# Monotonically convergent algorithms for solving quantum optimal control problems described by an integrodifferential equation of motion

Yukiyoshi Ohtsuki* and Yoshiaki Teranishi
*Department of Chemistry, Graduate School of Science, Tohoku University, Sendai 980-8578, Japan and CREST, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*

Peter Saalfrank
*Institut für Chemie, Universität Potsdam, Karl-Liebknecht-Strasse 24-25, D-14476 Potsdam-Golm, Germany*

Gabriel Turinici
*CEREMADE, Universite Paris Dauphine, Place du Marechal De Lattre De Tassigny, 75775 Paris Cedex 16, France*

Herschel Rabitz
*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*
(Received 6 January 2007; published 19 March 2007)

A family of monotonically convergent algorithms is presented for solving a wide class of quantum optimal control problems satisfying an inhomogeneous integrodifferential equation of motion. The convergence behavior is examined using a four-level model system under the influence of non-Markovian relaxation. The results show that high quality solutions can be obtained over a wide range of parameters that characterize the algorithms, independent of the presence or absence of relaxation.

DOI: 10.1103/PhysRevA.75.033407 PACS number(s): 32.80.Qk

## I. INTRODUCTION

Quantum optimal control theory provides a flexible way of designing a control to steer a dynamical system from the initial state to a desired target state [1]. The control design equations are derived by maximizing or minimizing a cost functional that measures control performance. The control design equations are nonlinear coupled equations, and the development of efficient solution algorithms is essential to numerically obtain the optimal control design [2–8].

The purpose of the present study is to extend the previous algorithms [8] for solving optimal control problems to now operate with a general inhomogeneous integrodifferential equation of motion. We adopt the non-Markovian master equation, which also connects the limiting case of no relaxation with that of Markovian relaxation. Relaxation effects on the convergence behavior are systematically examined. Through a case study, we will show the present algorithm can solve quantum optimal control problems with high quality, independent of the presence or absence of relaxation.

The systems of concern here are assumed to obey the following inhomogeneous integro-differential equation of motion:

$$\frac{\partial}{\partial t}|u(t)\rangle = [\alpha - \beta E(t)]|u(t)\rangle - \int_0^t d\tau \gamma(t,\tau)|u(\tau)\rangle + |i(t)\rangle. \tag{1}$$

With suitable interpretation, the vector $|u(t)\rangle$ specifying the system state can be a wave function, a quantum density operator, or a product state of a density operator [5,9]. The time-dependent control, $E(t)$, interacts with the system through the operator, $\beta$. The operators $\alpha$, $\beta$, and $\gamma$, which can be non-Hermitian, do not depend on the control. The inhomogeneous term $|i(t)\rangle$ appears, for example, in the master equation to capture an initial correlation between the relevant system and its environment.

The following two cost functionals describe a wide range of applications for designing controls [5,8]. For convenience, they are referred to as type I and type II, which are linear or bilinear in the system state, respectively. The type-I functional is expressed as

$$J_I = 2\,\mathrm{Re}\langle X|u(t_f)\rangle + 2\,\mathrm{Re}\int_0^{t_f} dt\langle Y(t)|u(t)\rangle - \int_0^{t_f} dt \frac{1}{A(t)}[E(t)]^2, \tag{2}$$

where Re refers to the real part. The state vector, $|X\rangle$, specifies the target state at the final time $t_f$ and $|Y(t)\rangle$ denotes the time-dependent target over the control period. The last term is a penalty over the integral of $[E(t)]^2$, in which $A(t)>0$ weighs the physical significance of the penalty. When $E(t)$ is an electric field, the penalty term is the pulse fluence. In the case of the type-II functional, the physical targets are specified by two positive, semidefinite Hermitian operators, $X$ and $Y(t)$, such that

$$J_{II} = \langle u(t_f)|X|u(t_f)\rangle + \int_0^{t_f} dt\langle u(t)|Y(t)|u(t)\rangle - \int_0^{t_f} dt \frac{1}{A(t)}[E(t)]^2. \tag{3}$$

Note that the type-II functional can be reduced to the type-I

*Electronic address: y-ohtsuki@mail.tains.tohoku.ac.jp

 

functional if we introduce a product space, in which a direct product $u_{\otimes}(t)=|u(t)\rangle\langle u(t)|$ specifies the system state. Introduction of the direct product state enables us to remove the positive constraints of the target operators; however, we have to deal with higher-dimensional dynamics that require time-consuming computation. The choice of either the type-I or type-II functional depends on target objectives and available computational resources.

An optimal control $E(t)$ is defined as one that maximizes a cost functional. The coupled control design equations are derived by applying the calculus of variations to the basic functionals, Eqs. (2) and (3), under the constraint of satisfying the equation of motion, Eq. (1). Then the optimal control is expressed as

$$E(t) = -A(t)\mathrm{Re}\langle\lambda(t)|\beta|u(t)\rangle, \tag{4}$$

where $|\lambda(t)\rangle$ is the Lagrange multiplier state that assures satisfaction of Eq. (1). In the type-I and type-II cases, the time evolution of the Lagrange multiplier is, respectively, given by

$$\frac{\partial}{\partial t}|\lambda(t)\rangle = -[\alpha^{\dagger} - \beta^{\dagger}E(t)]|\lambda(t)\rangle + \int_t^{t_f} d\tau\gamma^{\dagger}(\tau,t)|\lambda(\tau)\rangle - |Y(t)\rangle \tag{5}$$

with the final condition $|\lambda(t_f)\rangle=|X\rangle$, and

$$\frac{\partial}{\partial t}|\lambda(t)\rangle = -[\alpha^{\dagger} - \beta^{\dagger}E(t)]|\lambda(t)\rangle + \int_t^{t_f} d\tau\gamma^{\dagger}(\tau,t)|\lambda(\tau)\rangle - Y(t)|u(t)\rangle \tag{6}$$

with the final condition $|\lambda(t_f)\rangle=X|u(t_f)\rangle$. The optimal control design equations, Eqs. (1), (4), and (5) as well as Eqs. (1), (4), and (6), are derived by a local maximum condition of the cost functional. In this sense, the present algorithms are local optimization schemes.

## II. SOLUTION ALGORITHM AND ITS CONVERGENCE BEHAVIOR

The solution algorithm starts the iteration with $|u^{(0)}(t)\rangle$ using an appropriate initial trial control $E^{(0)}(t)$, and the type-I control design equations at the $k$th iteration step ($k\geqslant 1$) are

$$\frac{\partial}{\partial t}|\lambda^{(k)}(t)\rangle = -[\alpha^{\dagger} - \beta^{\dagger}\bar{E}^{(k)}(t)]|\lambda^{(k)}(t)\rangle + \int_t^{t_f} d\tau\gamma^{\dagger}(\tau,t)|\lambda^{(k)}(\tau)\rangle - |Y(t)\rangle \tag{7}$$

with the final condition $|\lambda^{(k)}(t_f)\rangle=|X\rangle$, and

$$\frac{\partial}{\partial t}|u^{(k)}(t)\rangle = [\alpha - \beta E^{(k)}(t)]|u^{(k)}(t)\rangle - \int_0^{t} d\tau\gamma(t,\tau)|u^{(k)}(\tau)\rangle + |i(t)\rangle \tag{8}$$

with the initial condition $|u^{(k)}(0)\rangle=|u_0\rangle$. The controls at the $k$th step are expressed as

$$\bar{E}^{(k)}(t) = (1-\eta_k)E^{(k-1)}(t) - \eta_k A(t)\mathrm{Re}\langle\lambda^{(k)}(t)|\beta|u^{(k-1)}(t)\rangle \tag{9}$$

and

$$E^{(k)}(t) = (1-\zeta_k)\bar{E}^{(k)}(t) - \zeta_k A\ \mathrm{Re}\langle\lambda^{(k)}(t)|\beta|u^{(k)}(t)\rangle, \tag{10}$$

where the convergence parameters, $\zeta_k$ and $\eta_k$, characterize the iteration algorithm [6,8]. As shown below, the parameters $\{\zeta_k,\eta_k\}\in[0,2]$ guarantee monotonic convergence behavior of the present algorithm. In the present algorithm, both the state vector and the Lagrange multiplier are improved at each iteration step. If we set one of the convergence parameters to zero, one of the possible means of improving the control is neglected. That is, we can set either $\bar{E}^{(k)}(t)=E^{(k-1)}(t)$ or $E^{(k)}(t)=\bar{E}^{(k)}(t)$; however, this does not lead to any significant reduction in the number of computational steps.

To reveal the convergence behavior, we consider the difference in the cost functionals between the $k$th and $(k-1)$th adjacent iteration steps:

$$\delta J_I^{(k,k-1)} = J_I^{(k)} - J_I^{(k-1)} = 2\ \mathrm{Re}\langle X|\delta u^{(k,k-1)}(t_f)\rangle + 2\ \mathrm{Re}\int_0^{t_f} dt\langle Y(t)|\delta u^{(k,k-1)}(t)\rangle - \int_0^{t_f} dt\frac{1}{A(t)}\{[E^{(k)}(t)]^2 - [E^{(k-1)}(t)]^2\}, \tag{11}$$

where $|\delta u^{(k,k-1)}(t)\rangle=|u^{(k)}(t)\rangle-|u^{(k-1)}(t)\rangle$. It is convenient to introduce the function, $P^{(k,k-1)}(t)=2\ \mathrm{Re}\langle\lambda^{(k)}(t)|\delta u^{(k,k-1)}(t)\rangle$, which satisfies $P^{(k,k-1)}(0)=0$ at $t=0$ and $P^{(k,k-1)}(t_f)=2\ \mathrm{Re}\langle X|\delta u^{(k,k-1)}(t_f)\rangle$ at $t=t_f$. Differentiating with respect to time produces

$$\begin{aligned}\frac{d}{dt}P^{(k,k-1)}(t) &+ 2\ \mathrm{Re}\langle Y(t)|\delta u^{(k,k-1)}(t)\rangle \\ &= -2\ \mathrm{Re}\langle\lambda^{(k)}(t)|\beta|u^{(k)}(t)\rangle[E^{(k)}(t)-\bar{E}^{(k)}(t)] \\ &\quad -2\ \mathrm{Re}\langle\lambda^{(k)}(t)|\beta|u^{(k-1)}(t)\rangle[\bar{E}^{(k)}(t)-E^{(k-1)}(t)] \\ &\quad +2\ \mathrm{Re}\int_t^{t_f} d\tau\langle\lambda^{(k)}(\tau)|\gamma(\tau,t)|\delta u^{(k,k-1)}(t)\rangle \\ &\quad -2\ \mathrm{Re}\int_0^{t} d\tau\langle\lambda^{(k)}(t)|\gamma(t,\tau)|\delta u^{(k,k-1)}(\tau)\rangle,\end{aligned} \tag{12}$$

where we have used the control design equations (7) and (8) at the $k$th iteration step. Substituting Eqs. (9) and (10) into Eq. (12) and integrating over $t\in[0,t_f]$ permits expressing Eq. (11) in terms of the controls

$$\delta J_I^{(k,k-1)} = \int_0^{t_f} dt\frac{1}{A(t)}\left\{\left(\frac{2}{\zeta_k}-1\right)[E^{(k)}(t)-\bar{E}^{(k)}(t)]^2 + \left(\frac{2}{\eta_k}-1\right)\times[\bar{E}^{(k)}(t)-E^{(k-1)}(t)]^2\right\}. \tag{13}$$

Summing both sides of Eq. (13) from $k=1$ to $N$, yields $\delta J_I^{(N,0)}=J_I^{(N)}-J_I^{(0)}\geqslant 0$, which establishes the monotonic con-

vergence of the cost functional for $\{\zeta_k, \eta_k\} \in [0,2]$ except for the two special cases of $\{\zeta_k = \eta_k = 0\}$ and $\{\zeta_k = \eta_k = 2\}$.

A similar algorithm can be developed for the type-II case, which again identifies monotonic convergence as long as the operators $X$ and $Y(t)$ are positive, semidefinite. The latter restriction arises from the final conditions in the type-II case. Finally, we note some possible extensions of the algorithms. The algorithms allow the equation of motion to include higher-order integral terms $\int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{n-1}} d\tau_n \gamma(t, \tau_1, \ldots, \tau_n) |u(\tau_n)\rangle\rangle$, while retaining their monotonic convergence behavior. The algorithms are also applicable to classical systems modeled as linear in the system state.

## III. NUMERICAL RESULTS

Numerical tests are implemented in a four-level model system having a time-dependent electric field, $E(t)$, coupled through a dipole interaction. In the double-space (Liouville-space) notation, the equation of motion for the reduced density operator, $|\rho(t)\rangle\rangle$, is given by

$$i\hbar \frac{\partial}{\partial t} |\rho(t)\rangle\rangle = [L_0 - ME(t)]|\rho(t)\rangle\rangle - i\hbar \int_0^t d\tau \Gamma(t,\tau)|\rho(\tau)\rangle\rangle. \tag{14}$$

The Liouvillians $L_0$ and $M$ correspond to the commutators $[H_0, \ldots]$ and $[\mu, \ldots]$, respectively, with $H_0$ and $\mu$ being the field-free system Hamiltonian and the electric dipole moment operator. The memory kernel is $\Gamma(t,\tau)$. In Eq. (14), we have neglected an inhomogeneous term assuming that the initial state of the system is uncoupled from the environment [10].

Consider the cost functional that aims at achieving the largest transition probability from the initial state $|1\rangle$ to an objective state $|\phi\rangle$ at a target time, $t_f$. The objective is specified by a Hermitian target operator, $W = |\phi\rangle\langle\phi|$, and the cost functional in type-I form is

$$J = \langle\langle W|\rho(t_f)\rangle\rangle - \int_0^{t_f} dt \frac{1}{A(t)}[E(t)]^2, \tag{15}$$

where $\langle\langle W|\rho(t_f)\rangle\rangle = \mathrm{tr}\{W^\dagger \rho(t_f)\}$.

The four-level system, $|n\rangle$ $(n=1,2,3,4)$, is characterized by the energy eigenvalues of $H_0$ as $\omega_1 = 0.0\ \mathrm{cm}^{-1}$, $\omega_2 = 69.6\ \mathrm{cm}^{-1}$, $\omega_3 = 139.2\ \mathrm{cm}^{-1}$, and $\omega_4 = 1811.0\ \mathrm{cm}^{-1}$, and by transition moment elements $\langle 4|\mu|m\rangle = \langle m|\mu|4\rangle = 1.0D$ $(m = 1,2,3)$. We assume phase relaxation processes, with nonzero matrix elements given by $\langle\langle m4|\Gamma(t,\tau)|m4\rangle\rangle = \langle\langle 4m|\Gamma(t,\tau)|4m\rangle\rangle = \frac{\Lambda}{T}\exp(-|t-\tau|/T)(m=1,2,3)$. Here $\Lambda$ and $T$ are the coupling strength and coherence time, respectively. The following simulations consider three cases with differing relaxation conditions: (a) no relaxation $(\Lambda = 0)$, (b) non-Markovian $(\Lambda = 15.0\ \mathrm{cm}^{-1}$ and $T = 20.0$ fs), and (c) Markovian $(\Lambda = 15.0\ \mathrm{cm}^{-1}$ and $T = 0)$ dynamics. The control objective is to transfer the initial state population in $|1\rangle$ to the superposition state $|\phi\rangle = \frac{1}{\sqrt{2}}[|2\rangle + |3\rangle]$. The amplitude weight is $A(t) = 5 \times 10^5$, and the final time is $t_f = 1000$ fs. For simplicity, the convergence parameters are taken as constant over the iteration, i.e., $\zeta_k = \zeta$ and $\eta_k = \eta$.

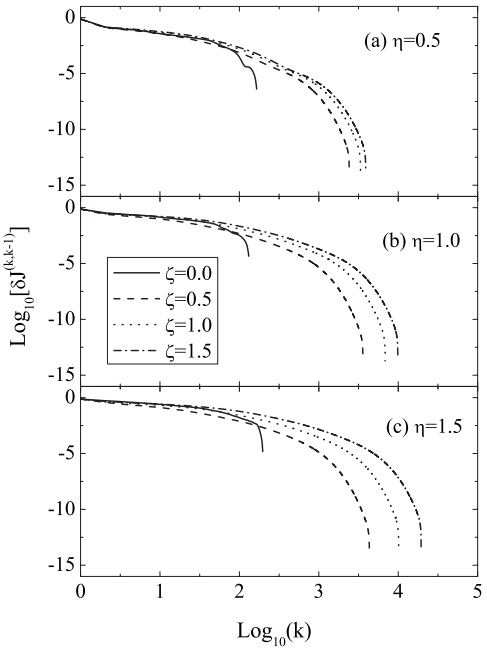


FIG. 1. Convergence behavior for several values of the convergence parameters, $\zeta$ and $\eta$, in the non-Markovian case.

Figure 1 shows the behavior of $\log_{10}(\delta J^{(k,k-1)})$ with respect to $\log_{10} k$ for several sets of convergence parameters $\{\zeta, \eta\}$ in the non-Markovian case. Here, the convergence limit is defined by the value of the cost functional at the final indicated iterative step, beyond which further monotonic convergence breaks down due to numerical limitations. More detailed discussions about the convergence limit will be given later.

The results in Fig. 1 confirm the monotonic convergence behavior up to high numerical accuracy, except when employing a value of $\zeta = 0$, i.e., the Krotov algorithm [12,13]. We have checked that an increase in the number of time grid points did not significantly change the convergence behavior [11]. When $\zeta \neq 0$, the converged values of the cost functionals (0.74518) are virtually independent of the choice of the convergence parameters, the difference among values being less than $5.3 \times 10^{-4}\%$. In these examples, smaller values of $\eta$ accelerate the convergence. As the present cost functional has a value of $\sim 0.75$, the difference in the finally attained cost functional, $\delta J^{(k,k-1)} < 10^{-16}$, cannot be evaluated due to the limited precision of the computer's floating-point capabilities. When $\zeta \neq 0$, the final values of $\delta J^{(k,k-1)}$ are $10^{-14} - 10^{-12}$. When $\zeta = 0$, the final values of $\delta J^{(k,k-1)}$ are $10^{-7} - 10^{-5}$, which is far from the computer's precision limit. Employment of the value of $\zeta = 0$ neglects either the improvement of the state vector or that of the Lagrange multiplier at each iteration step, which appears to explain the poor behavior in this case. Since the $\zeta = 0$ case leads to a larger final value of $\delta J^{(k,k-1)}$, it needed the smallest number of iteration steps to reach its own low quality convergence limit shown in Fig. 1.

Figure 2 shows the effects of relaxation when the convergence parameters are chosen to be (a) $(\zeta = 1.0,\ \eta = 1.0)$ and (b) $(\zeta = 1.0,\ \eta = 0.5)$. The solutions are characterized by small

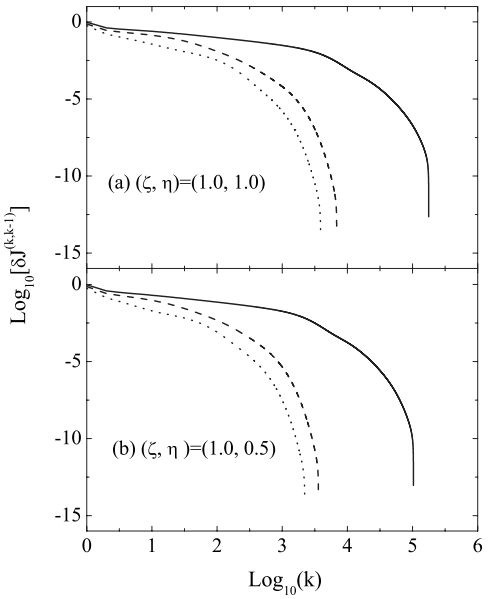


FIG. 2. Relaxation effects on the convergence behavior. The solid, dashed, and dotted line show the results for the cases of no relaxation, non-Markovian, and Markovian dynamics, respectively.

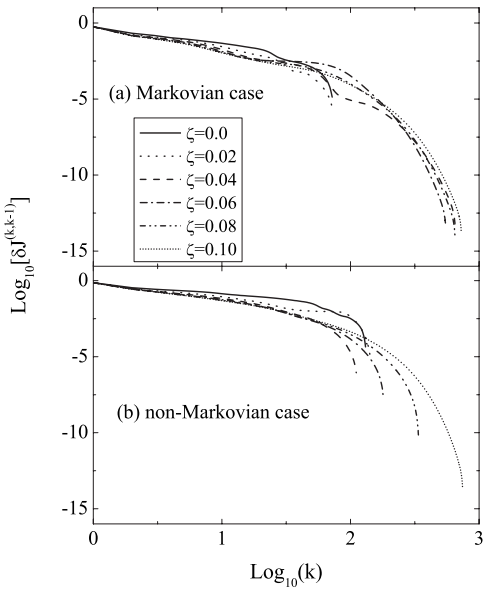


FIG. 3. Convergence behavior for small values of $\zeta$ (fixed value of $\eta=1.0$) in the (a) Markovian and (b) non-Markovian cases.

final converged intervals $\delta J^{(k,k-1)} \leqslant 10^{-12}$, indicating that high numerical accuracy is achieved independent of the presence or absence of relaxation. The presence of relaxation leads to rapid convergence without loss of numerical accuracy.

Figure 3 shows the convergence behavior for several values of $\zeta$ with a fixed value of $\eta=1.0$ to clarify the parameter range that yields high numerical accuracy. In the examples, the value of $\zeta=0.1$ produces high numerical accuracy even in the non-Markovian case. These results together with those in Fig. 1 suggest that the present iteration scheme can achieve high numerical accuracy over a wide range of convergence parameters. To further analyze this numerical observation, we rewrite Eq. (13) using Eqs. (9) and (10):

$$\delta J_I^{(k,k-1)} = \int_0^{t_f} dt \frac{1}{A(t)} \{[\bar{E}^{(k)}(t) - E^{(k-1)}(t)]^2 + \zeta(2-\zeta) \times [A(t)\mathrm{Re}\langle \lambda^{(k)}(t)|\beta|\delta u^{(k,k-1)}(t)\rangle]^2\}, \quad (16)$$

where $\zeta_k=\zeta$ and $\eta_k=1.0$. As $\zeta \to 0$, the first term in the integrand reduces to $[E^{(k)}(t)-E^{(k-1)}(t)]^2$, which is the integrand in the case of $\zeta=0$, and the second term goes to zero. Thus the second term gives a measure of how close the parameter, $\zeta$, is to the range in which high numerical accuracy is realized. For a small value of $\zeta$, the second term varies linearly with $\zeta$ through the prefactor $\zeta(2-\zeta) \sim 2\zeta$, which can explain the rapid recovery of numerical accuracy with an increase in the value of $\zeta$.

## IV. SUMMARY

We have presented a family of monotonically convergent algorithms for solving a wide class of quantum optimal control problems under dynamics described by an inhomogeneous integrodifferential equation of motion. The algorithm family is characterized by two convergence parameters, $\zeta, \eta \in [0,2]$. Numerical tests showed that high numerical accuracy can be realized over a wide range of parameter values, independent of the presence or absence of relaxation. Extremely precise solutions often require long computational time; however, the convergence behavior can be monitored by means of extrapolation [8]. If slow convergence is detected, the calculation can be restarted using more suitable convergence parameters.

## ACKNOWLEDGMENTS

One of the authors (Y.O.) acknowledges stimulating discussions with Professor Takamasa Momose and Professor Kenji Ohmori. This work was supported in part by a Grant-in-Aid from MEXT of Japan (Grant No. 17550005). P.S. acknowledges support by the Deutsche Forschungsgemeinschaft through Grant No. SFB 450. G.T. acknowledges support from INRIA Rocquencourt and GIP-ANR Grant No. C-QUID. H.R. acknowledges support form the NSF.

[1] R. J. Levis, G. M. Menkir, and H. Rabitz, Science **292**, 709 (2001).
[2] W. Zhu, J. Botina, and H. Rabitz, J. Chem. Phys. **108**, 1953 (1998).
[3] W. Zhu and H. Rabitz, J. Chem. Phys. **109**, 385 (1998).
[4] Y. Ohtsuki, W. Zhu, and H. Rabitz, J. Chem. Phys. **110**, 9825 (1999).
[5] Y. Ohtsuki and H. Rabitz, CRM Proc. Lecture Notes **33**, 163

(2003).
[6] Y. Maday and G. Turinici, J. Chem. Phys. **118**, 8191 (2003).
[7] Y. Ohtsuki, J. Chem. Phys. **119**, 661 (2003).
[8] Y. Ohtsuki, G. Turinici, and H. Rabitz, J. Chem. Phys. **120**, 5509 (2004).
[9] Y. Ohtsuki, K. Nakagami, Y. Fujimura, W. Zhu, and H. Rabitz, J. Chem. Phys. **114**, 8867 (2001).
[10] K. Blum, *Density Matrix Theory and Applications* (Plenum Press, New York, 1981).
[11] Y. Maday, J. Salomon, and G. Turinici, Numer. Math. **103**, 323 (2006).
[12] J. Somlói, V. A. Kazakov, and D. J. Tannor, Chem. Phys. **172**, 85 (1993).
[13] A. Bartana, R. Kosloff, and D. J. Tannor, J. Chem. Phys. **106**, 1435 (1997).