Computational Optimization and Applications, 26, 219–230, 2003

An Efficient Monte Carlo Method for Optimal Control Problems with Uncertainty

YANZHAO CAO
Department of Mathematics, Florida A & M University, Tallahassee, FL 32307, USA

M.Y. HUSSAINI
School of Computational Science and Information Technology, Florida State University, Tallahassee, FL 32306-4120, USA

T.A. ZANG
NASA Langley Research Center, Hampton, VA 23681-2199, USA

Received August 14, 2002; Revised February 10, 2003

Abstract. A general framework is proposed for what we call the sensitivity derivative Monte Carlo (SDMC) solution of optimal control problems with a stochastic parameter. This method employs the residual in the first-order Taylor series expansion of the cost functional in terms of the stochastic parameter rather than the cost functional itself. A rigorous estimate is derived for the variance of the residual, and it is verified by numerical experiments involving the generalized steady-state Burgers equation with a stochastic coefficient of viscosity. Specifically, the numerical results show that for a given number of samples, the present method yields an order of magnitude higher accuracy than a conventional Monte Carlo method. In other words, the proposed variance reduction method based on sensitivity derivatives is shown to accelerate convergence of the Monte Carlo method. As the sensitivity derivatives are computed only at the mean values of the relevant parameters, the related extra cost of the proposed method is a fraction of the total time of the Monte Carlo method.

Keywords: Monte Carlo method, optimal control, Burger's equation, uncertainty quantification, sensitivity derivatives

1. Introduction

Optimal control problems of fluid flow have a wide range of important applications. Examples include aerodynamic design of cars, airplanes and jet engines. Computational fluid dynamics (CFD) algorithms have achieved such a high level of maturity in terms of generality and efficiency that it is now feasible to implement sophisticated flow control and design optimization methods [2, 3, 8, 11]. In most optimal flow control and design optimization studies to date, the input data and parameters are assumed to be deterministic. However, in reality, due to modeling approximations, numerical errors and non-deterministic operating conditions, input data and parameters are subject to uncertainties. For instance, the viscosity of a fluid depends on temperature, but if temperature is not included as a state variable in a CFD model, it is appropriate to assume that the viscosity is a random variable rather than a precisely known constant. The uncertainty of the viscosity parameter affects the outcome of the fluid flow and must be taken into account in the corresponding optimal control problems.

The interest in optimization under uncertainty has grown dramatically in the past few years (see, e.g., [1, 4, 7]). Some of the early studies have been in the structural design area [5, 12, 14]. The work of Huyse [9] is one of the first attempts at solving aerodynamic optimization problems with uncertainty. In his study, the Mach number of the flow is treated as a random variable and the basic Monte-Carlo method is used to compute the risk function. The problem with the Monte-Carlo method is its slow convergence. It usually takes hundreds, if not thousands, of samples (solutions of the state equation) to obtain satisfactory approximation of the cost function, which often makes it prohibitively expensive. In Putko et al. [13] the CFD outputs are approximated using the first-order and second-order Taylor expansions around the means of the uncertain inputs. Although the approximation involves sensitivity derivatives, expensive Monte Carlo simulation is avoided since all variables and functions become deterministic after the approximation. Numerical examples in the paper show that the method is effective when the variances of the input variables are small (also see [10] for a similar approach). However, Taylor approximations of first or second order are inappropriate for non-Gaussian, especially highly skewed, distributions, and are always very inaccurate in the tails.

In this paper, we propose a Monte Carlo method with variance reduction based on the sensitivity derivative of the state variable with respect to the stochastic parameter. This method deals with the residual of the Taylor expansion instead of the function itself. The convergence of the method is guaranteed by the convergence property of the Monte Carlo method, and it is accelerated by the reduction of variance resulting from the Taylor approximation. To demonstrate the efficiency of the SDMC method, we apply it to the optimal control of the steady-state Burgers equation with viscosity as a random variable.

The paper is organized as follows. In Section 2, we present the formulation of optimal control problems with uncertain parameters and the description of the sensitivity derivative Monte Carlo method. Then in Section 3, we study the optimal boundary control problem of the steady state Burgers equation with viscosity as a random variable. Finally, in Section 4, we present the results of numerical experiments.

2. General framework

2.1. Optimal control with uncertain parameters

Consider an optimal control problem which involves the state equation

$$F(y, u, \xi) = 0, \tag{2.1}$$

where y is the state variable, u is the control variable and ξ is a random variable associated with the uncertainty of given data and the uncertainty of system parameters. We assume that y can be expressed as a function of ξ, that is, $y = y(\xi)$. Also assume that $J(y, u)$ is a cost function which measures a quantity of practical interest. The expectation of J is given by

$$\hat{J}(u) = \int J(y, u) p(\xi)\, d\xi, \tag{2.2}$$

where p is the probability density function (pdf) of ξ. The optimal control problem is then to find a control $\tilde{u}$ such that

$$\hat{J}(\tilde{u}) = \min \hat{J}(u). \tag{2.3}$$

A numerical solution of an optimal control problem involves three key tasks: (i) numerical solution of the state Eq. (2.1), (ii) Monte Carlo simulation to evaluate the cost function $\hat{J}$ and its gradient $\nabla\hat{J}$, and (iii) an optimization algorithm to find the minimizer of $\hat{J}$.

Many mature CFD codes are available to perform task 1. Gradient and non-gradient based optimization software can be easily found to perform task 3. An efficient Monte Carlo method to evaluate the cost function is crucial for optimal control problems of fluid flow as the solution of the state equation is usually computationally intensive. We propose to improve the performance of a Monte Carlo method by exploiting the sensitivity derivative of y with respect to ξ.

2.2. *Sensitivity derivative Monte Carlo method*

For a random variable η, we use $E(\eta)$ and $V(\eta)$ to denote its expectation and variance, respectively. We will also use $\bar{\eta}$ to denote the expectation of η. In a Monte Carlo method, the approximation of the integral (2.2) is given by

$$\hat{J}(u) \approx \frac{1}{N}\sum_{i=1}^{N} J(y(\xi_i), u), \tag{2.4}$$

where $\xi_1, \xi_2, \ldots, \xi_N$ is a sequence of samples of ξ. The convergence of (2.4) is, of course, guaranteed by the large number theorem. But the approximation error in (2.4) is proportional to $\frac{V(J)}{\sqrt{N}}$. One naturally looks for ways to reduce variance to improve convergence. The current effort exploits the information regarding the sensitivity of the state variable with respect to the stochastic parameter to achieve variance reduction.

For a fixed u, Let $J_1(\xi, u)$ be the linear Taylor expansion of J at $\bar{\xi}$, i.e.,

$$J_1(\xi, u) = J(y(\bar{\xi}), u) + J_y(y(\bar{\xi}), u)y_\xi(\bar{\xi})(\xi - \bar{\xi}),$$

where y_ξ the sensitivity of y with respect to ξ. Notice that

$$\int (J(y(\xi), u) - J_1(\xi, u))p(\xi)\,d\xi = \int (J(y(\xi), u))p(\xi)\,d\xi - J(y(\bar{\xi}), u).$$

This suggests the following Monte Carlo approximation of $\hat{J}(y, u)$:

$$\hat{J}(y, u) \approx \frac{1}{N}\sum_{i=1}^{N}(J(y(\xi_i), u) - J_1(\xi_i, u_i)) + J(y(\bar{\xi}), u). \tag{2.5}$$

The variance of $J(y(\xi), u) - J_1(\xi, u)$ is given by

$$\int (J(y(\xi), u) - \overline{J(y(\xi), u)} - J_y(y(\bar{\xi}), u) y_\xi(\bar{\xi})(\xi - \bar{\xi}))^2 p(\xi)\, d\xi, \tag{2.6}$$

where $\overline{J(y(\xi), u)}$ is the mean of $J(y(\xi), u)$. In the following theorem, we use the variance of ξ to estimate the variance of $J - J_1$. Without loss of generality, we assume that ξ is a scalar random variable.

Theorem 1. *Let $m = \max|\frac{d}{d\xi} J(y(\xi), u)|$ and $M = \max|\frac{d^2}{d\xi^2} J(y(\xi), u)|$. The following estimate holds*

$$V(J) \leq m V(\xi)$$

and

$$V(J - J_1) \leq \frac{M^2}{2}(V^2(\xi) + E(\xi - \bar{\xi})^4).$$

Proof: The proof of the first inequality is straightforward. We only provide a proof for the second inequality. By the Taylor remainder formula there exists ξ_1 such that

$$J(y(\eta), u) - J(y(\bar{\xi}), u) = J_y(y(\bar{\xi}), u) y_\xi(\bar{\xi})(\eta - \bar{\xi}) + \frac{1}{2}\frac{d^2}{d\xi^2} J(y(\xi), u)|_{\xi=\xi_1}(\eta - \bar{\xi})^2.$$

Since $\bar{\xi}$ is the expectation of ξ, we have that

$$\overline{J(y(\xi), u)} - J(y(\bar{\xi}), u) = \int \frac{1}{2}\frac{d^2}{d\xi^2} J(y(\xi), u)|_{\xi=\bar{\xi}}(\eta - \bar{\xi})^2 p(\eta)\, d\eta.$$

Thus

$$|\overline{J(y(\xi), u)} - J(y(\bar{\xi}), u)| \leq \frac{M}{2}\int (\eta - \bar{\xi})^2 p(\eta)\, d\eta = \frac{M}{2} V(\xi). \tag{2.7}$$

Using the Tayor remainder formula, we get

$$|J(y(\xi), u) - J(y(\bar{\xi}), u) - J_y(y(\bar{\xi}), u) y_\xi(\bar{\xi})(\xi - \bar{\xi})| \leq \frac{M}{2}(\xi - \bar{\xi})^2. \tag{2.8}$$

Combining (2.6), (2.7) and (2.8) yields

$$\begin{aligned} V(J - J_1) = \int & (J(y(\xi), u) - J(y(\bar{\xi}), u) - J_y(y(\bar{\xi}), u) y_\xi(\bar{\xi})(\xi - \bar{\xi}) \\ & - (\overline{J(y(\xi), u)} - J(y(\bar{\xi}), u))^2 d\xi \end{aligned}$$

$$
\begin{aligned}
&\leq 2\int (J(y(\xi),u) - J(y(\bar{\xi}),u) - J_y(y(\bar{\xi}),u)y_\xi(\bar{\xi})(\xi-\bar{\xi}))^2 p(\xi)\,d\xi \\
&\quad + 2\int (\overline{J(y(\xi),u)} - J(y(\bar{\xi}),u))^2 p(\xi)\,d\xi \\
&\leq 2\int \frac{M^2}{4}(\xi-\bar{\xi})^4 p(\xi)\,d\xi + \frac{M^2}{2}V^2(\xi) \\
&= \frac{M^2}{2}(V^2(\xi) + E(\xi-\bar{\xi})^4).
\end{aligned}
$$

This completes the proof. □

The theorem indicates that the sensitivity derivative method is effective when the variance of ξ is small. The numerical experiments in Section 4 confirm this result.

3. Application to the generalized Burgers equation

Consider the generalized steady-state Burgers equation [10],

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\left(\nu\frac{\partial y}{\partial x}\right) + g(x), \quad \text{where } a < x < b \tag{3.1}$$

$$y(a) = u_1, \quad \text{and} \quad y(b) = u_2, \tag{3.2}$$

where $f(y) = \frac{1}{2}y(1-y)$. We treat the viscosity ν as a random variable. As a result, the solution $y = y(\nu, u)$ of Burgers equation is also a random variable.

Assume that $p(\nu)$ is the probability density function of ν, and let $u = (u_1, u_2)$. Define a cost function

$$J(y,u) = \int_a^b \int (y(\nu) - y_0)^2 p(\nu)\,d\nu\,dx + \epsilon\left(u_1^2 + u_2^2\right) \tag{3.3}$$

where y_0 is the given target state and ϵ is a constant. Our goal is to find control $\tilde{u} = (\tilde{u}_1, \tilde{u}_2)$ such that

$$J(\tilde{y}, \tilde{u}) = \min J(y,u),$$

where $\tilde{y}$ is the solution of (3.1)–(3.2) with boundary value $\tilde{u}$.

3.1. *Numerical solution*

To solve the Burgers equations (3.1)–(3.2) numerically, we use Newton's method on a uniform grid:

$$
\begin{aligned}
\left(\frac{\partial r}{\partial u}\right)\Delta u^{(n)} &= -r\left(u^{(n)}\right), \\
\Delta u^{(n+1)} &= u^{(n)} + \Delta u^{(n)},
\end{aligned}
$$

where

$$r(u) = \frac{\partial f}{\partial x} - \frac{\partial}{\partial x}\left(\nu \frac{\partial y}{\partial x}\right) - g(x).$$

We use a central difference scheme to discretize (3.1).

3.2. Sensitivity derivatives

In any gradient based optimization subroutine the sensitivity derivatives $y_1 = \partial_{u_1} y$ and $y_2 = \partial_{u_2} y$ are needed. To obtain y_1 and y_2, we differentiate (3.1)–(3.2) with respect to u_1 and u_2:

$$\begin{aligned} &\partial_x f' y_1 = \partial_x(\nu \partial_x y_1), \quad \text{where } a < x < b, \\ &y_1(a) = 1, \quad y_1(b) = 0, \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} &\partial_x f' y_2 = \partial_x(\nu \partial_x y_2), \quad \text{where } a < x < b, \\ &y_2(a) = 0, \quad y_2(b) = 1. \end{aligned} \tag{3.5}$$

Once y is obtained through Newton's iteration, a central difference scheme can also be used to obtain numerical solutions for y_1 and y_2.

In the Monte Carlo simulation, we need the sensitivity of y versus ν, denoted by $y_\nu := \partial_\nu y$. We differentiate (3.1)–(3.2) with respect to ν to obtain y_ν:

$$\begin{aligned} &\partial_x f' y_\nu = \partial_x(\nu \partial_x y_\nu) + \partial_{xx} y, \\ &y_\nu(a) = 0, \quad y_\nu(b) = 0. \end{aligned} \tag{3.6}$$

3.3. Evaluation of the cost function and its gradient

Without loss of generality, we assume that y_0 in (3.3) is zero. Let $\nu_1, \nu_2, \ldots, \nu_N$ be samples of ν generated by a random number generator. The SDMC evaluation of the first term of the cost function is given by

$$\int_a^b \left(\frac{1}{N} \sum_{i=1}^{N} (y^2(\nu_i) - y^2(\bar{\nu}) - 2y(\bar{\nu}) y_\nu(\bar{\nu})(\nu_i - \bar{\nu})) + y^2(\bar{\nu}) \right) dx. \tag{3.7}$$

The gradients of the cost function are computed by finite differences in our numerical experiments.

4. Results and conclusions

4.1. Evaluating the expectation and variance of the solution of Burgers equation

To demonstrate the effectiveness of SDMC, we evaluate the first and second statistical moments, $E(y) = \int y(\nu)p(\nu)\,d\nu$ and $E(y^2) = \int y^2(\nu)p(\nu)\,d\nu$ by both the general Monte Carlo method and SDMC, and compare them with the exact solution. To this end, we address Eqs. (3.1) and (3.2) with $a = -0.5$, $b = 0.5$, $g = 0$, $u_1 = \frac{1}{2}(1 + \tanh(\frac{-1}{8\nu}))$ and $u_2 = \frac{1}{2}(1+\tanh(\frac{1}{8\nu}))$. The exact solution is given by $y(\nu) = \frac{1}{2}(1+\tanh(\frac{x}{4\nu}))$. The experiment is conducted for two different probability distributions of ν: a uniform distribution on the interval $[0.1, 0.3]$ and a Gaussian distribution with expectation $E(\nu) = 2$ and standard deviation $S(\nu) = 0.1$. Since the exact solution of the Burgers equation is known, both $E(y)$ and $E(y^2)$ can be evaluated with high accuracy by a quadrature formula, which enables us to calculate the L^2 errors of the general Monte Carlo simulation and of SDMC for comparison.

Tables 1 and 2 list the L^2 errors in the evaluation of $E(y^2)$ using the general Monte Carlo method and SDMC respectively. Here, N denotes the number of samples or simulations,

Table 1. Errors and cpu times of Monte Carlo simulation: Uniform distribution.

N	$\|E(y^2) - E_m(y^2)\|$	$\|E(y^2) - E_{sd}(y^2)\|$	CPU (MC)	CPU (SDMC)
100	0.7521E-02	0.1442E-02	1.110	1.156
200	0.6347E-02	0.2059E-02	1.896	2.000
400	0.1991E-02	0.4305E-03	3.467	3.676
800	0.3559E-02	0.1104E-02	6.613	7.031
1600	0.4262E-02	0.8747E-03	12.911	13.744
3200	0.2034E-02	0.1495E-03	25.533	27.195
6400	0.1818E-02	0.1446E-03	50.711	54.071
12800	0.7114E-03	0.9582E-04	101.213	108.00

Table 2. Errors of Monte Carlo simulation: Gaussian distribution.

N	$\|E(y^2) - E_m(y^2)\|$	$\|E(y^2) - E_{sd}(y^2)\|$
100	0.2577E-04	0.4010E-05
200	0.1500E-04	0.4498E-05
400	0.2450E-04	0.2912E-05
800	0.1009E-04	0.1188E-05
1600	0.1214E-04	0.8144E-06
3200	0.6341E-05	0.9881E-06
6400	0.7657E-05	0.6245E-06
12800	0.6650E-05	0.5476E-06

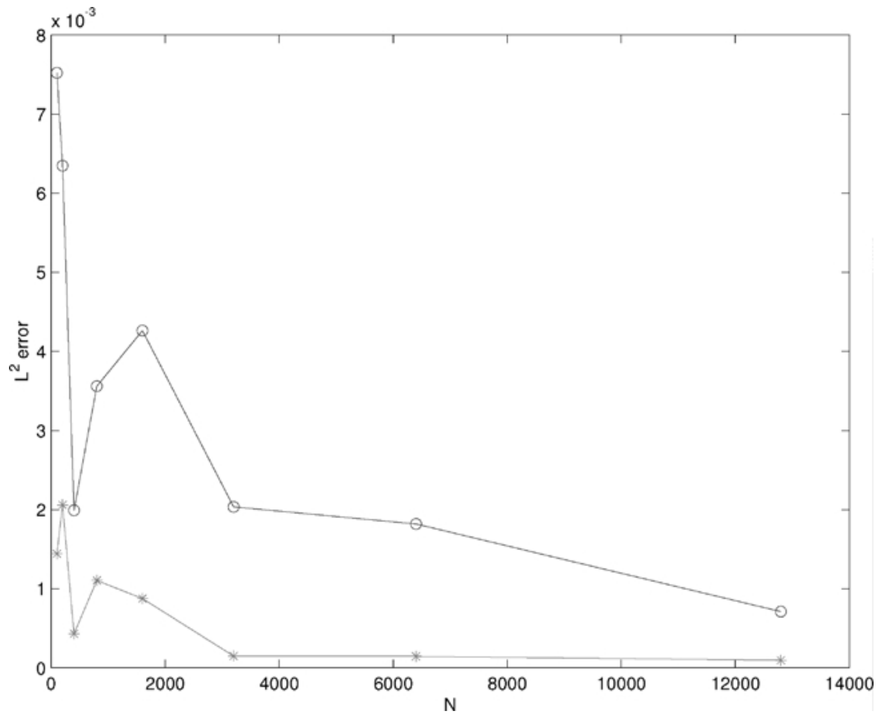

Figure 1. The L^2 error of the Monte Carlo approximations of $E(y^2)$: Uniform distribution of ν o: General Monte Carlo; $*$: SDMC.

$E_m(y^2)$ represents the general Monte Carlo simulation result, $E_{sd}(y^2)$ represents the SDMC result. The last two columns provide the cpu time for MC and SDMC methods respectively. For a fixed number of samples, note that the extra cost of SDMC (associated with the computation of sensitivity derivatives) is less than 7% of the cost of the MC method in terms of cpu time. A simple linear regression analysis shows that the convergence rates of both general Monte Carlo and SDMC are roughly $O(1/\sqrt{N})$. These tables show that the L^2 errors of SDMC are consistently smaller than the errors of general Monte Carlo by an order of magnitude. These errors are plotted in figure 1 (general Monte Carlo) and figure 2 (SDMC) as functions of N. The effectiveness of the variance reduction of SDMC can clearly be seen.

In figure 3, we plot the results of SDMC simulation for the solution y of Burgers equation and its confidence intervals. The upper and lower limits of the confidence intervals (shown as the upper and lower curves in figure 3) are defined to be $E_{sd}(y) - 2S_{md}(y)$ and $E_{md}(y) + 2S_{md}(y)$, respectively. Here S_{md} represents the approximation of the standard deviation of y using SDMC. In between the two curves lie the solutions of Burgers equation from 40 samples . Here, we choose $N = 200$ for the confidence intervals and a uniform distribution of ν. Figure 4 shows the analogous results for the Gaussian distribution of ν.

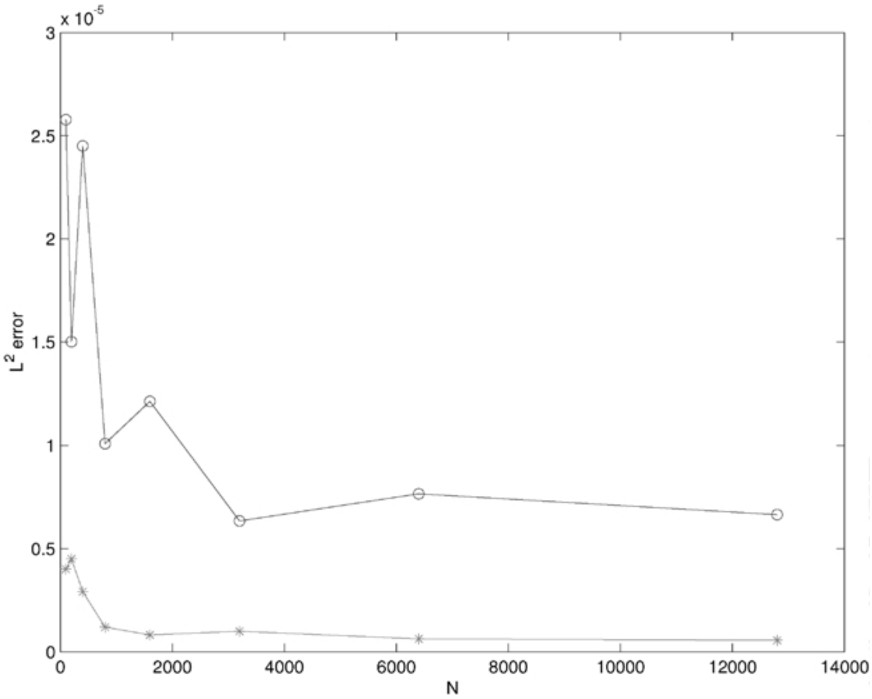

Figure 2. The L^2 error of the Monte Carlo approximations of $E(y^2)$: Gaussian distribution of ν o: General Monte Carlo; $*$: SDMC.

4.2. *Optimal control problem*

Uniform distribution of ν. We solve the optimal control problem described in Section 3 with g defined as $g(x) = 2 + \sin(x)$ in (3.1) and with the random variable ν having a uniform distribution on [0.1, 0.3]. Let $u^0 = (u_1^0, u_2^0)$ be an initial guess. The numerical algorithm for finding the minimizer of the objective function $\hat{J}$ involves 3 key steps.

Step 1: For boundary values $u^n = (u_1^n, u_2^n)$, solve the state Eq. (3.1) to obtain y^n and solve the sensitivity Eqs. (3.4)–(3.6) to obtain the sensitivity of y^n versus u_1, u_2 and ν.

Step 2: Evaluate the cost function $\hat{J}^n = \hat{J}(y^n, u^n)$ using the sensitivity derivative Monte Carlo method.

Step 3: Call the optimization subroutine [6] with $\hat{J}^n$ as input. This subroutine minimizes general unconstrained objective function using finite difference gradients and secant hessian approximations. The output of the subroutine is a new boundary value $u^{n+1} = h(u^n, \hat{J}^n)$. The subroutine checks if u^{n+1} is a satisfactory minimizer of J; if it is, it stops; if not, it goes back to Step 1.

After several iterations, the optimization subroutine finds 0.4620 as the minimum of $\hat{J}$ at $u = (0.3905, 0.1501)$.

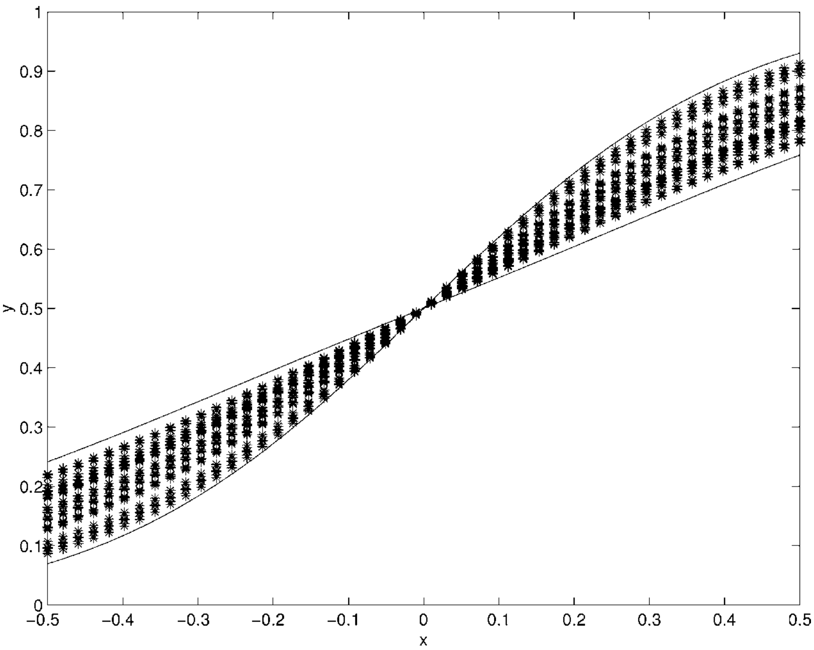

Figure 3. The confidence interval curves and the solutions y from simulations: Uniform distribution of ν.

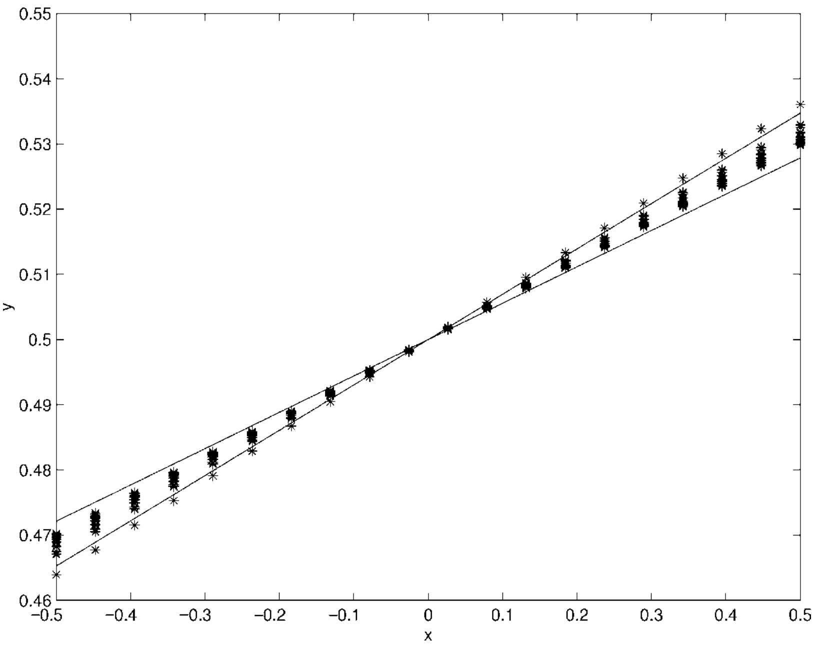

Figure 4. The confidence interval curves and the solutions y from simulations: Gaussian distribution of ν.

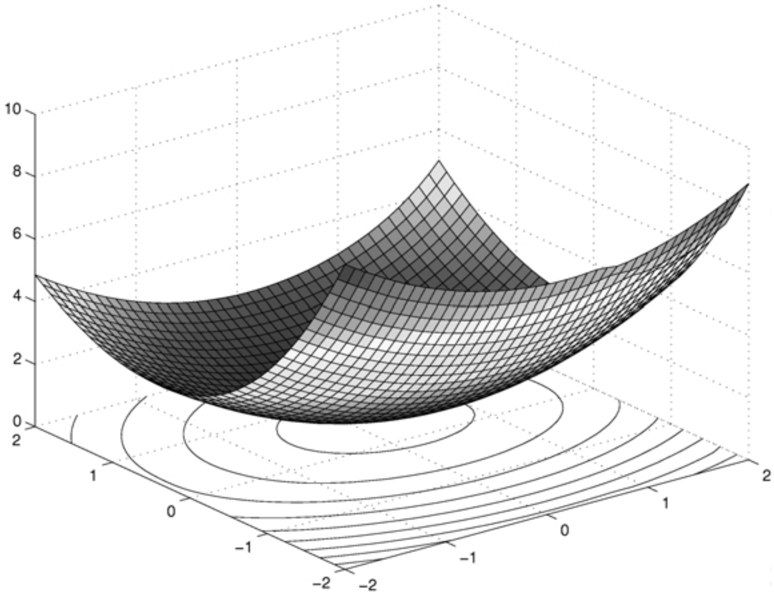

Figure 5. Cost function J of the optimal control problem: Uniform distribution of ν.

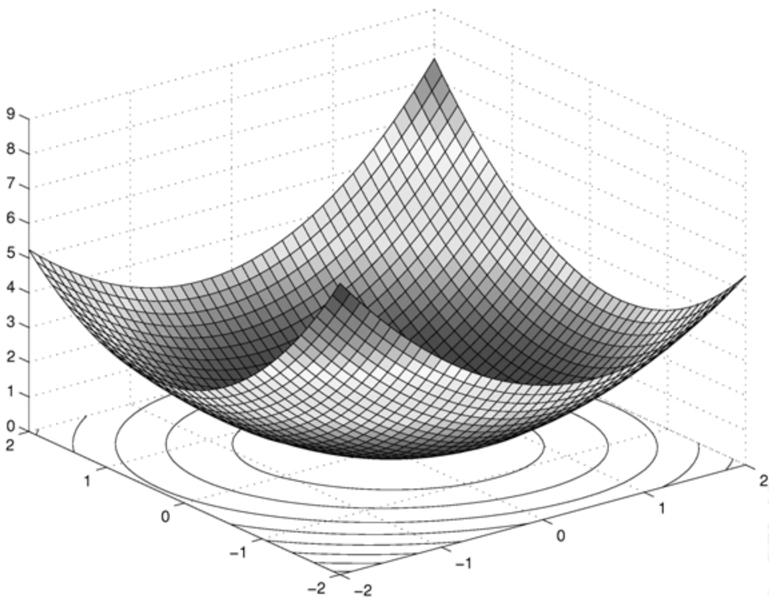

Figure 6. Cost function J of the optimal control problem: Gaussian distribution of ν.

Since the control variable u only has two components: u_1 and u_2, we can evaluate and graph the cost function and its contour (figure 5). The graph confirm that (i) the optimal control problem with uncertainty state variable has a unique solution, and (ii) SDMC algorithm indeed finds the global minimizer for the cost function $\hat{J}$.

Gaussian distribution of ν. This example is analogous to the previous one except that ν has a Gaussian distribution with $E(\nu) = 2$ and $S(\nu) = 0.1$. The optimization algorithm finds the least cost $\hat{J} = 0.0073$ at the boundary value $\tilde{u} = (0.05189, 0.05948)$. Figure 6 shows the graph and the contour of the cost function. Again the minimizer found by our optimization algorithm is consistent with the graph.

In conclusion, we have developed a sensitivity derivative based Monte Carlo method for solving optimal control problems with uncertain parameters. Theoretical analysis proves its effectiveness in variance reduction, and it is confirmed by the numerical results of a model problem governed by the generalized steady-state Burgers equation. Encouraged by these promising results, we are studying some shape optimization and boundary control problems (under uncertain conditions) governed by the Euler and Navier-Stokes equations.

References

1. J. Borggaard, D. Pelletier, and E. Turgeon, "Parametric uncertainty analysis for thermal fluid calculations," Journal of Nonlinear Analysis: Series A Theory and Methods, vol. 47, no. 7, pp. 4533–4543, 2001.
2. Y. Cao and M.Y. Hussaini, "On convergence of optimal control problem to exact control problem for systems governed by Navier-Stokes equations," in Proceedings of ICNPPA-2000, 2001, vol. 1, pp. 117–126.
3. Y. Cao and D. Stanescu, "Shape optimization for noise radiation problems," Comp. & Math. Appl., vol. 44, no. 12, pp. 1539–1556, 2002.
4. B. DeVolder, J. Glimm, J. Grove, Y. Kang, Y. Lee, K. Pao, D. Sharp, and K. Ye, "Uncertainty quantification for multiscale simulations," J. Fluids Eng., vol. 124, pp. 29–41, 2002.
5. X. Du and W. Chen, "Towards a better understanding of modeling feasibility robustness in engineering design," Amer. Soc. Mech. Eng., DAC-8565, Sept. 1999.
6. D. Gay, "Algorithm 611, subroutines for unconstrained minimization using a model/trust region approach," ACM Transactions on Mathematical Software, vol. 9, no. 4, pp. 503–524, 1983.
7. J. Glimm and H. Sharp, "Prediction and the quantification of uncertainty," Physica D, vol. 133, pp. 152–170, 1999.
8. M. Gunzburger, "Sensitivites, adjoints, and flow optimization," Inter. J. Num. Meth. Fluids., vol. 31, pp. 53–78, 1999.
9. L. Huyse, "Solving problems of optimization under uncertainty as statistical decision problem," AIAA Paper, 2001-1519, 2001.
10. L. Huyse and R. Walters, "Random field solutions including boundary condition uncertainty for the steady state generalized Burgers equation," ICASE Report, 2001-35, 2001.
11. R. Joslin, M. Gunzburger, R. Nicolaides, G. Erlebacher, and M.Y. Hussaini, "A self contained, automated methodology for optimal flow control validated for transition delay," AIAA Journal, vol. 35, pp. 816–824, 1997.
12. A. Parkinson, C. Sorenson, and N. Pourhassan, "A general approach for robust design," J. Mech. Design, vol. 115, no. 1, pp. 74–80, 1993.
13. M. Putko, P. Newman, A. Taylor III, and L. Green, "Approach for uncertainty propagation and robust design in CFD using sensitivity derivatives," AIAA Paper, 2001-2528, 2001.
14. P. Thandedar and S. Kodiyalam, "Structural optimization using probabilistic constraints," Structure Opt., vol. 4, pp. 236–240, 1991.