Numerical Solution of a Class of Parabolic Partial Differential Equations Arising in Optimal Control Problems with Uncertainty

M. Dahleh
Department of Mechanical Engineering, University of California at Santa Barbara, Santa Barbara, California 93106

A. P. Peirce
McMaster University, Hamilton, Ontario, Canada L8S 4K1

Received 5 October 1990

In this paper the optimal control of uncertain parabolic systems of partial differential equations is investigated. In order to search for controllers that are insensitive to uncertainties in these systems, an iterative optimization procedure is proposed. This procedure involves the solution of a set of operator valued parabolic partial differential equations. The existence and uniqueness of solutions to these operator equations is proved, and a stable numerical algorithm to approximate the uncertain optimal control problem is proposed. The viability of the proposed algorithm is demonstrated by applying it to the control of parabolic systems having two different types of uncertainty.

I. INTRODUCTION

Optimal control of infinite-dimensional systems is a rapidly growing field, with applications in many engineering disciplines. Modeling of most physical systems results in dynamical equations that evolve in infinite-dimensional spaces [1–4]. This is due to the fact that there is always an infinite number of modes in physical systems. These modes, if appropriately controlled, may not affect the overall performance of the system, and thus the standard procedure of reduced-order modeling is justified. On the other hand many systems of practical importance do not have this property, and to accurately model them the distributed nature of the dynamics ought to be taken into consideration.

In addition to the problem of infinite dimensionality all models of physical systems are plagued with uncertainties. The effect of uncertainties in control problems can be so severe as to completely degrade the performance, and even to destabilize the system. In this article we will address the issue of optimal control under the influence of uncertainty, and show how the problem reduces to a parameter optimization problem with dynamic constraints that are represented by partial differential equations.

In the process of designing optimal control systems for distributed parameter systems the solution of certain partial differential equations arise. These equations, which de-

Numerical Methods for Partial Differential Equations, 8, 77–95 (1992)
© 1992 John Wiley & Sons, Inc. CCC 0749-159X/92/010077-19$04.00

scribe the evolution of a certain performance index, form the basis of an iterative procedure for finding the optimal solution. Reliable numerical approximation of these equations is the crux of the optimal control problem. In this article we will formulate the problem, analyze the existence and uniqueness of the resulting partial differential equations, and present a numerical study of the equations that illustrate the effectiveness of the method.

II. OPTIMAL CONTROL PROBLEM

The systems considered in this article are modeled by linear, time-invariant, infinite-dimensional evolution equations in a real, separable Hilbert space X. The parametric uncertainty consists of a random variable α taking values in a closed and bounded set W in a Euclidean space R^m, according to a known probability density $\mu(d\alpha)$:

$$\frac{dx(t)}{dt} = A(\alpha)x(t) + B(\alpha)u(t),$$

where $x(t_0) = x_0$ is a given element in X [5, 6]. The control $u(t)$ is assumed to be of the form $u(t) = Kx(t)$, where K is a Hilbert-Schmidt operator mapping X into itself. In specific we will consider the case where ψ is a bounded open set in R^n with smooth boundary Γ, $X = L^2(\psi)$ and $A(\alpha)$ is the second-order elliptic operator

$$A(\alpha)\phi = \sum_{i,j=1}^{n} a_{ij}(\alpha)\frac{\partial^2\phi}{\partial x_i x_j} + a_0(\alpha)\phi$$

where $a_{ij}(\alpha), a_0(\alpha) \in R$ are continuous functions of α, and satisfy the following ellipticity requirement

$$\sum_{ij} a_{ij}(\alpha)b_i b_j \geq c\sum_{i} b_i b_i$$

and c is a positive constant independent of α. The parameter of uncertainty α belongs to W. The input $B(\alpha)$ is a bounded operator for each α which maps X into itself, and as a function of α it is assumed to be continuous. The initial condition is a Gaussian random functional on the dual space $(X)^*$, defined by

$$x_0(\phi^*) = \langle x_0, \phi^*\rangle_X .$$

The expectation of x_0 is defined by

$$E(x_0(\phi^*)) = \int_\Omega x_0(\phi^*)(\omega)\, dP(\omega)$$

where $(\Omega, \mathbf{B}, P)$ is an underlying probability space. Thus $E(x_0(\phi^*))$ is a bounded linear operator on the dual space X^*, and the Reisz representation theorem can be used to write

$$E(x_0(\phi^*)) = \langle m, \phi^*\rangle_X$$

for some m in X. In a similar fashion the covariance operator can be defined as

$$\Gamma(\phi_1^*, \phi_2^*) = E[(x_0(\phi_1^*) - E(x_0(\phi_1^*)))(x_0(\phi_2^*) - E(x_0(\phi_2^*)))].$$

From this it is clear that the covariance operator is a continuous bilinear form on $X \times X$, and thus can be written as

$$\Gamma(\phi_1^*, \phi_2^*) = \langle \Lambda\phi_1^*, \phi_2^* \rangle$$

where Λ is a positive self-adjoint operator.

The Hilbert space valued random variable x_0 induces a cylindrical measure on X given by the following finite-dimensional densities:

$$d\mu_{\phi_1^* \cdots \phi_n^*} = \frac{1}{(2\pi)^{n/2}} \exp(-\tfrac{1}{2}(x - m_n)^T \Lambda_n^{-1}(x - m_n))\, dx_1 \cdots dx_n$$

where

$$m_n = \begin{bmatrix} \langle m, \phi_1^* \rangle \\ \vdots \\ \langle m, \phi_n^* \rangle \end{bmatrix}$$

$(\Lambda_n)_{ij} = \langle \Lambda\phi_i^*, \phi_j^* \rangle$, and $\{\phi_i\}$ is an orthonormal basis for X. Any bounded, positive, and self-adjoint operator Λ induces a cylindrical measure on X, but the only class of covariance operators which induce regular measures on X is exactly the class of nuclear (trace-class) operators [7]. On the basis of this it is reasonable to assume that the initial condition has a nuclear operator for a covariance. Viewing the initial condition as a Hilbert space valued random variable enables us to remove the dependence of the optimal control on the initial conditions by formulating the following parameter optimization problem:

$$\text{minimize } J(K) = E_{\alpha,x_0}\left(\int_0^T \langle x(s) - r(s), Q(x(s) - r(s)) \rangle_X ds \right) + \langle K, K \rangle_{\text{H.S.}}$$

subject to

$$\frac{dx(t)}{dt} = A(\alpha)x(t) + B(\alpha)u(t) \tag{1}$$

over all $K \in X_{\text{H.S.}}$, the Hilbert space of Hilbert-Schmidt operators. The function $r(s)$ is an element of the Hilbert space $L([0,T]; X)$, and the notation $E_{x_0,\alpha}$ means expectation with respect to x_0, α. The inner product between any two elements of $X_{\text{H.S.}}$ say K and M, is given by [8]

$$\langle K, M \rangle_{\text{H.S.}} = \text{trace}(\text{KM}^*).$$

The initial condition $x(t_0) = x_0$ is a given element in X, and Q is a positive, self-adjoint, Hilbert-Schmidt operator.

III. NECESSARY CONDITIONS FOR OPTIMALITY

We will reformulate the problem by converting it to a final value optimization problem. This is done by obtaining a system of operator equations that evaluate the performance index $J_{\alpha,x_0}(K)$. The advantage of this approach is that the dependence of $J_{\alpha,x_0}(K)$ on the initial condition is displayed explicitly, and thus can be removed by averaging. By a standard calculation we can get

$$J_{\alpha,x_0}(K) = \langle x_0, P(0,\alpha)x_0 \rangle_X + \langle K, K \rangle_{\text{H.S.}} + \langle x_0, \nu(0) \rangle_X$$

where $P(0,\alpha)$ is evaluated by

$$\frac{dP(t,\alpha)}{dt} + (A(\alpha) + B(\alpha)K)^*P(t,\alpha) + P(t,\alpha)(A(\alpha) + B(\alpha)K) + Q = 0, \qquad P(T,\alpha) = 0, \tag{2a}$$

$$\frac{dv(t,\alpha)}{dt} + (A(\alpha) + B(\alpha)K)^*v(t,\alpha) - 2Qr = 0, \tag{2b}$$

$$\begin{aligned} J(K) &= E_{x_0,\alpha}\langle x_0, P(0,\alpha)x_0\rangle_X + \langle K,K\rangle_{\text{H.S.}} + E_{x_0,\alpha}\langle x_0, v(0)\rangle_X \\ &= \int_W d\mu(\alpha)(\langle P(0,\alpha),\Lambda\rangle_{\text{H.S.}} + \langle m, P(0,\alpha)m\rangle_X + \langle m, v(0,\alpha)\rangle_X) + \langle K,K\rangle_{\text{H.S.}} \end{aligned} \tag{3}$$

The new optimization problem will be

$$\min_{K\in X_{\text{H.S.}}} J(K) \tag{4}$$

subject to Eqs. (2a) and (2b).

To form an iterative numerical procedure to search for the optimal solution, we obtain a set of necessary conditions for the optimal solution via Lagrange multipliers. Let $N(t,\alpha)$ and $z(t,\alpha)$ denote the Lagrange multipliers,

$$N(t,\alpha) \in L^2([0,T]\times W; X_{\text{H.S.}})^* \simeq L^2([0,T]\times W; X_{\text{H.S.}})$$
$$z(t,\alpha) \in L^2([0,T]\times W; X)^* \simeq L^2([0,T]\times W; X).$$

Define the quantities $M_1(\alpha,K,P,v,t)$, $M_2(\alpha,K,P,v,t)$, $S_1(\alpha,K,P,v,t)$, and $S_2(\alpha,K,P,v,t)$ as follows:

$$M_1(\alpha,K,P,v,t) = \frac{dP(t,\alpha)}{dt} + (A(\alpha) + B(\alpha)K)^*P(t,\alpha) + P(t,\alpha)(A(\alpha) + B(\alpha)K) + Q,$$

$$M_2(\alpha,K,P,v,t) = \frac{dN(t,\alpha)}{dt} - (A(\alpha) + B(\alpha)K)N(t,\alpha) - N(t,\alpha)(A(\alpha) + B(\alpha)K)^*,$$

$$S_1(\alpha,K,P,v,t) = \frac{dv(t,\alpha)}{dt} + (A(\alpha) + B(\alpha)K)^*v(t,\alpha),$$

$$S_2(\alpha,K,P,v,t) = \frac{dz(t,\alpha)}{dt} - (A(\alpha) + B(\alpha)K)z(t,\alpha).$$

The first-order necessary conditions can be obtained by taking Frechet derivatives (with respect to P, N, v, z, and K) of the Lagrangian L defined by

$$\begin{aligned} L = &\int_W d\mu(\alpha)(\langle P(0,\alpha),\Lambda\rangle_{\text{H.S.}} + \langle m, P(0,\alpha)m\rangle_X + \langle m, v(0,\alpha)\rangle_X) + \langle K,K\rangle_{\text{H.S.}} \\ &+ \int_0^T\int_W d\mu(\alpha)\,dt(\langle N(t,\alpha), M_1(\alpha,K,P,v,t)\rangle_{\text{H.S.}} + \langle z(t,\alpha), S_1(\alpha,K,P,v,t)\rangle_X) \end{aligned}$$

yielding

$$M_1(\alpha,K,P,v,t) = 0 \qquad P(T,\alpha) = 0 \tag{5a}$$

$$M_2(\alpha,K,P,v,t) = 0 \qquad N(0,\alpha) = \Lambda + m_0 \otimes m_0 \tag{5b}$$

$$S_1(\alpha, K, P, v, t) = 0 \tag{6a}$$

$$v(T, \alpha) = 0$$

$$S_2(\alpha, K, P, v, t) = 0 \tag{6b}$$

$$z(0, \alpha) = 0$$

$$K = -\int_0^T \int_W d\mu(\alpha)\, dt((P(t,\alpha)B(\alpha))^* N(t,\alpha) - \tfrac{1}{2}(v(t,\alpha)^* B) z(t,\alpha)^*) . \tag{7}$$

Equations (5)–(7) constitute a system of transcendental equations that characterize the optimal solution for the parameter optimization problem.

IV. SYSTEMS GOVERNED BY PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

The main component in the numerical solution of the optimization problem is the solution of the operator Eqs. (5a) and (5b). In this section we will investigate the issue of existence and uniqueness of solutions to such equations. For this purpose will restrict ourselves to Eq. (2a) and assume, without loss of generality, that Q is equal to the identity operator denoted by I. Note that the situation in Eq. (6a) is identical. We will analyze the operator equation

$$\frac{dP}{dt} = (A + BK)^* P + P(A + BK) + I \tag{8}$$

$$P(0) = 0$$

by converting it to a parabolic partial equation using the Schwarz kernel theorem [9], and determine under what conditions it has a weak solution. By the Schwarz kernel theorem we can write

$$\langle \phi_1, P\phi_2 \rangle_X = \int_\psi \phi_1(\xi_1) P(t, \xi_1, \xi_2) \phi_2(\xi_2)\, d\xi_1\, d\xi_2 .$$

From this representation it follows that Eq. (8) leads to a partial differential equation describing the evolution of the kernel $P(t, \xi_1, \xi_2)$

$$\frac{\partial P(t,\xi_1,\xi_2)}{\partial t} = (A + BK)^*_{\xi_1} P(t,\xi_1,\xi_2) + (A + BK)^*_{\xi_2} P(t,\xi_1,\xi_2) + \delta(\xi_1 - \xi_2) \tag{9}$$

subject to

$$P(0, \xi_1, \xi_2) = 0$$

$$P(t, \xi_1, \xi_2) = 0 \quad \xi_1 \in \Gamma, \;\; \xi_2 \in \psi$$

$$P(t, \xi_1, \xi_2) = 0 \quad \xi_1 \in \psi, \;\; \xi_2 \in \Gamma .$$

The identity operator gives rise to a Dirac δ function with support along the diagonal of the region $\psi \times \psi$. This is exactly the point of difficulty which requires further analysis. However, we can still prove existence and uniqueness of a weak solution under a restrictive condition.

First we need some notation. Let $\Sigma = [0,T] \times \psi \times \psi$ and define the following Sobolev spaces

$$H_0^{1,1}(\Sigma) = \left\{ g \middle| g, \frac{\partial g}{\partial t}, \frac{\partial g}{\partial x_i} \in L^2(\Sigma), g|_{\partial(\psi\times\psi)} = 0 \right\}$$

$$H^{2,1}(\Sigma) = \left\{ g \middle| g, \frac{\partial g}{\partial t}, \frac{\partial g}{\partial x_i}, \frac{\partial^2 g}{\partial x_i x_j} \in L^2(\Sigma) \right\}$$

$$\tilde{H}_0^{1,1}(\Sigma) = \{ g | g \in H_0^{1,1}, \text{ and } g(T) = 0 \}$$

$P(t,\xi_1,\xi_2)$ will be called a weak solution of (9) if it satisfies the following integral equation

$$\int_\Sigma P(t,\xi_1,\xi_2)\left(-\frac{\partial G}{\partial t} - (A+BK)_{\xi_1}G - (A+BK)_{\xi_2}G\right) dt\, d\xi_1\, d\xi_2$$

$$= \int_0^T \int_\psi G(t,\xi_1,\xi_1)\, dt\, d\xi_1$$

for every $G(t,\xi_1,\xi_2) \in H^{2,1} \cap \tilde{H}_0^{1,1}$.

Theorem 1: *Under the assumptions stated above, if $\psi \subset R^n$ with $n \leq 3$ then there exists a unique weak solution $P(t,\xi_1,\xi_2) \in L^2(\Sigma)$.*

The proof of this proposition is based on the following two lemmas whose proofs are given in the appendix.

Lemma 1: Consider the map $L\colon H^{2,1} \cap \tilde{H}_0^{1,1} \to R$ defined by

$$L(G) = \int_0^T \int_\psi G(t,\xi_1,\xi_1)\, dt\, d\xi_1$$

L is a continuous linear functional on $H^{2,1} \cap \tilde{H}_0^{1,1}$.

Lemma 2: The map

$$M = \left(-\frac{\partial}{\partial t} - (A+BK)_{\xi_1} - (A+BK)_{\xi_2}\right): \tilde{H}_0^{1,1}(\Sigma) \cap H^{2,1}(\Sigma) \longrightarrow L^2(\Sigma)$$

is an isomorphism.

Proof of Theorem 1: Define a linear functional on $L^2(\Sigma)$ by

$$g(\phi) = L(M^{-1}(\phi)) \qquad \text{for all } \phi \in L^2(\Sigma)$$

then

$$|g(\phi)| = |L(M^{-1}(\phi))|.$$

By using Lemma 1 we get

$$|g(\phi)| \leq K\|M^{-1}(\phi)\|_{\tilde{H}_0^{1,1}\cap H^{2,1}}.$$

Applying Lemma 2 we get

$$|g(\phi)| \leq K\|\phi\|_{L^2}.$$

From the Reisz representation theorem there exists a unique $P \in L^2(\Sigma)$ such that

$$g(\phi) = \int_\Sigma P(t,\xi_1,\xi_2)\phi(t,\xi_1,\xi_2)\, dt\, d\xi_1\, d\xi_2.$$

Now for every $G \in \tilde{H}_0^{1,1} \cap H^{2,1}, M(G) \in L^2(\Sigma)$ and

$$g(M(G)) = L(M^{-1}(M(G))) = L(G) = \int_{\Sigma} P(t,\xi_1,\xi_2)M(G)\,dt\,d\xi_1\,d\xi_2.$$

Thus P is the desired weak solution. ■

V. NUMERICAL OPTIMAL CONTROL OF UNCERTAIN INFINITE-DIMENSIONAL SYSTEMS

In this section we illustrate the optimal design procedure outlined in Secs. II and III by considering the approximate numerical design of two uncertain infinite-dimensional controllers. Firstly, the optimal control of the heat equation in which the initial conditions (IC) are uncertain. Secondly, the optimal control of the heat equation in which the diffusion coefficient is considered to be an uncertain parameter. In both cases the performance of the optimal controller designed to cater for uncertainties (referred to as the insensitive control) is compared to the performance of the optimal control tailored to the mean initial condition/parameter value (which is referred to as the sensitive control).

A. Numerical Optimal Control of the Heat Equation with Uncertain Initial Conditions

In this example we consider the tracking optimization problem:

$$\text{minimize } J(K) = E_{x_0}\left(\int_0^T \langle x(s) - r, Q(x(s) - r)\rangle_X\, ds\right) + \gamma\int_0^T \langle K(s), K(s)\rangle_{\text{H.S.}}\, ds \quad (10)$$

$$\text{subject to: } \frac{dx(t)}{dt} = Ax + Bu, \qquad u = -Kx, \qquad x(0) = x_0. \quad (11)$$

Here we assume that $A = D(d^2/d\xi^2)$ and that $\xi \in (0, L)$. Following the procedure outlined in Sec. III we obtain the equations for the performance index:

$$J_{x_0}(K) = \langle x_0, P(0)x_0\rangle_X + \langle x_0, v(0)\rangle_X + \gamma\int_0^T \langle K(s), K(s)\rangle_{\text{H.S.}}\, ds \quad (12)$$

where $P(0)$ is determined by solving

$$\frac{dP(t)}{dt} + (A - BK)^*P(t) + P(t)(A - BK) + Q = 0, \qquad P(T) = 0 \quad (13)$$

and $v(0)$ is determined by solving

$$\frac{dv(t)}{dt} + (A - BK)^*v(t) - 2Qr = 0, \qquad v(T) = 0. \quad (14)$$

Making use of Eq. (12) and assuming that $x_0 \in X$ is a Hilbert-space-valued Gaussian random variable with mean $\bar{x}_0$ and covariance operator Λ, we obtain the new optimization problem:

$$\text{minimize } J(K) = \langle P(0), \Lambda\rangle_{\text{H.S.}} + \langle \bar{x}_0, P(0)\bar{x}_0\rangle_X + \langle \bar{x}_0, v(0)\rangle_X + \gamma\int_0^T \langle K(s), K(s)\rangle_{\text{H.S.}}\, ds$$

subject to Eqs. (13) and (14).

We define Lagrange multiplier functions $N(t) \in L^2([0,T], X_{\text{H.S.}})^*$, $z(t) \in L^2([0,T], X)^*$ and the corresponding Lagrangian L as was done in Sec. III. Taking Frechet derivatives of L with respect to P, v, N, z, and K we obtain Eqs. (13) and (14) as necessary conditions in addition to

$$\frac{dN(t)}{dt} = (A - BK)N(t) + N(t)(A - BK)^*; \qquad N(0) = \Lambda + \tilde{x}_0 x_0^* \tag{15}$$

$$\frac{dz(t)}{dt} = (A - BK)z(t), \qquad z(0) = \tilde{x}_0 \tag{16}$$

$$0 = \int_0^T \langle 2\gamma K(s) - 2(P(s)B)^* N(s) - (v^*(s)B)z^*(s), \delta K(s)\rangle_{\text{H.S.}} ds. \tag{17}$$

Given a trial controller $K^{(k)}$ we determine an approximate solution to Eqs. (13)–(16) numerically using finite differences. In the case of Eqs. (14) and (16), which only involve one spatial dimension, we divide the domain $[0, L]$ into N equal subintervals of length $\Delta\xi = L/N$ and denote the mesh points formed by the end points of these subintervals by $\xi_n = n\Delta\xi$. The numerical solution at such a mesh point is denoted by $v_n^{(k)} \approx v^{(k)}(t, \xi_n)$ and similarly for $z_n^{(k)}$. Using a central difference approximation to A, Eq. (14) reduces to the following system of ordinary differential equations (ODEs):

$$\frac{dv_i}{dt} + \hat{A}_i(t)v_i - 2\sum_m Q_{im} r_m = 0, \qquad v_i(T) = 0, \qquad i, m = 0, \ldots, N, \tag{18}$$

where $\hat{A}_i(t)v_i = \theta(v_{i-1} + v_{i+1}) + \phi_i(t)v_i$, $\theta = D\Delta\xi^{-2}$, and $\phi_i(t) = -2\theta - K_i^{(n)}(t)$. Here we have assumed that we have full observation so that $B = \delta_{ij}$, and that the operator K is diagonal so that $K_{ij} = K_i\delta_{ij}$. This assumption is a matter of computational convenience and will be made for the remainder of the numerical experiments performed in this article. A central difference approximation to Eq. (16) yields a similar system of ODEs for z_i to those given in Eq. (18).

In the case of the operator Eqs. (13) and (15) we divide the product domain $[0,L] \times [0,L]$ into the mesh of N^2 square cells that are formed by the Cartesian product of the subintervals used in Eq. (18). Using a central difference approximation to A, Eq. (13) reduces to the following system of ODEs:

$$\frac{dP_{ij}}{dt} + [\hat{A}_{1,i}(t) + \hat{A}_{2,j}(t)]P_{ij} + Q_{ij} = 0, \qquad P_{ij}(T) = 0, \qquad i, j = 0, \ldots, N, \tag{19}$$

where $\hat{A}_{1,i}(t)P_{ij} = \theta(P_{i-1j} + P_{i+1j}) + \phi_i(t)P_{ij}$, $\hat{A}_{2,j}(t)P_{ij} = \theta(P_{ij-1} + P_{ij+1}) + \phi_j(t)P_{ij}$, and θ and $\phi_i(t)$ are defined in Eq. (18). A central difference approximation to the operator equation (15) yields a similar system of ODEs for N_{ij} to those given in Eq. (19).

To solve the system of ODEs (18) we use the Crank-Nicholson (CN) procedure [10]. The time interval $[0, T]$ is divided into M subintervals of length $\Delta t = T/M$. The end points of these subintervals are denoted by $t_m = m\Delta t$, and superscripts are used to denote the time step at which a quantity is evaluated, e.g., $v_n^m \approx v(t_m, \xi_n)$. The CN procedure can be expressed in the form

$$\left[I - \frac{\Delta t}{2}\hat{A}_n^{(k),m-1}\right]v_n^{(k),m-1} = \left[I + \frac{\Delta t}{2}\hat{A}_n^{(k),m}\right]v_n^{(k),m} - 2\Delta t \sum_{l=0}^{N} Q_{nl} r_l. \tag{20}$$

The numerical solution $\{v_n^{(k),m}\}$ can be found by a marching process using Eq. (20), which involves an inversion of the tridiagonal matrix on the left-hand side of (20).

To solve the system of ODEs (19), which result from approximating the operator equation (13), we use an alternating direction implicit (ADI) scheme [10]:

$$\left[I - \frac{\Delta t}{2}\hat{A}_{1,i}^{(k),m-(1/2)}\right]P_{i,j}^{(k),m-(1/2)} = \left[I + \frac{\Delta t}{2}\hat{A}_{2,j}^{(k),m}\right]P_{ij}^{(k),m} + \frac{\Delta t}{2}Q_{ij}, \tag{21a}$$

$$\left[I - \frac{\Delta t}{2}\hat{A}_{2,j}^{(k),m-1}\right]P_{ij}^{(k),m-1} = \left[I + \frac{\Delta t}{2}\hat{A}_{1,i}^{(k),m}\right]P_{ij}^{(k),m-(1/2)} + \frac{\Delta t}{2}Q_{ij}. \tag{21b}$$

Assuming that $P_{ij}^{(k),m}$ is known, we invert the tridiagonal matrix on the left-hand side of Eq. (21a) to obtain $P_{ij}^{(k),m-(1/2)}$. The right-hand side of Eq. (21b) is now known and the tridiagonal matrix on the left-hand side is now inverted to yield $P_{ij}^{(k),m-1}$. This backward marching procedure is used to determine the approximate solution to Eq. (13). We assumed above that $K^{(k)}$ was known, and calculated the corresponding propagators for the performance index $P_{ij}^{(k)}$ and $v_i^{(k)}$ and the Lagrange multipliers $N_{ij}^{(k)}$ and $z_i^{(k)}$. All these quantities are then used to determine an approximate gradient $G_{ij}^{(k),m}$ from Eq. (17), which is used to set up a conjugate direction search procedure [11]. The reason for choosing the implicit CN procedure to solve Eq. (18) and the ADI scheme to solve Eq. (19) in preference to an explicit scheme, such as Euler's method, is that for explicit schemes the bound on the time step Δt to ensure stability depends on the unknown control K. For such explicit schemes it would be impossible to determine the magnitude of the time steps Δt without knowledge of K. Both the CN scheme (20) and the ADI scheme (21) are unconditionally stable and therefore do not require restrictions on the size of time step to ensure stability.

1. Numerical Results

We consider the design of an IC insensitive controller for the one-dimensional heat operator defined in Eq. (11) over a spatial domain of length $L = 1.0$ and a time interval of length $T = 0.02$. We assume: that the diffusion coefficient $D = 1.0$; that the reference function $r(\xi)$ to be tracked has the parabolic form $r(\xi) = \xi(1 - \xi)$; that the mean initial condition has the form

$$\bar{x}_0 = \begin{cases} 2\xi & \text{if } 0 < \xi < \frac{1}{2}, \\ 2(1-\xi), & \text{if } \frac{1}{2} < \xi < 1; \end{cases}$$

and that $Q_{ij} = \delta_{ij}$, the covariance matrix $\Lambda_{ij} = \delta_{ij}$, and $\gamma = 10^{-6}$. For the spatial mesh we assume that $N = 50$ while the time interval $[0,T]$ was divided into $M = 20$ time steps.

In Fig. 1 we plot the controller K designed to be insensitive to initial conditions with a Gaussian distribution around the mean $\bar{x}_0$. In Fig. 2 we plot the controller K that is tailored to the mean initial condition $\bar{x}_0$, i.e., designed to minimize the same performance index as (10) but without any averaging over the initial conditions x_0. Comparing Figs. 1 and 2 we see that both controllers resemble the mean initial condition $\bar{x}_0$, however the shape of the sensitive controller is much closer to the mean initial condition $\bar{x}_0$ than that of the insensitive controller. In Figs. 3 and 4 we plot the solutions $x(t, \xi)$ corresponding to the insensitive and sensitive controllers, respectively. For both these solutions the mean initial condition $x(0,\xi) = \bar{x}_0(\xi)$ was used. Both solutions tend to the specified function $r(\xi)$ as t increases. As a measure of the success of each of the controllers in forcing the

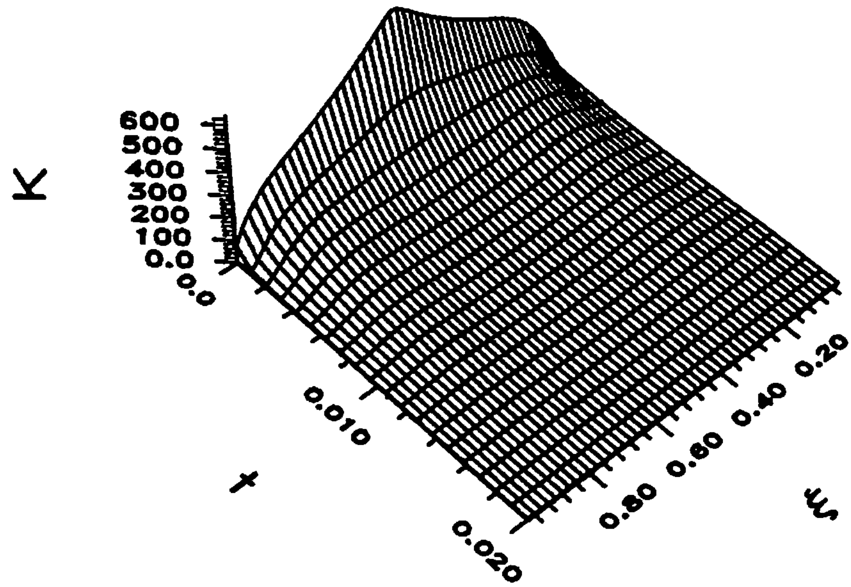

FIG. 1. IC insensitive control.

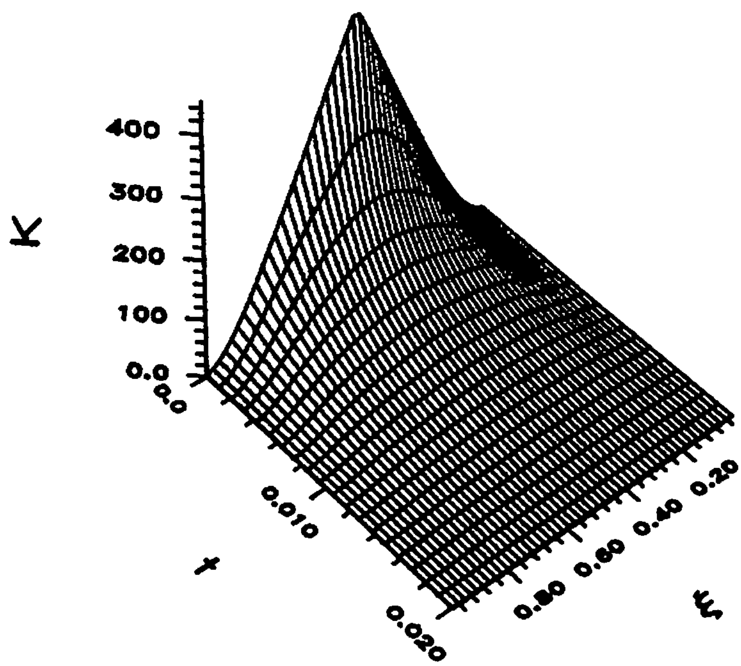

FIG. 2. IC sensitive control.

solution $x(\xi, t)$ to track the function $r(\xi)$ we consider the error functional

$$E(x) = \int_0^T \langle x(s) - r, Q(x(s) - r)\rangle_X \, ds \tag{22}$$

for the two solutions. In the case of the insensitive control $E_{\text{insens}} = 0.00791$, whereas in the case of the sensitive control $E_{\text{sens}} = 0.01110$.

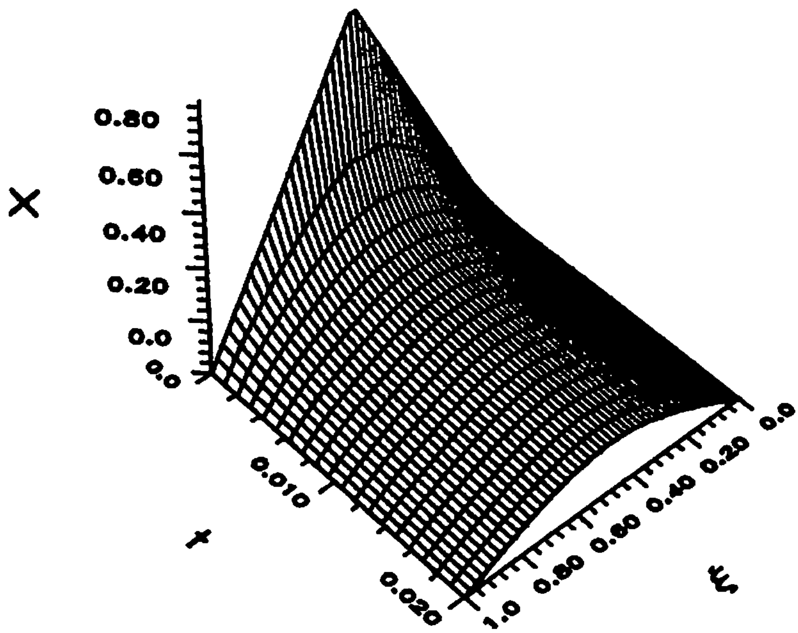

FIG. 3. State with insensitive control and mean IC.

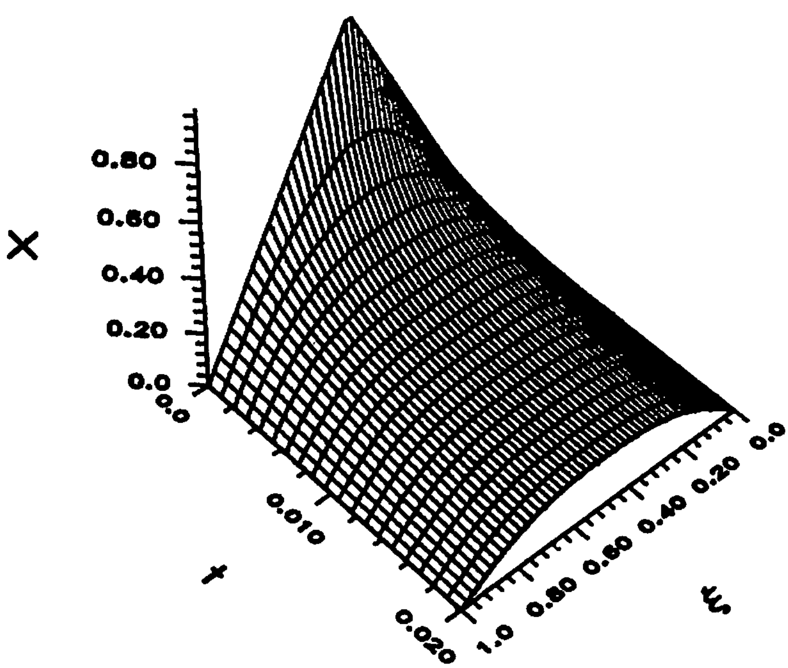

FIG. 4. State with sensitive control and mean IC.

To compare the performance of the two controllers when the initial condition is uncertain, we evaluated the error functional (22) for 5000 different initial conditions that are random normal perturbations of the mean initial condition $\bar{x}_0$. In particular $\bar{x}_0$ is discretized on the spatial mesh with N points to yield $\bar{x}_{0,i}$. The discrete perturbed initial conditions $x_{0,i}$ are defined to be

$$x_{0,i}(p) = x_{0,i} + R_i(p); \qquad i = 0,\ldots,N; \qquad p = 1,\ldots,5000, \tag{23}$$

where $R_i(p)$ is a family of normally distributed random N-vectors with covariance matrix δ_{ij}. In Fig. 5 the frequency distribution of the costs for the randomly perturbed initial conditions defined in Eq. (23) are given for the insensitive and sensitive controllers. It can be seen that the frequency distribution of the insensitive controller is narrower than that of the sensitive controller—the standard deviations are $\sigma_{\text{insens}} = 0.012\,307$, $\sigma_{\text{sens}} = 0.016\,349$. In addition the mean cost for the insensitive controller $\bar{E}_{\text{insens}} = 0.025\,636$ is lower than that for the sensitive controller $\bar{E}_{\text{sens}} = 0.032\,958$. Both the narrowing of the frequency distribution and the lowering of the mean cost by the insensitive controller clearly demonstrate that the insensitive controller is less sensitive to random perturbations in initial conditions than the sensitive controller.

B. Numerical Optimal Control of the Heat Equation with an Uncertain Parameter

In this problem we consider the tracking optimization problem:

$$\begin{aligned} \text{minimize } J(K) &= E_\alpha\left(\int_0^T \langle x(s) - r, Q(x(s) - r)\rangle_X\, ds\right) \\ &\quad + \gamma\int_0^T \langle K(s), K(s)\rangle_{\text{H.S.}}\, ds \\ &= \int_W \left(\int_0^T \langle x(s,\alpha) - r, Q(x(s,\alpha) - r)\rangle_X\, ds\right) d\mu(\alpha) \\ &\quad + \gamma\int_0^T \langle K(s), K(s)\rangle_{\text{H.S.}}\, ds \end{aligned} \tag{24}$$

subject to:

$$\frac{dx(t,\xi,\alpha)}{dt} = A(\alpha)x(t,\xi,\alpha) + Bu(t,\xi,\alpha), \qquad u = -Kx, \qquad x(0,\xi) = x_0(\xi). \tag{25}$$

For the purposes of demonstrating optimal control with parameter uncertainty we assume that $A(\alpha) = \alpha(d^2/d\xi^2)$, i.e., the diffusion coefficient is the uncertain parameter. We are not able in this case to exploit an operator approach to remove the dependence on α as we were in the case of uncertain initial conditions. Therefore our approach is to use numerical integration to evaluate the cost functional and gradients.

We define a Lagrange multiplier function $z(t) \in L^2([0,T],X)^*$ and the corresponding Lagrangian L as was done in Sec. III. Taking Frechet derivatives of L with respect to v, x, and K we obtain Eq. (25) as a necessary condition, in addition to

$$\frac{dz(t,\xi,\alpha)}{dt} = -(A(\alpha) - BK)z(t,\xi,\alpha) + Q(x(t,\xi,\alpha) - r(\xi)), \qquad z(T,\xi,\alpha) = 0, \tag{26}$$

$$\int_0^T \langle 2\gamma K(s) + \int_W (v^*(s,\alpha)B)x^*(s,\alpha)\, d\mu(\alpha), \delta K(s)\rangle_{\text{H.S.}}\, ds = 0. \tag{27}$$

Given a trial controller $K^{(k)}$ we determine an approximate solution to Eqs. (25) and (26) using finite differences. We use the same discretization procedure as that used above in the case of the one-dimensional equations (14) and (16). We solve the resulting systems of ODEs using the CN method described above. The numerical solutions $x_m^{(k)}$ and $v_m^{(k)}$ are

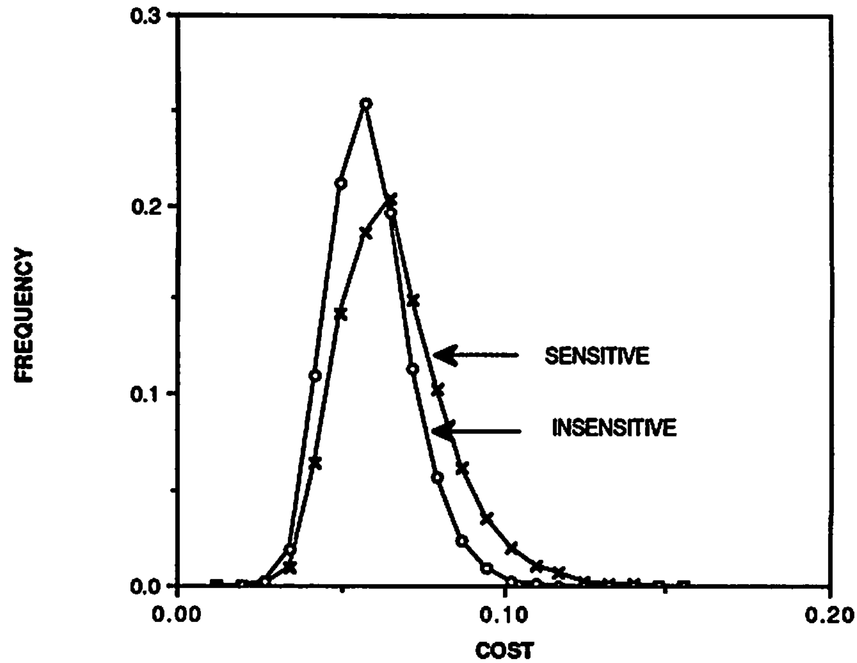

FIG. 5. Frequency distributions of sensitive and insensitive controllers.

used to determine an approximate gradient $G_{ij}^{(k),m}$ from Eq. (27), which is used to set up a conjugate direction search procedure.

1. Numerical Results

As above, we assume a spatial domain of length $L = 1.0$ and a time interval of length $T = 0.02$. We assume that $\gamma = 10^{-6}$ and that

$$d\mu(\alpha) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{(\alpha - \bar{\alpha})^2}{2\sigma^2}\right) d\alpha$$

where $\bar{\alpha} = 20$, $\sigma = 6$, and the width of the parametric domain W for the purposes of numerical integration was 3σ. As with the previous numerical example we assume that $B = \delta_{ij}$, that K is diagonal and that $r(\xi) = \xi(1 - \xi)$. The initial condition in this case was taken to be $x_0(\xi) = \sin^2(2\pi\xi)$.

In Fig. 6 we plot the controller K designed to be insensitive to the parameter α. In Fig. 7 we plot the controller that is tailored to the mean parameter value $\alpha = 20.0$, i.e., designed to minimize the same performance index as Eq. (24) but without any averaging over the parameter α. The two controllers are similar in shape; however the amplitude of the sensitive controller is larger than that of the insensitive controller. In Figs. 8 and 9 we plot the solutions $x(t, \xi)$ corresponding to the insensitive and sensitive controllers respectively. For both these solutions the mean parameter $\alpha = 20.0$ was used. The solution corresponding to the sensitive control can be seen to track $r(\xi)$ more closely than the insensitive controller. As a measure of the success of each of the controllers in forcing the solution $x(\xi, t)$ to track the function $r(\xi)$ we consider the error functional

$$E(\alpha) = \int_0^T \langle x(s, \alpha) - r, Q(x(s, \alpha) - r)\rangle_X \, ds \tag{28}$$

for the two solutions. In the case of the insensitive control $E_{\text{insens}}(20) = 0.020\,616$, whereas in the case of the sensitive control $E_{\text{sens}}(20) = 0.019\,367$. This measure demonstrates that the sensitive control is seen to perform better than the insensitive control for the mean parameter value—this is to be expected since the sensitive control is tailored to this parameter value. In Figs. 10 and 11 we plot the solutions $x(t, \xi)$ corresponding to the insensitive and sensitive controllers assuming a value of $\alpha = 2.0$, which is 3σ away from the mean. In this case the solution corresponding to the insensitive control can be seen to track $r(\xi)$ noticeably better than the sensitive controller. In the case of the insensitive control $E_{\text{insens}}(2) = 0.105\,071$, whereas in the case of the sensitive control $E_{\text{sens}}(2) = 0.522\,271$. Therefore for a large perturbation from the mean parameter value the insensitive control is seen to perform much better than the sensitive control. In Fig. 12 the cost function $E(\alpha)$ defined in Eq. (28) is plotted for a large range of values of α. Three regions of performance can be identified: Firstly in the immediate neighborhood [16, 25] of the mean diffusion coefficient $\bar{\alpha} = 20$ the sensitive controller outperforms the insensitive controller. The extent of this interval will depend upon the magnitude of the sensitivity gradient $dE/d\alpha$ at the nominal value $\bar{\alpha}$. Secondly, in the interval [2, 15] the insensitive controller performs progressively better than the sensitive controller as the value of α is decreased toward a region of higher sensitivity gradient. Thirdly, in the interval [25, 38] the insensitive controller marginally outperforms the sensitive controller due to a low sensitivity gradient in this region.

The above example clearly demonstrates the effectiveness of the parameter insensitive design procedure, particularly in regions where the sensitivity gradient of the cost is large.

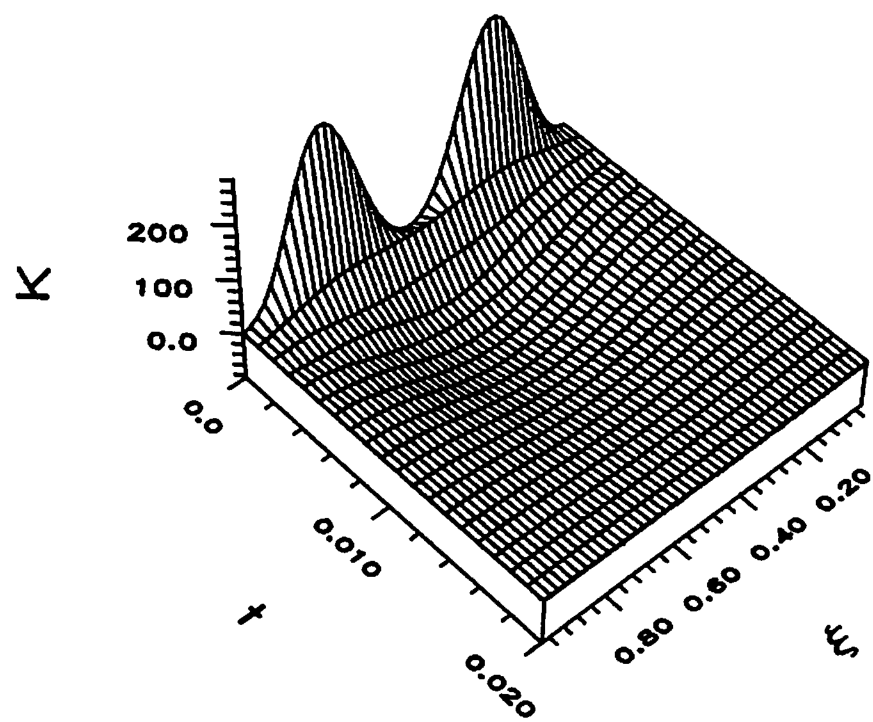

FIG. 6. Parameter insensitive control.

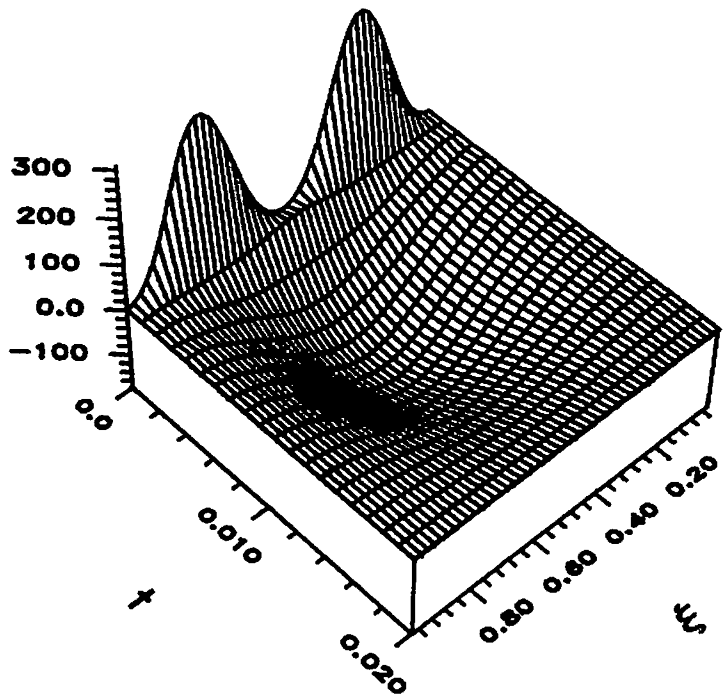

FIG. 7. Parameter sensitive control.

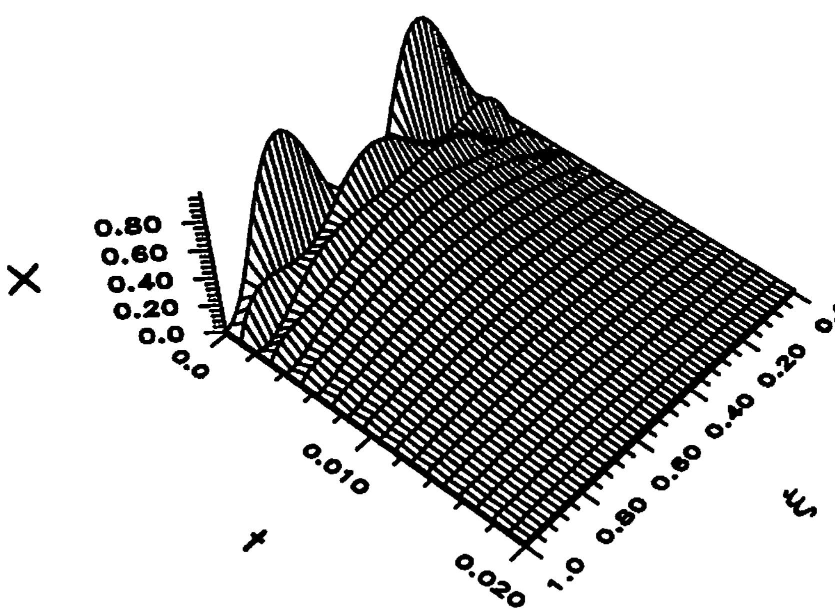

FIG. 8. State with insensitive control and mean parameter.

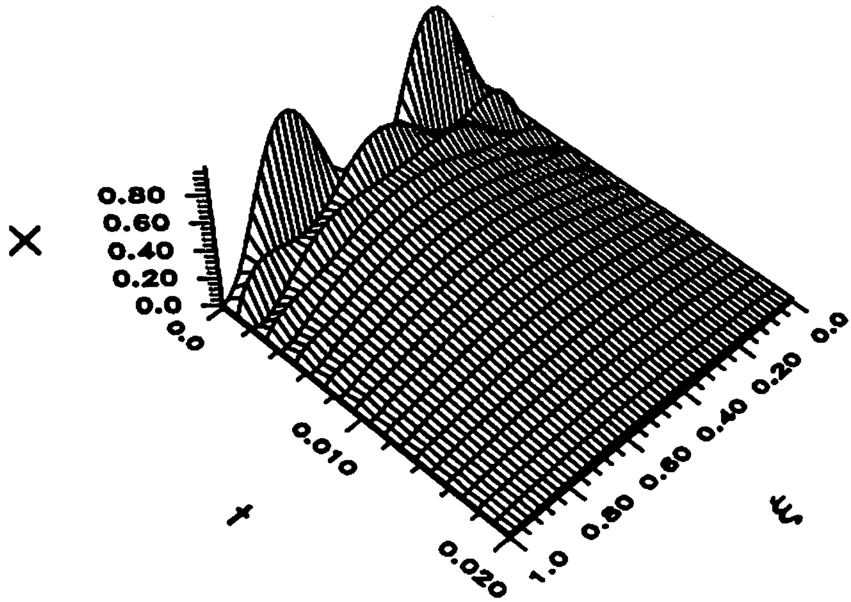

FIG. 9. State with sensitive control and mean parameter.

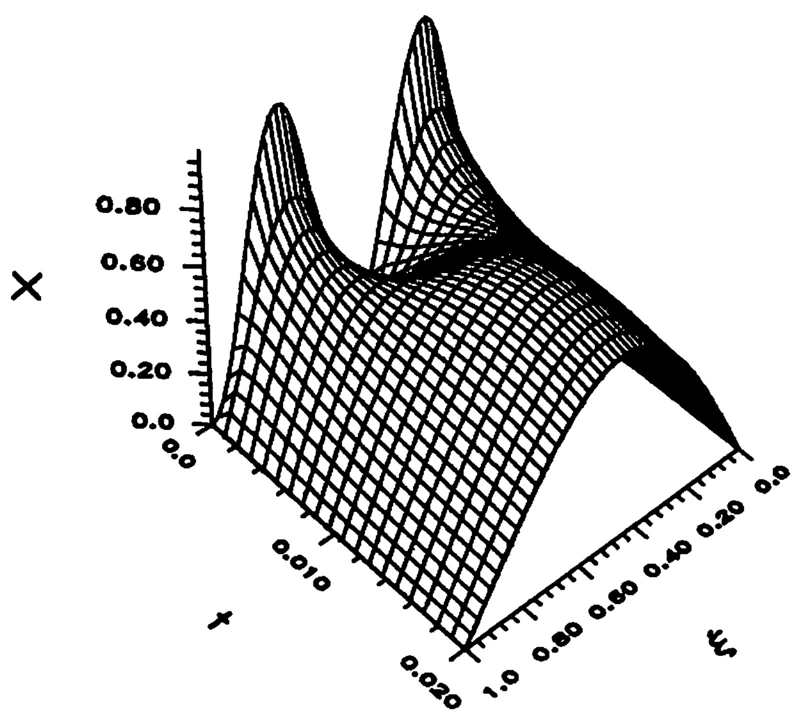

FIG. 10. State with insensitive control and perturbed parameter.

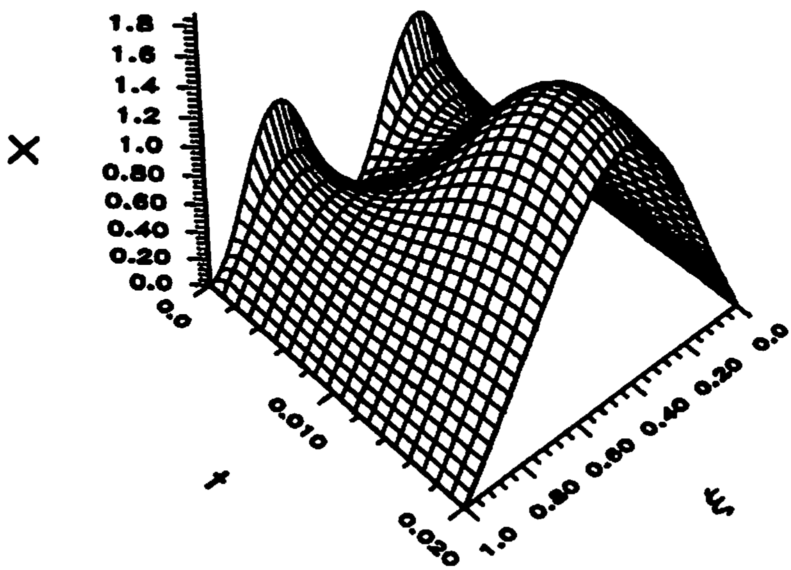

FIG. 11. State with sensitive control and perturbed parameter.

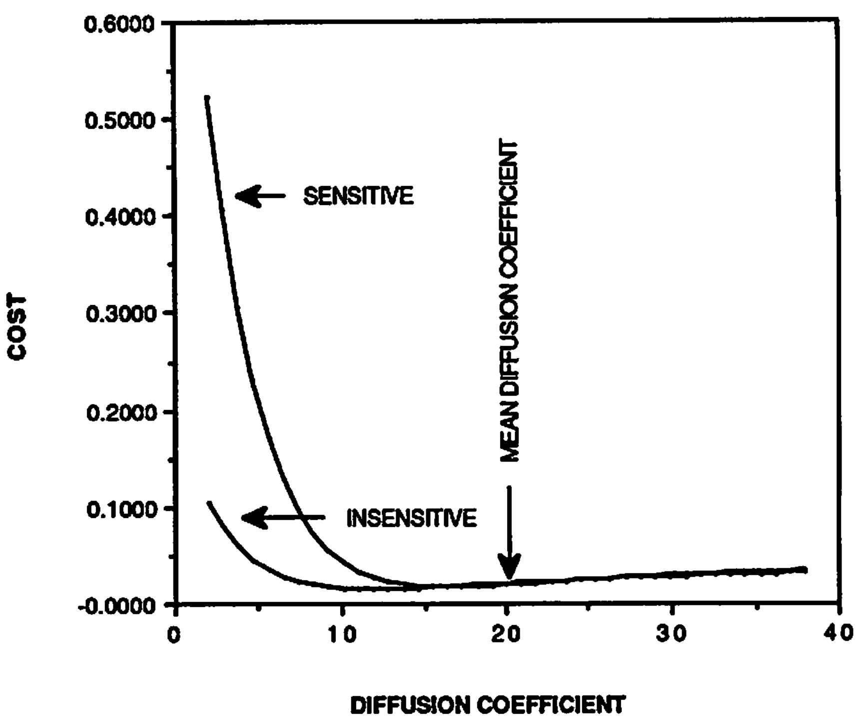

FIG. 12. Cost as a function of diffusion coefficient.

VI. CONCLUSION

In this article we showed how the problem of optimal control of uncertain parabolic systems leads to an iterative optimization procedure, involving the solution of a set of parabolic partial differential equations. The existence, uniqueness, and numerical approximation of the resulting partial differential equations were given. The method was shown to yield an effective way to compute optimal control laws for systems governed by uncertain parabolic systems. The positive conclusions of this article suggest that this method can be applied to many different control problems, arising in other engineering applications.

A.P.P. was partially supported by the National Science and Engineering Research Council of Canada. M. D. was supported by the National Science Foundation under Grant No. ECS-8914357.

APPENDIX

Proof of Lemma 1

$$L(G) = \int_0^T \int_\phi G(t,x,x)\,dt\,dx\,.$$

From a Sobolev imbedding theorem [12] we know that $H^m(\Omega^{n'})$ imbeds continuously in $C_0(\Omega^{n'})$ where $\Omega^{n'} \subset R^{n'}$ if $mp > n'$. Specializing to our case we have $m = 2, p = 2$, and $n' = 2n$. Therefore if $n = 1$ we have

$$|G(t,x,x)| \leq \|G\|_{C_0} \leq K\|G\|_{\tilde{H}_0^1 \cap H^2}$$

from which it follows that

$$|L(G)| \leq K\int_0^T \|G\|_{\tilde{H}_0^1 \cap H^2} \leq C\|G\|_{\tilde{H}_0^{1,1} \cap H^{2,1}}\,.$$

On the other hand, by another imbedding theorem [12] we know that $H^m(\Omega^{n'})$ imbeds continuously in $L^2(\Omega^{(n'/2)})$ if $mp \leq n' < 2mp$. Therefore if $n = 2$ or $n = 3$ the condition is satisfied and we have

$$\|G\|_{L^2(\Omega^{(n'/2)})} \leq \|G\|_{\tilde{H}_0^1 \cap H^2}.$$

The above inequality implies

$$|L(G)| \leq \int_0^T \|G\|_{\tilde{H}_0^1 \cap H^2}\,dt \leq K\|G\|_{\tilde{H}_0^{1,1} \cap H^{2,1}} \qquad \blacksquare$$

Proof of Lemma 2

We will use the fact that [13]

$$\left(-\frac{\partial}{\partial t} - (A)_x - (A)_y : \tilde{H}_0^{1,1} \cap H^{2,1} \longrightarrow L^2\right)$$

is an isomorphism and use a perturbation argument. Given any $\phi \in L^2$ there exists $v \in \tilde{H}_0^{1,1} \cap H^{2,1}$, which is a solution to the unperturbed problem. Using semigroups we can write

$$v = \int_0^t Z(t - s)\phi(s)\,ds.$$

From Ref. 13 we have

$$\|v\|_{\tilde{H}_0^1 \cap H^2} \leq K\int_0^t \|\phi\|_{L^2(\phi\times\phi)}\,ds \qquad \forall t > 0,$$

and K is independent of t. Now consider successive approximations

$$v^n(t) = \int_0^t Z(t-s)(\phi(s) + \tilde{F}v^{n-1}(s))\,ds \tag{A1}$$

where $\tilde{F} = (BK)_x + (BK)_y$ is a bounded operator on L^2 with $\|\tilde{F}\| = C$. Let $r = \int_0^T \|\phi(s)\|_{L^2} ds$, from Eq. (A1) we get

$$\|v^n\|_{\tilde{H}_0^1 \cap H^2} \le rK \sum_{i=0}^{n} \frac{(KCt)^i}{i!},$$

$$\|v^n\|_{\tilde{H}_0^1 \cap H^2} \le rK \exp(KCt).$$

This formula leads to

$$\|v^n\|_{\tilde{H}_0^{1,1} \cap H^{2,1}} \le C\|\phi\|_{L^2(\Sigma)}$$

which means that v^n is in fact a sequence in $\tilde{H}_0^{1,1} \cap H^{2,1}$. The next step is to observe that v^n satisfies

$$\|v^n - v^{n-1}\|_{\tilde{H}_0^1 \cap H^2} \le Kr\frac{(KCt)^n}{n!}.$$

Thus for $m \ge n + 1$ we have

$$\|v^m - v^n\|_{\tilde{H}_0^1 \cap H^2} \le \sum_{i=n+1}^{m} \frac{(KCt)^i}{i!}.$$

This in turn implies that for any $\epsilon > 0$ there exists $n > 0$ such that

$$\|v^m - v^n\|_{\tilde{H}_0^{1,1} \cap H^{2,1}} \le \epsilon \qquad \text{for } m, n > N.$$

v^n is a Cauchy sequence in the Hilbert space $\tilde{H}_0^{1,1} \cap H^{2,1}$, so it converges to an element v, which is a solution to the perturbed equation. ■

References

1. M. Balas, "Toward a More Practical Control Theory for Distributed Parameter Systems," in *Advances in Dynamics and Control: Theory and Applications,* Vol. 18, C.T. Leondes, Ed., Academic, New York, 1982.
2. H.T. Banks and K. Kunisch, "The linear regulator problem for parabolic systems," *SIAM J. Control Opt.* **22**, 684–698 (1984).
3. F.M. Callier and C.A. Desoer, "An algebra of transfer functions for distributed linear time-invariant systems," *IEEE Trans Circuits Syst.* **25**, 651–662 (1978).
4. R.F. Curtain and A.J. Pritchard, *Infinite-Dimensional Linear Systems Theory,* Springer-Verlag, New York, 1978.
5. W.E. Hopkins, Jr., "Optimal control of linear systems with parameter uncertainty," *IEEE Trans. Autom. Control* **31**, 72–74 (1986).
6. W.E. Hopkins, Jr., "Optimal stabilization of families of linear stochastic differential equations with jump coefficients and multiplicative noise," *SIAM J. Control Opt.* **25**, 1587–1600 (1987).
7. A. Bensoussan, *Filtrage Optimal des Systemes Lineaires,* Dunod, Paris, 1971.
8. A.V. Balakrishnan, *Applied Functional Analysis,* Springer-Verlag, New York, 1981.
9. Lars Hörmander, *The Analysis of Linear Partial Differential Operators I,* Springer-Verlag, New York, 1983, p. 128.
10. G.D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods,* 2nd ed., Oxford University Press, New York, 1978.
11. D.G. Luenberger, *Linear and Nonlinear Programming,* 2nd ed., Addison-Wesley, Reading, 1984.
12. R.A. Adams, *Sobolev Spaces,* Academic, New York, 1975, p. 97.
13. J.L. Lions and E. Magenes, *Nonhomogeneous Boundary Value Problems and Applications, 2,* Springer-Verlag, New York, 1972.