Optimal control of uncertain quantum systems

M. Dahleh
Texas A&M University, College Station, Texas 77843-3128

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

H. Rabitz
Princeton University, Princeton, New Jersey 08544

(Received 25 September 1989)

The design of optimal final-state controllers of quantum-mechanical systems that are insensitive to errors in the molecular Hamiltonian or to errors in the initial state of the system is considered. Control arises through the interaction of the system with an external field; the goal is optimal design of these latter fields for various physical objectives in the presence of system uncertainty. Sensitivity to modeling errors and other uncertainties in the molecular Hamiltonian is minimized by considering averaged costs for a family of Hamiltonian functions $H_0(\alpha)$ indexed by the random variable α taking values on a compact set in Euclidean space. Similarly, sensitivity of the optimal control to the initial state is minimized by viewing the initial condition as a Hilbert-space-valued random variable and considering an optimization problem with a cost functional that is averaged over the class of initial conditions. A precise formulation of the control problem is given, and its well-posedness is established. Cost propagators are defined to display the dependence of the performance index on the initial conditions explicitly, which allows analytic averaging of initial conditions. The constrained optimization problem is reduced to an unconstrained optimization problem by the introduction of Lagrange-multiplier operators. Necessary conditions for the unconstrained problem provide the basis for a gradient search for an optimal solution. Finite-difference schemes are utilized to provide a numerical approximation of the optimal control problem. Numerical examples are given for final-state control of a diatomic molecule represented by a Morse potential illustrating design for systems with initial-phase uncertainty and parametric uncertainty. The resultant insensitive controllers execute different strategies depending on the design requirements. The controller designed to be insensitive to errors in the initial phases adopts a strategy of phase imprinting during the initial stages of the control interval to compensate for a lack of knowledge of the initial phases. It is also shown that it is not possible to coerce a system from a state with completely random initial phases to a correlated state using the class of averaged controllers considered here. The controller designed to be insensitive to parametric Hamiltonian errors adopts a strategy of amplitude restraint to prevent the wave packet from taking significant excursions into the regions where the potential is uncertain. The interesting structure exhibited by the controllers in response to the different design requirements, and the superior performance of the insensitive controllers when compared with controllers designed at nominal phases and parameters, illustrate the usefulness of the cost-averaging technique for design in the presence of uncertainties.

I. INTRODUCTION

The search for the appropriate excitation of molecular-scale physical systems to achieve prescribed molecular motion has been an elusive objective over the past 20 years. Recently[1-3] the formulation of the problem in terms of optimal control theory has led to a hitherto unexplored theoretical framework to design optical-pumping strategies. Numerical field designs based on this framework have been achieved for linear classical,[2] semiclassical[3] and quantum-mechanical molecular[1] models. One drawback of the optimal design approach used in the work described above is that these controllers are likely to be sensitive to uncertainties in the molecular Hamiltonian and to uncertainties in the initial state of the system. This issue of sensitivity or robustness will have to be addressed if field designs are to be implemented in the laboratory.

In this paper we consider the design of optimal controllers for quantum systems that are minimally sensitive to errors in the molecular Hamiltonian or to the initial state of the system. To allow for modeling errors and other uncertainties in the molecular Hamiltonian we propose a family of Hamiltonian functions $H_0(\alpha)$ indexed by the random variable α taking values on a compact set in Euclidean space. Minimization with respect to the Hamiltonian uncertainty is achieved by considering a cost functional that is averaged over the random variable with respect to an appropriate measure. The dependence of the optimal control on the initial state is minimized by viewing the initial condition as a Hilbert-space-valued

random variable, and considering an optimization problem with a cost functional which is averaged over the class of initial conditions. We give a precise formulation of the above control problem and prove that it is well posed. We consider the numerical approximation of the uncertain optimal control problem and provide some one-dimensional examples. The performance of the resultant insensitive controllers is contrasted with that of sensitive controllers that are designed, at a nominal parameter or initial condition, assuming there is no uncertainty present.

In Sec. II we formulate the uncertain optimal control problem and prove that it is well posed. In Sec. III we reformulate the problem by converting it to a final-value optimization problem by introducing a system of operator equations that evaluate the performance index. This procedure is analogous to adopting the Heisenberg picture of the quantum system. Its importance in this context is that it allows the dependence of the performance index on the initial condition to be displayed explicitly and therefore removed by averaging. We introduce Lagrange-multiplier operators which are used to establish necessary conditions for an optimal solution. In Sec. IV we discuss the numerical approximation of the uncertain optimal control problem by means of the method of finite differences. In Sec. V we illustrate the proposed theory by providing the numerical solutions for the control of the one-dimensional Schrödinger equation in which we use the Morse potential to represent a diatomic molecule. The examples chosen here are quite simple to illustrate the methodology at the most basic level; the methodology directly encompasses more general quantum-mechanical systems but with a concomitant increase in computational cost. Both types of uncertainty catered for by the theory are considered separately in the illustrative examples to facilitate the interpretation of the results. The first example considers optimal control when there is uncertainty in the phases of the initial conditions. The second example illustrates optimal control when there is parametric uncertainty in the molecular Hamiltonian. The performance of the insensitive controllers is contrasted with those of controllers designed assuming no uncertainty. In Sec. VI we summarize the results and suggest future directions of research. In Appendix A we prove the well posedness of the operator equations whose solution evaluates the performance index. In Appendix B we demonstrate that the operator equations can be approximated.

II. UNCERTAIN OPTIMAL CONTROL PROBLEM

In this section we formulate the optimal control problem for an uncertain quantum system and prove that it is well posed. We remark that the results of Secs. II and III are general in that they apply to an arbitrary multiparticle quantum system.

Let $\Omega \subset R^n$ be the spatial domain under consideration and $[0,T]$ the finite time interval over which the problem is posed. We define the following Hilbert Spaces: $X \equiv L_2(\Omega)$, and $X_{\mathrm{HS}} \equiv$ the Hilbert space of Hilbert-Schmidt operators. Viewing the initial condition as a Hilbert-space-valued random variable enables us to minimize the dependence of the optimal control on the initial conditions by formulating the following parameter optimization problem:

$$\min J(K) = E_{\alpha,\psi_0}(\langle \psi(T)-r, Q(\psi(T)-r)\rangle X) + \beta \int_0^T \langle K,K\rangle_{\mathrm{HS}} ds \ ,$$

subject to

$$\frac{d\psi(t)}{dt} = \frac{-i}{\hbar} H(\alpha)\psi(t) = A(\alpha)\psi(t) \tag{1}$$

and

$$\psi(t=0) = \psi_0$$

over all $K \in X_{\mathrm{HS}}$. Here the initial condition $\psi_0 \in X$, and the function $r \in X$ is a specified reference state to which we desire to push the final state $\psi(T)$ of the system under the constraint of finite radiative fluence. By radiative fluence we refer to the term $\int_0^T \langle K,K\rangle_{\mathrm{HS}} ds$. The operators in (1) are defined as follows: $A(\alpha) = (-i/\hbar)H(\alpha)$, $H(\alpha) \equiv H_0(\alpha) - B(\alpha)K$, and $H_0(\alpha) = \tilde{H}_0 + V_0(\alpha)$. Here $\tilde{H}_0$ is the portion of the Hamiltonian that is independent of the uncertain parameters (e.g., the kinetic-energy operator), $B(\alpha)$ is the radiative coupling operator (e.g., the dipole moment), and K is the external field. We use the notation $E_{\psi_0,\alpha}$ to represent the expectation with respect to ψ_0, α, which will be defined more precisely below. The inner product between any two elements of X_{HS}, say K and M, is given by[4]

$$\langle K,M\rangle_{\mathrm{HS}} = \mathrm{tr}(K^* M) \ ,$$

where * denotes Hermitian conjugation. Finally, Q is a positive, self-adjoint, Hilbert-Schmidt operator.

The form chosen for J in one respect represents an extreme limiting case; in particular, the cost functional attempts to steer $\psi(x,T)$ to be as close as possible to a target state r—including its overall phase. This objective is a strong requirement and was chosen with the limiting characteristic in mind. Less demanding and also physically acceptable cost functionals could be chosen (e.g., requiring that the square of the projection of the final state onto the target be a maximum) and the illustrative results presented later in the paper would be expected to be even better. This latter comment applies equally well to quantum control problems without stochastic elements.[3]

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 $\mu(d\alpha)$. For each α we have the following assumptions on $H_0(\alpha)$ and $B(\alpha)$:

(1) $H_0(\alpha)$ is the infinitesimal generator of a C_0-semigroup $S_\alpha(t)$, for each α.

(2) The domain $\cap_\alpha D(H_0(\alpha))$ is dense in X.

(3) $S_\alpha(t)$ is a continuous function of α for each t, and uniformly continuous on bounded intervals of t.

(4) The input coupling term $B(\alpha)$ 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

$$\psi_0(\phi^*)=\langle \psi_0,\phi^* \rangle_X .$$

The expectation of ψ_0 is defined by

$$E(\psi_0(\phi^*))=\int_\Omega \psi_0(\phi^*)(\omega)dP(\omega) ,$$

where $(\Omega,\mathbf{B},P)$ is an underlying probability space. Thus $E(\psi_0(\phi^*))$ is a bounded linear operator on the dual space X^*, and the Reisz representation theorem can be used to write

$$E(\psi_0(\phi^*))=\langle m,\phi^* \rangle_X$$

for some $m \in X$.

The Hilbert-space-valued random variable ψ_0 induces a cylindrical measure on X given by the following finite-dimensional densities:

$$d\mu_{\phi_1^*\ldots\phi_n^*}=\frac{1}{(2\pi)^{n/2}}\exp[-\tfrac{1}{2}(\psi-m_n)^T\Lambda_n^{-1}(\psi-m_n)] \times d\psi_1\ldots d\psi_n ,$$

where

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

$(\Lambda_n)_{ij}=\langle \Lambda\phi_i^*,\phi_j^* \rangle$, 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.[5] On the basis of this it is reasonable to assume that the initial condition has a nuclear operator for a covariance.

The assumptions 1–3 on the generator $H_0(\alpha)$ are sufficient to insure a well-posed uncontrolled evolution equation. As for the controlled system, we have, for each α, a perturbation term $B(\alpha)K$. With this bounded perturbation it can be shown, in a standard way,[6] that the controlled system remains well posed and its semigroup satisfies $S_{H_0+BK}(t)\le Me^{(w+\|BK\|)t}$ for some M,w. The following theorem guarantees the existence of an optimal solution to the parameter optimization problem (1) which lies in the interior of X_{HS}.

Theorem 1. Under the assumptions 1–3, there exists a $K^* \in X_{\mathrm{HS}}$ that solves (1).

Proof. Since $J(K)$ is bounded from below by zero, we can take a sequence $K_n \in X_{\mathrm{HS}}$ such that $\lim_n J(K_n)=\inf_K J(K)$. The performance index satisfies $J(K)\ge\langle K,K \rangle_{\mathrm{HS}}$, which implies that $J(K_n)$ diverges to infinity if the norm $\langle K_n,K_n \rangle_{\mathrm{HS}}$ diverges to infinity. Therefore we conclude that there exists a constant $M>0$ such that $\|K_n\|_{\mathrm{HS}}\le M$ for all n. X_{HS} is a Hilbert space, so using the fact that the closed unit ball in X_{HS} is weakly compact, we can extract a subsequence K_n' of K_n that converges weakly to an element $K^* \in X_{\mathrm{HS}}$. The remainder of the proof is devoted to showing that K^* is in fact a minimizer. Let $\psi_n^\alpha(t)$ denote the solution of the evolution equation associated with the feedback K_n. From a previous argument we know that $\|\psi_n^\alpha(t)\|_X \le C(\alpha)\exp\{[w(\alpha)+\|B(\alpha)K_n\|]t\}$. Since $H_0(\alpha),B(\alpha)$ depend on α in a continuous manner (assumptions 3 and 4), it follows that $\|\psi_n^\alpha(t)\|_X \le C$ where C is a positive constant independent of n and α. This implies the existence of a positive constant R independent of α and n such that

$$E_{\alpha,\psi_0}\int_0^T\|\psi_n^\alpha(s)\|_X^2 ds \le R .$$

We denote by X' the space $L^2([0,T]\times\Omega\times W;X)$. This shows that $\|\psi_n^\alpha\|_{X'}\le R$, so we can extract a subsequence K_n^0 so that ψ_n^α converges weakly in X'. Finally, using the weak lower semicontinuity of the norm, we have

$$J(K^*)\le \liminf_n J(K_n^0)=\inf_K J(K) .$$

This implies that $J(K^*)=\inf_K J(K)$.

III. COST PROPAGATORS AND NECESSARY CONDITIONS FOR OPTIMALITY

We will reformulate the optimization problem by converting it to a final value optimization problem involving a system of operator equations that evaluate the cost $J_{\alpha,\psi_0}(K)$. The advantage of this approach is that the dependence of $J_{\alpha,\psi_0}(K)$ on the initial condition is displayed explicitly, and can thus be removed by averaging.

The quadratic term in the cost functional defined in (1) can be expressed as follows:

$$J_{2,\alpha,t}(K)=\langle \psi(T),Q\psi(T) \rangle_X \tag{2a}$$

$$=\langle \psi(t),P(t,\alpha)\psi(t) \rangle_X , \tag{2b}$$

where $P(t,\alpha)=S_{A(\alpha)}^*(T-t)QS_{A(\alpha)}(T-t)$ and $S_{A(\alpha)}$ is the C_0 semigroup generated by $A(\alpha)$. From (2a) it follows that $J_{2,\alpha,t}(K)$ is independent of t. Thus differentiating the term in (2b) with respect to t and equating the result to zero we obtain the following equation governing the evolution of $P(t,\alpha)$:

$$\frac{dP(t,\alpha)}{dt}+A(\alpha)^*P(t,\alpha)+P(t,\alpha)A(\alpha)=0 , \quad P(T,\alpha)=Q . \tag{3a}$$

Similarly, the linear term in the cost functional defined in (1) can be expressed as the evolution of a function $v(t,\alpha)=S_{A(\alpha)}^*(T-t)Qr$ governed by

$$\frac{dv(t,\alpha)}{dt}+A(\alpha)^*v(t,\alpha)=0; \quad v(T,\alpha)=Qr . \tag{3b}$$

An alternative to solving (3a) and (3b) directly to determine $P(t,\alpha)$ and $v(t,\alpha)$ would be to determine the more fundamental object $S_{A(\alpha)}$. This semigroup could then be used to determine $P(t,\alpha)$ and $v(t,\alpha)$ from their definitions. In terms of the evolution of the operator $P(t,\alpha)$ and the function $v(t,\alpha)$, the performance index in (1) can be expressed in the form

$$J(K)=E_{\psi_0,\alpha}(\langle \psi_0,P(0,\alpha)\psi_0 \rangle_X-2\mathrm{Re}\langle \psi_0,v(0) \rangle_X) +\langle r,Qr \rangle_X+\beta\int_0^T\langle K,K \rangle_{\mathrm{HS}}ds . \tag{4a}$$

The new final value optimization problem is

$$\min_{K\in X_{\mathrm{HS}}} J(K) , \tag{4b}$$

subject to equations (3a) and (3b).

In Appendix A we show directly without reference to the optimization problem that there exists a unique positive Hilbert-Schmidt operator $P(t,\alpha)$ which satisfies (3a). In the previous section we showed that there exists a minimizer which lies in the interior of X_{HS}. This justifies writing necessary conditions for the optimal solution. We define the residual quantities $M(\alpha,K,P,v,t)$, $w(\alpha,K,P,v,t)$ as follows:

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

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

in terms of which we define the Lagrangian L by

$$L=J(K)+\int_0^T\int_W d\mu(\alpha)\,dt[\langle N(t,\alpha),M(\alpha,K,P,v,t)\rangle_{\mathrm{HS}}+\langle z(t,\alpha),w(\alpha,K,P,v,t)\rangle_X] ,$$

where $N(t,\alpha)$ and $z(t,\alpha)$ are the Lagrange multipliers that belong to the following spaces:

$$N(t,\alpha)\in L^2([0,T]\times W;X_{\mathrm{HS}})^*\simeq L^2([0,T]\times W;X_{\mathrm{HS}}) ,$$

$$z(t,\alpha)\in L^2([0,T]\times W;X)^*\simeq L^2([0,T]\times W;X) .$$

From (3a) it follows that P is Hermitian (i.e., $P=P^*$), thus to ensure that the term $\langle N(t,\alpha),M(\alpha,K,P,v,t)\rangle_{\mathrm{HS}}$ in the Lagrangian L is real it is sufficient to require that the Lagrange multiplier N is also Hermitian.

The first-order necessary conditions can be obtained by taking Frechet derivatives of L with respect to P, N, v, z, and K, which yields (3a) and (3b) as necessary conditions, in addition to

$$\frac{dN(t,\alpha)}{dt}=A(\alpha)N(t,\alpha)+N(t,\alpha)A(\alpha)^*;\quad N(0,\alpha)=E_{\psi_0}(\psi_0\psi_0^*) , \tag{5a}$$

$$\frac{dz(t,\alpha)}{dt}=A(\alpha)z(t,\alpha);\quad z(0,\alpha)=-2E_{\psi_0}(\psi_0) , \tag{5b}$$

$$0=\int_0^T\left\langle 2\beta K-\mathrm{Re}\frac{i}{\hbar}\int_W d\mu(\alpha)[2B(\alpha)^*P(t,\alpha)N(t,\alpha)+B(\alpha)^*v(t,\alpha)z(t,\alpha)^*],\delta K\right\rangle_{\mathrm{HS}} dt . \tag{5c}$$

Equations (3) and (5) constitute a system of equations that characterize an optimal solution for the uncertain optimization problem. We observe from (5b) that N will be Hermitian, which ensures that L is always real.

IV. NUMERICAL APPROXIMATION OF THE UNCERTAIN OPTIMAL CONTROL PROBLEM

In Appendix B the validity of approximating the operator equations (3a) and (5a) is established. In this section we discuss the numerical solution of the uncertain optimal control problem using the method of finite differences. Since the necessary conditions (3) and (5) characterize the solution of the optimal control problem, we discuss the finite difference approximation of these equations. Although the above formulation applies for multiparticle quantum systems, we restrict our discussion in the remainder of the paper to a quantum system governed by the one-dimensional Schrödinger equation. In principle the procedure can be extended directly to higher dimensions, but a substantial increase in computational complexity can be expected.

Given a trial controller $K^{(k)}$ we determine an approximate solution to (3) and (5) numerically using finite differences. In the case of Eqs. (3b) and (5b), which only involve one spatial dimension, we divide the domain $\Omega=[0,L]$ into M equal subintervals of length $\Delta x=L/M$ and denote the mesh points formed by the end points of these subintervals by $x_m=m\Delta x$. The numerical solution at such a mesh point is denoted by $v_m^{(k)}\approx v^{(k)}(t,x_m)$ and similarly for $z_m^{(k)}$.

Using a central difference approximation to A, (3b) reduces to the following system of ordinary differential equations (ODE's):

$$\frac{dv_n}{dt}+\frac{i}{\hbar}H_n(t)v_n=0;\quad v_n(T)=\sum_m Q_{nm}r_m ;\qquad n,m=0,\ldots,M , \tag{6}$$

where $H_n(t)v_n=\theta(v_{n-1}+v_{n+1})+\phi_n(t)v_n$, $\theta=-\hbar^2/2m\Delta x^2$, and $\phi_n(t)=-2\theta+(i/\hbar)[V_{0n}-B_nK_n^{(k)}(t)]$. Here we have assumed that B is a diagonal operator (i.e., $B=\delta_{mn}B_n$), and that the operator K is spatially independent [i.e., $K_{mn}=K(t)\delta_{mn}$]. Since the operator K, for the purpose of the approximation, is finite dimensional and therefore Hilbert-Schmidt, the theory developed in Sec. II applies to this case. Physically, the operator B represents the dipole function, while $K(t)$ represents the externally applied electric field. A central difference approximation to (5b) yields a similar system of ODE's for z_n to those given in (6).

In the case of the operator equations (3a) and (5a) we divide the product domain $[0,L]\times[0,L]$ into the mesh of M^2 square cells that are formed by the Cartesian product of the subintervals used in (6). Using a central difference approximation to A, (3a) reduces to the following system of ODE's:

$$\frac{dP_{mn}}{dt}+\frac{i}{\hbar}[H_{1,m}(t)P_{mn}-H_{2,n}(t)P_{mn}]=0\ ;$$
$$P_{mn}(T)=Q_{mn};\quad m,n=0,\ldots,M\ , \tag{7}$$

where

$$H_{1,m}(t)P_{mn}=\theta(P_{m-1n}+P_{m+1n})+\phi_m(t)P_{mn}\ ,$$

$$H_{2,n}(t)P_{mn}=\theta(P_{mn-1}+P_{mn+1})+\phi_n(t)P_{mn}\ ,$$

and θ and $\phi_n(t)$ are defined in (6). A central difference approximation to the operator equation (5a) yields a similar system of ODE's for N_{mn} to those given in (7).

In order to solve the system of ODE's (6) we use the Crank-Nicholson (CN) procedure.[7] The time interval $[0,T]$ is divided into R subintervals of length $\Delta t=T/R$. The end points of these subintervals are denoted by $t_s=s\Delta t$, and superscripts are used to denote the time step at which a quantity is evaluated, e.g., $v_n^s\approx v(t_s,x_n)$. The CN procedure can be expressed in the form

$$\left[I-\frac{i\Delta t}{2\hbar}H_n^{(k),s-1}\right]v_n^{(k),s-1}=\left[I+\frac{i\Delta t}{2\hbar}H_n^{(k),s}\right]v_n^{(k),s}\ . \tag{8}$$

The numerical solution $\{v_n^{(k),s}\}$ can be found by a marching process using (8), which involves an inversion of the tridiagonal matrix on the left side of (8).

In order to solve the system of ODE's (7), which result from approximating the operator equation (3a), we use an alternating direction implicit (ADI) scheme:[7]

$$\left[I-\frac{i\Delta t}{2\hbar}H_{1,m}^{(k),s-1/2}\right]P_{mn}^{(k),s-(1/2)}=\left[I+\frac{i\Delta t}{2\hbar}H_{2,n}^{(k),s}\right]P_{mn}^{(k),s}\ , \tag{9a}$$

$$\left[I-\frac{i\Delta t}{2\hbar}H_{2,n}^{(k),s-1}\right]P_{mn}^{(k),s-1}=\left[I+\frac{i\Delta t}{2\hbar}H_{1,m}^{(k),s-(1/2)}\right]P_{mn}^{(k),s-(1/2)}\ . \tag{9b}$$

Assuming that $P_{mn}^{(k),s}$ is known, we invert the tridiagonal matrix on the left side of (9a) to obtain $P_{mn}^{(k),s-(1/2)}$. The right side of (9b) is now known and the tridiagonal matrix on the left side can be inverted to yield $P_{mn}^{(k),s-1}$. This backward marching procedure is used to determine the approximate solution to (3a).

We assumed above that $K^{(k)}$ was known, and calculated the corresponding propagators for the performance index $P_{mn}^{(k)}$ and $v_n^{(k)}$ and the Lagrange multipliers $N_{mn}^{(k)}$ and $z_n^{(k)}$. All these quantities are then used to determine an approximate gradient $G_{mn}^{(k),s}$ from (5c), which is used to set up a conjugate direction search procedure.[8]

The reason for choosing the implicit CN procedure to solve (6) and the ADI scheme to solve (7) in preference to an explicit scheme, such as the leapfrog 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 (8) and the ADI scheme (9) are unconditionally stable and therefore do not require restrictions on the size of time step to ensure stability.

V. NUMERICAL RESULTS

In this section we consider two numerical examples involving the optimal control of an uncertain quantum system described by the one-dimensional Schrödinger equation

$$i\hbar\frac{d\psi}{dt}=\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V_0(\alpha)-B(\alpha)K\right]\psi\ , \tag{10}$$
$$\psi(t=0)=\psi_0$$

in which we use the Morse potential

$$V_0(\alpha,x)=d(1-e^{-\gamma(x-x_0)})^2 \tag{11}$$

to represent a diatomic molecule. The issue of control in molecular systems is reduced to its most basic level in this example, but this simple case will serve to illustrate the nature of sensitive and insensitive quantum controllers in the presence of statistical uncertainty. At this time more complex systems (e.g., a triatomic oscillator) can be treated with the same formalism but at considerably increased computational expense. We reduce (10) to a dimensionless form by rescaling the length according to $x=\delta X$, where $\delta=(\hbar/m\omega)^{1/2}$, and $\omega=\gamma(2d/m)^{1/2}$ is the frequency of the harmonic oscillator that approximates (11) when $X\approx X_0$. In this case (10) and (11) reduce to

$$\frac{i}{\omega}\frac{d\psi}{dt}=\left[-\frac{1}{2}\frac{d^2}{dX^2}+\frac{D}{2}(1-e^{-D^{-1/2}(X-X_0)})^2-BK\right]\psi\ , \tag{12}$$

where $D=2d/\hbar\omega$, and $-BK$ is the dimensionless applied potential. As in Sec. IV, we assume that the operator B is diagonal and represents a dipole function that has the following form: $B(X)=(X-X_0)e^{-\epsilon X}$. The value of ϵ is chosen so that dipole function $B(X)$ is approximately linear on the interval $[0,X_0]$ and attenuates in the limit $X\to\infty$. Throughout this section we assume that $\omega=1$ and that control is sought over the interval $[0,T]$ in which $T=8\pi$, which constitutes four periods of oscillation at the fundamental frequency.

In the first example we assume that the phases of the initial conditions of the quantum system are uncertain, while the remaining parameters of the system are assumed to be known precisely. In the second example we assume that there is a parametric uncertainty in the molecular Hamiltonian, while the initial conditions are assumed to be known.

A. Uncertain initial phases

We assume that the initial condition ψ_0 in (10) is a superposition of eigenstates $\{\phi_k\}$ of the Morse potential[9] with random phases:

$$\psi_0(X)=\sum_k a_k e^{i\delta_k}\phi_k(X) , \tag{13}$$

where $\sum_k |a_k|^2=1$, $\langle \phi_k,\phi_m \rangle=\delta_{km}$, and the δ_k in (13) are assumed to be random variables having a given distribution. The purpose of introducing the cost propagators P and v is that it makes it possible to perform explicit averaging over the initial conditions in the cost functional defined in (1) by using the alternative expression of the performance index (4). We consider two types of distributions of the initial phases δ_k, for which it is possible to perform the averaging analytically.

First, we consider the initial phases δ_k to be uniformly distributed and uncorrelated. In this case the expectation in (4) is given by

$$E_{\psi_0}(\langle \psi_0,P(0)\psi_0\rangle_X-2\,\mathrm{Re}\langle \psi_0,v(0)\rangle_X) =\sum_k |a_k|^2\langle \phi_k,P(0)\phi_k\rangle . \tag{14}$$

In order to interpret (14) it is instructive to consider the case in which the weight operator Q in the cost (1) is the identity I. In this case it follows from (2) that $P\equiv I$ and the right-hand side (rhs) of (14) is unity for all possible initial conditions. The performance index (4) is thus reduced to $J(K)=2+\beta\int_0^T\langle K,K\rangle_{\mathrm{HS}}ds$, for which the optimal control is trivially $K\equiv 0$. Thus for uniformly distributed uncorrelated phases it is not possible to achieve any tracking or equivalently any reduction of the performance index. This extreme case is interesting, not only because it establishes a theoretical limit on the control in an environment of uncertain initial phases, but also because it indicates that it is not possible to coerce this quantum system from a completely random initial state to a correlated state as represented by the target function r. However, this strong conclusion is a result of the heavy demand of the cost functional in (1), which requires that the exact target state be achieved. Some element of control will still remain if we design to maximize the square of the projection of ψ onto the target r.

As a second case we consider a Gaussian distribution of initial phases, which would be consistent with some form of laboratory preparation of the initial states. In this case the averaging process involves evaluating integrals of the form

$$I_k=\frac{\int_{\bar\delta_k-\pi}^{\bar\delta_k+\pi}\exp\left[-\frac{(\delta-\bar\delta_k)^2}{2\sigma_k^2}+i\delta\right]d\delta}{\int_{-\pi}^{\pi}\exp\left[-\frac{s^2}{2\sigma_k^2}\right]ds}=e^{i\delta_k-\sigma_k^2/2}J_k ,$$

where

$$J_k=\frac{\mathrm{erf}[(-i\sigma_k^2+\pi)/\sqrt{2}\sigma_k]-\mathrm{erf}[(-i\sigma_k^2-\pi)/\sqrt{2}\sigma_k]}{2\mathrm{erf}(\pi/\sqrt{2}\sigma_k)} .$$

We note that if $\sigma_k \ll \pi$, then $J_k\approx 1$. In terms of these integrals the components of the expectation in (4) can be expressed in the following form:

$$E_{\psi_0}(\langle \psi_0,P(0)\psi_0\rangle_X)=\sum_{l\neq m}\bar a_l a_m\exp\left[i(\bar\delta_m-\bar\delta_l)-\frac{\sigma_l^2+\sigma_m^2}{2}\right]J_lJ_m\langle \phi_l,P(0)\phi_m\rangle_X+\sum_l |a_l|^2\langle \phi_l,P(0)\phi_l\rangle , \tag{15a}$$

$$E_{\psi_0}(\langle \psi_0,v(0)\rangle_X)=\sum_l \bar a_l e^{-i\bar\delta_l-(\sigma_l^2/2)}J_l\langle \phi_l,v(0)\rangle_X . \tag{15b}$$

Similar expressions can be obtained for the averaged initial conditions of the Lagrange multiplier functions in (5).

We note that in the limit of uncorrelated uniform phases $\sigma\to\infty$, and the first sum in (15a) and the sum in (15b) both vanish, so that (15) reduces to (14). The parameter σ thus connects the case in which the initial phases are known precisely ($\sigma\to 0$) with the case in which the initial phases are completely random ($\sigma\to\infty$). We shall explore the intermediate values $0<\sigma<\infty$ in the numerical experiments described below.

In the numerical experiments that follow we assume a spatial domain of length $L=50$, $X_0=6.0$, $D=10.0$, $\beta=10^{-6}$, $Q_{ij}=\delta_{ij}$, and the target state is taken to be a Gaussian of width $l=1$ centered at $X_f=8.0$:

$$r(X)=g(X,X_f,l)=\pi^{-1/4}l^{-1/2}\exp\left[-\frac{(X-X_f)^2}{2l^2}\right] .$$

We assume a nominal initial condition that is a superposition of the first two eigenstates:

$$\psi_0(X)=a_0e^{i\bar\delta_0}\phi_0(X)+a_1e^{i\bar\delta_1}\phi_1(X) , \tag{16}$$

where $\bar\delta_0=0$ and $\bar\delta_1=\pi/4$. The attenuation parameter ϵ in the dipole function $B(X)$ in this case was assumed to have the value $\epsilon=0.175$. The value of β chosen imposes a weak-fluence constraint.

We consider two distinct classes of controllers. The first class of controller is tailored to the nominal initial

conditions (i.e., $\sigma_k \to 0$; $k=0,1$), which we refer to as the *sensitive controller*. The second class of controllers comprise those that are designed to be insensitive to perturbations in the initial conditions by minimizing the average cost (4). The phases are assumed to have a Gaussian distribution so that the averaging can be performed analytically as outlined in (15). For the *insensitive controllers* we assume $\sigma_0=0.5$, and consider the cases with the following range of values of $\sigma_1=1.2,\ 1.5,\ 2.0$. These uncertainty parameters correspond to situations in which the phase of the ground state is known with more certainty than that of the second eigenstate. If the phases of both the eigenstates are known with equal confidence, i.e., $\sigma_0=\sigma_1$ and $Q_{ij}=\delta_{ij}$, then it can be seen from (15b) that the cost functional will assign them equal weight. In this case the factor $e^{-\sigma_1^2/2}$ that results from the averaging process can be removed from the sum in (15b). The cost for the insensitive controller can thus be obtained from that of the sensitive controller by merely adjusting the factor β. Therefore averaging in this case will not yield a controller that is any less sensitive to initial phases.

In Fig. 1(a) we plot the spatial distribution of the difference $|\psi(X,T)-r(X)|^2$ between the target state and the final state using the sensitive controller, and the same error distribution for the insensitive controller $(\sigma_0,\sigma_1)=(0.5,1.2)$. In both cases the nominal initial condition (16) was used. The sensitive controller can be seen to perform noticeably better than the insensitive controller, since it was tailored to the nominal initial condition. Even though the error of the insensitive controller is in this case substantially larger than that of the sensitive controller, the insensitive controller does manage to achieve an acceptable level of error and a close alignment with the target state as can be seen in Fig. 1(b). In Fig. 1(b) we plot the probability distributions of the target state and the final states using the same controllers as in Fig. 1(a). Consistent with the error distribution plotted in Fig. 1(a), the sensitive controller designed to the nominal initial condition achieves a final state that is impressively close to the target state. The insensitive control, though somewhat worse than the sensitive control due to its design constraints, also achieves a close alignment with the final target state. In Fig. 1(c) we plot the same pointwise difference as in Fig. 1(a) but using an initial state that has phase parameters $\delta_0=0$ and $\delta_1=\pi$ the second of which is far from its nominal phase value. In this case the action of the insensitive controller produces a final state that is much closer to the target state than that of the sensitive controller, which only moves the wave packet slightly away from the bottom of the well $X_0=4$ toward the target. Comparing Figs. 1(a) and 1(c) we observe a strong dependence of the performance index on the phase of the initial conditions, which is reduced to some extent by the use of the insensitive controllers.

In order to obtain a more complete picture of the performance of the two types of controllers over a range of values of the phase parameters, we define the following error functional:

$$\mathcal{E}[\psi(T)]=\langle \psi(T)-r, Q(\psi(T)-r)\rangle_X . \tag{17}$$

We note that this error functional is just the error component which is averaged when the performance index in (1) is calculated and represents the area under the curves shown in Figs. 1(a) and 1(c). For a particular choice of the phase parameters (δ_0,δ_1) we obtain for each controller a unique final state. The error functional $\mathcal{E}[\psi]$ can thus be used to define an error function $\mathcal{E}(\delta_0,\delta_1)$ that maps the parameter pair (δ_0,δ_1) onto the corresponding error.

In Fig. 1(d) we plot the error functions $\mathcal{E}(\delta_0,\delta_1)$ on the region $\delta_0=0$; $\delta_1\in[-\frac{3}{4}\pi,\frac{5}{4}\pi]$ corresponding to the sensitive controller $(\sigma_k=0;\ k=0,1)$ along with the insensitive controllers which share the same value of the ground-state uncertainty parameter $\sigma_0=0.5$ but with different values of the uncertainty parameter associated with the first eigenstate: $\sigma_1=1.2,\ 1.5;\ 2.0;\ 3.0$. For all these controllers the minimum error occurs at the nominal phase value $\delta_1=\pi/4$. All the controllers exhibit a similar sensitivity when the phase parameter δ_1 is perturbed away from its nominal value. In this case, however, the effectiveness of the insensitive design approach can be seen. In the immediate vicinity of the nominal phase value the sensitive controller performs better than the insensitive controllers. However, for large perturbations of the phase parameter δ_1, the insensitive controllers perform substantially better than the sensitive controller. As σ_1 is increased, the performance of the insensitive controllers is improved for large perturbations of δ_1 away from its nominal value; however, this improvement is accompanied by some reduction in performance in the vicinity of the nominal value of δ_1. This demonstrates the effect of the uncertainty parameter σ_1 in the design process. The insensitive controllers with larger values of σ_1 are designed to allow for wider distributions of initial phases than the insensitive controllers with lower values of σ_1. The flattening of the error curve as σ_1 increases indicates controllers can be designed that achieve a fairly uniform performance over the entire range $-\pi\le\delta_1\le\pi$. The compromise is the reduction in the best performance of these controllers at the nominal initial condition. However, depending on the design requirements these characteristics may be exploited in the design process.

The error functions $\mathcal{E}(\delta_0,\delta_1)$ on the region $\delta_0\in[-\pi,\pi]$; $\delta_1=\pi/4$ corresponding to the sensitive controller $(\sigma_k=0;\ k=0,1)$ and the insensitive controllers $\sigma_0=0.5$ and $\sigma_1=1.2;\ 1.5;\ 2.0;\ 3.0$ exhibit a similar phase sensitivity (not shown here) to that shown in Fig. 1(d) when the parameter δ_0 is perturbed away from its nominal value $\delta_0=0$. All the controllers exhibit a similar performance characteristic because the uncertainty parameter $\sigma_0=0.5$ for the insensitive controllers is close to zero. There is a slight degradation in the performance of the insensitive controllers with respect to the first phase parameter δ_0. This is compensated for by an increased performance with respect to the second phase parameter δ_1, which can be seen clearly in Fig. 1(d).

In Fig. 1(e) the external electric fields $K(t)$ for the sensitive controller and the insensitive controllers with $\sigma_0=0.5$ and $\sigma_1=1.2;\ 1.5;\ 2.0;\ 3.0$ are plotted for comparison. All the controllers execute a control strategy that

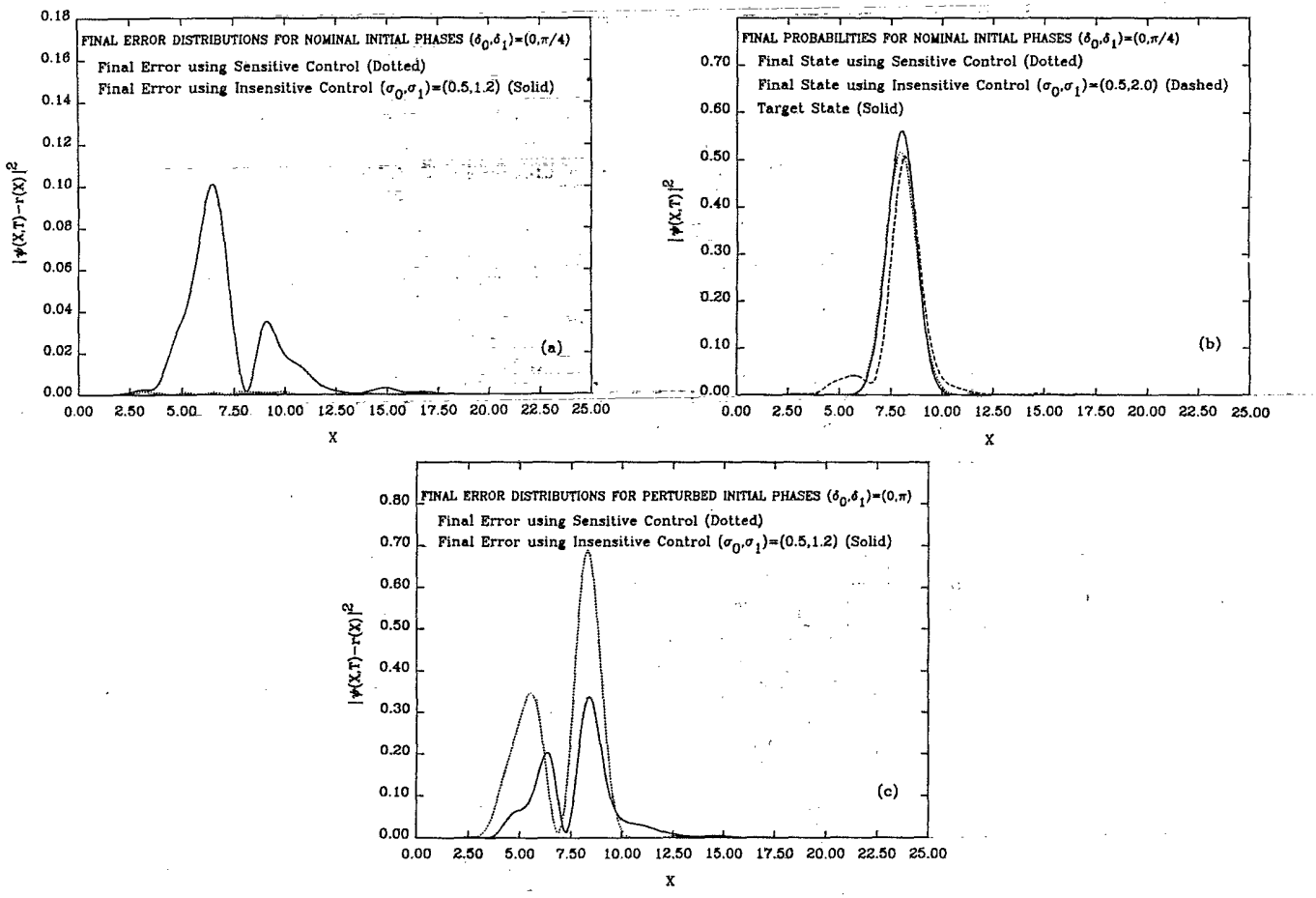
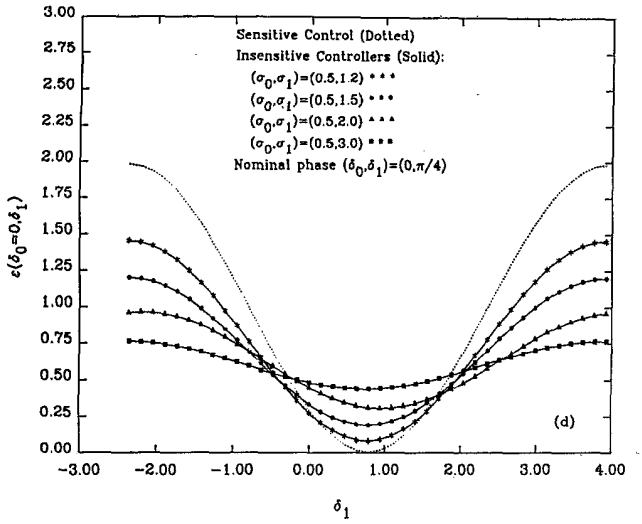


FIG. 1. (a) The spatial distribution of the difference $|\psi(X,T)-r(X)|^2$ between the target state and the final state using the sensitive controller, and the same error distribution for the insensitive controller $(\sigma_0,\sigma_1)=(0.5,1.2)$. In both cases the nominal initial condition (16) was used. The sensitive controller can be seen to perform better than the insensitive controller since it was tailored to the nominal initial condition. (b) The probability distributions of the target state, the final state using the sensitive controller, and the final state using the insensitive controller with $(\sigma_0,\sigma_1)=(0.5,1.2)$. In both cases the nominal initial condition (16) was used. The sensitive controller, which is designed to the nominal initial condition, achieves a final state that is impressively close to the target state. The insensitive control, though somewhat worse than the sensitive control due to its design constraints, also achieves a close alignment with the final target state. (c) The same error distributions as those plotted in (a), but using an initial state that has phase parameters $\delta_0=0$ and $\delta_1=\pi$, the second of which is far from its nominal phase value. In this case the action of the insensitive controller produces a final state that is much closer to the target state than that of the sensitive controller, which only moves the wave packet slightly away from the bottom of the well $X_0=4$ toward the target. (d) The error functions $\mathcal{E}(\delta_0,\delta_1)$ on the region $\delta_0=0$; $\delta_1\in[-\frac{3}{4}\pi,\frac{5}{4}\pi]$ corresponding to the sensitive controller $(\sigma_k=0;\ k=0,1)$; and the insensitive controllers which share the same value of the ground-state uncertainty parameter $\sigma_0=0.5$ but with different values of the uncertainty parameter associated with the first eigenstate: $\sigma_1=1.2$; 1.5; 2.0; 3.0. For all these controllers the minimum error occurs at the nominal phase value $\delta_1=\pi/4$. All the controllers exhibit sensitivity to perturbations of the phase parameter δ_1 away from its nominal value. The effectiveness of the insensitive design approach can be seen: in the immediate vicinity of the nominal phase value, the sensitive controller performs better than the insensitive controllers; however, for large perturbations of the phase parameter δ_1, the insensitive controllers perform substantially better than the sensitive controller. The insensitive controllers with larger σ_1 values are designed to allow for wider distributions of initial phases and thus perform better when large perturbations from the nominal phase value are considered. As a compromise, the insensitive controllers with larger σ_1 values have a reduced performance in the immediate vicinity of the nominal value of δ_1. (e) The external electric fields $K(t)$ for the sensitive controller and the insensitive controllers with $\sigma_0=0.5$ and $\sigma_1=1.2$; 1.5; 2.0; 3.0. All the controllers execute a control strategy that can be divided into roughly three regions. In the first region there is a relatively large amplitude pulse, which for the sensitive controller extends over the interval $[0,\pi]$. The duration of the initial pulse is longer for the insensitive controllers, namely on the interval $[0,2\pi]$, and the amplitude of this initial pulse for the insensitive controllers is noticeably larger than that for the sensitive controller. This initial pulse may possibly be interpreted as a "phase imprinting" stage. The insensitive controllers expend relatively more energy on ensuring that the wave packets adopt the appropriate phases for more intense pumping later on in the interval. These differences reflect the design requirement that the insensitive controllers be able to control wave packets with unknown phases, whereas the sensitive controller is certain of the initial phase and therefore expends relatively less energy on phase imprinting. On the second region, roughly $[\pi,5\pi]$ for the sensitive controller and $[2\pi,5\pi]$ for the insensitive controllers, all the controllers execute relatively small amplitude pulsing. The third region, roughly $[5\pi,8\pi]$ when most of the work is performed, is characterized by a large amplitude pulse having an approximately similar structure for all the controllers.



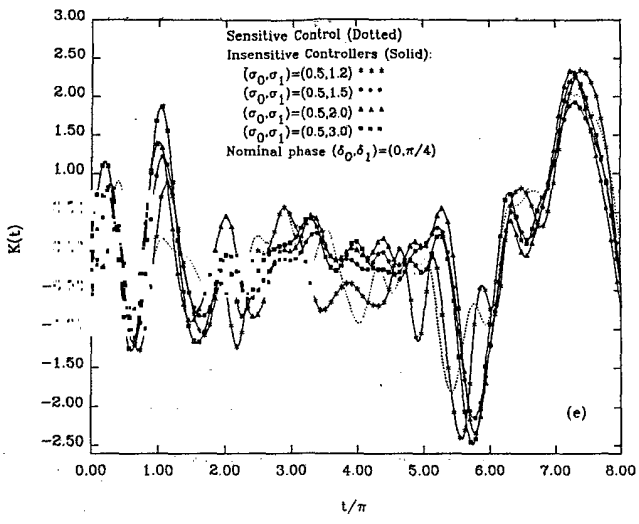


FIG. 1. (*Continued*).

can be divided into roughly three regions. In the first region there is a relatively large amplitude pulse, which for the sensitive controller extends over the interval $[0,\pi]$. The duration of the initial pulse is longer for the insensitive controllers—namely on the interval $[0,2\pi]$, and the amplitude of this initial pulse for the insensitive controllers is noticeably larger than that for the sensitive controller. This initial pulse may possibly be interpreted as a "phase-imprinting" stage. The insensitive controllers expend relatively more energy on ensuring that the wave packets adopt the appropriate phases for the more intense fields later on in the interval. On the second region, roughly $[\pi,5\pi]$ for the sensitive controller and $[2\pi,5\pi]$ for the insensitive controllers, all the controllers execute relatively small amplitude structure. The third region, roughly $[5\pi,8\pi]$, is characterized by a large amplitude pulse having an approximately similar structure for all the controllers.

Thus the major difference between the sensitive and insensitive controllers is that the insensitive controllers imprint phases on the wave packets relatively early in the interval $[0,8\pi]$ to achieve the desired phase at $t=5\pi$ (the beginning of the intense field region) and so compensate as best as possible for the imprecise knowledge of the phase at $t=0$. The sensitive controller contains information about how to go from a particular phase at $t=0$ to a desired phase at $t=5\pi$. These differences reflect the design requirement that the insensitive controllers be able to control wave packets with unknown phases, therefore they expend a greater portion of the available energy ensuring the phase is as correct as possible before the large pulse is executed at the end of the interval. Furthermore, the insensitive controller with large uncertainty in the phase [i.e., larger σ_1 in Fig. 1(e)] exhibits a more intense phase imprinting field. In contrast, the sensitive controller is certain of the initial phase and therefore expends relatively less energy on phase imprinting.

B. An uncertain parameter

Here we assume that the uncertain parameter α described in the theoretical development in Sec. II is the parameter D in the molecular Hamiltonian of the dimensionless Schrödinger equation (12). One may regard this problem as one of controlling a single molecule with uncertainty in D or an example of controlling a class of similar molecules having a distribution of D values. We are not able in this case to exploit an operator approach to remove the dependence on D analytically, as we were in the case of uncertain initial conditions. Therefore, our approach is to use numerical integration to evaluate the averaged cost functional and gradients.

We assume a spatial domain of length $L=24.0$, $X_0=6.0$, $\beta=10^{-6}$, $Q_{ij}=\delta_{ij}$, and that

$$d\mu(D)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(D-\bar{D})^2}{2\sigma^2}\right]dD\ ,$$

where $\bar{D}=14$ is the value of the nominal parameter, $\sigma=1$ or 2, and the width of the parametric domain W for the purposes of numerical integration was 3σ. The initial condition in this case was taken to be $\psi_0(X)=g(X,X_0=6,1)$, and the target state was taken to be $r=g(X,X_f=8,1)$. The attenuation parameter ϵ in the dipole function $B(X)$ in this case was assumed to have the value $\epsilon=0$.

In Fig. 2(a) we plot the spatial distribution of the difference $|\psi(X,T)-r(X)|^2$ between the target state and the final state using the sensitive controller, and the same error distributions for the insensitive controllers with $\sigma=1$ and 2. In all cases the nominal parameter $\bar{D}=14$ was used. The sensitive controller performs somewhat better than the insensitive controllers, since it was designed to be optimal for the nominal parameter. The controller designed to allow for more uncertainty in D by

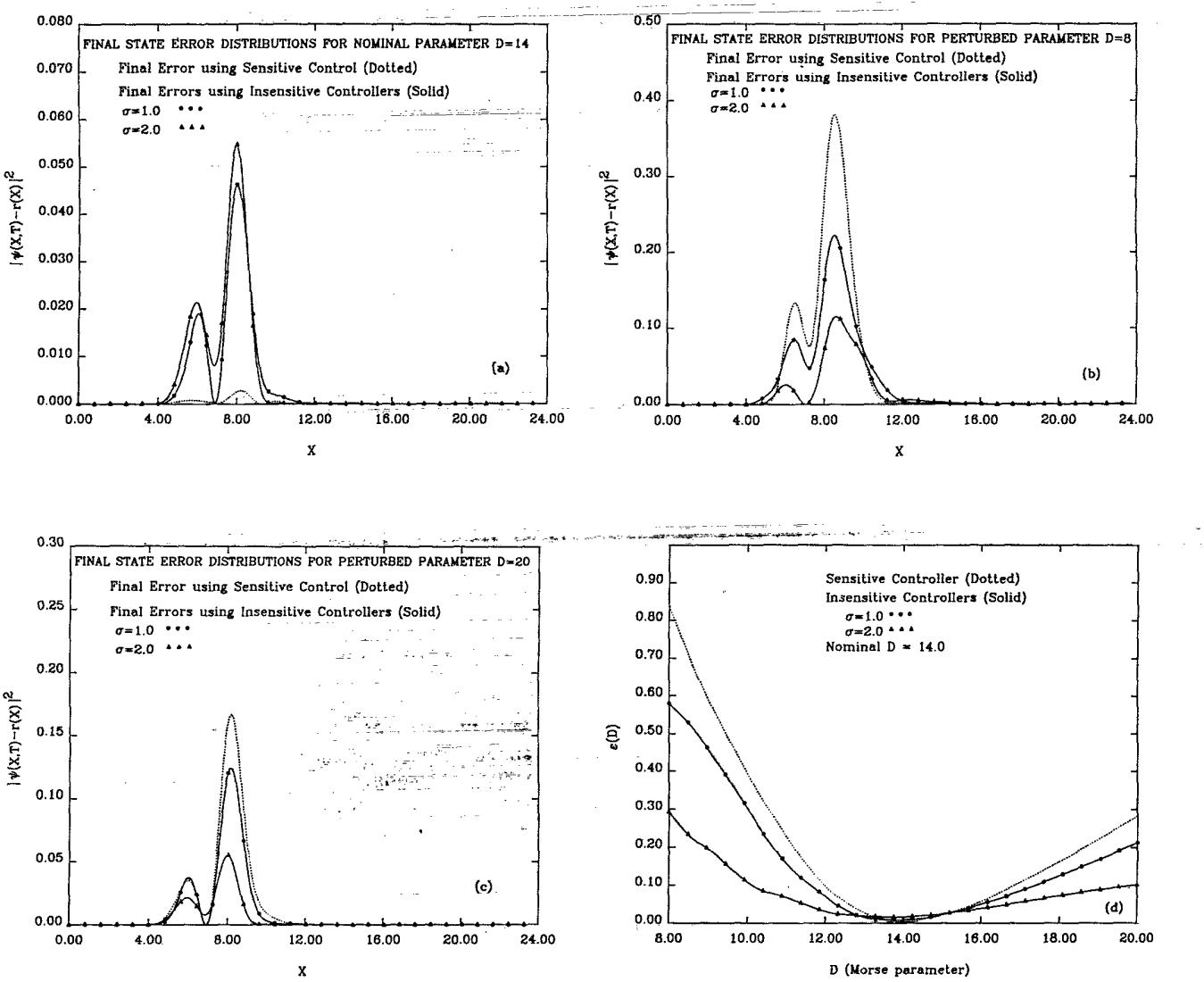


FIG. 2. (a) The spatial distribution of the difference $|\psi(X,T)-r(X)|^2$ between the target state and the final state using the sensitive controller, and the same error distributions for the insensitive controllers with $\sigma=1$ and 2. In all cases the nominal Morse parameter $\bar{D}=14$ was used. The sensitive controller performs somewhat better than the insensitive controllers, since it was designed to be optimal for the nominal parameter. (b) The same set of error distributions as in (a) for a substantial perturbation of the parameter D from the nominal value $\bar{D}=14$ to the smaller value $D=8$. In this case the insensitive controllers with $\sigma=1$ and 2 can be seen to perform noticeably better than the sensitive controller. (c) The same set of error distributions as in (a) and (b) for a substantial perturbation of the parameter D from the nominal value $\bar{D}=14$ to the larger value $D=20$. In this case the insensitive controllers still perform better than the sensitive controller. (d) The error function $\mathcal{E}(D)$ defined in (17) corresponding to the sensitive controller ($\sigma=0$) and that corresponding to the insensitive controllers ($\sigma=1$ and 2). The insensitive controllers achieve a substantial improvement in the performance over the sensitive controller away from the nominal parameter value $D=14$. (e) The external electric field $K(t)$ for the sensitive controller and the insensitive controllers with $\sigma=1$ and 2. Except for a small difference around $t=4\pi$, the two controllers have an almost identical phase structure. The insensitive controllers use subtle changes in amplitude to achieve a more robust design by ensuring that the wave packet avoids taking significant excursions into the regions $X\notin[5,7]$ [see (f)], where the perturbations to D have a large effect. This will prevent the wave packet from exploring uncertain regions of the potential, and therefore becoming uncontrollable. The sensitive controller does not exhibit this restraint, since it was not a design requirement that it cater to a large range of values of D. The strategy of depositing the majority of the energy in the last part of the interval minimizes the exposure time of the wave packet to the uncertain regions, and also the time over which the wave packet is able to respond. (f) The Morse potential for values $D=8$, 14, 20 representing the most extreme perturbations to D shown in (d) as well as the nominal value. Perturbations to the parameter D in the interval [8,20] leave the potential virtually unchanged in the region $X\in[5,7]$. This property is exploited by the insensitive controllers to avoid long excursions of the wave packet into regions of uncertainty in the potential.

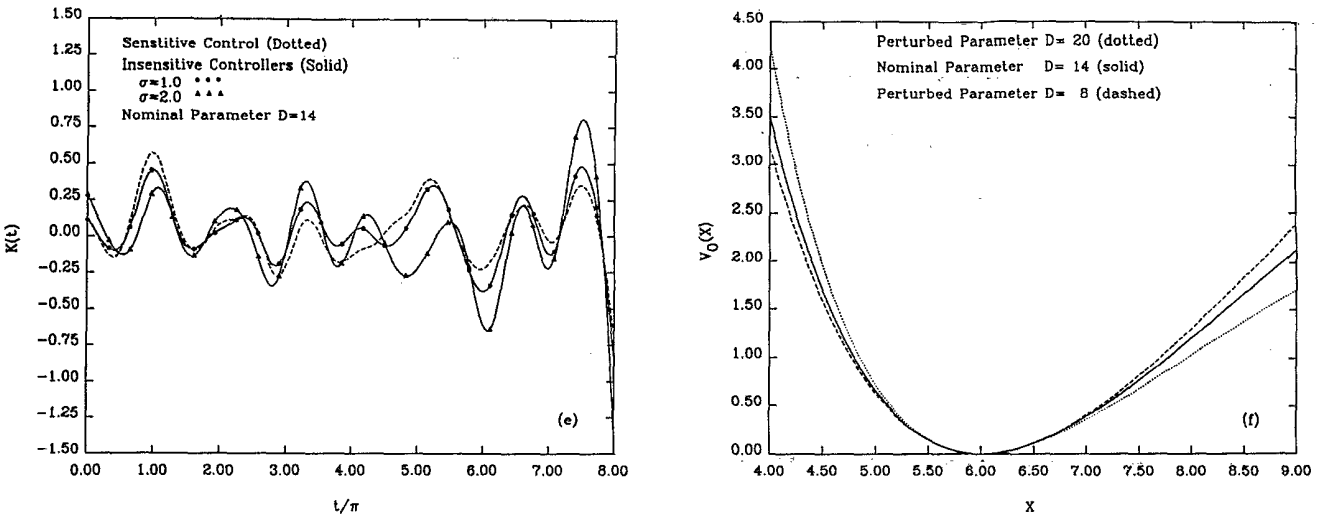


FIG. 2. (*Continued*).

increasing the parameter σ from 1 to 2 performs worse at the nominal parameter value. In Fig. 2(b) we plot the same set of error distributions as in Fig. 2(a) for a substantial perturbation of the parameter D from the nominal value $\bar{D}=14$ to the smaller value $D=8$. In this case the insensitive controllers with $\sigma=1$ and 2 can be seen to perform noticeably better than the sensitive controller. In Fig. 2(c) we plot the same set of error distributions as in Fig. 2(a) and 2(b) for a substantial perturbation of the parameter D from the nominal value $\bar{D}=14$ to the larger value $D=20$. In this case the insensitive controllers still perform better than the sensitive controller.

In order to obtain a more complete picture of the performance of the two types of controllers over a range of values of the parameter D, we use the error functional (17) to measure the performance of the two types of controllers. For a particular choice of parameter D we obtain for each controller a unique final state. We use the error functional $\mathcal{E}[(\psi)]$ to define an error function $\mathcal{E}(D)$ that maps the parameter D onto the corresponding error. In Fig. 2(d) we plot the error function $\mathcal{E}(D)$ corresponding to the sensitive controller ($\sigma=0$) and that corresponding to the insensitive controllers with $\sigma=1$ and 2. The sensitive controller performs marginally better than the insensitive controllers in the vicinity of the nominal parameter value $D=14$. However, as D is perturbed away from its nominal value, the insensitive controllers perform noticeably better than the sensitive controller. The performance of the insensitive controllers for large perturbations of D is enhanced by increasing the uncertainty design parameter σ from 1 to 2. All three controllers exhibit an asymmetric response in performance to perturbations in the parameter D about the nominal value $\bar{D}=14$.

In Fig. 2(e) the external electric field $K(t)$ for the sensitive controller and the insensitive controllers with $\sigma=1$ and 2 are plotted for comparison. Except for a small difference around $t=4\pi$, the controllers have an almost identical phase structure. The insensitive controllers use subtle changes in amplitude to achieve a more robust design. This may be explained in terms of the effect that the perturbations in the parameter D have on the potential. In Fig. 2(f) the Morse potential is plotted for values $D=8$, 14, and 20 representing the most extreme perturbations to D shown in Fig. 2(d) as well as the nominal value. The insensitive controllers initially pulse with amplitudes that are just small enough to have the wave packet avoid taking significant excursions into the regions $X\notin[5,7]$, where the perturbations to D have a large effect. This will keep the wave packet from exploring uncertain regions of the potential, and therefore becoming uncontrollable. The sensitive controller does not exhibit this restraint, since it was not a design requirement that it cater for a large range of values of D. On the interval $[6\pi,8\pi]$ the insensitive controllers choose to increase the pulse amplitude above that of the sensitive controller. The strategy of depositing the majority of the energy in the last part of the interval minimizes the exposure time of the wave packet to the uncertain regions and also the time over which the wave packet is able to respond.

The above example clearly demonstrates the effectiveness of the parameter insensitive design procedure, particularly in regions where the sensitive controller performs poorly. It also provides insight into how the insensitive controllers exploit the known regions of the potential to enhance controllability.

VI. CONCLUSIONS

We have proposed an optimal control methodology for design in an environment of parametric uncertainties in the molecular Hamiltonian and uncertainties in the initial state of the system. The technique involves defining the

optimal control problem involving averaged cost functionals, which minimize dependence of the optimal control on the uncertain parameters and uncertain initial conditions. We have given a precise formulation of the optimal control problem and proved that it is well posed. In order to be able to perform the averaging process explicitly in the case of uncertain initial conditions, we have introduced a class of cost propagators. By defining a corresponding set of Lagrange multipliers, we have derived necessary conditions for a minimum of the optimization problem. These necessary conditions form the basis for a gradient search procedure to search for a minimum. We have discussed the numerical approximation of the uncertain optimal control problem using central differences in space with Crank-Nicholson time stepping for the one-dimensional operators and ADI time stepping for the two-dimensional cost propagators.

Two distinct numerical experiments were performed to explore the performance of the proposed uncertain optimal controllers. The theory presented in this paper is generally applicable to a wide class of quantum systems, while the simplest example was chosen here to illustrate the theory. The first example considered the case in which the initial phases of the quantum system are uncertain while the remaining parameters of the system are assumed to be known precisely. We assumed various distributions of initial phases and performed the averaging analytically. We demonstrated that for uniformly distributed uncorrelated phases it is not possible to achieve final-state control. The result in this extreme case is interesting not only because it establishes a theoretical limit on the control in an environment of uncertain initial phases, but also because it indicates that it is not possible to coerce a quantum system from a completely random initial state to a correlated state, as represented by the target final state. However, as mentioned earlier, less demanding objectives will still leave an element of control even with fully random phases. For initial phases having a Gaussian distribution, it is possible to design controllers that are insensitive to perturbations in the nominal initial phases. The insensitive controllers were found to adopt a significantly different pulsing strategy when compared to controllers designed at the nominal initial phases. The insensitive controller expends a greater portion of the available energy, apparently ensuring the phase is correct before applying a large amplitude pulse toward the end of the control interval to achieve the final target state. This process we have named phase imprinting. In contrast, the sensitive controller is certain of the initial phase and therefore expends relatively less energy on phase adjustment.

In the second example we assumed that there was parametric uncertainty in the molecular Hamiltonian, while the initial conditions are assumed to be known. In this case we used numerical averaging over a range of parameter values having a Gaussian distribution. The parameter-sensitive controller was found to perform significantly worse when the parameter was perturbed away from the nominal value. The insensitive design technique considered in this paper yields controllers that perform much better in the regions of poor performance of the parameter-sensitive controller and whose performance is only marginally worse than that of the sensitive controller in the vicinity of the nominal parameter value. The insensitive controllers exploit the known regions of the potential to enhance controllability by pulsing with a delicate structure that prevents the wave packet from taking significant excursions into the regions where the potential is uncertain.

The system considered in this study exhibited a greater sensitivity to uncertainties in the phase of the initial condition than to uncertainties in the Morse parameter. The insensitive design technique using averaged costs proved to be effective in reducing the effects of these system uncertainties. Each system and type of uncertainty is expected to exhibit its own special characteristics, which the insensitive controllers will exploit to minimize the effect of uncertainties.

ACKNOWLEDGMENTS

The authors wish to thank Professor K. Lehmann for helpful comments. A. P. Peirce acknowledges the support of the National Science and Engineering Research Council of Canada. H. Rabitz acknowledges the support of the Army Research Office.

APPENDIX A: ANALYSIS OF THE OPERATOR EQUATIONS

In this appendix we will study the operator equation

$$\frac{dP(t,\alpha)}{dt}=A(\alpha)^*P(t,\alpha)+P(t,\alpha)A(\alpha)\ ,$$
$$P(T,\alpha)=Q\ , \tag{A1}$$

where $A(\alpha)$ and Q are defined in (1). The goal is to show that (A1) leads to a well-posed evolution equation in the space of Hilbert-Schmidt operators. Let $\{\phi_i\}$ be an orthonormal basis in X, and define the tensor product $\phi_i\otimes\phi_j$ as the bilinear form which acts on $X\times X$ by $(\phi_i\otimes\phi_j)(\psi_1,\psi_2)=\langle\psi_1,\phi_i\rangle\langle\psi_2,\phi_j\rangle$. Given a Hilbert-Schmidt operator P, it can be expanded in terms of $\phi_i\otimes\phi_j$ as follows:

$$P=\sum_{i,j}a_{ij}\phi_i\otimes\phi_j\ ,$$

where

$$\sum_{i,j}(a_{ij})^2<\infty\ .$$

As usual,[10] the tensor product of X with itself $X\otimes X$ is defined as the completion of the linear combinations $\sum_{i,j}^n a_{ij}\phi_i\otimes\phi_j$ with respect to the norm $[\sum_{ij}^n(a_{ij})^2]^{1/2}$. From these definitions it follows that $X\otimes X\simeq X_{\rm HS}$. We denote by U the isometric isomorphism between $X_{\rm HS}$ and $X\otimes X$. We will use this idea to write (A1) as an evolution equation in $X\otimes X$. To this end, define the tensor product $A\otimes B:X\otimes X\to X\otimes X$ of the operators $A,B:X\to X$ as

$$(A\otimes B)(v_1\otimes v_2)=Av_1\otimes Bv_2$$

for all $v_1\otimes v_2\in D(A\otimes B)\subset X\otimes X$. Thus we can write Eq.

(A1) as

$$\frac{d\tilde{P}(t,\alpha)}{dt}=\tilde{A}\tilde{P}(t,\alpha);\quad \tilde{P}(T,\alpha)=\tilde{Q}\ , \tag{A2}$$

where $\tilde{A}=I\otimes A(\alpha)^*+A(\alpha)^*\otimes I$, and $\tilde{P}$ and $\tilde{Q}$ are the images of P and Q under the isomorphism U. The well posedness of (A2) follows from the following theorem.

Theorem 2. $\tilde{A}$ generates a C_0 semigroup on the space $X\otimes X$.

Proof. The proof is a consequence of the following three lemmas and the fact that if $\tilde{A}$ is a closed, densely defined operator on $X\otimes X$ (Ref. 6) then $\tilde{A}$ generates a C_0-semigroup $T(t)$ such that $\|T(t)\|\leq Ce^{\gamma t}$ for all t, if and only if

$$\mathrm{Re}\langle \tilde{A}^*z,z\rangle_{X\otimes X}\leq\gamma\|z\|^2_{X\otimes X}\quad \text{for all } z\in D(\tilde{A}^*)\ ,$$

and

$$\mathrm{Re}\langle \tilde{A}z,z\rangle_{X\otimes X}\leq\gamma\|z\|^2_{X\otimes X}\quad \text{for all } z\in D(\tilde{A})\ .$$

Lemma 1. The domain of $\tilde{A}$, $D(\tilde{A})=\{P\in X\otimes X: \tilde{A}P\in X\otimes X\}$ is dense in $X\otimes X$.

Proof. This is an immediate consequence of the fact that $D(A(\alpha))$ is dense in X.

Lemma 2. $\tilde{A}$ is a closed operator.

Proof. Since A generates a C_0 semigroup, A is a closed operator.[6] The fact that $\tilde{A}$ is closed follows from its definition and the fact that A is closed.

Lemma 3. There exists a $\gamma\geq 0$ such that

$$\mathrm{Re}\langle \tilde{A}^*z,z\rangle_{X\otimes X}\leq\gamma\|z\|_{X\otimes X}\quad \text{for all } z\in D(\tilde{A}^*)$$

and

$$\mathrm{Re}\langle \tilde{A}z,z\rangle_{X\otimes X}\leq\gamma\|z\|_{X\otimes X}\quad \text{for all } z\in D(\tilde{A})\ .$$

Proof. Assumption (1) implies that there exists a β such that

$$\begin{aligned}&\mathrm{Re}\langle A(\alpha)z,z\rangle_X\leq\beta\|z\|^2_X\quad \text{for all } z\in D(A(\alpha))\ ,\\ &\mathrm{Re}\langle A(\alpha)^*z,z\rangle_X\leq\beta\|z\|^2_X\quad \text{for all } z\in D(A(\alpha)^*)\ .\end{aligned} \tag{A3}$$

For the sake of brevity we will present the remainder of the calculations assuming that we are dealing with real Hilbert spaces. Let $R\in D(\tilde{A}^*)$; this can be written as

$$R=\sum_{i,j}a_{ij}\phi_i\otimes\phi_j$$

and

$$\langle \tilde{A}^*R,R\rangle_{X\otimes X}=\mathrm{tr}(\tilde{A}^*RR^*)=\sum_k\langle \phi_k,\tilde{A}^*RR^*\phi_k\rangle_X\ ,$$

where $\{\phi_i\}$ is an orthonormal basis in X. From the definitions we have

$$\tilde{A}^*R=\sum_{i,j}a_{ij}[A(\alpha)\phi_i\otimes\phi_j+\phi_i\otimes A(\alpha)\phi_j]$$

and

$$\begin{aligned}\tilde{A}^*RR^*\phi_k=&\sum_{i,j,s,r}a_{ij}a_{sr}\langle\phi_j,\phi_r\rangle\langle\phi_s,\phi_k\rangle A(\alpha)\phi_i\\ &+a_{ij}a_{sr}\langle A(\alpha)\phi_j,\phi_r\rangle\langle\phi_s,\phi_k\rangle\phi_i\ .\end{aligned}$$

Therefore we have

$$\begin{aligned}\mathrm{tr}(\tilde{A}RR^*)=&\sum_j\left[\left\langle\sum_i a_{ij}A(\alpha)\phi_i,\sum_s a_{sj}\phi_s\right\rangle\right]\\ &+\sum_k\left[\left\langle\sum_j a_{kj}A(\alpha)\phi_j,\sum_r a_{kr}\phi_r\right\rangle\right]\ .\end{aligned}$$

At this point we can use the fact that $A(\alpha)$ itself generates a C_0 semigroup and thus satisfies (A3). This implies

$$\begin{aligned}\mathrm{tr}(\tilde{A}RR^*)\leq&\sum_j\left[\beta\|\sum_i a_{ij}\phi_i\|^2_X\right]\\ &+\sum_k\left[\beta\|\sum_i a_{ki}\phi_i\|^2_X\right]\\ \leq&2\beta\sum_{i,j}(a_{ij})^2=2\beta\|R\|^2_{X\otimes X}\ .\end{aligned}$$

Set $\gamma=2\beta$. Then

$$\langle \tilde{A}^*R,R\rangle_{X\otimes X}\leq\gamma\|R\|^2_{X\otimes X}\quad \text{for all } R\in D(\tilde{A})\ ,$$

and similarly we can show that

$$\langle \tilde{A}R,R\rangle_{X\otimes X}\leq\gamma\|R\|^2_{X\otimes X}\quad \text{for all } R\in D(\tilde{A})\ .$$

Therefore we can conclude that there exists a unique solution $P(t,\alpha)\in L^2([0,T];X_{\mathrm{HS}})$ for the operator equation. Note that the new state equation (3a) differs from the equation studied here in that its generator is $\tilde{A}$ plus a bounded operator. The extension of the results of this appendix to this case follows by a perturbation argument.[11]

APPENDIX B: APPROXIMATION

From a practical point of view, the question of whether the operator equations can be approximated is of extreme importance. In this appendix we will resolve this question by exhibiting conditions for the existence of a discrete semigroup approximating $T_{\tilde{A}}(t)$. The main device used here for this purpose is the Kato-Trotter theorem. We will restrict ourselves to the case where the generator A is self adjoint. The general case can be treated in the same way if A and its adjoint each satisfy the requirements of the Kato-Trotter theorem.[6]

Theorem (Kato-Trotter): Let $A_N\in G(M,w)$, $A\in G(M,w)$ and assume (a) As $N\to\infty$, $A_Nx\to Ax$ for every $x\in D(A)$, where D is a dense subset of X; (b) There exists a λ_0 with $\mathrm{Re}\lambda_0>w$ for which $(\lambda_0 I-A)D$ is dense in X. If $T_N(t)$ and $T(t)$ are the C_0 semigroups generated by A_N and A, respectively, then

$$\lim_{N\to\infty}T_N(t)x=T(t)x\quad \text{for all } t>0,\ x\in X\ ,$$

and the limit is uniform in t for bounded intervals. The notation $A\in G(M,w)$ means that A generates a C_0-semigroup $T(t)$ which satisfies $\|T(t)\|\leq Me^{wt}$.

We want to show that if the generator of the state equation admits an approximation, then the operator equation which is used to replace the state equation admits a similar approximation. For this purpose we will assume the following.

(1) There exists a continuous projection P_N such that

$$P_N: X \to X_N ,$$

where $X_N = \text{span}\{\phi_1, \ldots, \phi_N\}$, $\phi_i \in X$, and $A_N: X \to X_N$, $A_N \in G(M,w)$, $N = 1,2,\ldots$.

(2) $A_N z \to Az$ for all $z \in D(A)$.

(3) There exists a λ_0 with $\text{Re}(\lambda_0) > w$ such that $(A - \lambda_0 I)D$ is dense in X.

We define the following.

(d1) $\tilde{X}_N = \text{span}\{\phi_i \otimes \phi_j : i,j = 1, \ldots, N\}$;

(d2) $E_N: X \otimes X \to \tilde{X}_N$, defined by $E_N(v) = E_N(v_1 \otimes v_2) = P_N v_1 \otimes P_N v_2$ for all $v \in X \otimes X$;

(d) $\tilde{A}_N: X \otimes X \to \tilde{X}_N$, defined by

$$\tilde{A}_N(v) = \tilde{A}_N(v_1 \otimes v_2) = (I_N \otimes A_N + A_N \otimes I_N)(v_1 \otimes v_2) .$$

The Kato-Trotter theorem and the following three lemmas demonstrate that the discrete semigroup corresponding to $\tilde{A}_N$ converges to $T_{\tilde{A}}(t)$.

Lemma 4. E_N is a continuous projection onto $\tilde{X}_N$.

Proof. $v \in X \otimes X$:

$$E_N(v) = E_N\left[\sum_{i,j} v_{ij} \phi_i \otimes \phi_j\right] = \sum_{i,j}^{N} v_{ij} \phi_i \otimes \phi_j ,$$

$$\|E_N(v)\|_{\tilde{X}_N} = \sum_{i}^{N} \sum_{j}^{N} (v_{ij})^2 \le \sum_{i}^{\infty} \sum_{j}^{\infty} v_{ij}^2 ,$$

$$\|E_N(v)\|_{\tilde{X}_N} \le \|v\|_{X \otimes X} .$$

Lemma 5. $\tilde{A}_N v \to \tilde{A} v$ for all $v \in D(\tilde{A})$.

Proof. We define the following:

$$\tilde{D}(\tilde{A}) = \left\{ R \in X \otimes X : \tilde{A}R \in X \otimes X, \text{ and } R = \sum_{i,j} r_i s_j \phi_i \otimes \phi_j \right\} .$$

$\tilde{D}(\tilde{A})$ is dense in $X \otimes X$ since $D(A)$ is dense, and the functions v_{ij} can be approximated by products $v_i^1 v_j^2$. Let $v \in \tilde{D}(\tilde{A})$:

$$\tilde{A}_N(v) = \sum_{i,j}^{N} r_i s_j (I_N \otimes A_N + A_N \otimes I_N)(\phi_i \otimes \phi_j) ,$$

$$(\tilde{A}_N - \tilde{A})(v) = \sum_{i,j}^{N} r_i s_j \phi_i \otimes (A_N - A)\phi_j + \sum_{i,j}^{N} r_i s_j (A_N - A)\phi_i \otimes \phi_j - \sum_{i,j>N} (\phi_i \otimes A\phi_j + A\phi_i \otimes \phi_j) r_i s_j$$

$$= T(1) + T(2) + T(3) .$$

Now look at each of the terms separately. For the first term we have

$$\|T(1)\|_{X \otimes X} = \left\| \sum_{i,j}^{N} r_i s_j [\phi_i \otimes (A_N - A)\phi_j] \right\|_{X \otimes X} .$$

By defining $S = A_N - A$, $v_1 = \sum_i^N r_i \phi_i$, and $v_2 = \sum_j^N s_j \phi_j$, we get

$$\|v_1 \otimes S v_2\|_{X \otimes X}^2 = \sum_{i,j} |\langle v_1, \phi_i \rangle|^2 |\langle S v_2, \phi_j \rangle|^2 = \|v_1\|_X^2 \|S v_2\|_X^2 .$$

From assumption (2), it follows that

$$\|T(1)\|_{X \otimes X} \to 0 \quad \text{as} \quad N \to \infty .$$

Similarly, we can show that

$$\|T(2)\|_{X \otimes X} \to 0 \quad \text{as} \quad N \to \infty .$$

The third term gives

$$\|T(3)\|_{X \otimes X} = \left\| \sum_{i,j>N}^{\infty} r_i s_j (\phi_i \otimes A\phi_j + A\phi_i \otimes \phi_j) \right\|_{X \otimes X} .$$

We define $v' = \sum_{i>N}^{\infty} r_i \phi_i$ and $v'' = \sum_{j>N}^{\infty} s_j \phi_j$:

$$\|T(3)\|_{X \otimes X} = 2\|v' \otimes Av''\|_{X \otimes X}$$

$$= 2\|v'\|_X \|Av''\|_X ,$$

which implies that

$$\|T(3)\|_{X \otimes X} \to 0 \quad \text{as} \quad N \to \infty ,$$

since $v', v'' \in D(A)$.

Lemma 6. There exists a λ_0 with $\text{Re}(\lambda_0) > 2w$ such that $(\lambda_0 \tilde{I} - \tilde{A})D$ is dense in $X \otimes X$.

Proof. By assumption $A \in G(M,w)$ and for any $\mu > w$, $(\mu I - A)D(A)$ is dense in X. Now consider the equations:

$$(\lambda_0 \tilde{I} - \tilde{A})(\phi_1 \otimes \phi_2) = (\psi_1 \otimes \psi_2) ,$$

$$\lambda_0 \phi_1 \otimes \phi_2 - (\phi_1 \otimes A\phi_2 + A\phi_1 \otimes \phi_2) = \psi_1 \otimes \psi_2 ,$$

$$\left[\frac{\lambda_0}{2} I - A\right]\phi_1 \otimes \phi_2 + \phi_1 \otimes \left[\frac{\lambda_0}{2} I - A\right]\phi_2 = \psi_1 \otimes \psi_2 .$$

This leads to the following two equations:

$$\left[\frac{\lambda_0}{2}I - A\right]\phi_1 = \psi_1 , \tag{B1}$$

$$\left[\frac{\lambda_0}{2}I - A\right]\phi_2 = \psi_2 . \tag{B2}$$

From (B1) and (B2) and the fact that $\{\psi:[(\lambda_0/2)I - A]\phi = \psi, \phi \in D(A)\}$ is dense in X, it follows that $(\lambda_0 \tilde{I} - \tilde{A})D(\tilde{A})$ is dense $\in X \otimes X$.

With the use of lemmas (4), (5), and (6) we see that $\tilde{A}$ satisfies the conditions of the Kato-Trotter theorem. Therefore, if $\tilde{T}_N(t)$ and $\tilde{T}(t)$ are the semigroups generated by $\tilde{A}_N$ and $\tilde{A}$, respectively, we have $\tilde{T}_N(t)v \to \tilde{T}(t)v$ for all $v \in X \otimes X$.

[1]A. P. Peirce, M. Dahleh, and H. Rabitz, Phys. Rev. A **37**, 4950 (1988).

[2]S. Shi, A. Woody, and H. Rabitz, J. Chem. Phys. **88**, 6870 (1988).

[3]S. Shi and H. Rabitz, J. Chem. Phys. **92**, 364 (1990).

[4]A. V. Balakrishnan, *Applied Functional Analysis* (Springer-Verlag, New York, 1981).

[5]A. Bensoussan, *Filtrage Optimal des Systemes Lineaires* (Dunod, Paris, 1971).

[6]A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations* (Springer-Verlag, New York, 1983).

[7]G. D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods,* 2nd ed. (Oxford University, London, 1978).

[8]D. G. Luenberger, *Linear and Nonlinear Programming,* 2nd ed. (Addison-Wesley, Reading, MA, 1988).

[9]P. M. Morse, Phys. Rev. **34**, 57 (1929); N. Rosen, J. Chem. Phys. **1**, 319 (1933).

[10]M. Reed and B. Simon, *Methods of Modern Mathematical Physics 1: Functional Analysis* (Academic, New York, 1980).

[11]R. F. Curtain and A. J. Pritchard, *Infinite Dimensional Linear Systems Theory* (Springer-Verlag, New York, 1978).