Proceedings of the 1996 Winter Simulation Conference
ed. J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J. Swain

A COMPARISON OF PERTURBATION ANALYSIS TECHNIQUES

Michael C. Fu

College of Business and Management
University of Maryland at College Park
College Park, Maryland 20742, U.S.A.

Jian-Qiang Hu

Department of Manufacturing Engineering
Boston University
Boston, Massachusetts 02215, U.S.A.

ABSTRACT

Perturbation analysis (PA) is a technique for estimating gradients of performance measures, particularly applicable to the simulation of discrete-event systems. Over the past two decades, various "versions" have been developed. In this paper, we compare and contrast some of these perturbation analysis techniques by applying them to a simple example. This example also serves to highlight the issue of process representation that can play a very crucial role in the application of perturbation analysis.

1 INTRODUCTION

Sensitivity analysis and optimization of discrete-event systems via simulation has become an increasingly important area of research in the past two decades (cf. Fu 1994). One line of this research is perturbation analysis (cf. Ho and Cao 1991, Glasserman 1991). Even in the field of perturbation analysis, different techniques have been developed. For gradient estimation, the "original" technique is infinitesimal perturbation analysis (IPA), which remains the easiest PA technique to apply in practice. However, its ease of implementation is often offset by its relatively limited domain of applicability, which has led to the development of various other generalizations or alternatives, among them the following: smoothed perturbation analysis (SPA), rare perturbation analysis (RPA), discontinuous perturbation analysis (DPA), and augmented perturbation analysis (APA) . The purpose of this paper is to use a very simple example to illustrate these techniques and the role of process representation in their application. Much of this material is taken from Fu and Hu (1996).

The rest of the paper is organized as follows. We present the simple example in Section 2, including two different process representations. Section 3 will summarize briefly the techniques of IPA, SPA, DPA, RPA, and APA, and illustrate their application on the simple example. Obviously, full justice cannot be given due to space limitations, and Fu and Hu (1996) can be consulted for a more detailed exposition of the individual techniques. In Section 4, three specific cases of the examples are carried out to illustrate the specific form of each of the estimators. Some conclusions based on the example and the specific cases are provided in Section 5.

2 THE EXAMPLE

We consider a single random variable X, defined on an appropriate probability space, for which our goal is to estimate

$$\frac{dE[X]}{d\theta},$$

where θ is a scalar parameter of the distribution.

The example is the following:

$$X = \begin{cases} X_+ & \text{w.p. } \theta \\ X_- & \text{w.p. } 1-\theta \end{cases}. \tag{1}$$

Thus, we also allow X_+ and X_- to be random variables themselves with possible dependence on θ, but we impose the following conditions (cf. (4) and (5)):

$$E\left[\frac{dX_+}{d\theta}\right] = \frac{dE[X_+]}{d\theta}, \quad E\left[\frac{dX_-}{d\theta}\right] = \frac{dE[X_-]}{d\theta}.$$

This corresponds to the requirement that the components are sufficiently smooth so that IPA applies.

We consider two representations corresponding to two different methods for generating X from random numbers. Assume that each of X_+ and X_- can be expressed as functions of a single random variable $U \sim U(0,1)$, i.e., besides the possible implicit dependence on θ that we do not display, we write $X_+ = X_+(U)$ and $X_- = X_-(U)$. The first representation of (1) uses two independent random numbers, whereas the second uses a single random number.

Representation 1:

$$X = \begin{cases} X_+(U_2) & \text{if } U_1 \le \theta \\ X_-(U_2) & \text{if } U_1 > \theta \end{cases}. \quad (2)$$

Representation 2:

$$X = \begin{cases} X_+(U_1/\theta) & \text{if } U_1 \le \theta \\ X_-((1-U_1)/(1-\theta)) & \text{if } U_1 > \theta \end{cases}. \quad (3)$$

3 PA TECHNIQUES

In this section, we apply in turn the various PA techniques to our example after a very brief discussion of each. To keep things simple, we will limit the discussion of the techniques to a single random variable X, as in the example. Recall that our goal is to estimate

$$\frac{dE[X]}{d\theta}.$$

For discrete-event systems, θ can be more general. For example, it could represent a vector operating parameters such as the re-order point and order quantity in an inventory control system.

3.1 IPA

We will denote a sample point by ω with corresponding random variable value $X(\omega)$. The IPA estimator is simply the sample path derivative of the quantity of interest, defined by

$$\frac{dX(\theta,\omega)}{d\theta} = \lim_{\Delta\theta\to 0} \frac{X(\theta+\Delta\theta,\omega) - X(\theta,\omega)}{\Delta\theta}. \quad (4)$$

Thus, for the IPA estimator to be an unbiased gradient estimator, we need

$$\frac{dE[X]}{d\theta} = E\left[\frac{dX}{d\theta}\right], \quad (5)$$

i.e., it becomes a question of an interchange of expectation and limit:

$$\lim_{\Delta\theta\to 0} E\left[g_{\Delta\theta}\right] = E\left[\lim_{\Delta\theta\to 0} g_{\Delta\theta}\right],$$

$$\text{where} \quad g_{\Delta\theta} = \frac{X(\theta+\Delta\theta) - X(\theta)}{\Delta\theta}.$$

For our example, the IPA estimator is obtained by simply taking the derivatives in Equations (2) and (3) to obtain the following.

Representation 1:

$$\frac{dX}{d\theta} = \begin{cases} \frac{d}{d\theta}X_+(U_2) & \text{if } U_1 \le \theta \\ \frac{d}{d\theta}X_-(U_2) & \text{if } U_1 > \theta \end{cases}.$$

Representation 2:

$$\frac{dX}{d\theta} = \begin{cases} \frac{d}{d\theta}X_+(U_1/\theta) & \text{if } U_1 \le \theta \\ \frac{d}{d\theta}X_-((1-U_1)/(1-\theta)) & \text{if } U_1 > \theta \end{cases}.$$

In general, these sample derivatives are biased, i.e., they do not satisfy Equation (5). We now establish necessary and sufficient conditions for their unbiasedness.

Rewriting the representations in terms of indicator functions, we have

$$X = X_+(U_2)\mathbf{1}\{U_1 \le \theta\} + X_-(U_2)\mathbf{1}\{U_1 > \theta\},$$

$$X = X_+(U_1/\theta)\mathbf{1}\{U_1 \le \theta\} + X_-((1-U_1)/(1-\theta))\mathbf{1}\{U_1 > \theta\},$$

where $\mathbf{1}\{\cdot\}$ is the set indicator function. In the first representation, we can uncouple to get

$$E[X] = \theta E[X_+(U_2)] + (1-\theta)E[X_-(U_2)],$$

so that differentiation yields

$$\begin{aligned} \frac{dE[X]}{d\theta} &= \theta\frac{dE[X_+]}{d\theta} + E[X_+] \\ &\quad +(1-\theta)\frac{dE[X_-]}{d\theta} - E[X_-] \\ &= E\left[\frac{dX}{d\theta}\right] + E[X_+ - X_-], \end{aligned}$$

the latter following from our assumption on X_+ and X_-. Thus, we have the following result:

Proposition 1 For Representation 1, a necessary and sufficient condition for unbiasedness of IPA is given by:

$$E[X_+ - X_-] = 0.$$

In addition, the unbiasedness of subsequent SPA estimators imply the following corresponding result for Representation 2:

Proposition 2 For Representation 2, a necessary and sufficient condition for unbiasedness of IPA is given by:

$$X_+(1) = X_-(1).$$

i.e., X is a.s. continuous across its breakpoint.

Note that $X_+(1)$ and $X_-(1)$ correspond to setting $U_1 = \theta$ for Representation 2.

The techniques SPA, RPA, DPA, and APA can be applied even when the interchange equation (5) is not satisfied. We now consider each of these in turn.

3.2 SPA

SPA uses conditional expectation to derive alternative estimators for which a modified version of the interchange equation is satisfied. Introduced by Gong and Ho (1987), this technique has been developed into a very general methodology in the generalized semi-Markov process (GSMP) framework, e.g., Fu and Hu (1992) and Dai and Ho (1995). The resulting estimator contains an IPA term (possibly 0) and a conditional contribution that is the product of a probability rate term and a conditional performance difference term. We illustrate the estimator on the example.

Let $Y = \mathbf{1}\{U_1 \leq \theta\}$. Then, Y is a Bernoulli random variable with parameter θ, i.e., $P(Y = 1) = \theta$ and $P(Y = 0) = 1 - \theta$. We first consider the right-hand (RH) derivative, i.e., $\Delta\theta > 0$. If $Y = 1$, then there is no possibility of a change to $Y = 0$ from the perturbation $\Delta\theta$, since

$$P(U_1 > \theta + \Delta\theta | U_1 \leq \theta) = 0, \quad \Delta\theta > 0.$$

However, if $Y = 0$, then we have

$$P(U_1 \leq \theta + \Delta\theta | U_1 > \theta) = \frac{\Delta\theta}{1-\theta},$$

and the probability rate term is given by

$$\lim_{\Delta\theta \to 0^+} \frac{P(U_1 \leq \theta + \Delta\theta | U_1 > \theta)}{\Delta\theta} = \frac{1}{1-\theta}.$$

Similarly, for the left-hand (LH) derivative, $\Delta\theta < 0$, the probability rate term is given by $1/\theta$.

The conditional difference term is given by

$$X(U_1 \to \theta^-) - X(U_1 \to \theta^+),$$

which depends on the representation used.

Thus, for our example, we have the following SPA derivative estimators:

Representation 1 (LH and RH, respectively):

$$\frac{dX}{d\theta} + \frac{1}{\theta}(X_+(U_2) - X_-(U_2))\,\mathbf{1}\{U_1 \leq \theta\}, \qquad (6)$$

$$\frac{dX}{d\theta} + \frac{1}{1-\theta}(X_+(U_2) - X_-(U_2))\,\mathbf{1}\{U_1 > \theta\}. \qquad (7)$$

Representation 2 (LH and RH, respectively):

$$\frac{dX}{d\theta} + \frac{1}{\theta}(X_+(1) - X_-(1))\,\mathbf{1}\{U_1 \leq \theta\}, \qquad (8)$$

$$\frac{dX}{d\theta} + \frac{1}{1-\theta}(X_+(1) - X_-(1))\,\mathbf{1}\{U_1 > \theta\}. \qquad (9)$$

3.3 RPA

Intuitively, IPA corresponds to the case where small perturbations in the parameter cause small perturbations in the output random variable. In contrast, RPA corresponds to the case where small perturbations in the parameter cause no change in the output random variable much of the time, and the "rare" times when a change does occur results in a relatively large change in the output random variable. First introduced by Brémaud and Vázquez-Abad (1992), a maximal coupling interpretation was provided in Brémaud (1993), which is presented here.

Letting f_θ represent the p.d.f. of X and defining

$$\begin{aligned}
g_\theta(\Delta\theta, x) &= \frac{\min(f_\theta(x), f_{\theta+\Delta\theta}(x))}{1 - \delta_\theta(\Delta\theta)}, \\
g_\theta^1(\Delta\theta, x) &= \frac{(f_\theta(x) - f_{\theta+\Delta\theta}(x))^+}{\delta_\theta(\Delta\theta)}, \\
g_\theta^2(\Delta\theta, x) &= \frac{(f_{\theta+\Delta\theta}(x) - f_\theta(x))^+}{\delta_\theta(\Delta\theta)}, \\
\text{where } \delta_\theta(\Delta\theta) &= \frac{1}{2}\int |f_{\theta+\Delta\theta}(x) - f_\theta(x)|dx,
\end{aligned}$$

the main result is the following. If

$$\delta_\theta' = \lim_{\Delta\theta \to 0^+} \frac{\delta_\theta(\Delta\theta)}{\Delta\theta} \quad \text{exists,}$$

and the c.d.f.'s corresponding to densities $g_\theta(\Delta\theta, \cdot), g_\theta^1(\Delta\theta, \cdot), g_\theta^2(\Delta\theta, \cdot)$, converge weakly to c.d.f.'s $F_\theta(\cdot), G_\theta^1(\cdot), G_\theta^2(\cdot)$, respectively, as $\Delta\theta \to 0^+$, then

$$\delta_\theta'[V^{(2)} - V^{(1)}] \qquad (10)$$

is an unbiased estimator for $dE[X]/d\theta$, where $V^{(1)}$ and $V^{(2)}$ are independent r.v.'s with respective c.d.f.'s G_θ^1, G_θ^2. This is the two-sided RPA estimator. One-sided (left and right) RPA estimators are also available, but will not be presented here. For our particular example, we cannot write out a more explicit estimator without specification of the distributions involved in X_- and X_+ in Equation (1). This will be taken up in Section 4 for the three specific cases.

3.4 DPA

The main idea in DPA is to represent the jumps in the sample performance function by using step functions (Shi 1996). For example, if the random variable X depends on θ through N random numbers as follows:

$$X = X_i^k \text{ if } P_i^{k-1}(\theta) < U_i \leq P_i^k(\theta),$$

$$k = 1, ..., M,\ 0 = P_i^0 < P_i^1 < ... < P_i^k < ... < P_i^M = 1,$$

where the P_i^k are cumulative probabilities, then the DPA estimator is given by

$$\sum_{i=1}^{N}\sum_{k=1}^{M-1}\left(L_i^k - L_i^{k+1}\right)\frac{dP_i^k(\theta)}{d\theta}.$$

For our example, we obtain the following estimators.

Representation 1:

$$\frac{dX}{d\theta} + \left(X_+(U_2) - X_-(U_2)\right). \qquad (11)$$

Representation 2:

$$\frac{dX}{d\theta} + \left(X_+(1) - X_-(1)\right). \qquad (12)$$

3.5 APA

The APA method is an extension of IPA that attempts to partition the sample space into a finite number of sets with the same event sequence (Gaivoronski et al. 1992). Here, for the case of a single random variable, we will simplify the exposition considerably. Assume that the sample space can be partitioned into a finite number of sets $\{E_1, ..., E_M\}$ (events in the probability space) such that for each E_j, there is a corresponding $(P_j^{\min}, P_j^{\max}] \subset (0,1]$ on which X is uniformly differentiable with respect to θ. Then, the APA estimator is given by

$$\frac{dX}{d\theta} + \sum_{j=1}^{M}\frac{X(P_j^{\max})\frac{d}{d\theta}P_j^{\max} - X(P_j^{\min})\frac{d}{d\theta}P_j^{\min}}{P_j^{\max} - P_j^{\min}}\mathbf{1}\{E_j\}.$$

For our example, we obtain

Representation 1:

$$\frac{dX}{d\theta} + \frac{X_+(U_2)}{\theta}\mathbf{1}\{U_1 \le \theta\} - \frac{X_-(U_2)}{1-\theta}\mathbf{1}\{U_1 > \theta\}. \qquad (13)$$

Representation 2:

$$\frac{dX}{d\theta} + \frac{X_+(1)}{\theta}\mathbf{1}\{U_1 \le \theta\} - \frac{X_-(1)}{1-\theta}\mathbf{1}\{U_1 > \theta\}. \qquad (14)$$

4 SPECIFIC CASES

We now consider three specific cases of (1). In the first case, IPA works for neither representation. In the second case, IPA works for Representation 2, but not for 1. In the third case, IPA works for Representation 1 (but not for 2), in spite of the fact that the function is a.s. *discontinuous*. Note that this cannot happen for Representation 2 (across the breakpoint). The corresponding SPA, RPA, DPA, and APA estimators obtained via Equations (6) through (14) are then presented for each of the two representations (RPA is discussed first, because it is independent of the representation). Note that even when the IPA estimator is unbiased, the other estimators may be different (though equal in expectation).

4.1 CASE 1

$$X = \begin{cases} U(0,\theta) & \text{w.p. } \theta \\ 0.5 & \text{w.p. } 1-\theta \end{cases}.$$

Then, $E[X] = 0.5(\theta^2 - \theta + 1)$, $dE[X]/d\theta = \theta - 0.5$. Since $X_+(U) = \theta U$, $X_-(U) = 0.5$, applying Propositions 1 and 2, IPA fails for both representations. For this case, the RPA "estimator" turns out to be the exact answer $(\theta - 0.5)$, since $\delta_\theta(\Delta\theta) = \Delta\theta$, and $g_\theta^1(\Delta\theta, \cdot)$ and $g_\theta^2(\Delta\theta, \cdot)$ converge to masses at θ and 0.5, respectively. The other estimators follow.

Representation 1:

$$X = \begin{cases} \theta U_2 & \text{if } U_1 \le \theta \\ 0.5 & \text{if } U_1 > \theta \end{cases}.$$

$$\frac{dX}{d\theta} = \begin{cases} U_2 & \text{if } U_1 \le \theta \\ 0 & \text{if } U_1 > \theta \end{cases}.$$

The IPA estimator has expectation $E[dX/d\theta] = 0.5\theta \ne dE[X]/d\theta$, and hence is biased.

SPA (LH):

$$\frac{dX}{d\theta} + \frac{1}{\theta}(\theta U_2 - 0.5)\mathbf{1}\{U_1 \le \theta\}.$$

SPA (RH):

$$\frac{dX}{d\theta} + \frac{1}{1-\theta}(\theta U_2 - 0.5)\mathbf{1}\{U_1 > \theta\}.$$

DPA:

$$\frac{dX}{d\theta} + (\theta U_2 - 0.5).$$

APA:

$$\frac{dX}{d\theta} + \frac{\theta U_2}{\theta}\mathbf{1}\{U_1 \le \theta\} - \frac{0.5}{1-\theta}\mathbf{1}\{U_1 > \theta\}.$$

Representation 2:

$$X = \begin{cases} U_1 & \text{if } U_1 \le \theta \\ 0.5 & \text{if } U_1 > \theta \end{cases}.$$

$$\frac{dX}{d\theta} = 0.$$

IPA is clearly biased, since it yields identically zero.

SPA (LH):

$$\frac{1}{\theta}(\theta - 0.5)\mathbf{1}\{U_1 \le \theta\}.$$

SPA (RH):

$$\frac{1}{1-\theta}(\theta-0.5)\,\mathbf{1}\{U_1>\theta\}.$$

DPA:

$$\theta-0.5.$$

APA:

$$\mathbf{1}\{U_1\le\theta\}-\frac{0.5}{1-\theta}\mathbf{1}\{U_1>\theta\}.$$

From this case, we note the following:

- the DPA estimator can be obtained from the SPA estimators via appropriate unconditioning;
- the DPA estimator for Representation 2 coincides with the RPA estimator; however, this is not true for Representation 1;
- the RPA estimator is independent of the representation.

4.2 CASE 2

$$X=\begin{cases} U(0,\theta) & \text{w.p. } \theta \\ \theta & \text{w.p. } 1-\theta \end{cases}.$$

Then, $E[X]=\theta-0.5\theta^2$, $dE[X]/d\theta=1-\theta$. Since $X_+(U)=\theta U$, $X_-(U)=\theta$, applying Propositions 1 and 2, IPA works for Representation 2, but fails for 1. For this example, RPA does not apply, as the limit defining δ_θ' is infinity. The other estimators follow.

Representation 1:

$$X=\begin{cases} \theta U_2 & \text{if } U_1\le\theta \\ \theta & \text{if } U_1>\theta \end{cases}.$$

$$\frac{dX}{d\theta}=\begin{cases} U_2 & \text{if } U_1\le\theta \\ 1 & \text{if } U_1>\theta \end{cases}.$$

The IPA estimator has expectation $E[dX/d\theta]=1-0.5\theta\neq dE[X]/d\theta$, and hence is biased.

SPA (LH):

$$\frac{dX}{d\theta}+(U_2-1)\,\mathbf{1}\{U_1\le\theta\}.$$

SPA (RH):

$$\frac{dX}{d\theta}+\frac{\theta}{1-\theta}(U_2-1)\,\mathbf{1}\{U_1>\theta\}.$$

DPA:

$$\frac{dX}{d\theta}+\theta\,(U_2-1).$$

APA:

$$\frac{dX}{d\theta}+U_2\mathbf{1}\{U_1\le\theta\}-\frac{\theta}{1-\theta}\mathbf{1}\{U_1>\theta\}.$$

Representation 2:

$$X=\begin{cases} U_1 & \text{if } U_1\le\theta \\ \theta & \text{if } U_1>\theta \end{cases}.$$

$$\frac{dX}{d\theta}=\begin{cases} 0 & \text{if } U_1\le\theta \\ 1 & \text{if } U_1>\theta \end{cases}.$$

Here the IPA estimator is unbiased, since $E[dX/d\theta]=1-\theta$. Furthermore, we have

$$X_+(1)=\theta=X_-(1),$$

so both SPA estimators and the DPA estimator reduce to the IPA estimator, which is unbiased. The additional term in the APA estimator, however, is nonzero, but it clearly must have zero expectation.

APA:

$$\frac{dX}{d\theta}+\mathbf{1}\{U_1\le\theta\}-\frac{\theta}{1-\theta}\mathbf{1}\{U_1>\theta\}.$$

4.3 CASE 3

$$X=\begin{cases} U(0,\theta) & \text{w.p. } \theta \\ 0.5\theta & \text{w.p. } 1-\theta \end{cases}.$$

Then, $E[X]=0.5\theta$, $dE[X]/d\theta=0.5$. Since $X_+(U)=\theta U$, $X_-(U)=0.5\theta$, applying Propositions 1 and 2, IPA works for Representation 1, but fails for 2. Again, for this example, RPA does not apply. The other estimators follow.

Representation 1:

$$X=\begin{cases} \theta U_2 & \text{if } U_1\le\theta \\ 0.5\theta & \text{if } U_1>\theta \end{cases}.$$

$$\frac{dX}{d\theta}=\begin{cases} U_2 & \text{if } U_1\le\theta \\ 0.5 & \text{if } U_1>\theta \end{cases}.$$

We note that IPA works for this case, $E[dX/d\theta]=0.5$, even though it is not a.s. continuous. In fact, it is continuous when $U_2=0.5$, which occurs w.p. 0. Even though IPA works, the other estimators still have additional terms that clearly must have expectation zero.

SPA (LH):

$$\frac{dX}{d\theta}+(U_2-0.5)\mathbf{1}\{U_1\le\theta\}.$$

SPA (RH):

$$\frac{dX}{d\theta}+\frac{\theta}{1-\theta}(U_2-0.5)\mathbf{1}\{U_1>\theta\}.$$

DPA:

$$\frac{dX}{d\theta}+\theta(U_2-0.5).$$

APA:

$$\frac{dX}{d\theta} + U_2 \mathbf{1}\{U_1 \le \theta\} - \frac{0.5\theta}{1-\theta}\mathbf{1}\{U_1 > \theta\}.$$

Representation 2:

$$X = \begin{cases} U_1 & \text{if } U_1 \le \theta \\ 0.5\theta & \text{if } U_1 > \theta \end{cases}.$$

$$\frac{dX}{d\theta} = \begin{cases} 0 & \text{if } U_1 \le \theta \\ 0.5 & \text{if } U_1 > \theta \end{cases}.$$

The IPA estimator has expectation $E[dX/d\theta] = 0.5(1-\theta) \neq dE[X]/d\theta$, and hence is biased.

SPA (LH):

$$\frac{dX}{d\theta} + 0.5\mathbf{1}\{U_1 \le \theta\}.$$

SPA (RH):

$$\frac{dX}{d\theta} + \frac{0.5}{1-\theta}\mathbf{1}\{U_1 > \theta\}.$$

DPA:

$$\frac{dX}{d\theta} + 0.5\theta.$$

APA:

$$\frac{dX}{d\theta} + \mathbf{1}\{U_1 \le \theta\} - \frac{0.5\theta}{1-\theta}\mathbf{1}\{U_1 > \theta\}.$$

5 CONCLUSIONS

For this simple example, we have developed necessary and sufficient conditions for unbiasedness of the IPA estimator. The conditions clearly depend on the representation, as Propositions 1 and 2 indicate. When IPA fails, the other techniques (SPA, DPA, RPA, APA) can almost always be applied. All but RPA can be considered extensions or generalizations of IPA, and this is evident in the example where each contains the IPA estimator as a component of the total estimator. Furthermore, RPA is independent of the process representation. In fact, it has been pointed out by a number of researchers that RPA is more closely akin to the so-called weak derivative estimators (Pflug 1992). On the other hand, in certain problems, the RPA estimator coincides with the DPA estimator (Shi 1996), but the latter depends on the representation. The example also illuminates the fact that DPA and SPA can be viewed as conditional forms of one another. The simple example also showed that aside from IPA, the APA estimator is the only one that can always be estimated directly from the original sample. In fact, the APA estimator often corresponds to the likelihood ratio/score function method estimator when the IPA contribution is zero.

ACKNOWLEDGMENTS

Michael Fu was supported in part by a 1996 Summer Research Grant from the Maryland Business School and by the National Science Foundation under grant EEC-9402384. Jian-Qiang Hu was supported in part by the National Science Foundation under grants EEC-9527422 and DDM-9212368.

REFERENCES

Brémaud, P. 1993. Maximal coupling and rare perturbation sensitivity analysis, *Queueing Systems: Theory and Applications* 11: 307-333.

Brémaud, P. and F.J. Vázquez-Abad. 1992. On the pathwise computation of derivatives with respect to the rate of a point process: the phantom RPA method, *Queueing Systems: Theory and Applications* 10: 249-270.

Dai, L., and Y.C. Ho. 1995. Structural infinitesimal perturbation analysis for derivative estimation in discrete event dynamic systems, *IEEE Transactions on Automatic Control* 40: 1154-1166.

Fu, M.C. 1994. Optimization via simulation: a review. *Annals of Operations Research* 53: 199-248.

Fu, M.C. and J. Q. Hu. 1992. Extensions and generalizations of smoothed perturbation analysis in a generalized semi-Markov process framework, *IEEE Transactions on Automatic Control* 37: 1483-1500.

Fu, M.C. and J. Q. Hu. 1996. *Conditional Monte Carlo: Gradient Estimation and Optimization Applications*, Kluwer Academic Publishers, forthcoming.

Gaivoronski, A., L.Y. Shi, and R.S. Sreenivas. 1992. Augmented infinitesimal perturbation analysis: an alternate explanation, *Discrete Event Dynamic Systems: Theory and Applications*, 2: 121-138.

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

Gong, W.B. and Y.C. Ho. 1987. Smoothed perturbation analysis of discrete-event dynamic systems, *IEEE Transactions on Automatic Control* 32: 858-867.

Ho, Y.C. and X.R. Cao. 1991. *Discrete Event Dynamic Systems and Perturbation Analysis,* Kluwer Academic.

Pflug, G. Ch. 1992. Gradient estimates for the performance of Markov chains and discrete event processes, *Annals of Operations Research* 39: 173-194.

Shi, L. Y. 1996. Discontinuous perturbation analysis of discrete event dynamic systems, to appear in *IEEE Transactions on Automatic Control.*

AUTHORS BIOGRAPHIES

MICHAEL C. FU is an Associate Professor of Management Science & Statistics in the College of Business and Management, with a joint appointment in the Institute for Systems Research, at the University of Maryland at College Park. He received a bachelor's degree in mathematics and bachelor's and master's degrees in electrical engineering from MIT in 1985, and M.S and Ph.D. degrees in applied mathematics from Harvard University in 1986 and 1989, respectively. His research interests include simulation optimization and sensitivity analysis, particularly with applications towards manufacturing systems, inventory control, and the pricing of financial derivatives. He teaches courses in applied probability, stochastic processes, discrete-event simulation, and operations management, and in 1995 was awarded the Maryland Business School's annual Allen J. Krowe Award for Teaching Excellence. He is a member of **IEEE** and **INFORMS**, and was on the Program Committee for the Spring 1996 INFORMS National Meeting in Washington, D.C.

JIAN-QIANG HU is an Associate Professor in the Department of Manufacturing Engineering at Boston University. He received a B.S. degree in applied mathematics from Fudan University, China in 1985, and M.S. and Ph.D. degrees in applied mathematics from Harvard University in 1987 and 1990, respectively. His research interests include modeling, sensitivity analysis, simulation, optimization, and control of discrete event stochastic systems and queueing networks with applications to manufacturing systems and communication networks. He is a member of **IEEE** and **INFORMS**, and is on the Organizing Committee for the 1997 INFORMS Applied Probability Section Conference in Boston.