

INTEGRATED VARIANCE REDUCTION STRATEGIES

Athanassios N. Avramidis

American Airlines Decision Technologies
P.O. Box 619616, M.D. 4462
Dallas/Fort Worth Airport, TX, 75261, U.S.A.

James R. Wilson

Department of Industrial Engineering
North Carolina State University
Raleigh, North Carolina 27695-7906, U.S.A.

ABSTRACT

In this paper we develop strategies for integrating certain well-known variance reduction techniques to estimate a mean response in a finite-horizon simulation experiment. Our building blocks are the techniques of conditional expectation, correlation induction, and control variates. Under some mild assumptions, we explain how each integrated strategy yields a smaller response variance than its constituent variance reduction techniques yield individually. We also provide asymptotic variance comparisons for integrated strategies involving the correlation-induction technique of Latin hypercube sampling. Our Monte Carlo results show that in the simulation of stochastic activity networks, large efficiency gains can be achieved by using these integrated variance reduction strategies.

1 INTRODUCTION

A diversity of variance reduction techniques (VRTs) have been developed to improve the efficiency of simulations—that is, to reduce the computing effort necessary to obtain some specified precision. For a survey of VRTs, see Wilson (1984) and Nelson (1987). There has been relatively little work on integrating these well-known VRTs into an overall variance reduction strategy that can exploit various sources of efficiency improvement simultaneously. Moreover, few attempts have been made either to quantify the efficiency improvement resulting from integrated variance reduction strategies or to establish general conditions under which these integrated strategies are preferable to direct simulation or standard VRTs used alone. See Avramidis and Wilson (1992) for a review of the literature on integrating variance reduction techniques.

This paper is organized as follows. In Section 2 we define our notation, formulate the variance reduction problem, and review some basic VRTs. In Section 3

we develop and study some integrated strategies for variance reduction. In Section 4 we provide asymptotic variance comparisons for integrated strategies involving the correlation-induction technique of Latin hypercube sampling. Section 5 contains the results of a Monte Carlo study designed to gauge the efficiency gains due to the integrated variance reduction strategies in the context of activity network simulation. In Section 6 we summarize the main findings of this work, and we recommend directions for future research.

2 NOTATION AND BACKGROUND

The problem is to estimate the expected value θ of a target response variable Y . This includes estimating noncentral moments and probabilities, but it does not include estimating, for example, central moments or quantiles. We assume throughout that $E[Y^2] < \infty$ so that $\theta \equiv E[Y]$ and $\sigma_Y^2 \equiv \text{Var}(Y)$ are both finite. The response is assumed to have the form $Y = f(V_1, \dots, V_p)$, where the function $f(\cdot)$ has a fixed number of inputs; and the *input random variates* $\{V_1, \dots, V_p\}$ have a known probabilistic structure. By this we merely mean that we have a way of generating the random vector $\mathbf{V} \equiv (V_1, \dots, V_p)$ so that it has the correct distribution. The input random variates are generated as $\mathbf{V} = \mathbf{H}(\mathbf{U})$, where: $\mathbf{U} \equiv (U_1, \dots, U_d)$ is composed of d independent *random numbers* that are uniformly distributed on the unit interval $(0, 1)$; and $\mathbf{H}(\cdot)$ is a *sampling plan*, corresponding to the random-variate generation scheme used. In some parts of this paper, it is convenient to view Y as a function of the input random variates, whereas elsewhere we prefer to view Y as a function of the input random numbers. In the latter situation, we write $Y = f[\mathbf{H}(\mathbf{U})] \equiv y(\mathbf{U})$. Throughout this paper, the word *function* will mean a Borel measurable function, taking either real scalar values or real vector values. Vectors and matrices will be denoted by

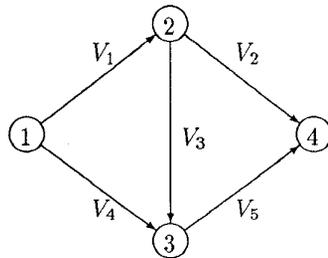


Figure 1: A Directed Network

boldface type.

EXAMPLE 1. Consider the directed network in Figure 1. The input random variates are $\{V_1, \dots, V_5\}$, where V_i is the (random) duration of arc i for $i = 1, \dots, 5$. Thus $p = 5$ in this example. Let

$$T = \max\{V_1 + V_2, V_1 + V_3 + V_5, V_4 + V_5\}$$

be the longest directed path from node 1 to node 4, and suppose the response is $Y = 1_{\{T \leq t\}}$, the indicator function of the event $\{T \leq t\}$ for a given cutoff time t . We assume that the random variates V_1, V_4 , and V_5 are mutually independent with known distributions and that the random vector (V_2, V_3) is independent of V_1, V_4 , and V_5 with a known bivariate distribution. Consider the following sampling plan:

$$\begin{aligned} V_1 &= H_1(U_1, U_2), & V_2 &= H_2(U_3), & V_3 &= H_3(U_3, U_4), \\ V_4 &= H_4(U_5), & V_5 &= H_5(U_6), \end{aligned} \tag{1}$$

where $\{U_1, \dots, U_6\}$ are random numbers and $H_1(\cdot), \dots, H_5(\cdot)$ are given functions that can be evaluated readily. Here we do not use the method of inversion to generate all input variates; instead V_1 is generated by some other method that requires two random numbers. Moreover, V_3 is generated conditional on V_2 ; and thus V_3 is also a function of two random numbers. With the sampling plan above, we have $d = 6$, and

$$y(u_1, \dots, u_6) = \begin{cases} 1 & \text{if } \max\{H_1(u_1, u_2) + H_2(u_3), \\ & H_1(u_1, u_2) + H_3(u_3, u_4) + H_5(u_6), \\ & H_4(u_5) + H_5(u_6)\} \leq t, \\ 0 & \text{otherwise. } \blacksquare \end{cases}$$

In a direct simulation experiment, we perform n independent replications that yield independent and

identically distributed (i.i.d.) observations of the target response $\{Y_i : i = 1, \dots, n\}$. The direct simulation estimator is the corresponding sample mean $\bar{Y}(n)$, which is unbiased and has variance $n^{-1}\sigma_Y^2$. The aim of variance reduction techniques is to identify an alternative estimator $\hat{\theta}(n)$ based on n replications (which are not necessarily i.i.d.) such that

$$E[\hat{\theta}(n)] = \theta, \quad \text{Var}[\hat{\theta}(n)] < \text{Var}[\bar{Y}(n)].$$

Even when the estimator $\hat{\theta}(n)$ is based on n dependent simulation runs, typically a central limit theorem (CLT) holds so that

$$n^{1/2}[\hat{\theta}(n) - \theta] \xrightarrow{\mathcal{D}} N(0, \sigma^2) \text{ as } n \rightarrow \infty, \tag{2}$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution and $N(\mu, \sigma^2)$ denotes a normal random variable with mean μ and variance σ^2 . We then say that $\hat{\theta}(n)$ has asymptotic mean θ and asymptotic variance parameter σ^2 .

Suppose we have two estimators $\hat{\theta}_1(n)$ and $\hat{\theta}_2(n)$ satisfying CLTs of the form (2) with respective variance parameters σ_1^2 and σ_2^2 such that $\sigma_1^2 \leq \sigma_2^2$. We then say that $\hat{\theta}_1$ asymptotically dominates $\hat{\theta}_2$. For any given finite replication count n , this does not guarantee that either the bias or the variance of $\hat{\theta}_1(n)$ has smaller magnitude than the corresponding characteristic of $\hat{\theta}_2(n)$. However, asymptotic dominance is a reasonable criterion for comparing estimators when it is difficult to obtain exact expressions for the bias and variance of each estimator at each sample size n . For simplicity, we will occasionally suppress the argument n in the discussion of alternative simulation-based estimators $\hat{\theta}_1, \hat{\theta}_2$ when no confusion can result from this usage.

2.1 Conditional Expectation

Suppose we can identify an auxiliary random vector \mathbf{X} such that we can evaluate the conditional expectation $h(\mathbf{x}) \equiv E[Y|\mathbf{X} = \mathbf{x}]$ analytically or numerically for each possible value of \mathbf{x} . Thus the random variable $Z \equiv h(\mathbf{X})$ is an alternative estimator of θ based on a single replication. From the results $\{\mathbf{X}_i : i = 1, \dots, n\}$ of n independent simulation runs, we compute the corresponding random sample $\{Z_i = h(\mathbf{X}_i) : i = 1, \dots, n\}$; and the conditional expectation (CE) estimator of θ is

$$\hat{\theta}_{\text{CE}}(n) \equiv n^{-1} \sum_{i=1}^n Z_i.$$

Since

$$E[Z] = E[E(Y|\mathbf{X})] = \theta$$

and

$$\sigma_Z^2 \equiv \text{Var}(Z) = \sigma_Y^2 - E[\text{Var}(Y|\mathbf{X})],$$

we see immediately that $\hat{\theta}_{\text{CE}}(n)$ is an unbiased estimator of θ and that $\text{Var}[\hat{\theta}_{\text{CE}}(n)] \leq \text{Var}[\bar{Y}(n)]$, with equality holding if and only if Y is a function of \mathbf{X} .

EXAMPLE 1 (CONTINUED). To evaluate the conditional expectation of $Y = 1_{\{T \leq t\}}$ analytically, we condition on $\mathbf{X} \equiv (X_1, X_2) \equiv (V_1, V_5)$. If $F_{2,3}(\cdot, \cdot)$ denotes the joint cumulative distribution function (c.d.f.) of the input random vector (V_2, V_3) , and if $F_4(\cdot)$ denotes the c.d.f. of V_4 , then it is easy to check that

$$\begin{aligned} h(\mathbf{x}_1, \mathbf{x}_2) &= \\ &= \Pr\{T \leq t | X_1 \equiv V_1 = \mathbf{x}_1, X_2 \equiv V_5 = \mathbf{x}_2\} \\ &= \Pr\{V_2 \leq t - \mathbf{x}_1, V_3 \leq t - \mathbf{x}_1 - \mathbf{x}_2, V_4 \leq t - \mathbf{x}_2\} \\ &= F_{2,3}(t - \mathbf{x}_1, t - \mathbf{x}_1 - \mathbf{x}_2) \cdot F_4(t - \mathbf{x}_2). \quad \blacksquare \quad (3) \end{aligned}$$

2.2 Correlation Induction

We give a general method for obtaining negatively correlated observations of an arbitrary random output W observed in the simulation. (Since we will apply the following development not only to the target response Y but also to other simulation-generated outputs, we let the symbol W denote a “generic” simulation output to which a correlation-induction strategy will be applied.) We view W as a function of the input random numbers:

$$W = w(U_j : j \in I_W),$$

where I_W is a subset of $\{1, \dots, d\}$, and the function $w(\cdot)$ is defined by the simulation code. In the sequel, we assume that the arguments of the function $w(U_j : j \in I_W)$ always occur in increasing order of the index j .

An easy-to-check condition that guarantees negative correlation induction is based on the notion of negative quadrant dependence proposed by Lehmann (1966). We say that the distribution of the bivariate random vector (A_1, A_2) is negatively quadrant dependent (n.q.d.) if

$$\begin{aligned} &\Pr\{A_1 \leq a_1, A_2 \leq a_2\} \\ &\leq \Pr\{A_1 \leq a_1\} \cdot \Pr\{A_2 \leq a_2\} \quad \text{for all } a_1, a_2. \end{aligned}$$

We exploit this concept in Result 1 below to provide the desired sufficient condition for induced negative correlations. Moreover, we use the concept of negative quadrant dependence to define a special class \mathcal{G} of distributions for the random-number inputs. Every distribution $G \in \mathcal{G}$ must have the following properties:

CI₁ For some $k \geq 2$, G is a k -variate distribution with univariate marginals that are uniform on the unit interval $(0, 1)$.

CI₂ Each bivariate marginal of G is n.q.d.

CI₃ All bivariate marginals of G are equal.

Throughout this paper, we let $G_0(k)$ denote the distribution of k mutually independent random numbers. It is clear that $G_0(k)$ satisfies conditions CI₁-CI₃ so that $G_0(k) \in \mathcal{G}$.

Using a k -variate distribution G selected from \mathcal{G} , we induce negative correlations between k replications of the simulation output W according to the following scheme. Let L_W denote an arbitrary subset of I_W consisting of the indices of the random-number inputs to $w(\cdot)$ that are used for correlation induction. We perform k dependent replications yielding outputs

$$W^{(i)} = w(U_j^{(i)} : j \in I_W), \quad i = 1, \dots, k, \quad (4)$$

where the input column vectors

$$\mathbf{U}_j \equiv [U_j^{(1)}, \dots, U_j^{(k)}]^T, \quad j \in I_W,$$

are sampled as follows:

- (a) For every index $j \in L_W$, the random vector \mathbf{U}_j has distribution G ;
- (b) For every index $j \in I_W - L_W$, the random vector \mathbf{U}_j has distribution $G_0(k)$; and
- (c) The column vectors $\mathbf{U}_1, \dots, \mathbf{U}_d$ are mutually independent.

Condition (a) specifies that we induce dependence between the outputs $\{W^{(i)} : i = 1, \dots, k\}$ by arranging a negative quadrant dependence between the j th random numbers sampled on each pair of replications, provided $j \in L_W$. Condition (b) specifies that for each $j \notin L_W$, the j th random number should be sampled independently on different replications. Finally condition (c) requires mutual independence of the random numbers used within the i th replication to generate the output $W^{(i)}$; and this guarantees that each $W^{(i)}$ has the correct distribution. We then define the average across the k replications

$$W_{\text{CI}}(G, L_W) \equiv k^{-1} \sum_{i=1}^k W^{(i)}, \quad (5)$$

where we make explicit the dependence of W_{CI} on the distribution G and the index set L_W where G applies. The dependence of W_{CI} on k is not shown explicitly, but is implicit through G .

The mean and variance of $W_{CI}(G, L_W)$ are easily derived. Clearly, for any $G \in \mathcal{G}$ and $L_W \subseteq I_W$, the statistic $W_{CI}(G, L_W)$ is an unbiased estimator of θ with variance

$$\begin{aligned} \text{Var}[W_{CI}(G, L_W)] &= \\ &= k^{-1} \text{Var}(W) \left[1 + (k-1)\rho_W(G, L_W) \right], \end{aligned}$$

where

$$\rho_W(G, L_W) \equiv \text{Corr}[W^{(1)}, W^{(2)}]. \quad (6)$$

The next result follows easily from Lemma 3 of Lehmann (1966).

RESULT 1. *If G satisfies condition CI_2 and $w(\cdot)$ is a monotone function of each argument with index in L_W , then $\text{Cov}[W^{(i)}, W^{(j)}] \leq 0$ for $i \neq j$, with equality holding if and only if $W^{(i)}$ and $W^{(j)}$ are independent.*

Thus $W_{CI}(G, L_W)$ has smaller variance than $\bar{W}(k)$, the average of k independent replications of W , whenever $w(\cdot)$ is a monotone function of each random-number input U_j with index $j \in L_W$; no assumption is needed with respect to the behavior of $w(\cdot)$ as a function of U_ℓ for $\ell \notin L_W$.

Using definitions (4) and (5), we formulate $\hat{\theta}_{CI}(G, n)$, the *correlation-induction* (CI) estimator of θ based on the k -variate distribution G and n simulation runs. Specifically, $\hat{\theta}_{CI}(G, n)$ is obtained by averaging $m = n/k$ i.i.d. replications of the statistic $Y_{CI}(G, L_Y)$, where we take $L_Y = I_Y = \{1, \dots, d\}$:

$$\hat{\theta}_{CI}(G, n) \equiv m^{-1} \sum_{j=1}^m \tilde{Y}_j,$$

where $m = n/k$ and $\{\tilde{Y}_j\}_{j=1}^m \stackrel{\text{i.i.d.}}{\sim} Y_{CI}(G, \{1, \dots, d\})$. To obtain a single observation of $Y_{CI}(G, \{1, \dots, d\})$, we average k negatively correlated responses, where all d random-number inputs are used for correlation induction. To simplify the exposition, we assume throughout the paper that n is an integral multiple of k . Next we review two important special cases of the method of correlation induction.

2.2.1. Antithetic Variates (AV). Here $k = 2$, and correlation is induced by using complementary random numbers so that we have

$$U_j^{(2)} = 1 - U_j^{(1)}, \quad j = 1, \dots, d.$$

We let G_{AV} denote the distribution of $[U_1^{(1)}, U_1^{(2)}]$ and we observe that $G_{AV} \in \mathcal{G}$.

2.2.2. Latin Hypercube Sampling (LHS). Here we sample in a stratified scheme from each of the marginal input distributions so that we have

$$U_j^{(i)} = \frac{\pi_j(i) - 1 + U_{ij}}{k}, \quad i = 1, \dots, k, \quad j = 1, \dots, d, \quad (7)$$

where

- (a) $\pi_1(\cdot), \dots, \pi_d(\cdot)$ are independent random permutations of the integers $\{1, \dots, k\}$; and
- (b) $\{U_{ij} : j = 1, \dots, d, i = 1, \dots, k\}$ are random numbers sampled independently of each other and of the permutations $\pi_1(\cdot), \dots, \pi_d(\cdot)$.

We let $G_{LH}(k)$ denote the distribution of each k -dimensional column vector of input random numbers generated in this way—that is, $\mathbf{U}_j \sim G_{LH}(k)$ if and only if $\mathbf{U}_j = [U_j^{(1)}, \dots, U_j^{(k)}]^T$ is generated according to (7). In Avramidis (1993) it is shown that $G_{LH}(k) \in \mathcal{G}$ for any $k \geq 2$.

Our definition of LHS is more general than the usual one, introduced by McKay, Beckman, and Conover (1980) and followed by Stein (1987). These authors assume that the input random variates are independent, and each of these variates is generated by the method of inversion. We do not require either of these assumptions. For any $\mathcal{K} \subseteq \{1, \dots, p\}$, we say that *strict LHS is used on \mathcal{K}* if the input random variates $\{V_i : i \in \mathcal{K}\}$ are sampled using the method of inversion.

2.3 Control Variates

Suppose we can identify a $1 \times q$ vector of concomitant random variables $\mathbf{C} = (C_1, \dots, C_q)$ having known, finite expectation $\boldsymbol{\mu}_C \equiv E[\mathbf{C}]$ and a strong linear association with Y . We try to predict the unknown deviation $Y - \theta$ as a linear combination of the known deviation $\mathbf{C} - \boldsymbol{\mu}_C$ in order to adjust the response accordingly; this yields the “controlled” response

$$Y_{CV} \equiv Y - \mathbf{b}(\mathbf{C} - \boldsymbol{\mu}_C)^T.$$

For any constant $1 \times q$ vector \mathbf{b} , the controlled response Y_{CV} is an unbiased estimator of θ . Let $\boldsymbol{\sigma}_{YC} \equiv \text{Cov}(Y, \mathbf{C})$ be the $1 \times q$ vector of covariances $[\text{Cov}(Y, C_1), \dots, \text{Cov}(Y, C_q)]$ and let $\boldsymbol{\Sigma}_C \equiv \text{Cov}(\mathbf{C})$ be the $q \times q$ covariance matrix of \mathbf{C} , where we assume that $\boldsymbol{\Sigma}_C$ is positive definite. The variance of Y_{CV} is minimized by the optimal control coefficient vector $\mathbf{b}^* = \boldsymbol{\sigma}_{YC} \boldsymbol{\Sigma}_C^{-1}$. Even though in some applications $\boldsymbol{\Sigma}_C$ may be known, $\boldsymbol{\sigma}_{YC}$ is almost always unknown, and therefore \mathbf{b}^* must be estimated.

Suppose we have available n i.i.d. observations $\{(Y_i, \mathbf{C}_i) : i = 1, \dots, n\}$. The most commonly used control coefficient vector is the sample analog of \mathbf{b}^* , $\hat{\mathbf{b}} = \mathbf{S}_{YC} \mathbf{S}_C^{-1}$, where \mathbf{S}_{YC} is the $1 \times q$ vector of sample covariances $[\hat{\text{Cov}}(Y, C_1), \dots, \hat{\text{Cov}}(Y, C_q)]$ and \mathbf{S}_C is the sample covariance matrix of \mathbf{C} . The *control-variate* (CV) estimator based on the sample $\{(Y_i, \mathbf{C}_i) : i = 1, \dots, n\}$ is then defined as

$$\hat{\theta}_{CV}(n) \equiv \bar{Y} - \hat{\mathbf{b}}(\bar{\mathbf{C}} - \boldsymbol{\mu}_C)^T,$$

where \bar{Y} and $\bar{\mathbf{C}}$ are the sample means of $\{Y_i : i = 1, \dots, n\}$ and $\{\mathbf{C}_i : i = 1, \dots, n\}$ respectively.

Although the basic variance-reducing properties of $\hat{\theta}_{CV}$ have been established under relatively stringent assumptions about the joint distribution of the response Y and the control vector \mathbf{C} , Nelson (1990) pointed out that irrespective of the distribution of (Y, \mathbf{C}) ,

$$n^{1/2}[\hat{\theta}_{CV}(n) - \theta] \xrightarrow{D} N[0, \sigma_Y^2(1 - R_{YC}^2)] \text{ as } n \rightarrow \infty,$$

where $R_{YC}^2 = \sigma_{YC} \boldsymbol{\Sigma}_C^{-1} \boldsymbol{\sigma}_{YC}^T / \sigma_Y^2$ is the squared coefficient of multiple correlation between Y and \mathbf{C} . Thus $\hat{\theta}_{CV}(n)$ asymptotically dominates $\bar{Y}(n)$.

3 INTEGRATED STRATEGIES FOR VARIANCE REDUCTION

3.1 Conditional Expectation and Correlation Induction

We begin by expressing the conditioning vector \mathbf{X} as a function of the input random numbers:

$$\mathbf{X} = x(U_j : j \in I_X) \text{ for some } I_X \subseteq \{1, \dots, d\}, \quad (8)$$

where I_X is the set of indices of the random numbers on which \mathbf{X} depends. As seen in §2.1, the random variable $Z \equiv E[Y|\mathbf{X}] \equiv h(\mathbf{X})$ is an alternative estimator of θ based on a single replication of the simulation; and Z has no larger variance than Y . Thus we may view Z as the new response of interest, and we can obtain an even more precise estimator by applying the technique of correlation induction to the random variable Z . For this purpose, we express Z as a function of the input random numbers

$$Z = h(\mathbf{X}) = h[x(U_j : j \in I_X)] \equiv z(U_j : j \in I_X).$$

EXAMPLE 1 (CONTINUED). In §2.1 we took $\mathbf{X} = (V_1, V_5)$. In view of (1), we have $I_X = \{1, 2, 6\}$. Moreover, (3) implies that

$$\begin{aligned} z(u_1, u_2, u_6) = & F_{2,3}[t - H_1(u_1, u_2), t - H_1(u_1, u_2) - H_5(u_6)] \\ & \times F_4[t - H_5(u_6)]. \quad \blacksquare \end{aligned}$$

Given an arbitrary k -dimensional distribution $G \in \mathcal{G}$, we perform k dependent replications of the simulation using the distribution G to sample the random numbers with indices in I_X . The random numbers with indices in $\{1, \dots, d\} - I_X$ need not be sampled, since Z does not depend on them. Following the notation in (5), we define the *conditional expectation-correlation induction* (CE+CI) estimator based on the distribution G and n replications as

$$\hat{\theta}_{CE+CI}(G, n) \equiv m^{-1} \sum_{i=1}^m \tilde{Z}_i,$$

where $m = n/k$ and $\{\tilde{Z}_i\}_{i=1}^m \stackrel{\text{i.i.d.}}{\sim} Z_{CI}(G, I_X)$.

The following result is proved in Avramidis and Wilson (1992).

PROPOSITION 1. For any $G \in \mathcal{G}$, the estimator $\hat{\theta}_{CE+CI}(G, n)$ is an unbiased estimator of θ , with $\text{Var}[\hat{\theta}_{CE+CI}(G, n)] \leq \text{Var}[\hat{\theta}_{CI}(G, n)]$. If $z(\cdot)$ is a monotone function of each of its arguments, then $\text{Var}[\hat{\theta}_{CE+CI}(G, n)] \leq \text{Var}[\hat{\theta}_{CE}(n)]$.

3.2 Control Variates and Correlation Induction

Our approach to the joint application of the methods of control variates and correlation induction is based on the observation that the control vector \mathbf{C} usually depends only on a proper subset $(U_i : i \in I_C)$ of the input random numbers $(U_i : i = 1, \dots, d)$, so that we may write

$$\mathbf{C} = c(U_j : j \in I_C) \text{ for some } I_C \subset \{1, \dots, d\}, \quad (9)$$

where $I'_C \equiv \{1, \dots, d\} - I_C \neq \emptyset$.

EXAMPLE 1 (CONTINUED). Suppose $\mathbf{C} = V_4 + V_5 = H_4(U_5) + H_5(U_6)$, a scalar. Then $I_C = \{5, 6\}$, $I'_C = \{1, 2, 3, 4\}$, and $c(u_5, u_6) = H_4(u_5) + H_5(u_6)$. ■

Our development is in the same spirit as the approach of Tew and Wilson (1993) for integrating the Schruben-Margolin strategy with the method of control variates. The key idea is to induce the desired negative correlation between the responses by sampling dependently only on the coordinates that do not affect the control vector, thus preserving the dependency structure between the response and the control vector on each simulation run.

Given an arbitrary k -dimensional distribution $G \in \mathcal{G}$, we perform k dependent replications of the simulation using the distribution G to sample the random numbers with indices in I'_C . The random numbers

with indices in I_C are sampled independently according to $G_0(k)$. Following the notation in (5), we define the auxiliary quantities

$$\begin{aligned} \tilde{Y} &\equiv Y_{CI}(G, I'_C), \\ \tilde{\mathbf{C}} &\equiv [C_{1,CI}(G, I'_C), \dots, C_{q,CI}(G, I'_C)]. \end{aligned} \quad (10)$$

To simplify the notation, we will take $\rho(G) \equiv \rho_Y(G, I'_C)$, where $\rho_Y(G, I'_C)$ is defined as in (6). We define $\hat{\theta}_{CV+CI}(G, n)$, the *control variate–correlation induction* (CV+CI) estimator based on the distribution G and n replications, as the control variate estimator

$$\hat{\theta}_{CV+CI}(G, n) \equiv \frac{1}{m} \sum_{i=1}^m \tilde{Y}_i - \hat{\mathbf{b}} \left(\frac{1}{m} \sum_{i=1}^m \tilde{\mathbf{C}}_i - \boldsymbol{\mu}_C \right)^T,$$

with $m = n/k$, $\{(\tilde{Y}_i, \tilde{\mathbf{C}}_i)\}_{i=1}^m \stackrel{\text{i.i.d.}}{\sim} (\tilde{Y}, \tilde{\mathbf{C}})$ as in display (10), and $\hat{\mathbf{b}} = \mathbf{S}_{\tilde{Y}\tilde{\mathbf{C}}} \mathbf{S}_{\tilde{\mathbf{C}}}^{-1}$, where $\mathbf{S}_{\tilde{Y}\tilde{\mathbf{C}}}$ denotes the $1 \times q$ vector of sample covariances between \tilde{Y} and the components of $\tilde{\mathbf{C}}$ in (10) and $\mathbf{S}_{\tilde{\mathbf{C}}}$ denotes the $q \times q$ sample covariance matrix of $\tilde{\mathbf{C}}$.

The following result is proved in Avramidis and Wilson (1992).

PROPOSITION 2. *Suppose that \mathbf{C} is of the form (9), $G \in \mathcal{G}$, and $y(\cdot)$ is a monotone function of each argument with index in I'_C . If each of (Y, \mathbf{C}) and $(\tilde{Y}, \tilde{\mathbf{C}})$ has a multivariate Normal distribution, then $\hat{\theta}_{CV+CI}(G, n)$ is an unbiased estimator of θ ; and ignoring terms of order $O(1/n)$, we have*

$$\frac{\text{Var}[\hat{\theta}_{CV+CI}(G, n)]}{\text{Var}[\hat{\theta}_{CV}(n)]} = 1 + \frac{(k-1)\rho(G)}{1 - R_{Y,C}^2} \leq 1.$$

Moreover, irrespective of the distributions of (Y, \mathbf{C}) and $(\tilde{Y}, \tilde{\mathbf{C}})$, $\hat{\theta}_{CV+CI}(G)$ asymptotically dominates $\hat{\theta}_{CV}$.

Unfortunately, a variance comparison between $\hat{\theta}_{CV+CI}(G, n)$ and $\hat{\theta}_{CI}(G, n)$ is not possible in general. With $\hat{\theta}_{CI}(G, n)$ we sample dependently on all the coordinates, which might induce more correlation between replications of Y than when we only sample dependently on some of the coordinates, as with $\hat{\theta}_{CV+CI}(G, n)$; and this extra correlation might outweigh the benefit of the term $-R_{Y,C}^2$ that arises from the use of control variates.

3.3 Conditional Expectation and Control Variates

To combine the methods of conditional expectation and control variates, we must select a control vector \mathbf{C} and an auxiliary random vector \mathbf{X} such that

we can evaluate the conditional expectations $h(\mathbf{x}) \equiv E[Y|\mathbf{X} = \mathbf{x}]$ and $g(\mathbf{x}) \equiv E[\mathbf{C}|\mathbf{X} = \mathbf{x}]$ analytically or numerically for every possible value of \mathbf{x} . Define the auxiliary random variables $Z \equiv h(\mathbf{X})$ and $\mathbf{D} \equiv g(\mathbf{X})$, and note that $E[Z] = \theta$ and $E[\mathbf{D}] = \boldsymbol{\mu}_C$.

EXAMPLE 1 (CONTINUED). With the previously defined vectors $\mathbf{X} \equiv (X_1, X_2) \equiv (V_1, V_5)$ and $\mathbf{C} \equiv V_4 + V_5$, we have

$$\begin{aligned} g(x_1, x_2) &= E[V_4 + V_5 | V_1 = x_1, V_5 = x_2] \\ &= E[V_4] + x_2. \quad \blacksquare \end{aligned}$$

We view Z and \mathbf{D} as the new response and control vector respectively, and we use the control-variate technique to further reduce the variance of Z . By analogy with the standard control-variate methodology, we assume that the vector \mathbf{D} has a positive definite (p.d.) covariance matrix $\boldsymbol{\Sigma}_D$. Moreover, we assume that

$$\text{There is no } 1 \times q \text{ vector } \mathbf{d} \neq \mathbf{0} \text{ such that } \mathbf{d}\mathbf{C}^T \text{ is a function of } \mathbf{X} \text{ alone.} \quad (11)$$

Let $\{\mathbf{X}_i : i = 1, \dots, n\}$ be i.i.d. observations of \mathbf{X} . In terms of the auxiliary observations

$$Z_i = h(\mathbf{X}_i), \quad \mathbf{D}_i = g(\mathbf{X}_i), \quad i = 1, \dots, n, \quad (12)$$

we define $\hat{\theta}_{CE+CV}(n)$, the *conditional expectation–control variate* (CE+CV) estimator based on n replications

$$\hat{\theta}_{CE+CV}(n) \equiv \bar{Z} - \hat{\mathbf{b}} (\bar{\mathbf{D}} - \boldsymbol{\mu}_C)^T,$$

where: \bar{Z} and $\bar{\mathbf{D}}$ are the sample means of $\{Z_i : i = 1, \dots, n\}$ and $\{\mathbf{D}_i : i = 1, \dots, n\}$ respectively;

$$\hat{\mathbf{b}} \equiv \mathbf{S}_{ZD} \mathbf{S}_D^{-1};$$

\mathbf{S}_{ZD} denotes the $1 \times q$ vector of sample covariances between Z and the components of \mathbf{D} in (12); and \mathbf{S}_D denotes the $q \times q$ sample covariance matrix of \mathbf{D} .

Let R_{ZD} denote the coefficient of multiple correlation between Z and \mathbf{D} . The following result is proved in Avramidis and Wilson (1992).

PROPOSITION 3. *Suppose that $\boldsymbol{\Sigma}_D$ is positive definite and (11) holds. If each of (Y, \mathbf{C}) and (Z, \mathbf{D}) has a multivariate Normal distribution, then $\hat{\theta}_{CE+CV}(n)$ is an unbiased estimator of θ , and $\text{Var}[\hat{\theta}_{CE+CV}(n)] \leq \min\{\text{Var}[\hat{\theta}_{CV}(n)], \text{Var}[\hat{\theta}_{CE}(n)]\}$ for $n \geq q/R_{ZD}^2 + 2$. Moreover, irrespective of the distributions of (Y, \mathbf{C}) and (Z, \mathbf{D}) , $\hat{\theta}_{CE+CV}$ asymptotically dominates $\hat{\theta}_{CV}$ and $\hat{\theta}_{CE}$.*

EXAMPLE 2. Suppose that $(Y, \mathbf{C}, \mathbf{X})$ is nonsingular multivariate Normal. Then each of (Y, \mathbf{C}) and (Z, \mathbf{D}) is multivariate Normal, $\boldsymbol{\Sigma}_D$ is positive definite, and (11) holds. \blacksquare

4 SOME ASYMPTOTIC COMPARISONS

In this section we show that Latin hypercube sampling, a special case of correlation induction, is asymptotically more efficient than the method of control variates when a certain class of controls is used. We also establish some general conditions under which a combined variance reduction strategy based on the methods of Latin hypercube sampling and conditional expectation is asymptotically more efficient than many of the other strategies discussed in this paper.

Suppose the control vector \mathbf{C} has the following form in terms of the input random variates:

$$C_i = \sum_{j \in J_C} \phi_{i,j}(V_j) \text{ for } i = 1, \dots, q, \quad (13)$$

for some $J_C \subseteq \{1, \dots, p\}$, and where

- (a) For each $j \in J_C$, $\phi_{i,j}(\cdot)$ is an arbitrary univariate function;
- (b) The input random variates $\{V_j : j \in J_C\}$ are mutually independent; and
- (c) The remaining input variates $(V_j : j \in J'_C)$, where $J'_C \equiv \{1, \dots, p\} - J_C$, are independent of $(V_j : j \in J_C)$.

That is, each component C_i of the control vector \mathbf{C} is a *separable* function of a set of independent input variates; and although the remaining set of input variates may be stochastically interdependent, the latter set is independent of the former set. This setup often occurs in practice since many input variates are generated independently of each other and control variates are usually taken to be sums of selected input variates. For example, in queueing simulations, sums or averages of service times observed at selected service centers are frequently used as controls (Wilson 1984); and in simulations of stochastic activity networks, sums of activity times along selected paths are often used as controls.

EXAMPLE 1 (CONTINUED). Using again $\mathbf{C} = V_4 + V_5$, we see that (13) holds with $J_C = \{4, 5\}$, $q = 1$, and $\phi_{1,4}(\mathbf{x}) \equiv \phi_{1,5}(\mathbf{x}) \equiv x$ for all real x . ■

We define the Latin hypercube sampling estimator $\hat{\theta}_{LH}(n)$ to be $\hat{\theta}_{CI}(G_{LH}(n), n)$, where strict LHS is used on J_C ; and for simplicity in the following development, we may assume that $I_C = J_C$ without loss of generality (recall the definition of I_C in (9)). To be explicit, we use the following sampling plan for $\hat{\theta}_{LH}(n)$:

$$V_j = \begin{cases} F_j^{-1}(U_j), & j \in J_C \\ H_j(U_i : i \in I'_C), & j \in J'_C, \end{cases}$$

where $F_j^{-1}(\cdot)$ is the inverse c.d.f. of V_j , $I'_C \equiv \{1, \dots, d\} - I_C$ as in (9), and the functions $\{H_j(\cdot) : j \in J'_C\}$ are the remaining part of the sampling plan, whose form we do not need to make explicit.

The following result is proved in Avramidis and Wilson (1992).

PROPOSITION 4. *If the response Y is bounded and the control vector \mathbf{C} has components of the form (13), then $\hat{\theta}_{LH}$ asymptotically dominates $\hat{\theta}_{CV}$.*

In the rest of this section we examine the asymptotic efficiency of a combined variance reduction strategy based on the methods of Latin hypercube sampling and conditional expectation. Our result depends on the observation that usually the conditioning vector is a subset of the input random variates, so that we may write

$$\mathbf{X} = (V_j : j \in J_X) \text{ for some } J_X \subseteq \{1, \dots, p\}. \quad (14)$$

We define the conditional expectation–Latin hypercube sampling estimator $\hat{\theta}_{CE+LH}(n)$ to be $\hat{\theta}_{CE+CI}(G_{LH}(n), n)$, where strict LHS is used on the index-set $J_X \cap J_C$, and again we may assume without loss of generality that $J_X \cap J_C = I_X \cap I_C$. For concreteness, we state that the sampling plan corresponding to $\theta_{CE+LH}(n)$ is

$$V_j = \begin{cases} F_j^{-1}(U_j), & j \in J_X \cap J_C \\ H_j[U_i : i \in I_X - (J_X \cap J_C)], & j \in J_X - J_C, \end{cases}$$

where I_X is defined in (8), and the functions $\{H_j(\cdot) : j \in J_X - J_C\}$ are the remaining part of the sampling plan, whose form we do not need to make explicit.

The following result is proved in Avramidis and Wilson (1992).

PROPOSITION 5. *If the response Y is bounded, if the control vector \mathbf{C} has components of the form (13), if the conditioning vector \mathbf{X} is of the form (14), if Σ_D is positive definite, and if (11) holds, then $\hat{\theta}_{CE+LH}$ asymptotically dominates $\hat{\theta}_{CE+CV}$, $\hat{\theta}_{CE}$, $\hat{\theta}_{CV}$, and $\hat{\theta}_{LH}$.*

5 AN APPLICATION

5.1 Integrated Strategies for Variance Reduction

The graph-theoretic structure of a stochastic activity network (SAN) is described by the pair $(\mathcal{N}, \mathcal{A})$, where $\mathcal{N} = \{1, \dots, \nu\}$ is the set of nodes (vertices) in the network and $\mathcal{A} = \{(a_j, b_j) : \text{activity } j \text{ has start node } a_j \in \mathcal{N} \text{ and end node } b_j \in$

$\mathcal{N}, j = 1, \dots, p$. The network is assumed to be acyclic, with source node $r \in \mathcal{N}$ and sink node $s \in \mathcal{N}$. Each activity j has a random duration V_j , so the input random variates are $\{V_j : j = 1, \dots, p\}$, and the probabilistic structure of the network is described by the joint distribution of the random vector (V_1, \dots, V_p) . Let ξ denote the number of directed r - s paths, and let $A(\ell)$ denote the index set of activities on the ℓ th path, so $A(\ell) = \{j : \text{activity } j \text{ is on the } \ell\text{th directed } r\text{-}s \text{ path}\}$ for $\ell = 1, \dots, \xi$. The duration of the ℓ th path is the random variable

$$P_\ell = \sum_{j \in A(\ell)} V_j,$$

and the *network completion time* is

$$T \equiv \max\{P_1, \dots, P_\xi\}.$$

We consider the problem of estimating the cumulative distribution function $F_T(\cdot)$ of the network completion time. Let \mathcal{T} be the set of cutoff values at which the c.d.f. is to be estimated. Then, for each $t \in \mathcal{T}$, the response of interest is $Y \equiv 1_{\{T \leq t\}} \equiv f(V_1, \dots, V_p)$, where 1_B denotes the indicator function of the event B . Here we view the overall estimation problem as a set of univariate estimation problems—that is, each value in \mathcal{T} corresponds to a single estimand of interest.

We assume that the activity durations $\{V_j : j = 1, \dots, p\}$ are independent, each with a known distribution. We use the method of inversion to generate all random variates, so the sampling plan is

$$V_j = F_j^{-1}(U_j), \quad j = 1, \dots, p.$$

Thus, in the previously established notation, $d = p$, $I_X = J_X$, and $I_C = J_C$. The variance reduction techniques discussed in the previous sections are applied as follows.

As a conditional-expectation estimator, we use an adaptation of an estimator developed for stochastic shortest route problems by Sigal, Pritsker, and Solberg (1980). A *uniformly directed cutset* (UDC) \mathcal{L} is a set of activities such that any directed r - s path contains exactly one activity in \mathcal{L} . See Sigal, Pritsker, and Solberg (1980) for properties of a UDC and the derivation of their estimator of $E[Y]$; and see Provan and Kulkarni (1984) for an efficient algorithm to identify a “good” UDC. For our purposes, we only need to note that (i) the conditioning vector is of the form (14) with $J_X = \{1, \dots, p\} - \mathcal{L}$, where \mathcal{L} is the selected UDC; and (ii) the conditional-expectation estimator $Z \equiv E[Y|\mathbf{X}] \equiv h(\mathbf{X})$ is nonincreasing in each component of \mathbf{X} .

To form a control-variate estimator, we use the same approach as in Avramidis, Bauer, and Wilson (1991). Ranking the directed r - s paths in decreasing order of expected duration, we let ℓ_1, ℓ_2, ℓ_3 be the first three such paths. We use as control variables the durations of these three paths:

$$C_i \equiv P_{\ell_i} = \sum_{j \in A(\ell_i)} V_j, \quad i = 1, 2, 3.$$

Observe that for $i = 1, 2, 3$, $E[C_i]$ can be computed as sums of mean activity durations, which will either be known as part of the input to the simulation, or will have to be evaluated from the known distributions of activity durations. Also note that the controls are of the form (13), with $J_C = \cup_{i=1}^3 A(\ell_i)$ and

$$\phi_{i,j}(x) \equiv \begin{cases} x & \text{if } j \in A(\ell_i) \\ 0 & \text{otherwise.} \end{cases}$$

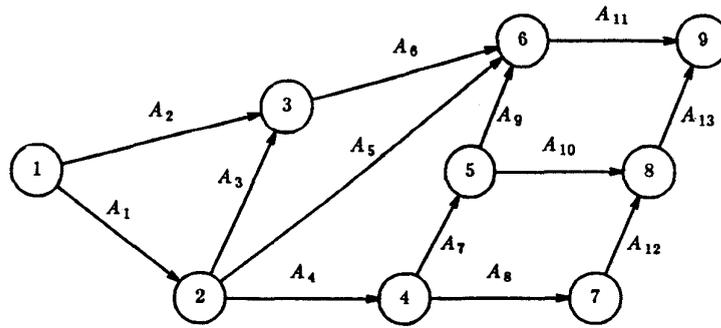
We consider two correlation-induction techniques: (a) antithetic variates (see §2.2.1); and (b) Latin hypercube sampling (see §2.2.2). Note that since we use inversion, we are in fact using strict LHS on the entire set of input random variates (V_1, \dots, V_p) .

5.2 Monte Carlo Results

The SAN we used for this study was taken from Elmaghraby (1977), page 275; and it is depicted in Figure 2. For each activity duration V_i , the associated distribution was taken to be either (a) a normal distribution with a specified mean μ_i and standard deviation $\sigma_i = \mu_i/4$ whose tail was truncated below the value 0; or (b) an exponential distribution with a specified mean μ_i . We chose the exponential distribution as the nonnormal alternative for reasons elaborated in Avramidis, Bauer, and Wilson (1991). The set of activities with durations as in (a) was taken to be $\{(1,2), (1,3), (2,4), (6,9), (7,8)\}$. As a uniformly directed cutset, we chose $\mathcal{L} = \{(3,6), (2,6), (5,6), (5,8), (4,7)\}$.

The purpose of the Monte Carlo study was to estimate the variance reductions achieved by the following estimators: (i) $\hat{\theta}_{CE}(n)$, the conditional expectation estimator; (ii) $\hat{\theta}_{AV}(n) \equiv \hat{\theta}_{CI}(G_{AV}, n)$, the antithetic variate estimator; (iii) $\hat{\theta}_{LH}(n) \equiv \hat{\theta}_{CI}(GLH(n), n)$, the Latin hypercube sampling estimator; (iv) $\hat{\theta}_{CV}(n)$, the control variate estimator; (v) $\hat{\theta}_{CE+CV}(n)$, the conditional expectation–control variate estimator; and (vi) $\hat{\theta}_{CE+LH}(n)$, the conditional expectation–Latin hypercube sampling estimator.

Table 1 shows the resulting variance ratios with respect to the direct-simulation estimator $\hat{Y}(n)$ for



$\mu_1 = 5.5, \mu_2 = 13.0, \mu_3 = 7.0, \mu_4 = 5.2,$
 $\mu_5 = 16.5, \mu_6 = 14.7, \mu_7 = 6.0, \mu_8 = 10.3,$
 $\mu_9 = 20.0, \mu_{10} = 4.0, \mu_{11} = 3.2, \mu_{12} = 3.2, \mu_{13} = 16.5$

Figure 2: Network used in the Monte Carlo study.

Table 1: Estimated Variance Ratios $\text{Var}[\tilde{Y}(n)]/\text{Var}[\hat{\theta}(n)]$ for various estimators $\hat{\theta}$ and sample sizes n .

Estimator	Estimand			
	$F_T(30)$	$F_T(50)$	$F_T(70)$	$F_T(90)$
$\hat{\theta}_{\text{CE}}(32)$	11.1	4.7	4.2	4.4
$\hat{\theta}_{\text{AV}}(32)$	1.0	1.2	1.1	1.0
$\hat{\theta}_{\text{LH}}(32)$	1.2	2.2	2.9	2.2
$\hat{\theta}_{\text{CV}}(32)$	1.2	1.4	1.3	1.2
$\hat{\theta}_{\text{CE+CV}}(32)$	16.4	19.9	11.0	7.7
$\hat{\theta}_{\text{CE+LH}}(32)$	42.0	74.0	51.7	18.2
$\hat{\theta}_{\text{CE+LH}}(64)$	40.7	83.6	75.7	46.2
$\hat{\theta}_{\text{CE+LH}}(128)$	46.5	87.4	77.9	54.3

network 1. The CE technique appears to be the most effective of the individual VRTs, followed by LHS and CV. The AV technique was of little benefit, so LHS was investigated more thoroughly as the correlation-induction technique of choice. With the exception of $\hat{\theta}_{\text{AV}}$, as expected from the results of Section 4, $\hat{\theta}_{\text{CE+LH}}$ is the best of the six estimators considered here. (No asymptotic comparison of $\hat{\theta}_{\text{AV}}$ against the other estimators seems possible in general.) Thus we chose to include results for several sample sizes for $\hat{\theta}_{\text{CE+LH}}$. Observe the large improvement that the integrated variance reduction strategies yield over the individual VRTs.

6 CONCLUSIONS AND RECOMMENDATIONS

Both our theoretical and experimental results strongly suggest that integrated variance reduction strategies have the potential to be highly effective in a large class of simulation experiments. Although our development is limited to finite-horizon simulations with a fixed-dimensional vector of random-number inputs, we believe that much of this development can ultimately be extended to infinite-horizon simulations with an infinite-dimensional vector of random-number inputs.

In light of the demonstrated effectiveness of the joint application of Latin hypercube sampling (LHS) and the method of conditional expectations (CE), we believe that particular emphasis should be given to this combined variance reduction strategy in future research. Currently the key properties of LHS are limited to the case of a bounded simulation response; these properties should be extended to handle an unbounded simulation response. Moreover as mentioned in the previous paragraph, a version of LHS should be developed for infinite-dimensional random-number inputs.

Finally, we observe that all of the results presented in this paper are limited to independent replications of a univariate simulation response. These results should be extended to multiresponse simulations. In addition, the combined variance reduction strategies should be adapted to responses generated within a single replication of a simulation model in steady-state operation—that is, covariance-stationary simulation output processes.

We believe that our development provides a frame-

work for effective application of combined variance reduction strategies in many contexts. Beyond the theoretical comparisons of the asymptotic efficiencies of the various combined strategies, our experimental results for moderately complex stochastic activity networks provide substantial evidence of the practical value of using this approach to improving the efficiency of large-scale simulations.

ACKNOWLEDGMENTS

This work was partially supported by NSF Grant No. DMS-8717799 and by a David Ross Grant from the Purdue Research Foundation. The authors thank Bruce Schmeiser for many enlightening discussions on this paper.

REFERENCES

- Avramidis, A. 1993. Variance reduction techniques for simulation with applications to stochastic networks. Ph.D. diss., School of Industrial Engineering, Purdue University, West Lafayette, Indiana.
- Avramidis, A. N., K. W. Bauer, and J. R. Wilson. 1991. Simulation of stochastic activity networks using path control variates. *Naval Research Logistics* 38 (2): 183-201.
- Avramidis, A., and J. R. Wilson. 1992. Integrated variance reduction strategies for simulation. Technical Report SMS 92-7, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.
- Elmaghraby, S. 1977. *Activity networks: Project planning and control by network models*. New York: Wiley-Interscience.
- Fishman, G. S. 1989. Monte Carlo, control variates, and stochastic ordering. *SIAM Journal on Scientific and Statistical Computing* 10 (1): 187-204.
- Lehmann, E. L. 1966. Some concepts of dependence. *Annals of Mathematical Statistics* 37:1137-1153.
- McKay, M. D., R. J. Beckman, and W. J. Conover. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21 (2): 239-245.
- Nelson, B. L. 1987. A perspective on variance reduction in dynamic simulation experiments. *Communications in Statistics—Simulation and Computation* 16 (2): 385-426.
- Nelson, B. L. 1990. Control variate remedies. *Operations Research* 38 (6): 974-992.
- Provan, J. S., and V. G. Kulkarni. 1984. On the use of exact cuts in stochastic networks. Technical Report TR-84-3, Curriculum in Operations Research and Systems Analysis, University of North Carolina, Chapel Hill, North Carolina.
- Sigal, C. E., A. A. B. Pritsker, and J. J. Solberg. 1980. The stochastic shortest route problem. *Operations Research* 28 (5): 1122-1129.
- Stein, M. 1987. Large sample properties of simulations using Latin hypercube sampling. *Technometrics* 29 (2): 143-151.
- Tew, J. D., and J. R. Wilson. 1993. Estimating simulation metamodels using integrated variance reduction techniques. *IIE Transactions* (to appear).
- Wilson, J. R. 1984. Variance reduction techniques for digital simulation. *American Journal of Mathematical and Management Sciences* 4 (3 & 4): 277-312.

AUTHOR BIOGRAPHIES

ATHANASSIOS N. AVRAMIDIS is an Operations Research Consultant with American Airlines Decision Technologies. He received a diploma in mechanical engineering from the University of Thessaloniki (Greece) in 1987, and he received M.S. and Ph.D. degrees in industrial engineering from Purdue University in 1989 and 1993 respectively. His research interests are in probabilistic and statistical issues in the design and analysis of simulation experiments, particularly in input modeling, output analysis, and variance reduction techniques. He is also interested in applied probability and operations research applications. He is a member of ORSA, TIMS, and IIE.

JAMES R. WILSON is a Professor in the Department of Industrial Engineering at North Carolina State University. He received a B.A. degree in mathematics from Rice University, and he received M.S. and Ph.D. degrees in industrial engineering from Purdue University. His current research interests are focused on the design and analysis of simulation experiments. He also has an active interest in applications of operations research techniques to all areas of industrial engineering. From 1988 to 1992, he served as Departmental Editor of *Management Science* for Simulation. He was *Proceedings* Editor for WSC '86, Associate Program Chair for WSC '91, and Program Chair for WSC '92. He has also held various offices in TIMS/College on Simulation. He is a member of ASA, ACM/SIGSIM, IIE, ORSA, SCS, and TIMS.