton showed that in some specific cases large k is preferable to small k for minimizing $\mathrm{Var}[\bar{Y}]$. However, if initial condition effects are reduced by discarding, say, the first d outputs from each replication, then one drawback to large k is that it may be prohibitively expensive to discard kd outputs.

Received August 1986; revised July 1987

In this paper we add an additional 'degree of freedom' to Kelton's approach by considering the possibility of splitting each independent replication at the point of output truncation into $s \geqslant 2$ possibly dependent (antithetic) replications. The goal is to reduce both the point estimator variance and the number of outputs discarded. To facilitate the analysis, we use the autoregressive order 1 (AR(1)) process as a surrogate for the simulation output process. Specifically, we assume

$$Y_{ij} = \theta + \phi(Y_{i-1,j} - \theta) + X_{ij} \tag{2}$$

where for fixed j, the $\{X_{ij};\ i = 1, 2, \ldots, m\}$ are i.i.d. random variables with mean 0 and variance $\sigma^2 < \infty$; we take $0 < \phi < 1$ so that the outputs within the j-th replication, $\{Y_{ij};\ i = 1, 2, \ldots, m\}$, are positively correlated. The AR(1) was introduced as a model for simulation output processes by Fishman (1972) because it shares many characteristics observed in these processes, including autocorrelations that decline exponentially with increasing lag. Kelton and Law (1984) and Snell and Schruben (1985) used the AR(1) surrogate to examine alternative methods for mitigating the effects of initial conditions, and Kelton (1986) used it in the study outlined above.

The original motive for investigating antithetic-variate splitting was the need to precisely estimate the steady-state expected number of customers in $GI^x/G^y/c/q$ bulk arrival-bulk service queues to validate a diffusion approximation (Lee, 1986). Mitigating the effects of initial conditions can require substantial output truncation for these queues. The antithetic-variate splitting approach with $s = 2$ was extremely effective, in terms of both cost and variance reduction, in this application. In Section 5 we argue that antithetic-variate splitting is useful for practical applications as well.

Sections 2 and 3 below give analytic results for applying antithetic-variate splitting ($s = 2$) to stationary and nonstationary AR(1) output processes. These two cases represent discarding enough initial output to consider the process to be in steady state, and directly quantifying the initial condition bias as a function of d, respectively. Section 4 extends the idea to $s \geqslant 2$ splits, and presents analytic and simulation results. Section 5 summarizes the results and gives some practical guidelines. An outline of how the results were derived is given in the Appendix.

2. Stationary output process

We first assume that sufficient outputs are discarded from the beginning of each replication to insure that the remainder of the replication is a covariance stationary process. For the AR(1) this means that for fixed j and all i, $E[Y_{ij}] = \theta$, $\gamma_0 \equiv \mathrm{Var}[Y_{ij}] = \sigma^2/(1 - \phi^2)$, and $\mathrm{Cov}[Y_{ij}, Y_{i+h,j}] = \phi^h \gamma_0$, for $h = 1, 2, \ldots, m - i$. The joint distribution of the $\{Y_{ij}\}$ depends on the distribution of the $\{X_{ij}\}$, but these are the moments of the steady-state distribution regardless. We do not make the common assumption that the $\{X_{ij}\}$ are normally distributed, but only that the output process is covariance stationary.

Let d be the number of outputs discarded from the beginning of each replication; if the number of outputs discarded is a random variable, then let d be its expected value. In this section, we do not include the discarded outputs as part of the sampling budget $n = km$, since d is not under our control. In the next section, where we consider nonstationary output processes, d is treated as a decision variable and its effect on both bias and variance is quantified.

Let the d-th output from the j-th replication be denoted Y_{0j}, which has mean θ and variance γ_0 by the assumptions above. When all k of the replications are independent, Kelton (1986) showed that

$$\mathrm{Var}[\bar{Y}] = \frac{(m - m\phi^2 - 2\phi + 2\phi^{m+1})\sigma^2}{km^2(1-\phi)^3(1+\phi)} \equiv \frac{\eta}{k}.$$

The cost of the experiment is $km + kd = n + kd$, where we take the cost of obtaining a single output Y_{ij} to be 1. These are the baseline results against which we compare new procedures. We assume throughout that n is divisible by k.

When the $\{X_{ij}\}$ are i.i.d. for all i and j we get the result above. We now consider the possibility of using antithetic-variate sampling to cause pairs of replications to be dependent. For the AR(1) surrogate process we represent this possibility by assuming

$$\mathrm{Corr}[X_{i,2j-1}, X_{h,2j}] = \begin{cases} \rho & \text{if } i = h, \\ 0 & \text{otherwise,} \end{cases}$$

where $-1 \leqslant \rho \leqslant 0$ and $j = 1, 2, \ldots, k/2$. For fixed j, the $\{X_{ij};\ i = 1, 2, \ldots, m\}$ remain i.i.d. For the moment, we are not concerned with how the dependence between input processes (the $\{X_{ij}\}$ here) is induced, but we will return to that issue in Section 4. However, inducing dependence between input processes to realize dependence between output processes is typical of the way antithetic-variate sampling is done.

There are two cases to consider: (1) we independently initialize k replications, meaning $\{Y_{01}, \ldots, Y_{0k}\}$ are i.i.d., then induce dependence between pairs of replications following truncation, or (2) we initialize $k/2$ independent replications, but at the point of truncation we split each one into $s = 2$ dependent replications of length m starting at the same point. We can represent case 2 by taking $Y_{0,2j-1} = Y_{0,2j}$, where $j = 1, 2, \ldots, k/2$. The point estimator is still the overall sample mean (1), but we denote it $\tilde{Y}$ for case 1 and $\hat{Y}$ for case 2. The estimator $\tilde{Y}$ is one form of the classical antithetic-variate estimator, while $\hat{Y}$ is the new antithetic-variate splitting estimator. Under the assumptions all three estimators $\overline{Y}$, $\tilde{Y}$, and $\hat{Y}$ are unbiased, so their mean squared errors (MSEs) are their variances, which are different.

Result 1. Under the assumptions above,

$$\mathrm{Var}[\tilde{Y}] = \frac{(1+\rho)\eta}{k} - \frac{\phi^2(1-\phi^m)^2\rho\sigma^2}{km^2(1-\phi)^3(1+\phi)},$$

$$\mathrm{Var}[\hat{Y}] = \frac{(1+\rho)\eta}{k} + \frac{\phi^2(1-\phi^m)^2(1-\rho)\sigma^2}{km^2(1-\phi)^3(1+\phi)}.$$

When $-1 \leqslant \rho \leqslant 0$, we have $(1+\rho)\eta/k \leqslant \eta/k$; however, for both estimators the second term on the right is nonnegative. While $\mathrm{Var}[\tilde{Y}] \leqslant \mathrm{Var}[\hat{Y}]$, the cost of $\tilde{Y}$ is $n + kd$, the same as $\overline{Y}$, while the cost of $\hat{Y}$ is $n + kd/2$. (We do not consider the potential savings from generating fewer random numbers to be significant.) Thus, to make a fair comparison among $\overline{Y}$, $\tilde{Y}$, and $\hat{Y}$ we need to make the sampling budget of $\hat{Y}$ equal to $n + kd$. There are at least two approaches: (1) increase the length of each of the replications from m to $m' = m + d/2$, or (2) increase the number of split replications from $k/2$ to $k'/2 = k(m+d)/(2m+d)$. Of course, m' and $k'/2$ would be rounded down to the nearest integer. Increasing the length of each replication results in all of the savings from splitting contributing to $\hat{Y}$, while increasing the number of replications results in some of the savings being discarded. In all of the cases considered, the first approach leads to smaller point estimator variance, so we only discuss increasing the replication length from here on.

Figure 1 shows the effect on $\sqrt{\mathrm{Var}[Q]}$, for $Q = \overline{Y}$, $\tilde{Y}$, and $\hat{Y}$, of different k when $\phi = 0.9$, $\rho = -0.5$, $n = 4000$, and $d = 0$, 50 or 100. As d gets larger, meaning that more outputs must be discarded before reaching steady-state conditions, $\hat{Y}$ begins to dominate $\tilde{Y}$, and both dominate $\overline{Y}$. For antithetic-variate splitting the replications must be long enough (k small enough) or d large enough so that the effect of the induced negative dependence between pairs of split replications can overcome the positive dependence from starting them at the same point. Figure 2 shows $\sqrt{\mathrm{Var}[Q]}$ with the same AR(1) parameters, but d fixed at 50 and $n = 1000$ or 8000. For $\tilde{Y}$, the effect of starting the dependent pairs of replications at independent, randomly selected points becomes more pronounced as k increases. However, as the budget n is increased this effect is diminished.

3. Nonstationary output process

In this section, we explicitly model initial condition effects by starting all replications of the AR(1) surrogate process at a fixed initial state, y_0. Thus, the output process is no longer stationary. The length of a replication is m, with the first $d < m$ outputs discarded, and the total sampling budget is $n = km$. In this section the discarded outputs are part of the total sampling budget since d will be under our control. The point estimator is the truncated sample mean

$$\overline{Y}(d) \equiv k^{-1} \sum_{j=1}^{k} (m-d)^{-1} \sum_{i=d+1}^{m} Y_{ij} = k^{-1} \sum_{j=1}^{k} \overline{Y}_j(d) \quad (3)$$

where $\overline{Y}_j(d)$ is the sample mean of the j-th truncated replication. When all k of the replications are independent, Kelton and Law (1984) showed that

$$\mathrm{Var}[\bar{Y}(d)] = \frac{\sigma^2}{k(m-d)(1-\phi)^2} \cdot \left[1 - \frac{\phi(1-\phi^{m-d})(2+\phi^{2d+1}(1-\phi^{m-d}))}{(m-d)(1-\phi^2)}\right] \equiv \frac{\eta(d)}{k}.$$

We again consider two estimators analogous to the previous section: (1) $\tilde{Y}(d)$, the classical antithetic-variate estimator, where we induce dependence between pairs of replications and discard d outputs from the beginning of each one, and (2) $\hat{Y}(d)$, the antithetic-variate splitting estimator, where we initialize $k/2$ independent replications, but at the point of truncation split each one into $s = 2$ dependent replications of length $m - d$. Functionally, both of the estimators are still the truncated sample mean. We can represent $\hat{Y}(d)$ by taking $Y_{d,2j-1} = Y_{d,2j}$, where $j = 1, 2, \ldots, k/2$. The important differences from Section 2 are that dependence is induced beginning from the initial state for the classical antithetic-variate estimator, rather than after trunction, and that we explicitly account for the effect of d on the bias and the

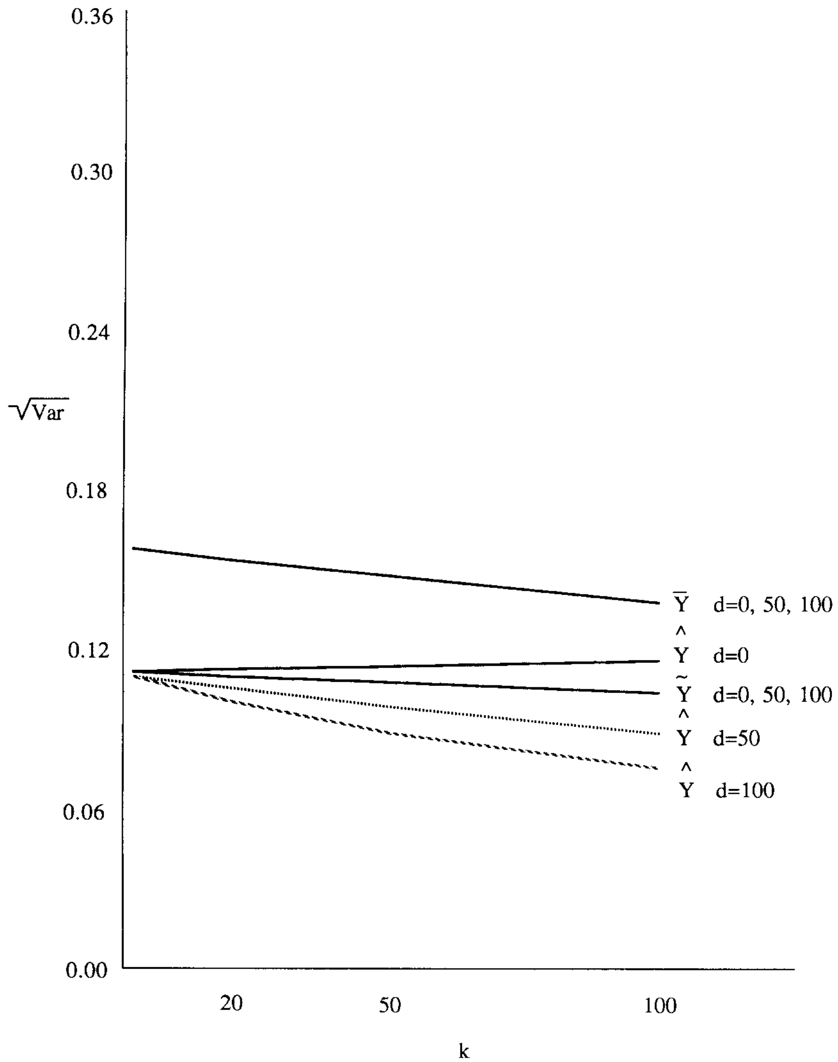

Figure 1. $\sqrt{\mathrm{Var}}$ as a function of the number of independent replications k and the truncation amount d for the stationary case with $\phi = 0.9$, $\rho = -0.5$, $n = 4000$ and $\sigma^2 = 1$

variance of all three estimators, rather than assuming that d is large enough to approximate steady-state conditions.

Result 2. Under the assumptions stated above, $\mathrm{Var}[\tilde{Y}(d)] = (1+\rho)\eta(d)/k$, and

$$\mathrm{Var}\big[\hat{Y}(d)\big] = \frac{(1+\rho)\eta(d)}{k} + \frac{\phi^2(1-\phi^{2d})(1-\phi^{m-d})^2(1-\rho)\sigma^2}{k(m-d)^2(1-\phi)^3(1+\phi)}.$$

If $-1 \leqslant \rho \leqslant 0$, then $\mathrm{Var}[\tilde{Y}(d)] \leqslant \mathrm{Var}[\bar{Y}(d)]$; for $\mathrm{Var}[\hat{Y}(d)]$ the second term on the right is positive. However, $\bar{Y}(d)$ and $\tilde{Y}(d)$ both cost n, while $\hat{Y}(d)$ costs $k(m-d)+kd/2 = n - kd/2$. Thus, to make a fair comparison we consider making the sampling budget of $\hat{Y}(d)$ equal to n by increasing the length of each replication to $m' = m + d/2$. Again, this approach dominates increasing the number of split replications in terms of point estimator variance in the cases considered. An additional benefit of increasing the replication length is discussed below.

Figure 3 shows the effect on $\sqrt{\mathrm{Var}[Q(d)]}$, for $Q(d) = \bar{Y}(d)$, $\tilde{Y}(d)$, and $\hat{Y}(d)$, of different k when $\phi = 0.9$, $\rho = -0.5$, $n = 16000$, and $d = 0$ or 125. The variance of all three estimators increases

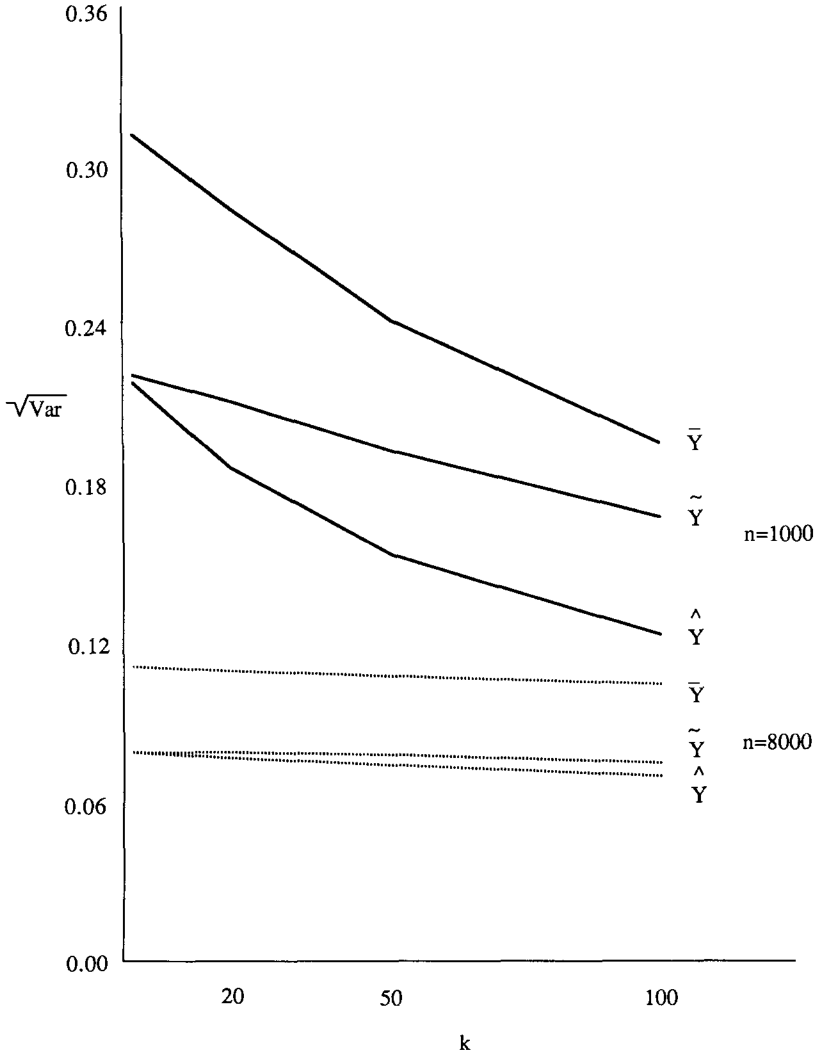

Figure 2. $\sqrt{\mathrm{Var}}$ as a function of the number of independent replications k and the sampling budget n for the stationary case with $\phi = 0.9$, $\rho = -0.5$, $d = 50$ and $\sigma^2 = 1$

with increasing d and k, but $\sqrt{\mathrm{Var}[\hat{Y}(d)]}$ does not increase as rapidly. Figure 4 shows the effect on $\sqrt{\mathrm{Var}[Q(d)]}$ of increasing n when d is fixed at 50. The relative benefit of antithetic-variate splitting is greater when the sampling budget is tight.

If we do not equate sampling budgets, then all three estimators $\bar{Y}(d)$, $\tilde{Y}(d)$, and $\hat{Y}(d)$ are identically biased. The bias is (Kelton and Law, 1984)

$$E[Q(d)] - \theta = \frac{(y_0 - \theta)\phi^{d+1}(1 - \phi^{m-d})}{(m-d)(1-\phi)}.$$

The bias decreases with increasing d and increasing m. Thus, increasing the length of each replication in $\hat{Y}(d)$ to $m' = m + d/2$ makes the antithetic-variate splitting estimator less biased than $\bar{Y}(d)$ and $\tilde{Y}(d)$ for the same budget, decreasing its MSE relative to the other estimators. Increasing the number of split replications in $\hat{Y}(d)$ has no effect on the bias.

4. Multiple splitting

In Sections 2 and 3 we considered splitting $k/2$ independent replications at the point of truncation into $s = 2$ possibly dependent replications. The natural next step is to consider $s \geqslant 2$ splits on each of k/s independent replications. Suppose first that, with the exception of starting from the same point, the split replications are also independent. Then for the AR(1) process we have

Result 3. For $s \geqslant 2$ independent splits of k/s independent replications

$$\mathrm{Var}[\hat{Y}] = \frac{\eta}{k} + \frac{(s-1)\phi^2(1-\phi^m)^2\sigma^2}{km^2(1-\phi)^3(1+\phi)},$$

$$\mathrm{Var}[\hat{Y}(d)] = \frac{\eta(d)}{k} + \frac{(s-1)\phi^2(1-\phi^{2d})(1-\phi^{m-d})^2}{k(m-d)^2(1-\phi)^3(1+\phi)}$$

where $\hat{Y}$ denotes the stationary case and $\hat{Y}(d)$ the nonstationary case. While the variances increase as s increases, the cost of both estimators goes down. The cost of $\hat{Y}$ is $n + kd/s$ compared to $n + kd$ for $\bar{Y}$, and the cost of $\hat{Y}(d)$ is $n - (s-1)kd/s$ compared to n for $\bar{Y}(d)$. Thus, the length of each replication could be increased to $m' = m + (s-1)d/s$ in both cases.

Extending the idea of dependent splits to $s > 2$ is not straightforward. It is well known that if X_{ij} has cumulative distribution function (cdf) F, then the pair $\{X_{i,2j-1}, X_{i,2j}\}$ can be generated with minimal possible covariance by letting $X_{i,2j-1} = F^{-1}(U)$ and $X_{i,2j} = F^{-1}(1-U)$, where $U \sim$ U(0, 1), the uniform distribution on the interval (0, 1). This is one method for realizing the negative correlation we assumed for the AR(1) process. However, there is no corresponding general results for making $\{X_{ij};\ j = 1, 2, \ldots, s\}$ negatively correlated.

One possibility is the 'rotation sampling' scheme proposed by Fishman and Huang (1983). In rotation sampling, we let $X_{ij} = F^{-1}(U * (j-1)/s)$, $j = 1, 2, \ldots, s$, where

$$U * \frac{j-1}{s} = \begin{cases} U + (j-1)/s & \text{if } 0 \leqslant U < 1 - (j-1)/s, \\ U + (j-1)/s - 1 & \text{if } 1 - (j-1)/s \leqslant U < 1. \end{cases}$$

For several distributions F (e.g. exponential), rotation sampling achieves the minimal possible average correlation among the $\{X_{ij}:\ j = 1, 2, \ldots, s\}$, and for many other distributions it achieves greatly reduced average correlation. Results in Fishman and Huang suggest that rotation sampling will become more effective as s increases. Unfortunately, these results do not permit derivation of explicit expressions for the variance of alternative estimators using the AR(1) surrogate process.

To investigate the performance of both antithetic ($s = 2$) and rotation ($s > 2$) splitting, we simulated GI/G/1 queues with Weibull distributed interarrival and service times. Thus, the M/M/1 queue is a special case. The performance measures considered were θ_1, the steady-state expected delay in the queue, and θ_2, the steady-state expected number of customers in the system; θ_1 is the parameter of a discrete-time process, like θ for the AR(1), while θ_2 is the parameter of a continuous-time process. Unlike the AR(1), a replication j of the GI/G/1 process is not a function of a single input process $\{X_{ij};\ i = 1, 2, \ldots, m\}$, but rather two input processes: the interarrival times and service times. Thus, synchronization of random number streams is important to achieve the desired negative dependence. A separate stream of U(0, 1) random numbers for each input process works well in this simple queue, but in complex simulations more sophistication may be required. This point is discussed further in the next section.

The estimators $\bar{Y}(d)$, $\tilde{Y}(d)$, and $\hat{Y}(d)$ were investigated with $m' = m + (s-1)d/s$ for $\hat{Y}(d)$. In the experiments, each independent replication was initialized with the queue empty and idle. At the point of truncation, which was dth customer waiting time in the queue, the entire state of the simulation, including the current random number seeds for each input process, was recorded. This information was used to restart split replications.

The sample variances were computed using the formula

$$\hat{\text{Var}}[Q(d)] = (k/s(k/s-1))^{-1} \times \sum_{j=1}^{k/s} \{Q_j(d) - Q(d)\}^2$$

for $Q(d) = \bar{Y}(d)$, $\tilde{Y}(d)$ or $\hat{Y}(d)$, $s = 1$ for independent sampling, $s = 2$ for antithetic variates and antithetic splitting, and $s = 4$ or 8 for rotation splitting; $Q_j(d)$ denotes the average of all outputs in an independent replication. Thus, when antithetic sampling is used, all dependent replications are averaged to form $\tilde{Y}_j(d)$ or $\hat{Y}_j(d)$ for $j =$

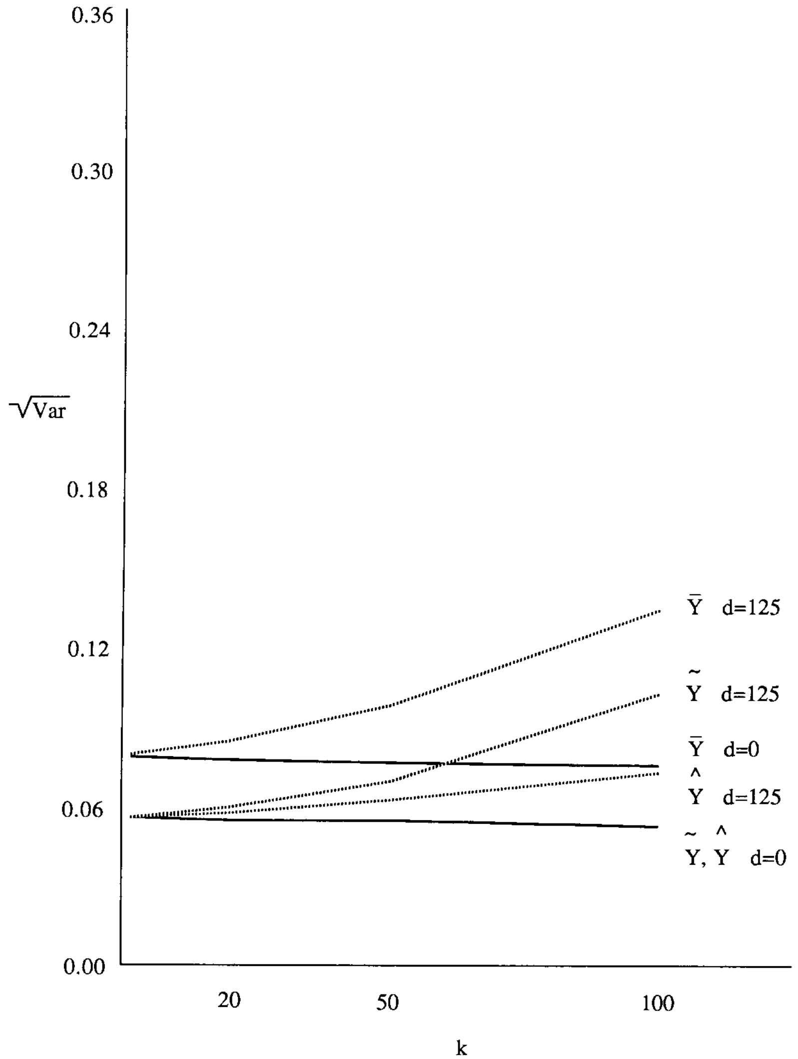

Figure 3. $\sqrt{\text{Var}}$ as a function of the number of independent replications k and the truncation amount d for the nonstationary case with $\phi = 0.9$, $\rho = -0.5$, $n = 16000$ and $\sigma^2 = 1$

1, 2,..., k/s. Notice that antithetic sampling results in a loss of degrees of freedom from $k-1$ to $k/s-1$.

Tables 1 and 2 show results for M/M/1 queues with traffic intensities 0.5 and 0.8, respectively. Tables 3 and 4 also show results for traffic intensities 0.5 and 0.8, but with Weibull parameters such that the interarrival and service time distributions are more bell shaped (the specific parameters are given in the tables). The basic experiment had a sampling budget $n = 16000$ and trunction point $d = 240$; the value of d was selected based on results in Kelton and Law (1984) for M/M/1 queues. The basic experiment was replicated 25 times, each replication providing an estimated variance using the formula above. The results reported in the tables are the average and standard error of the average of the 25 replications.

Variance reductions of more than 50% are clearly achievable. Rotation splitting was generally more effective for $s = 8$ splits, and sometimes variance increases relative to $s = 2$ splits occurred for $s = 4$. Also notice that larger variance reductions were obtained for estimating θ_2, the expected number of customers in the system. Although these results are encouraging, they are not definitive since this is only one experiment. However, antithetic-variate splitting does seem to be more

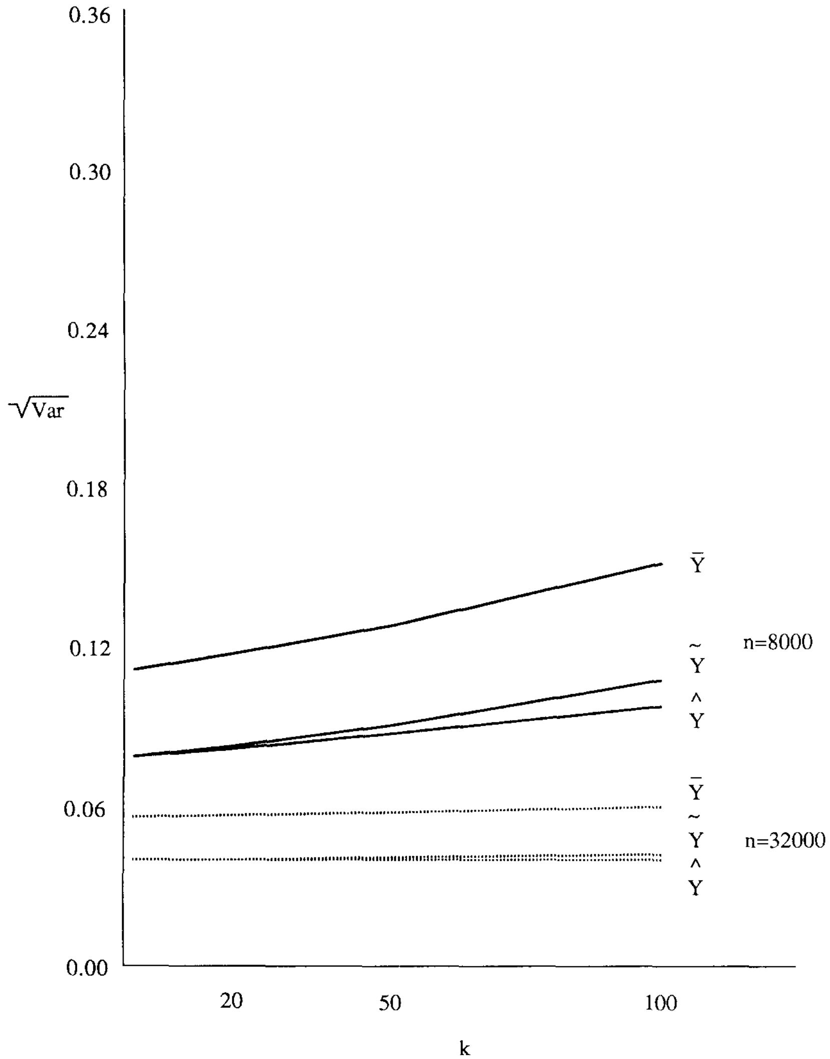

Figure 4. $\sqrt{\mathrm{Var}}$ as a function of the number of independent replications k and the sampling budget n for the nonstationary case with $\phi = 0.9$, $\rho = -0.5$, $d = 50$ and $\sigma^2 = 1$

Table 1
Results for 25 replications of GI/G/1 simulation with $n=16000$, $k=40$, $d=240$, and Weibull parameters $\alpha_1=1.0$, $\beta_1=1.0$, $\alpha_2=1.0$, $\beta_2=0.5$ (standard errors in parentheses) [a]

	$\hat{\mathrm{Var}}[\bar{Y}(d)]$	$\hat{\mathrm{Var}}[\tilde{Y}(d)]$	$\hat{\mathrm{Var}}[\hat{Y}(d)]$		
			$s=2$	$s=4$	$s=8$
θ_1	9.35	8.51	5.20	4.30	3.72
	(0.62)	(0.90)	(0.55)	(0.55)	(0.54)
θ_2	17.11	11.29	6.62	6.43	4.97
	(1.06)	(1.29)	(0.78)	(0.90)	(0.71)

[a] All results $\times 10^{-4}$.

Table 2
Results for 25 replications of GI/G/1 simulation with $n=16000$, $k=40$, $d=240$, and Weibull parameters $\alpha_1=1.0$, $\beta_1=1.0$, $\alpha_2=1.0$, $\beta_2=0.8$ (standard errors in parentheses) [a]

	$\hat{\mathrm{Var}}[\bar{Y}(d)]$	$\hat{\mathrm{Var}}[\tilde{Y}(d)]$	$\hat{\mathrm{Var}}[\hat{Y}(d)]$		
			$s=2$	$s=4$	$s=8$
θ_1	1.14	0.84	0.75	0.92	0.57
	(0.15)	(0.10)	(0.11)	(0.16)	(0.10)
θ_2	1.49	1.05	0.87	1.02	0.64
	(0.20)	(0.12)	(0.14)	(0.19)	(0.11)

[a] All results $\times 10^{-1}$.

Table 3
Results for 25 replications of GI/G/1 simulation with $n=16000$, $k=40$, $d=240$, and Weibull parameters $\alpha_1=2.0$, $\beta_1=1.8482$, $\alpha_2=2.0$, $\beta_2=0.5642$ (standard errors in parentheses) [a]

	$\hat{\mathrm{Var}}[\bar{Y}(d)]$	$\hat{\mathrm{Var}}[\tilde{Y}(d)]$	$\hat{\mathrm{Var}}[\hat{Y}(d)]$		
			$s=2$	$s=4$	$s=8$
θ_1	3.03	3.01	1.64	1.40	1.14
	(0.18)	(0.28)	(0.13)	(0.17)	(0.16)
θ_2	11.54	1.32	0.74	3.32	1.77
	(0.56)	(0.12)	(0.04)	(0.38)	(0.22)

[a] All results $\times 10^{-6}$.

Table 4
Results for 25 replications of GI/I/1 simulation with $n=16000$, $k=40$, $d=240$, and Weibull parameters $\alpha_1=2.0$, $\beta_1=1.8284$, $\alpha_2=2.0$, $\beta_2=0.9027$ (standard errors in parentheses) [a]

	$\hat{\mathrm{Var}}[\bar{Y}(d)]$	$\hat{\mathrm{Var}}[\tilde{Y}(d)]$	$\hat{\mathrm{Var}}[\hat{Y}(d)]$		
			$s=2$	$s=4$	$s=8$
θ_1	4.45	3.97	2.27	2.01	2.05
	(0.22)	(0.26)	(0.15)	(0.19)	(0.28)
θ_2	6.23	1.63	0.92	2.08	1.37
	(0.28)	(0.11)	(0.06)	(0.18)	(0.19)

[a] All results $\times 10^{-5}$.

effective than classical antithetic variates, as expected from the AR(1) results. For an example of a successful application of rotation sampling in the simulation of Markov chains, see Fishman (1983ab).

5. Conclusions

The results in this paper indicate that antithetic-variate splitting can be more effective than the classical antithetic-variate estimator in steady-state simulations where antithetic sampling is effective. As in all applications of antithetic sampling, it is essential to design the simulation experiment so that negative correlation induced between input processes is preserved in the output processes. There are three key factors in the design: The method of input process generation, the synchronization of the random number streams, and the monotonicity of the input–output transformation. We discuss these factors briefly as they pertain to antithetic-variate splitting; for in-depth discussions see Kleijnen (1974) and Bratley, Fox and Schrage (1983).

Dependence induction, such as antithetic-variate sampling, usually requires the inverse cdf method of random-variate generation as described in Section 4. Other variate generation methods that are employed when no closed-form expression for the inverse cdf exists are not designed to induce dependence between input processes by manipulating the random number streams. While these other methods are fast relative to numerically inverting the cdf, the variance reduction achieved by the antithetic-splitting estimator may make numerical inversion worth the extra effort. Recently, Schmeiser and Kachitvichyanukul (1986) proposed fast, noninverse cdf algorithms that permit dependence induction.

The antithetic-variate splitting estimator facilitates synchronization of the random number streams. Unlike the AR(1) surrogate process used here, the dependence induced between outputs in practical simulations may decrease as the replication length increases because of lack of synchronization. Inducing dependence after the truncation point helps to preserve as much dependence as possible in the outputs of interest. Thus, the antithetic-splitting estimator may actually perform better than indicated by the AR(1) results in practical problems relative to the classical antithetic-variate estimator.

Factors that are important in deciding how to apply antithetic-variate splitting in practical problems include the total sampling budget, the degree of negative correlation that can be induced, and the degree of positive correlation within each replication. Unfortunately, the relevant correlations are rarely known in practice. Also, the experiment design may need to reflect other criteria in addition to point estimator variance. For example, if interval estimates are desired then degrees of freedom, point estimator bias, and point estimator normality play a role (Nelson, 1987b). In its favor, the antithetic-splitting estimator is less biased than either the sample mean or the antithetic-variate estimator for the same number of replications and truncation point.

Directly extrapolating the AR(1) results to practical simulations is problematical. Even if the optimal design (in terms of point estimator variance) is determined for the AR(1) process, we do not know how well the AR(1) represents the simulation. The AR(1) and GI/G/1 simulation results suggest the following tentative recommendation: If initial condition effects are moderate with respect to the available budget, use antithetic splitting ($s = 2$). If initial condition effects are severe so that d is large, use rotation splitting with $s \geqslant 8$. Overall, the antithetic-variate splitting approach appears promising and practical for steady-state simulations requiring large initial truncations.

We have used the term 'splitting' in a sense not too different from the VRT by the same name (Kahn, 1956; Nelson, 1987a). In classical splitting, when the state of a dynamic simulation enters an important subset of states (e.g. a subset of states from which a rare event is likely to occur), s independent replications are started and their results averaged. The idea is to precisely estimate a conditional expection; i.e. the expected system response given the system is in a certain subset of states. Antithetic-variate splitting could further sharpen the estimate, and all the preceding results apply directly.

Appendix

Kelton and Law (1984) show that for the AR(1) process (2)

$$\bar{Y}_j(d) = \theta + \frac{(y_{0j} - \theta)\phi^{d+1}(1 - \phi^{m-d})}{(m-d)(1-\phi)} + \left\{(1-\phi^{m-d})\Sigma_{i=1}^{d+1}\phi^{d-i+1}X_{ij} + \Sigma_{i=d+2}^{m}(1-\phi^{m-i+1})X_{ij}\right\} \times \{(m-d)(1-\phi)\}^{-1}.$$

The key to the results in this paper is to notice that if $\mathrm{Corr}[X_{i,2j-1}, X_{i,2j}] = \rho$, then $\mathrm{Var}[(X_{i,2j-1} + X_{i,2j})/2] = (1+\rho)\sigma^2/2$. All the results can be derived by direct calculation of the appropriate variance, noting which X_{ij} terms are common, independent, and dependent in the calculation. For the stationary cases, we take Y_{0j} to be a random variable with mean θ and variance γ_0.

Acknowledgment

The author gratefully acknowledges the assistance of J.P.C. Kleijnen, Editorial Associate, and two anonymous referees who suggested an improved approach for comparing alternative estimators. The author benefited from discussions with Charles H. Reilly, Jane M. Fraser, and John B. Neuhardt of the Ohio State University.

References

Bratley, P., Fox, B.L., and Schrage, L.E. (1983), *A Guide to Simulation*, Springer, New York.

Fishman, G.S. (1972), "Bias considerations in simulation experiments", *Operations Research* 20, 785–790.

Fishman, G.S. (1983a), "Accelerated accuracy in the simulation of Markov chains", *Operations Research* 31/3, 466–487.

Fishman, G.S. (1983b), "Accelerated convergence in the simulation of countably infinite state Markov chains", *Operations Research* 31/6, 1074–1089.

Fishman, G.S., and Huang, B.D. (1983), "Antithetic variates revisited", *Communications of the ACM* 26/11, 964–971.

Kahn, H. (1956), "Use of different Monte Carlo sampling techniques", in: H. Meyer (ed.), *Symposium on Monte Carlo Methods*, Wiley, New York.

Kelton, W.D. (1986), "Replication splitting and variance for simulating discrete-parameter stochastic processes", *Operations Research Letters* 4/6, 275–279.

Kelton, W.D., and Law, A.M. (1984), "An analytical evaluation of alternative strategies in steady-state simulation", *Operations Research* 32/1, 169–184.

Kleijnen, J.P.C. (1974), *Statistical Techniques in Simulation, Part I*, Dekker, New York.

Lee, H. (1986), "A diffusion approximation for multi-server finite-capacity bulk queues", unpublished Ph.D. dissertation, Department of Industrial and Systems Engineering, The Ohio State University.

Nelson, B.L. (1987a), "A perspective on variance reduction in dynamic simulation experiments", *Communications in Statistics* B16/2, 385–426.

Nelson, B.L. (1987b), "Some properties of simulation interval estimators under dependence induction", *Operations Research Letters* 6/4, 169–176.

Schmeiser, B., and Kachitvichyanukul, V. (1986), "Correlation induction without the inverse transform", Wilson, Henrikson and Roberts (eds.), *Proceedings of the Winter Simulation Conference*, 266–274.

Snell, M., and Schruben, L. (1985), "Weighting simulation data to reduce initialization effects", *IIE Transactions* 17/4, 354–363.