Efficient algorithm for optimal control of mixed-state quantum systems

S. G. Schirmer

Department of Mathematics and Institute of Theoretical Science, University of Oregon, Eugene, Oregon 97403

M. D. Girardeau

Department of Physics and Institutes of Theoretical Science and Chemical Physics, University of Oregon, Eugene, Oregon 97403

J. V. Leahy

Department of Mathematics and Institute of Theoretical Science, University of Oregon, Eugene, Oregon 97403

Zhu and Rabitz [J. Chem. Phys. **109**, 385 (1998)] presented a rapidly convergent iterative algorithm for optimal control of the expectation value of a positive-definite observable in a pure-state quantum system. In this paper we generalize this algorithm to a quantum-statistical mechanics setting and show that it is both efficient in the mixed-state case and effective in achieving the control objective of maximizing the ensemble average of arbitrary observables in the cases studied.

I. INTRODUCTION

Much work has recently been done on control of pure-state quantum systems using the traditional wave-function formalism [1–3]. This work is most important; however, many physical systems, such as systems initially in thermal equilibrium or otherwise described by an ensemble of states, or systems where dissipative processes are significant, cannot be treated using this approach. Therefore, a development of optimal control for mixed-state quantum systems is necessary. In this paper we shall focus on generalizing an efficient iterative algorithm for quantum control [1] to a quantum-statistical mechanics setting used in previous work [4–8]. This work is closely related to recently published, independently developed work by Yukiyoshi, Zhu, and Rabitz [9] on quantum optimal control for systems with dissipation. However, in our work we do not consider dissipation terms since those terms are represented by non-Hermitian operators resulting in nonunitary evolution of the system. Unfortunately, the very accurate numerical implementation of the algorithm we propose depends on unitary evolution, as do the results on kinematical bounds [4] and controllability [10], which we use to show that the actual global maximum is reached by this algorithm.

II. MATHEMATICAL SETUP

As in our previous work, we consider a quantum-mechanical system whose state space $\mathcal{H}$ is a separable Hilbert space. Any mixed state of the system can be represented by a density operator $\hat{\rho}(t)$ (acting on $\mathcal{H}$) with eigenvalue decomposition

$$\hat{\rho}(t) = \sum_k w_k |\Psi_k(t)\rangle\langle\Psi_k(t)|, \qquad (1)$$

where w_k are the eigenvalues, and $|\Psi_k(t)\rangle$ the corresponding normalized eigenstates of $\hat{\rho}(t)$, which evolve in time according to the time-dependent Schrödinger equation. The eigenvalues satisfy

$$0 \le w_k \le 1, \quad \sum_k w_k = 1, \qquad (2)$$

i.e., they can be ordered in a (possibly finite) nonincreasing sequence

$$w_1 \ge w_2 \ge \cdots \ge w_k \ge \cdots \ge 0.$$

Unless otherwise mentioned, the word state will in the following refer to a mixed state represented by a density operator $\hat{\rho}(t)$.

The dynamical law for the system is given by the *quantum Liouville equation*

$$\frac{\partial}{\partial t}\hat{\rho}(t) = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}(t)], \qquad (3)$$

where $\hat{H}$ is the (total) Hamiltonian of the system and $\hat{\rho}(t_0) = \hat{\rho}_0$ defines the *initial state* of the system (at time $t = t_0$). Observables are represented by Hermitian operators $\hat{A}$ on $\mathcal{H}$ and we define their expectation value to be the *ensemble average*

$$\langle\hat{A}(t)\rangle = \mathrm{Tr}[\hat{A}\hat{\rho}(t)]. \qquad (4)$$

The set of bounded linear operators $\hat{A}$ on $\mathcal{H}$ forms itself a Hilbert space, usually called *Liouville space* and it is convenient to assign to each operator $\hat{A}$ (on $\mathcal{H}$) a Liouville ket $|A\rangle\rangle$ denoting its representation in Liouville space. The dual of $|A\rangle\rangle$ will be denoted by the Liouville bra $\langle\langle A|$. The inner product in Liouville space is defined by

$$\langle\langle A|B\rangle\rangle = \mathrm{Tr}(\hat{A}^\dagger\hat{B}). \qquad (5)$$

Thus, an arbitrary mixed state of the system is represented by a Liouville ket $|\rho(t)\rangle\rangle$ that satisfies

$$\frac{\partial}{\partial t}|\rho(t)\rangle\rangle = -\frac{i}{\hbar}\mathcal{L}(t)|\rho(t)\rangle\rangle \qquad (6)$$

with some initial condition $|\rho(t_0)\rangle\rangle = |\rho_0\rangle\rangle$. $\mathcal{L}$ is the *Liouville operator* defined by the dual correspondence

$$\mathcal{L}|\rho(t)\rangle\rangle \leftrightarrow [\hat{H}, \hat{\rho}(t)]. \qquad (7)$$

The expectation value $\langle\hat{A}(t)\rangle$ of the observable $\hat{A}$ is given by the Liouville inner product $\langle\langle A|\rho(t)\rangle\rangle$.

III. CONTROLLING THE DYNAMICS

If the number M of external *control functions*

$$\mathbf{f}(t) = (f_1(t), f_2(t), \ldots, f_M(t)) \qquad (8)$$

acting on the system is finite and the system is *control-linear* then the total Hamiltonian of the system can be decomposed as follows:

$$\hat{H} = \hat{H}_0 + \sum_{m=1}^{M} f_m(t)\hat{H}_m. \qquad (9)$$

In this case, the corresponding Liouville operator also decomposes:

$$\mathcal{L} = \mathcal{L}_0 + \sum_{m=1}^{M} f_m(t)\mathcal{L}_m. \qquad (10)$$

The restrictions imposed on the controls depend on the particular system studied. However, a reasonable minimal requirement for the control functions $f_m(t)$ is that they should be bounded, measurable, real-valued functions defined on a time interval $[t_0, t_F]$ that depends on the application.

In the remainder of this paper we shall furthermore assume that there is only one control $f(t)$ acting on the system, which is sufficient for many applications of laser control. However, we would like to point out that it is possible to generalize the algorithm to the case where there are multiple controls, such as two laser fields with perpendicular polarization driving the system.

Our goal is to maximize the expectation value (ensemble average) of a given observable, e.g., the population of a particular energy level or subspace of quantum states, the energy of a molecular bond, etc., at some fixed target time $t = t_F$ subject to certain constraints.

More precisely, we define a functional [7,6,11]

$$W(f, \rho_v, A_v) = W_1(\rho_v) - W_2(f, \rho_v, A_v) - W_3(f), \qquad (11)$$

whose value at a certain target time t_F we would like to maximize. W_1 is the expectation value of $\hat{A}$ that we wish to maximize at the target time t_F,

$$W_1(f) = \langle A(t_F)\rangle = \langle\langle A|\rho_v(t_F)\rangle\rangle; \qquad (12)$$

W_2 and W_3 are constraint functionals, which we define as follows:

$$W_2(f, \rho_v, A_v) = \int_{t_0}^{t_F}\left\langle\left\langle A_v(t)\left|\frac{\partial}{\partial t} + \frac{i}{\hbar}\mathcal{L}(t)\right|\rho_v(t)\right\rangle\right\rangle dt, \qquad (13)$$

$$W_3(f) = \frac{\lambda}{2}\int_{t_0}^{t_F} f^2(t)dt. \qquad (14)$$

W_2 ensures that the quantum Liouville equation is satisfied. W_3 constrains the fluence, i.e., the total energy of the pulse.

$\rho_v(t)$ and $A_v(t)$ are variational trial functions that must satisfy the boundary conditions

$$\rho_v(t_0) = \rho(t_0) = \rho_0, \quad A_v(t_F) = A. \qquad (15)$$

For simplicity we shall in the following choose units such that $\hbar = 1$ and define $\partial_t = \partial/\partial t$.

Equations (12)–(15) are the generalization to Liouville space of the Hilbert space formulation in [1]. The details of the connection with this paper will be discussed in Appendix A. The solution of this control problem requires finding an admissible control $f(t)$ such that W and thus $\langle\hat{A}(t)\rangle$ will attain its global maximum at time $t = t_F$.

IV. ALGORITHM

We start by guessing an initial control $f^{(0)}(t)$ and determining an initial $|\rho_v^{(0)}(t)\rangle\rangle$ by solving

$$\partial_t|\rho_v^{(0)}(t)\rangle\rangle = -i[\mathcal{L}_0 + f^{(0)}(t)\mathcal{L}_1]|\rho_v^{(0)}(t)\rangle\rangle$$

with initial condition $|\rho_v^{(0)}(t_0)\rangle\rangle = |\rho_0\rangle\rangle$. For $n \geq 1$ and $k = 0,1$ we define

$$f^{(n,k)}(t) \equiv -\frac{i}{\lambda}\langle\langle A_v^{(n)}(t)|\mathcal{L}_1|\rho_v^{(n-k)}(t)\rangle\rangle, \qquad (16)$$

$$\mathcal{L}^{(n,k)}(t) \equiv \mathcal{L}_0 + f^{(n,k)}(t)\mathcal{L}_1, \qquad (17)$$

and solve iteratively

$$\partial_t|A_v^{(n)}(t)\rangle\rangle = -i\mathcal{L}^{(n,1)}(t)|A_v^{(n)}(t)\rangle\rangle, \qquad (18)$$

$$\partial_t|\rho_v^{(n)}(t)\rangle\rangle = -i\mathcal{L}^{(n,0)}(t)|\rho_v^{(n)}(t)\rangle\rangle, \qquad (19)$$

with the boundary conditions

$$|A_v^{(n)}(t_F)\rangle\rangle = |A\rangle\rangle, \quad |\rho_0^{(n)}(t_0)\rangle\rangle = |\rho_0\rangle\rangle.$$

We observe that $f^{(n,k)}(t)$ is real. Hence $\mathcal{L}^{(n,k)}$ is Hermitian and the time evolution of both $|A_v^{(n)}(t)\rangle$ and $|\rho_v^{(n)}(t)\rangle$ is unitary, i.e.,

$$\|A_v^{(n)}(t)\|_2 = \|A\|_2, \quad \|\rho_v^{(n)}(t)\|_2 = \|\rho_0\|_2 \qquad (20)$$

for all $t \in [t_0, t_F]$ and any n. Furthermore,

$$\|\rho_0\|_2^2 = \mathrm{Tr}(\hat{\rho}_0^\dagger\hat{\rho}_0) = \mathrm{Tr}(\rho_0^2) \leq 1. \qquad (21)$$

This algorithm can be shown to converge quadratically and monotonically as does the pure-state version due to Zhu and Rabitz. The details of the proof can be found in Appendix B.

However, we have no guarantee that $W_1(f)$ indeed assumes its *global* maximum for this $f(t)$. Additional criteria, such as kinematical bounds and knowledge about controllability of the system are necessary to decide if the control the algorithm produced is indeed optimal in the sense of steering the system to a global maximum of $W_1(f)$.

V. NUMERICAL IMPLEMENTATION

The differential equations arising from this feedback algorithm must be solved numerically. While there are many methods of integrating differential equations numerically, we employ a symmetric split operator method [1,12]. The main advantage of this method is that it preserves the norm of the operators involved, which is of great importance in this problem.

We divide the time interval $[t_0, t_F]$ in subintervals $[t_j, t_{j+1}]$ of a fixed length $\Delta t = t_{j+1} - t_j$. On each subinterval $[t_j, t_{j+1}]$ we approximate $f^{(n,k)}(t)$ by the constant $f^{(n,k)}(\tau_j)$ where

$$\tau_j = t_j + \Delta t/2 = t_{j+1} - \Delta t/2. \tag{22}$$

With this approximation the propagator can be written as

$$U^{(n,k)}(t_{j+1}, t_j) = \exp[-i\Delta t(\mathcal{L}_0 + f^{(n,k)}(\tau_j)\mathcal{L}_1)]. \tag{23}$$

For arbitrary matrices A and B we have

$$e^{-i\alpha(A+B)} = e^{-i(\alpha/2)A}e^{-i\alpha B}e^{-i(\alpha/2)B}$$

up to second-order terms in A and B. Thus Eq. (23) agrees to second order with

$$e^{-i(\Delta t/2)\mathcal{L}_0}e^{-i\Delta t f^{(n,k)}(\tau_j)\mathcal{L}_1}e^{-i(\Delta t/2)\mathcal{L}_0}. \tag{24}$$

This symmetric splitting is numerically favorable since it allows us to reduce the matrix exponentials to a simple linear combination of complex exponentials:

$$U_0 \equiv \exp(-i\Delta t\mathcal{L}_0/2) = \sum_{a=1}^{N} |a\rangle\rangle e^{-ia\Delta t/2}\langle\langle a|, \tag{25}$$

$$U_1^{(n,k)}(\tau_j) \equiv \exp[-i\Delta t f^{(n,k)}(\tau_j)\mathcal{L}_1]$$
$$= \sum_{b=1}^{N} |b\rangle\rangle e^{-i\Delta t f^{(n,k)}(\tau_j)b}\langle\langle b|, \tag{26}$$

where $|a\rangle\rangle$ and $|b\rangle\rangle$ are the eigenkets of $\mathcal{L}_0$ and $\mathcal{L}_1$, respectively; a and b are the corresponding (real) eigenvalues. This leads to

$$U^{(n,k)}(\tau_j) \equiv \sum_{a,b=1}^{N} |\langle\langle a|b\rangle\rangle|^2 e^{-i\Delta t[a+bf^{(n,k)}(\tau_j)]}|a\rangle\rangle\langle\langle a|. \tag{27}$$

$U^{(n,k)}(\tau_j)$ agrees up to second order with $U^{(n,k)}(t_{j-1}, t_j)$. Since $\mathcal{L}_0$ and $\mathcal{L}_1$ do not depend on $f^{(n,k)}$, the eigenvalue decomposition needs to be done only once, i.e., the only quantities that need to be computed in each step of the itera-

tion are the complex exponentials $e^{-i\Delta t[a+bf^{(n,k)}(\tau_j)]}$ for all possible values of a and b. In order to compute $f(\tau_j)$, we note that

$$f(t \pm \Delta t) \approx f(t) \pm \Delta t \frac{df}{dt}(t) \tag{28}$$

to first order, and hence we have

$$f^{(n,0)}(\tau_j) = f^{(n,0)}(t_j) + \frac{\Delta t}{2\lambda}\langle\langle A_v^{(n)}(t_j)|[\mathcal{L}_0, \mathcal{L}_1]\rho_v^{(n)}(t_j)\rangle\rangle, \tag{29}$$

$$f^{(n,1)}(\tau_{j-1}) = f^{(n,1)}(t_j) - \frac{\Delta t}{2\lambda}\langle\langle A_v^{(n)}(t_j)|[\mathcal{L}_0, \mathcal{L}_1]\rho_v^{(n-1)}(t_j)\rangle\rangle. \tag{30}$$

VI. ILLUSTRATIVE COMPUTATIONS

As an example for molecular quantum control, we consider a Morse oscillator model for a diatomic molecule with N discrete energy levels E_n corresponding to independent vibrational eigenstates $|n\rangle$ of the system. The unperturbed Hamiltonian is thus

$$\hat{H}_0 = \sum_{n=1}^{N} E_n|n\rangle\langle n|. \tag{31}$$

The interaction Hamiltonian of the driven system can be approximated by $\hat{H}_1 = f(t)\hat{V}$, where $f(t)$ is an external laser field that serves as control function and $\hat{V}$ is the transition operator, which we choose to be of the dipole form

$$\hat{V} = \sum_{n=1}^{N-1} d_n(|n\rangle\langle n+1| + |n+1\rangle\langle n|). \tag{32}$$

This system is completely controllable, which can easily be verified using an algorithm described in [10]. Thus, the global minima and maxima of any observable are determined by the kinematical bounds, and these extrema are dynamically attainable.

For the sake of illustration we choose $N=4$. The corresponding energy levels are $E_1 = 0.4843$, $E_2 = 1.4214$, $E_3 = 2.3691$, and $E_4 = 3.2434$ in units of $\hbar\omega_0$, where $\omega_0 = 7.8 \times 10^{14}$ s^{-1} for HF.

Let us first assume that the system is initially in the ground state, i.e., $\hat{\rho}_0 = |1\rangle\langle 1|$ and that our goal is to maximize the vibrational energy of the bond, i.e., $\hat{A} = \hat{H}_0$. In this case, the results on kinematical bounds in [4] give

$$1.4214 \leq \langle \hat{A}(t)\rangle \leq 3.2434. \tag{33}$$

The lower bound is attained exactly if the population of level 1 (ground state) is 1. The upper bound is attained exactly if the population of level 4 (highest state) is 1. Figures 1–3 show the results of our computations using the algorithm described above. Starting with a randomly generated function f of sufficiently small magnitude and $\lambda = 4$, the observ-
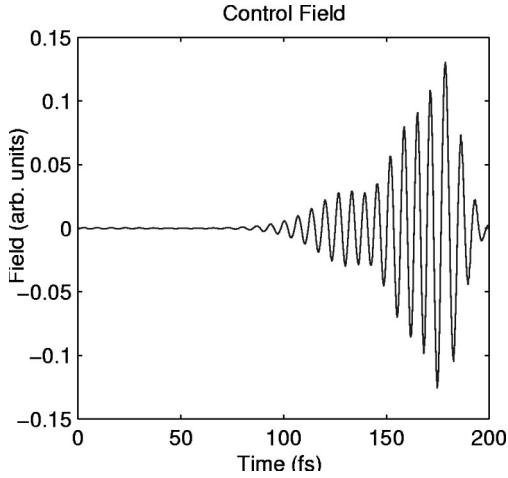
FIG. 1. Optimal pulse for a four-level Morse oscillator with $\hat{\rho}_0 = |1\rangle\langle 1|$.

FIG. 3. Evolution of the vibrational energy for a four-level Morse oscillator with $\hat{\rho}_0 = |1\rangle\langle 1|$.

able rapidly approaches its converged value within only a few iterations. Figure 1 shows the final pulse $f(t)$, Fig. 2 the corresponding evolution of the populations of energy levels 1 through 4, and Fig. 3 shows the evolution of the expectation value of the observable. At the target time $t_F = 200$ fs, we observe a nearly complete inversion of the populations, with the population of level 4 being close to 97%. $\langle \hat{A}(t_F) \rangle$ is about 98% of the theoretical maximum.

Second, we assume that the system is initially in thermal equilibrium, i.e.,

$$\hat{\rho}_0 = \sum_{n=1}^{N} w_n |n\rangle\langle n|$$

with weights

$$w_n = C \exp[-E_n/(E_4 - E_1)].$$

This is a Boltzmann distribution with $kT = E_4 - E_1$;

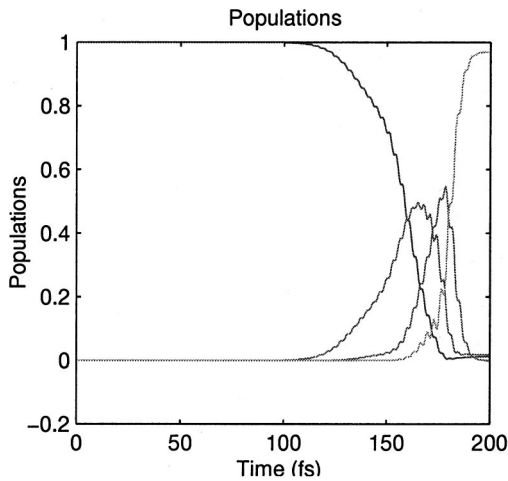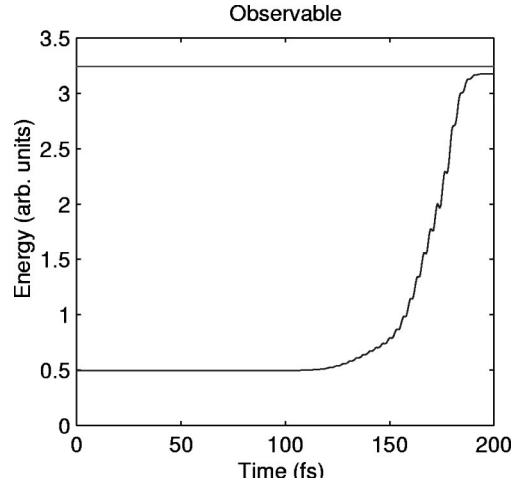$$C = (e^{-E_1/kT} + e^{-E_2/kT} + e^{-E_3/kT} + e^{-E_4/kT})^{-1}$$

is the normalization constant. Concretely, $w_1 = 0.3850$, $w_2 = 0.2758$, $w_3 = 0.1976$, and $w_4 = 0.1416$. According to [4],

$$1.5059 \leqslant \langle \hat{A} \rangle \leqslant 2.2592. \tag{34}$$

The lower bound is attained in thermal equilibrium. The upper bound is attained exactly if the populations are inverted, i.e., the most energetic state (here $n = 4$) has the highest population, the second most energetic state has the second highest population, etc. Figures 4–6 show the results of our computations using the algorithm described above. Again, we started with a randomly generated function f of sufficiently small magnitude and $\lambda = 4$. Figure 4 shows the final pulse $f(t)$, Fig. 5 the corresponding evolution of the populations of energy levels 1 through 4, and Fig. 6 shows the evolution of the expectation value of the observable. At the target time $t_F = 200$ fs we observe a nearly complete inversion of the populations with $\langle \hat{A}(t_F) \rangle$ being 99% of the theoretical maximum.
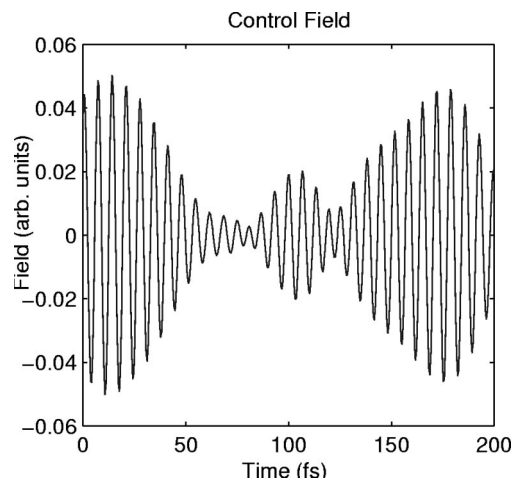

FIG. 2. Evolution of the populations for a four-level Morse oscillator with $\hat{\rho}_0 = |1\rangle\langle 1|$.

FIG. 4. Optimal pulse for a four-level Morse oscillator with $\hat{\rho}_0 = \Sigma_{n=1}^{4} |n\rangle\langle n|$.
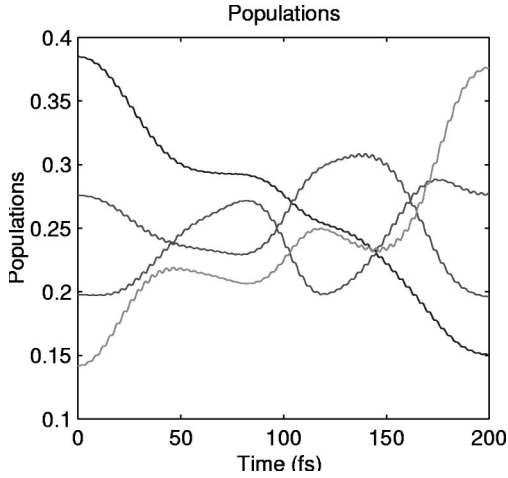
FIG. 5. Evolution of the populations for a four-level Morse oscillator with $\hat{\rho}_0 = \Sigma_{n=1}^4 |n\rangle\langle n|$.

VII. CONCLUSION

In this paper we demonstrated that an efficient algorithm for optimal control of quantum systems can be applied in a quantum-statistical mechanics setting and that this algorithm is also highly effective in realizing the control objective of maximizing the ensemble average of an observable.

APPENDIX A: RELATION TO WORK OF RABITZ *ET AL.*

Our variational functional and Euler-Lagrange equations are equivalent to the ones used in [1] in the pure-state limit, i.e., if $\hat{\rho}_v(t) = |\psi_v(t)\rangle\langle\psi_v(t)|$, where $|\psi_v(t)\rangle$ is a normalized state, then
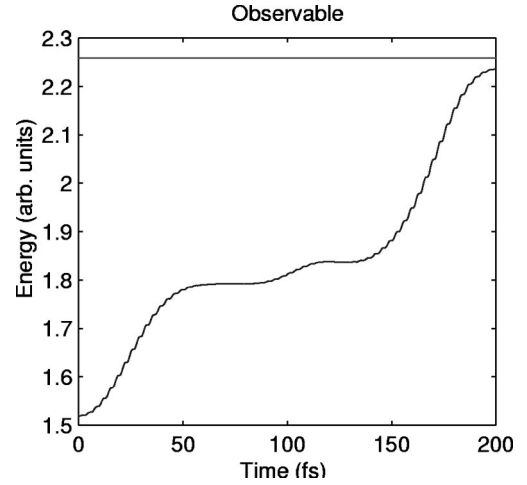

FIG. 6. Evolution of the vibrational energy for a four-level Morse oscillator with $\hat{\rho}_0 = \Sigma_{n=1}^4 |n\rangle\langle n|$.

$$W = \langle \psi_v(t_F) | A | \psi_v(t_F) \rangle - \alpha_0 \int_{t_0}^{t_F} f^2(t)\,dt$$

$$- 2\,\mathrm{Re} \int_{t_0}^{t_F} \langle \chi_v(t) | [\partial_t + i\hat{H}(f,t)] | \phi_v(t) \rangle\,dt. \quad (A1)$$

Choose a (time-dependent) complete orthonormal set $\{|n\rangle : n = 1, 2, \ldots \}$, such that, $|\psi_1(t)\rangle = |\psi_v(t)\rangle$ for all t. Then we have

$$W_1 = \mathrm{Tr}(\hat{A}\hat{\rho}_v(t_F)) = \sum_n \langle \psi_n(t_F) | \hat{A}_v(t_F) | \psi_v(t_F) \rangle$$

$$\times \langle \psi_v(t_F) | \psi_n(t_F) \rangle = \langle \psi_v(t_F) | A | \psi_v(t_F) \rangle.$$

Furthermore, setting $|\chi_v(t)\rangle = \hat{A}_v(t)|\psi_v(t)\rangle$ we obtain

$$\langle\langle A_v(t) | \partial_t \rho_v(t) \rangle\rangle = \mathrm{Tr}[\hat{A}_v(t)\partial_t\hat{\rho}_v(t)]$$

$$= \sum_n \langle \psi_n(t) | \hat{A}_v(t) [\partial_t | \psi_v(t)\rangle] \langle \psi_v(t) | \psi_n(t)\rangle + \sum_n \langle \psi_n(t) | \hat{A}_v(t) | \psi_v(t)\rangle [\partial_t \langle \psi_v(t) |] | \psi_n(t)\rangle$$

$$= \langle \psi_v(t) | \hat{A}_v(t) \partial_t | \psi_v(t)\rangle + \sum_n [\partial_t \langle \psi_v(t)|] | \psi_n(t)\rangle \langle \psi_n(t) | \hat{A}_v(t) | \psi_v(t)\rangle$$

$$= \langle \psi_v(t) | \hat{A}_v(t) \partial_t | \psi_v(t)\rangle + [\partial_t \langle \psi_v(t)|] \hat{A}_v(t) | \psi_v(t)\rangle$$

$$= \langle \psi_v(t) | \hat{A}_v(t) \partial_t | \psi_v(t)\rangle + [\langle \psi_v(t) | \hat{A}_v(t) \partial_t | \psi_v(t)\rangle]^* = 2\,\mathrm{Re}\langle \psi_v(t) | \hat{A}_v(t) \partial_t | \psi_v(t)\rangle$$

$$= 2\,\mathrm{Re}\langle \chi_v(t) | \partial_t \psi_v(t)\rangle$$

and

$$\langle\langle A_v(t) | i\mathcal{L}(f,t)\rho_v(t) \rangle\rangle = i\,\mathrm{Tr}\{\hat{A}_v(t)[\hat{H}(f,t),\hat{\rho}_v(t)]\}$$

$$= \sum_n i\langle \psi_n(t) | \hat{A}_v(t)\hat{H}(f,t) | \phi_v(t)\rangle\langle \psi_v(t) | \psi_n(t)\rangle - \sum_n i\langle \psi_n(t) | \hat{A}_v(t) | \phi_v(t)\rangle\langle \psi_v(t) | \hat{H}(f,t) | \psi_n(t)\rangle$$

$$= i\langle \psi_v(t) | \hat{A}_v(t)\hat{H}(f,t) | \phi_v(t)\rangle - \sum_n i\langle \psi_v(t) | \hat{H}(f,t) | \psi_n(t)\rangle\langle \psi_n(t) | \hat{A}_v(t) | \phi_v(t)\rangle$$

$$= \langle \chi_\nu(t) | i\hat{H}(f,t) | \phi_\nu(t) \rangle - \langle \psi_\nu(t) | i\hat{H}(f,t) | \chi_\nu(t) \rangle$$

$$= \langle \chi_\nu(t) | i\hat{H}(f,t) | \phi_\nu(t) \rangle + [\langle \chi_\nu(t) | i\hat{H}(f,t) | \phi_\nu(t) \rangle]^*$$

$$= 2 \operatorname{Re} \langle \chi_\nu(t) | i\hat{H}(f,t) | \phi_\nu(t) \rangle.$$

Hence, we have

$$W_2 = \int_{t_0}^{t_F} \langle\langle A_\nu(t) | \partial_t + i\mathcal{L}(t) | \rho_\nu(t) \rangle\rangle dt$$

$$= 2 \operatorname{Re} \int_{t_0}^{t_F} \langle \chi_\nu(t) | [\partial_t + i\hat{H}(f,t)] | \phi_\nu(t) \rangle dt$$

in the pure-state case. W_3 remains essentially the same, i.e., we simply set $\alpha_0 = \lambda/2$. The equivalence of the Euler-Lagrange equations follows.

APPENDIX B: PROOF OF CONVERGENCE PROPERTIES

After the nth iteration step, the objective functional is

$$W^{(n)} = W_1^{(n)} - W_3^{(n)} = \langle\langle A | \rho_\nu^{(n)}(t_F) \rangle\rangle - \frac{\lambda}{2} \int_{t_0}^{t_F} [f^{(n,0)}(t)]^2 dt, \tag{B1}$$

since $W_2^{(n)} = W_2(f^{(n,0)}, \rho_\nu^{(n)}, A_\nu^{(n)}) = 0$ according to Eqs. (13) and (19).

Lemma. $W^{(n)}$ is *uniformly bounded.*

Proof. Cauchy-Schwarz's inequality and Eqs. (20) and (21) give

$$|\langle\langle A | \rho_\nu^{(n)}(t) \rangle\rangle|^2 \leq \|A\|_2^2 \|\rho_\nu^{(n)}(t)\|_2^2 \leq \|A\|_2^2,$$

as well as

$$|f(t)|^2 = \left| -\frac{i}{\lambda} \langle\langle A_\nu^{(n)}(t) | \mathcal{L}_1 \rho_\nu^{(n)}(t) \rangle\rangle \right|^2$$

$$\leq \frac{1}{\lambda^2} \|A_\nu^{(n)}(t)\|_2^2 \|\mathcal{L}_1 \rho_\nu^{(n)}(t)\|_2^2$$

$$\leq \frac{1}{\lambda^2} \|A_\nu^{(n)}(t)\|_2^2 \|\mathcal{L}_1\| \|\rho_\nu^{(n)}(t)\|_2^2$$

$$\leq \frac{1}{\lambda^2} \|A_\nu^{(n)}(t)\|_2^2 \|\mathcal{L}_1\|,$$

where $\|\mathcal{L}_1\|$ is the usual operator norm. Thus,

$$|W^{(n)}| \leq |W_1^{(1)}| + |W_3^{(n)}| \leq \|A\|_2 + \frac{t_F - t_0}{2\lambda} \|A_\nu^{(n)}(t)\|_2^2 \|\mathcal{L}_1\|$$

for all n, which establishes the claim.

Lemma. If $U(t,t_0)$ satisfies

$$\partial_t U(t,t_0) = -i\mathcal{L}(t) U(t,t_0),$$

then

$$|\rho(t)\rangle\rangle = U(t,t_0) \int_{t_0}^{t} U^\dagger(t',t_0) |\phi(t')\rangle\rangle dt'$$

is a solution of

$$\partial_t |\rho(t)\rangle\rangle = -i\mathcal{L}(t) |\rho(t)\rangle\rangle + |\phi(t)\rangle\rangle.$$

Proof. Using the product rule and

$$\partial_t \int_{t_0}^{t} U^\dagger(t',t_0) |\phi(t')\rangle\rangle dt' = U^\dagger(t,t_0) |\phi(t)\rangle\rangle$$

to differentiate $|\rho(t)\rangle\rangle$ gives

$$\partial_t |\rho(t)\rangle\rangle = [-i\mathcal{L}(t) U(t,t_0)] \int_{t_0}^{t} U^\dagger(t',t_0) |\phi(t')\rangle\rangle dt'$$

$$+ U(t,t_0) U^\dagger(t,t_0) |\phi(t)\rangle\rangle$$

$$= -i\mathcal{L}(t) |\rho(t)\rangle\rangle + |\phi(t)\rangle\rangle.$$

Theorem 1. Convergence. The sequence $\{W^{(n)}\}$ converges monotonically and quadratically in the control, i.e.,

$$0 = \lim_{n \to \infty} W^{(n+1)} - W^{(n)}$$

$$= \lim_{n \to \infty} \frac{\lambda}{2} \int_{t_0}^{t_F} [\delta f^{(n+1)}(t)]^2 + [\delta f^{(n+1,n)}(t)]^2 \, dt. \tag{B2}$$

Proof. Setting

$$|\delta \rho_\nu^{(n)}(t)\rangle\rangle = |\rho_\nu^{(n+1)}(t)\rangle\rangle - |\rho_\nu^{(n)}(t)\rangle\rangle, \tag{B3}$$

$$\delta W^{(n+1,n)} = W^{(n+1)} - W^{(n)}$$

$$= \langle\langle A | \delta \rho_\nu^{(n)}(t_F) \rangle\rangle - \frac{\lambda}{2} \int_{t_0}^{t_F} [f^{(n+1,0)}(t)]^2$$

$$- [f^{(n,0)}]^2 dt. \tag{B4}$$

During the iteration

$$\partial_t |\rho_\nu^{(n)}(t)\rangle\rangle = -i[\mathcal{L}_0 + f^{(n,0)}(t)\mathcal{L}_1] |\rho_\nu^{(n)}(t)\rangle\rangle. \tag{B5}$$

Hence, setting

$$\delta f^{(n+1,n)} = f^{(n+1,1)}(t) - f^{(n,0)}(t), \tag{B6}$$

$$\delta f^{(n)} = f^{(n,0)}(t) - f^{(n,1)}(t), \qquad (B7)$$

and noting that

$$\mathcal{L}_1 f^{(n+1,1)} | \delta\rho_V^{(n)}(t) \rangle\rangle$$

$$+ \mathcal{L}_1 | \delta f^{(n+1)} \rho_V^{(n+1)}(t) + \delta f^{(n+1,n)} \rho_V^{(n)}(t) \rangle\rangle$$

$$= \mathcal{L}_1 f^{(n+1,1)} | \rho_V^{(n+1)}(t) \rangle\rangle - \mathcal{L}_1 f^{(n+1,1)} | \rho_V^{(n)}(t) \rangle\rangle$$

$$+ \mathcal{L}_1 f^{(n+1,0)} | \rho_V^{(n+1)}(t) \rangle\rangle - \mathcal{L}_1 f^{(n+1,1)} | \rho_V^{(n+1)}(t) \rangle\rangle$$

$$+ \mathcal{L}_1 f^{(n+1,1)} | \rho_V^{(n)}(t) \rangle\rangle - \mathcal{L}_1 f^{(n,0)} | \rho_V^{(n)}(t) \rangle\rangle$$

$$= \mathcal{L}_1 f^{(n+1,0)} | \rho_V^{(n+1)}(t) \rangle\rangle - \mathcal{L}_1 f^{(n,0)} | \rho_V^{(n)}(t) \rangle\rangle,$$

we obtain

$$\partial_t | \delta\rho_V^{(n)}(t) \rangle\rangle = -i\mathcal{L}^{(n+1,1)} | \delta\rho_V^{(n)}(t) \rangle\rangle$$

$$-i\mathcal{L}_1 | (\delta f^{(n+1)} \rho_V^{(n+1)} + \delta f^{(n+1,n)} \rho_V^{(n)})(t) \rangle\rangle. \qquad (B8)$$

Setting

$$U(t,t_0,f^{(n+1,1)}) = \exp_+\left[-i \int_{t_0}^t \mathcal{L}^{(n+1,1)}(\tau) d\tau \right], \quad (B9)$$

where $\exp_+$ denotes the time-ordered exponential, the formal solution of Eq. (B8) is (according to the previous lemma) given by

$$| \delta\rho_V^{(n)}(t) \rangle\rangle = -i U(t,t_0,f^{(n+1,1)})$$

$$\times \int_{t_0}^t U^\dagger(t',t_0,f^{(n+1,1)}) \mathcal{L}_1 | (\delta f^{(n+1)} \rho_V^{(n+1)}$$

$$+ \delta f^{(n+1,n)} \rho_V^{(n)})(t') \rangle\rangle dt'. \qquad (B10)$$

Observing that

$$| A_V^{(n)}(t) \rangle\rangle = U(t,t_0,f^{(n+1,1)}) U^\dagger(t_F,t_0,f^{(n+1,1)}) | A \rangle\rangle,$$

and thus

$$\langle\langle A_V^{(n)}(t) | = \langle\langle A | U(t_F,t_0,f^{(n+1,1)}) U^\dagger(t,t_0,f^{(n+1,1)}),$$

we arrive at

$$\langle\langle A | \delta\rho_V^{(n)}(t_F) \rangle\rangle = -i \int_{t_0}^{t_F} \langle\langle A | U(t_F,t_0,f^{(n+1,1)})$$

$$\times U^\dagger(t,t_0,f^{(n+1,1)}) \mathcal{L}_1$$

$$\times | (\delta f^{(n+1)} \rho_V^{(n+1)} + \delta f^{(n+1,n)} \rho_V^{(n)})(t) \rangle\rangle dt$$

$$= -i \int_{t_0}^{t_F} \langle\langle A_V^{(n+1)}(t) | \mathcal{L}_1 | (\delta f^{(n+1)} \rho_V^{(n+1)}$$

$$+ \delta f^{(n+1,n)} \rho_V^{(n)})(t) \rangle\rangle dt$$

$$= \int_{t_0}^{t_F} -i \delta f^{(n+1)}(t)$$

$$\times \langle\langle A_V^{(n+1)}(t) | \mathcal{L}_1 \rho_V^{(n+1)}(t) \rangle\rangle dt$$

$$+ \int_{t_0}^{t_F} -i \delta f^{(n+1,n)}(t)$$

$$\times \langle\langle A_V^{(n+1)}(t) | \mathcal{L}_1 \rho_V^{(n)}(t) \rangle\rangle dt$$

$$= \lambda \int_{t_0}^{t_F} \delta f^{(n+1)}(t) f^{(n+1,0)}(t)$$

$$+ \delta f^{(n+1,n)}(t) f^{(n+1,1)}(t) dt$$

$$= \lambda \int_{t_0}^{t_F} [f^{(n+1,0)}(t)]^2 - f^{(n+1,1)}(t) f^{(n+1,0)}(t)$$

$$+ [f^{(n+1,1)}(t)]^2 - f^{(n,0)}(t) f^{(n+1,1)}(t) dt, \qquad (B11)$$

$$\delta W^{(n+1,n)} = \frac{\lambda}{2} \int_{t_0}^{t_F} [f^{(n+1,0)}(t)]^2 + 2 f^{(n+1,1)}(t) f^{(n+1,0)}(t)$$

$$+ 2[f^{(n+1,1)}(t)]^2 - 2 f^{(n,0)}(t) f^{(n+1,1)}(t)$$

$$+ [f^{(n,0)}(t)]^2 dt$$

$$= \frac{\lambda}{2} \int_{t_0}^{t_F} [\delta f^{(n+1)}(t)]^2 + [\delta f^{(n+1,n)}(t)]^2 dt, \qquad (B12)$$

and thus the total variation from $n=0$ to n_F is

$$\delta W^{(n_F,0)} = W^{(n_F)} - W^{(0)} = \sum_{n=0}^{n_F-1} \delta W^{(n+1,n)}$$

$$= \sum_{n=0}^{n_F-1} \frac{\lambda}{2} \int_{t_0}^{t_F} [\delta f^{(n+1)}(t)]^2 + [\delta f^{(n+1,n)}(t)]^2 dt. \qquad (B13)$$

Since $W^{(n)}$ is uniformly bounded, $W^{(n_F)} - W^{(0)}$ is also uniformly bounded for all n_F and thus the sequence $\{ \delta W^{(n_F,0)} : n_F \in \mathbf{N}_0 \}$ is uniformly bounded:

$$\frac{\lambda}{2} \int_{t_0}^{t_F} [\delta f^{(n+1)}(t)]^2 + [\delta f^{(n+1,n)}(t)]^2 dt > 0$$

for any n implies furthermore that $\delta W^{(n_F,0)}$ is an increasing sequence. Hence,

$$\lim_{n_F \to \infty} \delta W^{(n_F,0)}$$

exists and is finite. Consequently,

$$\lim_{n \to \infty} \frac{\lambda}{2} \int_{t_0}^{t_F} [\delta f^{(n+1)}(t)]^2 + [\delta f^{(n+1,n)}(t)]^2 dt = 0.$$

[1] W. Zhu and H. Rabitz, J. Chem. Phys. **109**, 385 (1998).

[2] W. Zhu and H. Rabitz, J. Chem. Phys. **110**, 7142 (1999).

[3] W. Zhu, M. Smit, and H. Rabitz, J. Chem. Phys. **110**, 1905 (1999).

[4] M. D. Girardeau, S. G. Schirmer, J. V. Leahy, and R. M. Koch, Phys. Rev. A **58**, 2684 (1998).

[5] M. D. Girardeau, M. Ina, S. Schirmer, and T. Gulsrud, Phys. Rev. A **55**, R1565 (1997).

[6] K. G. Kim and M. D. Girardeau, Phys. Rev. A **52**, R891 (1995).

[7] P. Gaspard and P. van Ede van der Pals, in *Towards the Harnessing of Chaos*, edited by M. Yamaguti (Elsevier Science, B.V., Amsterdam, 1994), p. 205.

[8] Y. Yan *et al.*, J. Chem. Phys. **97**, 2320 (1993).

[9] Y. Ohtsuki, W. Zhu, and H. Rabitz, J. Chem. Phys. **110**, 9825 (1999).

[10] V. Ramakrisna *et al.*, Phys. Rev. A **51**, 960 (1995).

[11] R. Balian and M. Vénéroni, Phys. Rev. Lett. **47**, 1353 (1981).

[12] M. Suzuki, Phys. Lett. A **146**, 319 (1990).