Optimal control theory for unitary transformations

José P. Palao[1,2] and Ronnie Kosloff [1]
[1]*Department of Physical Chemistry and the Fritz Haber Research Center for Molecular Dynamics, Hebrew University, Jerusalem 91904, Israel*
[2]*Departamento de Física Fundamental II, Universidad de La Laguna, La Laguna 38204, Spain*
(Received 31 August 2003; published 9 December 2003)

The dynamics of a quantum system driven by an external field is well described by a unitary transformation generated by a time-dependent Hamiltonian. The inverse problem of finding the field that generates a specific unitary transformation is the subject of study. The unitary transformation which can represent an algorithm in a quantum computation is imposed on a subset of quantum states embedded in a larger Hilbert space. Optimal control theory is used to solve the inversion problem irrespective of the initial input state. A unified formalism based on the Krotov method is developed leading to a different scheme. The schemes are compared for the inversion of a two-qubit Fourier transform using as registers the vibrational levels of the $X\ ^1\Sigma_g^+$ electronic state of Na_2. Raman-like transitions through the $A\ ^1\Sigma_u^+$ electronic state induce the transitions. Light fields are found that are able to implement the Fourier transform within a picosecond time scale. Such fields can be obtained by pulse-shaping techniques of a femtosecond pulse. Of the schemes studied, the square modulus scheme converges fastest. A study of the implementation of the Q qubit Fourier transform in the Na_2 molecule was carried out for up to five qubits. The classical computation effort required to obtain the algorithm with a given fidelity is estimated to scale exponentially with the number of levels. The observed moderate scaling of the pulse intensity with the number of qubits in the transformation is rationalized.

DOI: 10.1103/PhysRevA.68.062308 PACS number(s): 03.67.Lx, 82.53.Kp, 33.90.+h, 32.80.Qk

I. INTRODUCTION

Coherent control was initiated to steer a quantum system to a final objective via an external field [1,2]. If the initial and final objective states are pure, the method can be termed state-to-state coherent control. By generalizing, the problem of steering simultaneously a *set* of initial pure states to a set of final states can be formulated. Such a possibility has direct applicability in quantum computing where an algorithm implemented as a unitary transformation operating on a set of states has to be carried out irrespective of the input. In this application both input and output are encoded as a superposition of these states [3].

To implement such a control, the external driving field that induces a prespecified unitary transformation has to be found. Different methods have been suggested for this task. Some rely on factorizing the algorithm encoded as a unitary transformation, to a set of elementary gates and then finding a control solution for the elementary unitary evolution of a single gate [3,4]. The inherent difficulty in such an approach is that in general the field addresses many levels simultaneously. Therefore, when a particular single-gate operation is carried out other levels are affected. This means that the ideal single-gate unitary transformation has to be implemented so that all other possible transitions are avoided. The problem is simpler when each allowed transition is selectively addressable [5]. However, in general the problem of undesired coupling has to be corrected. A specific solution has been suggested [6] but a general solution is not known.

The presence of a large number of levels coupled to the external driving field is especially relevant in the implementation of quantum computing in molecular systems [7–9]. The use of optimal control theory (OCT) has been proposed as a possible solution [9,10]. In recent years OCT for quantum systems [1] has received considerable attention, leading to effective methods for obtaining the driving field which will induce a desired transition between preselected initial and final states. To address the control problem of inducing a particular unitary transformation the state-to-state OCT has to be augmented. For example, if the unitary transformation is to relate the initial states $|\varphi_{ik}\rangle$ with the final states $|\varphi_{fk}\rangle$, the state-to-state OCT derives an optimal field ϵ_k for each pair of initial and final states $(\varphi_{ik},\varphi_{fk})$. But the fields ϵ_k obtained are in general different so that the evolution induced by ϵ_k is not appropriate for a different set k' of initial and final states. In order to implement a given unitary transformation a single field ϵ that relates simultaneously to all the relevant pairs $(\varphi_{ik},\varphi_{fk})$ is needed.

Two approaches have been suggested to generalize OCT for unitary transformations. The first approach is formulated directly on the evolution operator [11]. An alternative approach uses the simultaneous optimization of several state-to-state transitions [10,12]. The present paper develops a comprehensive framework for constructing an OCT solution for the unitary transformation. The study explores various approaches. A common framework for an iterative solution based on the Krotov approach [13] is developed. As a result, the numerical implementation of the methods are almost identical, enabling an unbiased assessment. The implementation of the Fourier-transform algorithm in a molecular environment is chosen as a case study. The performance of the various OCT schemes is compared in a realistic setup. A crucial demand in quantum computing is obtaining high fidelity of the solution. The present OCT scheme can be viewed as an iterative classical algorithm which finds a field that implements the quantum algorithm. The obvious questions are the following.

(1) What are the computational resources required to obtain a high fidelity result?

(2) How do these computational resources scale with the number of qubits in the quantum algorithm?

(3) How do the actual physical resources, i.e., the integrated power of the field scale with the number of qubits in the quantum algorithm?

The paper is organized as follows. In Sec. II the problem is formulated, introducing different objectives devoted to the optimization of a given unitary transformation. In Sec. III the application of the Krotov method of optimization of the given objectives is described. Expressions obtained for the optimal field are formulated as well as the implementation of the method. The variational method to derive the optimization equations is commented on in Appendix. The results are used to study the implementation of a unitary transformation in a molecular model in Sec. IV. Finally, in Sec. V the results are discussed.

II. IMPLEMENTATION OF A UNITARY TRANSFORMATION

A. Description of the problem

The objective of the study is to devise a method to find the driving field that executes a unitary transformation on a subsystem embedded in a larger Hilbert space. The unitary transformation is to be applied in a Hilbert space $\mathcal{M}$ of dimension M, expanded by an orthonormal basis of states $\{|m\rangle\}$ $(m=1,\ldots,M)$. The selected unitary transformation is imposed on the subspace $\mathcal{N}$ of N levels of the system ($N \leqslant M$). In the context of quantum computation, the N levels correspond to the physical implementation of the qubit(s) embedded in a larger system. The additional levels $(m=N+1,\ldots,M)$, considered as ''spurious levels,'' are only indirectly involved in the target unitary transformation.

In any realistic implementation of quantum computing, spurious levels always exist. One reason is that the system is never completely isolated from the environment. In addition, the control lever, which in the present case is the dipole operator, connects directly only part of the primary levels. An example is the implementation of quantum computation using rovibronic molecular levels. The transition dipole connects two electronic surfaces [11]. The primary states reside on one surface, so that Raman-like transitions are used to implement the unitary transformation. The advantage of this setup is that the transition frequencies between the electronic surfaces are in the visible region, for which the pulse shaping technology is well developed [14]. Other levels residing on both of the electronic surfaces become spurious in the sense that any leakage to them destroys the desired final result. However at intermediate times these levels constitute a temporary storage space which facilitates the execution of the transformation.

The objective is to implement a selected unitary transformation in the relevant subspace $\mathcal{N}$ at a given final time T. The target unitary transformation is represented by an operator in the Hilbert space of the primary system and is denoted by $\hat{\mathbf{O}}$. For $N<M$, the matrix representation of $\hat{\mathbf{O}}$ in the basis $\{|m\rangle\}$ has two blocks of dimension $N\times N$ and $(M-N)\times(M-N)$. The operator $\hat{\mathbf{O}}$ is block diagonal, therefore the elements connecting the diagonal blocks are zero. This structure means that population is not transferred between the two subspaces at the target time T, but can take place at intermediate times. Only the $N\times N$ block is relevant for the optimization procedure, while the other remains arbitrary.

The dynamics of the system is generated by the Hamiltonian $\hat{\mathbf{H}}$,

$$\hat{\mathbf{H}}(\boldsymbol{\epsilon})=\hat{\mathbf{H}}_0-\hat{\boldsymbol{\mu}}\,\boldsymbol{\epsilon}(t), \tag{1}$$

where $\hat{\mathbf{H}}_0$ is the free Hamiltonian, $\boldsymbol{\epsilon}(t)$ is the driving field, and $\hat{\boldsymbol{\mu}}$ is a system operator describing the coupling between system and field. In the molecular systems, this coupling corresponds to the transition dipole operator and the driving field becomes radiation. In some cases more than one independent driving field can be considered. An example is when two components of the polarization of an electromagnetic field are separately controlled [15]. The generalization of the formalism in such a case is straightforward.

The system dynamics is fully specified by the evolution operator $\hat{\mathbf{U}}(t,0;\boldsymbol{\epsilon})$. An optimal field $\boldsymbol{\epsilon}_{opt}$ induces the target unitary transformation $\hat{\mathbf{O}}$ at time T when

$$\hat{\mathbf{U}}(T,0;\boldsymbol{\epsilon}_{opt})=e^{-i\phi(T)}\,\hat{\mathbf{O}}. \tag{2}$$

Equation (2) implies a condition on only the $N\times N$ block of the matrix representation of $\hat{\mathbf{U}}$. The phase $\phi(T)$ is introduced to point out that the target unitary transformation $\hat{\mathbf{O}}$ can be implemented only up to an arbitrary global phase. The phase ϕ can be decomposed into two terms, $\phi_1(T)+\phi_2(T)$. The first, ϕ_1, originates from the arbitrary choice of the origin of the energy levels. Formally, a term proportional to the identity operator can always be added to the Hamiltonian. When the states $|m\rangle$ correspond to the eigenstates of $\hat{\mathbf{H}}_0$, the phase ϕ_1 is

$$\phi_1(T)=\frac{\sum_{m=1}^{M} E_m\,T}{M\,\hbar}, \tag{3}$$

where E_m is the energy of the level m. The phase ϕ_2 reflects the arbitrariness of the unitary transformation for the levels $m=N+1,\ldots,M$ which are not part of the target.

The method to determine the optimal field is based on maximizing a real functional of the field that depends on both the target unitary transformation and the evolution generated by the Hamiltonian, fulfilling Eq. (2). The problem of unitary transformation optimization is then reduced to a functional optimization. However a variety of formulations of the problem can be stated, leading to different functionals and then, in principle, to different results. In the present context two formulations have been proposed, one based on the evolution operator [11] and the other on simultaneous N

state-to-state transitions [12]. These formulations are closely related. A similar optimization procedure is described in Sec. III.

B. Evolution operator formulation

The optimization formulation is based on the definition of a complex functional τ that depends on the evolution operator at time T [11]. The following functional is introduced:

$$\tau(\hat{\mathbf{O}};T;\epsilon)=\mathrm{Tr}\{\hat{\mathbf{O}}^{\dagger}\hat{\mathbf{U}}(T,0;\epsilon)\hat{\mathbf{P}}_N\}=\sum_{n=1}^{N}\langle n|\hat{\mathbf{O}}^{\dagger}\hat{\mathbf{U}}(T,0;\epsilon)|n\rangle, \tag{4}$$

where the projection $\hat{\mathbf{P}}_N=\Sigma_{n=1}^{N}|n\rangle\langle n|$ is used. $\{|n\rangle\}$ denotes an orthonormal basis of the subspace $\mathcal{N}$. As $\hat{\mathbf{O}}$ is a unitary transformation in the relevant subspace, the functional τ is a complex number restricted to the interior of a circle in the complex plane of radius N centered at the origin. The modulus of τ is equal to N only for an optimal field fulfilling Eq. (2). τ can then be interpreted as an indicator of the fidelity of the implementation on the target unitary transformation [11]. When τ approaches N, the transformation imposed by the field converges to the target objective.

Since τ is complex, several different real functionals can be associated with it. In Ref. [11] the optimization of the real part of τ, or the imaginary part, or a linear combination of both was suggested to find the optimal field. It was found that all these possibilities show a similar performance. For this reason, the present paper employs the optimization of the real part chosen as a representative case. The functional is therefore defined as

$$F_{re}=-\mathrm{Re}[\,\tau(\hat{\mathbf{O}};T;\epsilon)]=-\mathrm{Re}\left[\sum_{n=1}^{N}\langle n|\hat{\mathbf{O}}^{\dagger}\hat{\mathbf{U}}(T,0;\epsilon)|n\rangle\right]. \tag{5}$$

The functional reaches its minimum value, $F_{re}=-N$, when the driving field induces the target unitary transformation but with the additional condition that the phase term $\exp[-i\phi(T)]$ is equal to one.

Other functionals based on τ but without any condition on the phase can be defined. In this work the squared modulus of τ with a negative sign is studied:

$$\begin{aligned}F_{sm}&=-|\tau(\hat{\mathbf{O}};T;\epsilon)|^2\\&=-\sum_{n=1}^{N}\sum_{n'=1}^{N}\langle n|\hat{\mathbf{O}}^{\dagger}\hat{\mathbf{U}}(T,0;\epsilon)|n\rangle\langle n'|\hat{\mathbf{U}}(T,0;\epsilon)^{\dagger}\hat{\mathbf{O}}|n'\rangle,\end{aligned} \tag{6}$$

with minimum value $F_{sm}=-N^2$ for any field satisfying Eq. (2).

C. Formulation of the simultaneous N state-to-state transitions

This formulation is based on the simultaneous optimization of N transitions between a set of initial states $|l\rangle$ and the corresponding final states $\hat{\mathbf{O}}|l\rangle$ $(l=1,\ldots,N)$ [12]. For this purpose the following functional is defined:

$$\begin{aligned}\eta(\hat{\mathbf{O}};T;\epsilon)&=\mathrm{Tr}\left\{\sum_{l=1}^{N}\hat{\mathbf{O}}^{\dagger}\hat{\mathbf{U}}(T,0;\epsilon)|l\rangle\langle l|\hat{\mathbf{U}}(T,0;\epsilon)^{\dagger}\hat{\mathbf{O}}|l\rangle\langle l|\right\}\\&=\sum_{l=1}^{N}\langle l|\hat{\mathbf{O}}^{\dagger}\hat{\mathbf{U}}(T,0;\epsilon)|l\rangle|^2.\end{aligned} \tag{7}$$

Note that while τ is defined as the sum of amplitudes, η is defined as the sum of overlaps at the final time T. The parameter η is a positive real number and its maximum value N is reached when all the initial states $|l\rangle$ are driven by the field to the final target states $\hat{\mathbf{O}}|l\rangle$, except for a possible arbitrary phase associated with each transition. The arbitrariness of these phases implies that the set of initial states $|l\rangle$ must be chosen carefully. In order to account for all the possible transitions, the states $|l\rangle$ have to faithfully represent all the relevant subspace $\mathcal{N}$, i.e., constitute a complete basis set. However, a choice of an orthonormal basis could produce undesired results. For example, the ambiguity of using an orthonormal basis $\{|n\rangle\}$ in the relevant subspace and an arbitrary unitary transformation $\hat{\mathbf{D}}$, diagonal in that basis. The product $\hat{\mathbf{O}}\,\hat{\mathbf{D}}$ will also be a unitary transformation. If ϵ_O and ϵ_{OD} are fields that generate $\hat{\mathbf{O}}$ and $\hat{\mathbf{O}}\,\hat{\mathbf{D}}$ at time T, respectively, they both will have the same fidelity η,

$$\eta_{\perp}(\hat{\mathbf{O}};T;\epsilon_O)=\eta_{\perp}(\hat{\mathbf{O}};T;\epsilon_{OD}), \tag{8}$$

where $\perp$ denotes that η was evaluated using an orthonormal basis. Then any algorithm based on optimizing η that uses an orthonormal basis could find a solution for the field corresponding to the implementation of an arbitrary unitary transformations of the form $\hat{\mathbf{O}}\,\hat{\mathbf{D}}$. ($\hat{\mathbf{O}}$ is a particular case when $\hat{\mathbf{D}}$ is the identity operator). The reason for this discrepancy is that η is only sensitive to the overlap of each pair of initial $|l\rangle$ and final $\hat{\mathbf{O}}|l\rangle$ states, leaving undetermined the relative phases between them. For the optimization procedure to succeed a careful choice of the initial set of states is necessary. A simple solution is to compose the last N states as a superposition of all states in the basis $\Sigma_{l=1}^{N}|n\rangle/\sqrt{N}$, and to keep as is the first $N-1$ states of an orthonormal basis. For this set of states, the maximum condition is achieved only when the field induces the target unitary transformation up to a possible global phase.

To summarize the functional $-\eta$ is used,

$$F_{ss}=-\eta(\hat{\mathbf{O}};T;\epsilon), \tag{9}$$

with a minimum value $F_{ss}=-N$. The optimal field satisfies Eq. (2), subject to a choice of the set of states $|l\rangle$ which determines the relative phases.

D. Initial- to final-state optimization

The present formulations of quantum control assume that the target unitary transformation $\hat{\mathbf{O}}$ is explicitly known, at

least in the subspace $\mathcal{N}$. In most previous applications of optimal control theory, the objective was specified as the maximization of the expectation value of a given observable at time T subject to a predefined initial state [1]. Both mixed and pure initial states were considered [16,17]. A particular case is the determination of an optimal field to drive the system from a given pure initial state $|\varphi_i\rangle$ to a target pure final state $|\varphi_f\rangle$ at time T. This state-to-state objective optimization can be derived from the present formulation if the target unitary transformation becomes $\hat{\mathbf{O}}|\varphi_i\rangle=|\varphi_f\rangle$. The evolution operator formulation is then obtained by setting the projector $\hat{\mathbf{P}}_N=|\varphi_i\rangle\langle\varphi_i|$ in Eq. (4), obtaining the functional τ,

$$\tau(\varphi_i,\varphi_f;T;\boldsymbol{\epsilon})=\langle\varphi_f|\hat{\mathbf{U}}(T,0;\boldsymbol{\epsilon})|\varphi_i\rangle. \tag{10}$$

The real functionals $F_{re}=-\mathrm{Re}[\tau]$ and $F_{sm}=-|\tau|^2$ are to be used in the study. As only a state-to-state transition is involved, the formulation is obtained by choosing $|l\rangle\equiv|\varphi_i\rangle$. In this case $\eta=|\tau|^2$ and $F_{ss}=F_{sm}$. Note that this result is valid only when there is a single term in the sum in Eqs. (4) and in (7).

III. OPTIMIZATION

A common optimization procedure for all the functionals F as defined in the preceding section is developed. The notation $|n\rangle$ for the states and n, its index, will be used in the evolution operator formulation. The notation $|l\rangle$ and l will be used in the simultaneous N state-to-state transitions formulation. The notation $|\varphi_{ik}\rangle$ and k, where $k=1,\ldots,N$, will be used when the results are valid for both cases. An evaluation of any of the functionals requires the knowledge of the states $|\varphi_k(T)\rangle=\hat{\mathbf{U}}(T,0;\boldsymbol{\epsilon})|\varphi_{ik}\rangle$ and $|\varphi_{fk}\rangle=\hat{\mathbf{O}}|\varphi_{ik}\rangle$. The operation of the evolution equation $\hat{\mathbf{U}}(t,0;\boldsymbol{\epsilon})|\varphi_{ik}\rangle$ can be calculated by solving the time-dependent Schrödinger equation

$$\frac{d}{dt}|\varphi_k(t)\rangle=-\frac{i}{\hbar}\hat{\mathbf{H}}(\boldsymbol{\epsilon})|\varphi_k(t)\rangle, \tag{11}$$

with an initial condition $|\varphi_k(0)\rangle=|\varphi_{ik}\rangle$. Since $\hat{\mathbf{H}}=\hat{\mathbf{H}}(\boldsymbol{\epsilon})$ the state evolution will depend on the particular field. An alternative to Eq. (11) is the evolution equation for the unitary transformation itself [11].

The method of optimization depends on the availability of the states of the system $|\varphi_k(t)\rangle$ at intermediate times $0<t<T$.

Experimental realizations of OCT are typical examples where only initial and final knowledge of the states exist. For such cases feedback control and evolutionary methods are effective [1]. Such methods however require a large number of iterations to achieve convergence. A simulation of such a process requires repeated propagation of the N states by the Schrödinger equation. Thus they are computationally intensive.

Computationally more effective methods are based on the knowledge of the states $|\varphi_k(t)\rangle$ at intermediate times. Additional constrains on the evolution are included that allow a modification of the field at intermediate times consistent with the improvement of the objective at the target time T. Some examples are the local-in-time optimization method [18,19], the conjugate gradient search method [20], the Krotov method [13], and the variational approach [21,22]. A review of these common methods can be found in Ref. [1]. In the present study the Krotov method has been adopted. A brief description of the alternative variational method is given in Appendix.

A. Krotov method of optimization

The Krotov method is utilized to derive an iterative algorithm to minimize a given functional that depends on both final and intermediate times, cf. Ref. [23].

For convenience, the equations are stated using real functions: $\alpha_{km}(t)\equiv\mathrm{Re}[\langle m|\varphi_k(t)\rangle]$ and $\beta_{km}(t)\equiv\mathrm{Im}[\langle m|\varphi_k(t)\rangle]$. The notation $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ is used to describe the M-dimensional vectors with components α_{km} and β_{km}. Using such a notation, the evolution equation (11) becomes

$$\frac{d}{dt}\boldsymbol{\alpha}_k(t)=\boldsymbol{\Omega}_R(t,\boldsymbol{\epsilon})\cdot\boldsymbol{\alpha}_k(t)-\boldsymbol{\Omega}_I(t,\boldsymbol{\epsilon})\cdot\boldsymbol{\beta}_k(t),$$

$$\frac{d}{dt}\boldsymbol{\beta}_k(t)=\boldsymbol{\Omega}_I(t,\boldsymbol{\epsilon})\cdot\boldsymbol{\alpha}_k(t)+\boldsymbol{\Omega}_R(t,\boldsymbol{\epsilon})\cdot\boldsymbol{\beta}_k(t), \tag{12}$$

where $\boldsymbol{\Omega}_R$ and $\boldsymbol{\Omega}_I$ are real matrices with the corresponding components composed of the real and imaginary parts of $\boldsymbol{\Omega}_{ij}=\langle i|(-i\hat{\mathbf{H}}/\hbar)|j\rangle$ where $|i\rangle$ and $|j\rangle$ are states from the basis set $\{|m\rangle\}$. The initial conditions are given by the vectors $\boldsymbol{\alpha}_k(0)$ and $\boldsymbol{\beta}_k(0)$ with components composed of the real and imaginary part of the amplitudes $\langle m|\varphi_{ik}\rangle$. $\boldsymbol{\alpha}_{fk}$ and $\boldsymbol{\beta}_{fk}$ denote the vectors corresponding to the amplitudes $\langle m|\varphi_{fk}\rangle$.

The formalism considers $\boldsymbol{\alpha}_k$, $\boldsymbol{\beta}_k$, and the field $\boldsymbol{\epsilon}$ to be independent variables. A necessary consistency between them will be required in the final step of the algorithm. The vectors $\mathbf{f}_\alpha$ and $\mathbf{f}_\beta$ constitute the right-hand side of Eq. (12):

$$\mathbf{f}_\alpha(t,\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k,\boldsymbol{\epsilon})\equiv\boldsymbol{\Omega}_R(t,\boldsymbol{\epsilon})\cdot\boldsymbol{\alpha}_k(t)-\boldsymbol{\Omega}_I(t,\boldsymbol{\epsilon})\cdot\boldsymbol{\beta}_k(t),$$

$$\mathbf{f}_\beta(t,\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k,\boldsymbol{\epsilon})\equiv\boldsymbol{\Omega}_I(t,\boldsymbol{\epsilon})\cdot\boldsymbol{\alpha}_k(t)+\boldsymbol{\Omega}_R(t,\boldsymbol{\epsilon})\cdot\boldsymbol{\beta}_k(t). \tag{13}$$

The vectors $\mathbf{f}_\alpha$ ($\mathbf{f}_\beta$) are equal to the total time derivative of $\boldsymbol{\alpha}$ ($\boldsymbol{\beta}$) only when the state is consistent with the field through the evolution equation (12). The dependence of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ on t will be made explicit only when necessary. An important property of the problems under study is that $\mathbf{f}_\alpha$ and $\mathbf{f}_\beta$ are linear in the functions $\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$ and $\boldsymbol{\epsilon}$. This choice simplifies the optimization problem, the nonlinear case has been studied in Ref. [23].

A "process" $w=w[t,\{\boldsymbol{\alpha},\boldsymbol{\beta}\},\boldsymbol{\epsilon}]$ is defined as the set $\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$ of N vectors $\boldsymbol{\alpha}_k$ and N vectors $\boldsymbol{\beta}_k$ related to the field $\boldsymbol{\epsilon}$ through the evolution equations with the initial conditions $\boldsymbol{\alpha}_k(0)$ and $\boldsymbol{\beta}_k(0)$. A functional of the process can be defined as

$$J[w]=F(\{\boldsymbol{\alpha}(T),\boldsymbol{\beta}(T)\})+\int_0^T g(\boldsymbol{\epsilon})\,dt. \tag{14}$$

For the present applications F can be any of the functionals F_{re}, F_{sm}, F_{ss} as introduced in Sec. II. The optimal field is found by a minimization of the functional J. The integral term represents additional constraints originating from the evolution equation of the system. For simplicity only the case where g is a function of the field $\boldsymbol{\epsilon}$ is presented, but a generalization to the more general case in which g depends on $\boldsymbol{\alpha}(t)$ and $\boldsymbol{\beta}(t)$ is straightforward. The particular dependence of g on the field will be discussed later.

The main idea in the Krotov method is to introduce a new functional that mixes the separate dependences on intermediate and final times in the original functional (14). Using the new functional it is possible to derive an iterative procedure that modifies the field at intermediate times in a way consistent with the minimization of F at time T. The new functional is defined as

$$L[w,\Phi]=G(\{\boldsymbol{\alpha}(T),\boldsymbol{\beta}(T)\})-\Phi(0,\{\boldsymbol{\alpha}(0),\boldsymbol{\beta}(0)\}) - \int_0^T R(t,\{\boldsymbol{\alpha},\boldsymbol{\beta}\},\boldsymbol{\epsilon})\,dt, \tag{15}$$

where

$$G(\{\boldsymbol{\alpha}(T),\boldsymbol{\beta}(T)\})=F(\{\boldsymbol{\alpha}(T),\boldsymbol{\beta}(T)\})+\Phi(T,\{\boldsymbol{\alpha}(T),\boldsymbol{\beta}(T)\}), \tag{16}$$

and

$$R(t,\{\boldsymbol{\alpha},\boldsymbol{\beta}\},\boldsymbol{\epsilon})=-g(\boldsymbol{\epsilon})+\frac{\partial\Phi}{\partial t}(t,\{\boldsymbol{\alpha},\boldsymbol{\beta}\}) + \sum_{k=1}^{N}\frac{\partial\Phi}{\partial\boldsymbol{\alpha}_k}(t,\{\boldsymbol{\alpha},\boldsymbol{\beta}\})\cdot\mathbf{f}_\alpha(t,\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k,\boldsymbol{\epsilon}) + \sum_{k=1}^{N}\frac{\partial\Phi}{\partial\boldsymbol{\beta}_k}(t,\{\boldsymbol{\alpha},\boldsymbol{\beta}\})\cdot\mathbf{f}_\beta(t,\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k,\boldsymbol{\epsilon}). \tag{17}$$

$\Phi(t,\{\boldsymbol{\alpha},\boldsymbol{\beta}\})$ denotes an arbitrary continuously differentiable function. The partial derivatives of Φ, $\partial\Phi/\partial\boldsymbol{\alpha}_k$ and $\partial\Phi/\partial\boldsymbol{\beta}_k$, form a vector with m components. In the following, t, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are considered to be independent variables in Φ.

When $\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$ and the field are related by Eq. (12), R can be written as $R=-g+d\Phi/dt$. Introducing this result in Eq. (15), it can be shown [23] that for any scalar function Φ and any process w, $L[w,\Phi]=J[w]$. Then the minimization of J is completely equivalent to the minimization of L.

1. Iterative algorithm to minimize L

The advantage of the definition of the functional $L[w]$ is the complete freedom in the choice of Φ. This property is used to derive from an arbitrary process $w^{(0)}[t,\{\boldsymbol{\alpha}^{(0)},\boldsymbol{\beta}^{(0)}\},\boldsymbol{\epsilon}^{(0)}]$ a new process $w^{(1)}[t,\{\boldsymbol{\alpha}^{(1)},\boldsymbol{\beta}^{(1)}\},\boldsymbol{\epsilon}^{(1)}]$ such that $L[w^{(1)},\Phi]\leq L[w^{(0)},\Phi]$. The procedure can be summarized as follows.

(i) Φ is constructed so that the functional $L[w^{(0)}]$ is a maximum with respect to any possible choice of the set $\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$. This condition gives a complete freedom to change $\boldsymbol{\epsilon}$. The related changes of the states are consistent with the system evolution, Eq. (12), and therefore, will not interfere with the minimization of L.

(ii) A new field $\boldsymbol{\epsilon}^{(1)}$ is derived with the condition of maximizing R, decreasing then the value of L with respect to the process $w^{(0)}$. In this step the consistency between the new field and the new states of the system $\{\boldsymbol{\alpha}^{(1)},\boldsymbol{\beta}^{(1)}\}$ must be maintained.

The new field $\boldsymbol{\epsilon}^{(1)}$ becomes the starting point of a new iteration, and steps (i) and (ii) are repeated until the desired convergence is achieved.

2. The linear problem: Construction of Φ to first order

The difficult task in the Krotov method is the construction of Φ so that L is maximum for $\{\boldsymbol{\alpha}^{(0)},\boldsymbol{\beta}^{(0)}\}$. The maximum condition on L is equivalent to imposing a maximum on G and a minimum on R. However, in some cases the maximum and minimum conditions can be relaxed to extreme conditions for G and R, which simplifies the problem.

The extreme conditions for R with respect to $\{\boldsymbol{\alpha}^{(0)},\boldsymbol{\beta}^{(0)}\}$ are given by

$$\frac{\partial R}{\partial\boldsymbol{\alpha}_k}[t,\{\boldsymbol{\alpha}^{(0)},\boldsymbol{\beta}^{(0)}\},\boldsymbol{\epsilon}^{(0)}]=0,$$

$$\frac{\partial R}{\partial\boldsymbol{\beta}_k}(t,\{\boldsymbol{\alpha}^{(0)},\boldsymbol{\beta}^{(0)}\},\boldsymbol{\epsilon}^{(0)})=0. \tag{18}$$

The following vectors are introduced:

$$\boldsymbol{\gamma}_k(t)=\frac{\partial\Phi}{\partial\boldsymbol{\alpha}_k}(t,\{\boldsymbol{\alpha}^{(0)},\boldsymbol{\beta}^{(0)}\}),$$

$$\boldsymbol{\delta}_k(t)=\frac{\partial\Phi}{\partial\boldsymbol{\beta}_k}(t,\{\boldsymbol{\alpha}^{(0)},\boldsymbol{\beta}^{(0)}\}). \tag{19}$$

$\boldsymbol{\gamma}_k$ and $\boldsymbol{\delta}_k$ are only functions of t, as the partial derivatives are evaluated in the specific set $\{\boldsymbol{\alpha}^{(0)},\boldsymbol{\beta}^{(0)}\}$. Using Eq. (12) the extreme conditions can be written as

$$\frac{d}{dt}\boldsymbol{\gamma}_k(t)=-\boldsymbol{\Omega}_R^T(t,\boldsymbol{\epsilon}^{(0)})\cdot\boldsymbol{\gamma}_k(t)-\boldsymbol{\Omega}_I^T(t,\boldsymbol{\epsilon}^{(0)})\cdot\boldsymbol{\delta}_k(t),$$

$$\frac{d}{dt}\boldsymbol{\delta}_k(t)=\boldsymbol{\Omega}_I^T(t,\boldsymbol{\epsilon}^{(0)})\cdot\boldsymbol{\gamma}_k(t)-\boldsymbol{\Omega}_R^T(t,\boldsymbol{\epsilon}^{(0)})\cdot\boldsymbol{\delta}_k(t), \tag{20}$$

where $\boldsymbol{\Omega}^T$ denotes the transpose of the matrix $\boldsymbol{\Omega}$. The extreme conditions for G are

$$\frac{\partial G}{\partial\boldsymbol{\alpha}_k(T)}[\{\boldsymbol{\alpha}^{(0)}(T),\boldsymbol{\beta}^{(0)}(T)\}]=0,$$

$$\frac{\partial G}{\partial\boldsymbol{\beta}_k(T)}[\{\boldsymbol{\alpha}^{(0)}(T),\boldsymbol{\beta}^{(0)}(T)\}]=0. \tag{21}$$

Using Eqs. (16) and (19),

$$\boldsymbol{\gamma}_k(T) = -\frac{\partial F}{\partial \boldsymbol{\alpha}_k(T)}[\{\boldsymbol{\alpha}(0)(T), \boldsymbol{\beta}^{(0)}(T)\}],$$

$$\boldsymbol{\delta}_k(T) = -\frac{\partial F}{\partial \boldsymbol{\beta}_k(T)}[\{\boldsymbol{\alpha}^{(0)}(T), \boldsymbol{\beta}^{(0)}(T)\}]. \quad (22)$$

The above conditions at time T, together with Eq. (20), determine completely the set $\{\boldsymbol{\gamma}(t), \boldsymbol{\delta}(t)\}$. As they are defined as the partial derivatives of Φ with respect to α_{km} and β_{km}, Φ is expanded to first order (denoted as Φ^*),

$$\Phi^*(t, \{\boldsymbol{\alpha}, \boldsymbol{\beta}\}) = \sum_{k=1}^{N} [\boldsymbol{\gamma}_k(t) \cdot \boldsymbol{\alpha}_k(t) + \boldsymbol{\delta}_k(t) \cdot \boldsymbol{\beta}_k(t)]. \quad (23)$$

By employing Φ^*, the functions G^* and R^* can also be constructed to first order using Eqs. (16) and (17), respectively. This completes the first step in the iterative algorithm.

To accomplish the second step R is maximized with respect to the field. Again in some cases the maximum condition can be relaxed to the extreme condition $\partial R/\partial \boldsymbol{\epsilon} = 0$. Using the expression for R^* leads to

$$\frac{\partial g}{\partial \boldsymbol{\epsilon}}(\boldsymbol{\epsilon}^{(1)}) = \sum_{k=1}^{N} \boldsymbol{\gamma}_k(t) \cdot \frac{\partial \mathbf{f}_\alpha}{\partial \boldsymbol{\epsilon}}(t, \boldsymbol{\alpha}_k^{(1)}, \boldsymbol{\beta}_k^{(1)}, \boldsymbol{\epsilon}^{(1)}) + \sum_{k=1}^{N} \boldsymbol{\delta}_k(t) \cdot \frac{\partial \mathbf{f}_\beta}{\partial \boldsymbol{\epsilon}}(t, \boldsymbol{\alpha}_k^{(1)}, \boldsymbol{\beta}_k^{(1)}, \boldsymbol{\epsilon}^{(1)}). \quad (24)$$

Equation (24) is used to derive the new field $\boldsymbol{\epsilon}^{(1)}$ in each iteration. This equation must be solved in a way consistent with Eq. (12) describing the system dynamics.

Due to the use of extreme instead of maximum or minimum conditions, it must be checked that the new process $w^{(1)}$ improves the original objective in each iteration $J[w^{(0)}] - J[w^{(1)}] \geqslant 0$:

$$J[w^{(0)}] - J[w^{(1)}] = L[w^{(0)}, \Phi^*] - L[w^{(1)}, \Phi^*] = \Delta_1 + \int_0^T \Delta_2(t)\, dt, \quad (25)$$

where

$$\Delta_1 = G^*(\{\boldsymbol{\alpha}^{(0)}(T), \boldsymbol{\beta}^{(0)}(T)\}) - G^*(\{\boldsymbol{\alpha}^{(1)}(T), \boldsymbol{\beta}^{(1)}(T)\}), \quad (26)$$

$$\Delta_2(t) = R^*(t, \{\boldsymbol{\alpha}^{(1)}, \boldsymbol{\beta}^{(1)}\}, \boldsymbol{\epsilon}^{(1)}) - R^*(t, \{\boldsymbol{\alpha}^{(1)}, \boldsymbol{\beta}^{(1)}\}, \boldsymbol{\epsilon}^{(0)}). \quad (27)$$

The above relation is obtained when $\mathbf{f}_\alpha$ and $\mathbf{f}_\beta$ are linear in $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$,

$$R^*(t, \{\boldsymbol{\alpha}, \boldsymbol{\beta}\}, \boldsymbol{\epsilon}^{(0)}) = -g(t, \boldsymbol{\epsilon}^{(0)}), \quad (28)$$

for any value of $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$.

A sufficient condition for $J[w^{(0)}] - J[w^{(1)}] \geqslant 0$ is $\Delta_1, \Delta_2(t) \geqslant 0$. Δ_1 depends on the choice of F and $\Delta_2(t)$ on the choice of g so that each case must be analyzed separately. These conditions imply that the Krotov iterative algorithm converges monotonically to the final objective.

3. Dependence on F

The dependence of G^* on F can be made explicit by introducing Φ^* and using Eqs. (23) and (16),

$$G^*(\{\boldsymbol{\alpha}(T), \boldsymbol{\beta}(T)\}) = F(\{\boldsymbol{\alpha}(T), \boldsymbol{\beta}(T)\}) - \sum_{k=1}^{N} \frac{\partial F}{\partial \boldsymbol{\alpha}_k(T)} \cdot \boldsymbol{\alpha}_k(T) + \frac{\partial F}{\partial \boldsymbol{\beta}_k(T)} \cdot \boldsymbol{\beta}_k(T). \quad (29)$$

When F is linear in $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$ $G^* \equiv 0$ and then $\Delta_1 \equiv 0$. In this case all the improvement towards the original objective in the iteration is due to the term $\Delta_2(t)$. When F is nonlinear in $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$ the condition $\Delta_1 \leqslant 0$ must be checked in each case.

An additional difficulty is that the conditions (19) for $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ depend on the particular choice of F. In all the cases under study (F_{re}, F_{sm}, and F_{ss}),

$$\boldsymbol{\gamma}_k(T) = c_k\, \boldsymbol{\alpha}_{fk}(T),$$

$$\boldsymbol{\delta}_k(T) = d_k\, \boldsymbol{\beta}_{fk}(T), \quad (30)$$

where the coefficients c_k and d_k depend on the sets $\{\boldsymbol{\alpha}^{(0)}(T), \boldsymbol{\beta}^{(0)}(T)\}$ and $\{\boldsymbol{\alpha}_f, \boldsymbol{\beta}_f\}$. Defining the vectors $\tilde{\boldsymbol{\gamma}}_k = c_k^{-1}\, \boldsymbol{\gamma}_k$ and $\tilde{\boldsymbol{\delta}}_k = d_k^{-1}\, \boldsymbol{\delta}_k$, the conditions (22) for all the cases under consideration are

$$\tilde{\boldsymbol{\gamma}}_k(T) = \boldsymbol{\alpha}_{fk}(T),$$

$$\tilde{\boldsymbol{\delta}}_k(T) = \boldsymbol{\beta}_{fk}(T). \quad (31)$$

Their evolution is given by Eq. (20). Changing $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ to $\tilde{\boldsymbol{\gamma}}$ and $\tilde{\boldsymbol{\delta}}$ Eq. (24) can be written as

$$\frac{\partial g}{\partial \boldsymbol{\epsilon}}(\boldsymbol{\epsilon}^{(1)}) = \sum_{k=1}^{N} c_k\, \tilde{\boldsymbol{\gamma}}_k(t) \cdot \frac{\partial \mathbf{f}_\alpha}{\partial \boldsymbol{\epsilon}}(t, \boldsymbol{\alpha}_k^{(1)}, \boldsymbol{\beta}_k^{(1)}, \boldsymbol{\epsilon}^{(1)}) + \sum_{k=1}^{N} d_k\, \tilde{\boldsymbol{\delta}}_k(t) \cdot \frac{\partial \mathbf{f}_\beta}{\partial \boldsymbol{\epsilon}}(t, \boldsymbol{\alpha}_k^{(1)}, \boldsymbol{\beta}_k^{(1)}, \boldsymbol{\epsilon}^{(1)}). \quad (32)$$

The different choices of F imply different coefficients (c_k and d_k) and a possible different set of initial $|\varphi_{ik}\rangle$ and final $|\varphi_{fk}\rangle$ states. Nevertheless, the iterative procedure is identical in all the cases.

4. Dependence on g(ε)

A delicate point is the choice of the function $g(\boldsymbol{\epsilon})$ in $J[w]$. The time integral in the functional should be bounded from below, otherwise the additional constraint will dominate over the original objective F in the functional J. In addition $\Delta_2(t) \geqslant 0$ is required in order to guarantee the monotonic convergence of the optimization method.

A consequence of the linear dependence of $\mathbf{f}_\alpha$, $\mathbf{f}_\beta$, and Eq. (32) is that the function R^* for the new process $w^{(1)}$ has the simple form

$$R^*(t,\{\boldsymbol{\alpha}^{(1)},\boldsymbol{\beta}^{(1)}\},\epsilon^{(1)}) = -g(\epsilon^{(1)}) + (\epsilon^{(1)}-\epsilon^{(0)})\frac{\partial g}{\partial \epsilon}(\epsilon^{(1)}). \tag{33}$$

Using this expression together with Eq. (28) leads to

$$\Delta_2(t) = -g(\epsilon^{(1)}) + g(\epsilon^{(0)}) + (\epsilon^{(1)}-\epsilon^{(0)})\frac{\partial g}{\partial \epsilon}(\epsilon^{(1)}). \tag{34}$$

A choice of $g(\epsilon)$ fulfilling the requirement $\Delta_2(t)\geqslant 0$ is

$$g(\epsilon) = \lambda(t)[\epsilon(t)-\tilde{\epsilon}(t)]^2, \tag{35}$$

where $\tilde{\epsilon}$ is a reference field and $\lambda(t)$ is a positive function of t. Using Eq. (34) and for any field $\tilde{\epsilon}$

$$\Delta_2(t) = \lambda(t)[\Delta\epsilon(t)]^2 \geqslant 0, \tag{36}$$

where $\Delta\epsilon(t)\equiv\epsilon^{(1)}(t)-\epsilon^{(0)}(t)$. The method therefore will converge monotonically. Using Eqs. (32) and (35) the field in the new iteration becomes

$$\epsilon^{(1)}(t) = \tilde{\epsilon}(t) + \frac{1}{2\lambda(t)}\sum_{k=1}^{N}\left\{ c_k\,\tilde{\boldsymbol{\gamma}}_k(t)\cdot\frac{\partial \mathbf{f}_\alpha}{\partial \epsilon}(t,\boldsymbol{\alpha}_k^{(1)},\boldsymbol{\beta}_k^{(1)},\epsilon^{(1)}) + d_k\,\tilde{\boldsymbol{\delta}}_k(t)\cdot\frac{\partial \mathbf{f}_\beta}{\partial \epsilon}(t,\boldsymbol{\alpha}_k^{(1)},\boldsymbol{\beta}_k^{(1)},\epsilon^{(1)})\right\}. \tag{37}$$

The result of the iterative algorithm depends strongly on the choices of the reference field $\tilde{\epsilon}$ and on the function $\lambda(t)$.

Two possible choices of $\tilde{\epsilon}$ are analyzed. The first choice, $\tilde{\epsilon}=0$, is the one commonly used in OCT applications [1]. In this case, the additional constraint in $J[w]$ has the physical meaning that the total energy of the field in the time interval $[0,T]$ is limited. This however presents a problem when the iterative procedure reaches the optimal field. The iterative method is found to reduce the total objective J by reducing the total pulse energy, slowing and even spoiling the convergence toward the original objective F. The usual remedy is to stop the iterative algorithm before this difficulty is reached. However, such a procedure could prevent the optimization algorithm from obtaining high fidelity.

A different possibility is $\tilde{\epsilon}=\epsilon^{(0)}$ can avoid this problem [10,16,23]. In this iterative algorithm $\epsilon^{(0)}$ must be interpreted as the field in the previous iteration. Now the additional constraint in $J[w]$ has the physical interpretation that the change of the pulse energy in each iteration is limited. When the iterative procedure approaches the optimal solution the change in the field vanishes. Therefore, the convergence to the original objective is guaranteed. In the rest of the study $\tilde{\epsilon}=\epsilon^{(0)}$ was chosen.

The function $\lambda(t)$ introduces the shape function $s(t)$, i.e., $\lambda(t)=\lambda_0/s(t)$. The purpose of $s(t)$ is to turn the field on and off smoothly at the boundaries of the interval [24]. λ_0 is a scaling parameter which determines the optimization strategy. When λ_0 is small the additional constraint on the field in the functional becomes insignificant, resulting in large modifications in the field in each iteration. This is equivalent to a bold search strategy where large excursions in the functional space of the field take place with each iteration. Large values of λ_0 imply small modifications in the field in each iteration, slowing the convergence process. Using large values of λ_0 is a conservative search strategy which is advantageous when a good initial guess field can be found. A possible mixed strategy is to initially use a bold optimization with small values of λ_0. This leads to a guess field for a new optimization with a large value of λ_0 [25].

B. Application to the functionals F_{re}, F_{sm}, and F_{ss}

Based on the derivation of the Krotov method it is possible to connect directly the minimization F_{re}, F_{sm}, and F_{ss} to the correction to the field. Equation (20) corresponds to the evolution of a set of states $\{|\chi_k(t)\rangle\}$,

$$\frac{d}{dt}|\chi_k(t)\rangle = -\frac{i}{\hbar}\hat{\mathbf{H}}^\dagger(\epsilon^{(0)})|\chi_k(t)\rangle, \tag{38}$$

with the conditions (31), $|\chi_k(T)\rangle=|\varphi_{fk}\rangle$. The formal solution of the equation is given by $|\chi_k(t)\rangle = \hat{\mathbf{U}}(t,T;\epsilon^{(0)})\hat{\mathbf{O}}|\varphi_{ik}\rangle$. Using Eq. (37) the correction to the field in each iteration becomes

$$\Delta\epsilon(t) = -\frac{s(t)}{\lambda_0\hbar}\mathrm{Im}\left[\sum_{k=1}^{N} a_k(\epsilon^{(0)}) \times\langle\varphi_{ik}|\hat{\mathbf{O}}^\dagger\hat{\mathbf{U}}^\dagger(t,T;\epsilon^{(0)})\hat{\boldsymbol{\mu}}\hat{\mathbf{U}}(t,0;\epsilon^{(1)})|\varphi_{ik}\rangle\right]. \tag{39}$$

The coefficients a_k will depend on the particular choice of the functional F and are related to c_k and d_k defined in Eq. (19). For F_{re} and F_{sm}, the states $\{|\varphi_{ik}\rangle\}$ denote an orthonormal basis $\{|n\rangle\}$ of the relevant subspace $\mathcal{N}$. For F_{re} the coefficients are $a_{n,re}=1/2$, and as this functional is linear on the states $\Delta_{1,re}=0$. For F_{sm} the coefficients are

$$a_{n,sm} = \sum_{n'=1}^{N}\langle n'|\hat{\mathbf{U}}(0,T;\epsilon^{(0)})\hat{\mathbf{O}}|n'\rangle, \tag{40}$$

and then are equal for all the states in the basis of $\mathcal{N}$. In addition,

$$\Delta_{1,sm} = \left|\sum_{n=1}^{N}\langle n|[\hat{\mathbf{U}}(T,0;\epsilon^{(0)})-\hat{\mathbf{U}}(T,0;\epsilon^{(1)})]\hat{\mathbf{O}}|n\rangle\right|^2. \tag{41}$$

Therefore, $\Delta_{1,sm}\geqslant 0$. For the F_{ss} functional the set $\{|\varphi_{ik}\rangle\}$ for which the states are denoted by $\{|l\rangle\}$ the coefficients a_l are

$$a_{l,ss} = \langle l|\hat{\mathbf{U}}(0,T;\epsilon^{(0)})\hat{\mathbf{O}}|l\rangle, \tag{42}$$

depending on the index l corresponding to each state. In this case

$$\Delta_{1,ss}=\sum_{l=1}^{N}|\langle l|[\hat{\mathbf{U}}(T,0;\boldsymbol{\epsilon}^{(0)})-\hat{\mathbf{U}}(T,0;\boldsymbol{\epsilon}^{(1)})]\hat{\mathbf{O}}|l\rangle|^2, \quad (43)$$

and then $\Delta_{1,ss}\geqslant 0$.

The results $\Delta_1\geqslant 0$ and $\Delta_2(t)\geqslant 0$ guarantee the monotonic convergence of the iterative algorithm based on the Krotov method for the three functionals.

C. The optimal field

The optimal field has the property that the field correction in the next iteration Eq. (39) should vanish. Defining

$$C(t;\boldsymbol{\epsilon})=\mathrm{Im}\left[\sum_{k=1}^{N} a_k(\boldsymbol{\epsilon})\langle\varphi_{ik}|\hat{\mathbf{O}}^{\dagger}\hat{\mathbf{U}}^{\dagger}(t,T;\boldsymbol{\epsilon})\hat{\boldsymbol{\mu}}\hat{\mathbf{U}}(t,0;\boldsymbol{\epsilon})|\varphi_{ik}\rangle\right], \quad (44)$$

when $\bar{\boldsymbol{\epsilon}}$ is an arbitrary solution of the iterative algorithm $C(t;\bar{\boldsymbol{\epsilon}})\equiv 0$.

The first question to be addressed is whether any optimal field, defined by Eq. (2), is a possible solution of the iterative algorithm. $\bar{\boldsymbol{\epsilon}}_{opt}$ denotes a field that generates the target unitary transformation up to a global phase, $\hat{\mathbf{U}}(T,0;\bar{\boldsymbol{\epsilon}}_{opt})=e^{-i\bar{\phi}}\hat{\mathbf{O}}$, using the relation

$$\hat{\mathbf{U}}(t,0;\boldsymbol{\epsilon})=\hat{\mathbf{U}}(t,T;\boldsymbol{\epsilon})\,\hat{\mathbf{U}}(T,0;\boldsymbol{\epsilon}). \quad (45)$$

In addition, the relation $|\Psi_k(t)\rangle=\hat{\mathbf{U}}(t,T;\bar{\boldsymbol{\epsilon}}_{opt})\hat{\mathbf{O}}|\varphi_{ik}\rangle$ implying that the term $\langle\Psi_k(t)|\hat{\boldsymbol{\mu}}|\Psi_k(t)\rangle$ is real, simplifies Eq. (44) to

$$C(t;\bar{\boldsymbol{\epsilon}}_{opt})=\sum_{k=1}^{N}\langle\Psi_k(t)|\hat{\boldsymbol{\mu}}|\Psi_k(t)\rangle\mathrm{Im}[a_k(\bar{\boldsymbol{\epsilon}}_{opt})e^{-i\bar{\phi}}]. \quad (46)$$

Using Eq. (40) for the functional F_{sm} leads to $a_{n,sm}=N\exp(i\bar{\phi})$ in Eq. (44). A similar result is found for the functional F_{ss}, $a_{l,ss}=\exp(i\bar{\phi})$, given by Eq. (42). Therefore any field generating the target unitary transformation is a possible solution of the iterative algorithm based on any of the functionals F_{sm} and F_{ss}. This result does not imply that when initializing the different iteration schemes with the same guess field the same solution will be obtained.

The analysis is more complex for the functional F_{re}. The coefficients are now real $a_{n,re}=1/2$, and are independent of the state index. This leads to $C_{re}=\mathrm{Im}[\exp(-i\bar{\phi})]\Sigma_{n=1}^{N}\langle\Psi_k(t)|\hat{\boldsymbol{\mu}}|\Psi_k(t)\rangle$. The sum is generally different from zero and the solutions to the algorithm are fields with a phase term $\exp(-i\bar{\phi})=\pm 1$. Only the case $+1$ minimizes the original functional F_{re}, but the relaxation to extreme conditions in the Krotov method allows one to obtain other physically valid solutions. In the special case in which the unitary transformation is imposed on all the Hilbert space ($N=M$), any optimal field is a possible solution regardless of the global phase. The reason is that the sum in C_{re} is zero since $\hat{\boldsymbol{\mu}}$ is a traceless operator.

The phase sensitivity of the functional F_{re} can be demonstrated in the state-to-state optimization. The iterative algorithm in this case will converge to a field that drives the system to the final state $+|\varphi_f\rangle$ or $-|\varphi_f\rangle$, while the optimization of F_{sm} or F_{ss} will converge to the final state up to an arbitrary global phase. There is no *a priori* advantage however to any of the three functionals in the convergence rate or in the simplicity of the solution. The solutions are physically equivalent since they differ only in a global phase.

In addition to the desired optimal fields, the algorithm could also generate spurious solutions. An example is given by the functional F_{sm} employed to implement a unitary transformation $\hat{\mathbf{O}}_D$ with a matrix representation diagonal in the basis of the free Hamiltonian eigenstates $|e_n\rangle$. In such a case $C(t,\boldsymbol{\epsilon}=0)$ is proportional to the diagonal matrix elements $\langle e_n|\hat{\boldsymbol{\mu}}|e_n\rangle$. When these matrix elements are zero, $\boldsymbol{\epsilon}=0$ is a solution of the iterative algorithm, but it does not implement the desired unitary transformation. A simple remedy to overcome this difficulty is to use a different initial guess to start the algorithm.

D. Discrete implementation of the optimization algorithm

A numerical solution of the iterative optimization algorithm requires a discretization scheme for the time axes. The correction to the field $\Delta\boldsymbol{\epsilon}$ is implicit and appears on both sides of Eq. (39). To implement the procedure, two interleaved grid points in time were used. The first grid was used to propagate the states. The second grid was used to evaluate the field. The grid describing the states has N_t+1 points separated by $\Delta t=T/N_t$, from $t=0$ to $t=T$. The grid representing the field has N_t points separated by Δt and starting at $t=\Delta t/2$. The initial set of states $|\varphi_{ik}\rangle$ was used for the target unitary transformation $\hat{\mathbf{O}}$ optimization with the functionals F_{re}, F_{sm}, or F_{ss}. The numerical implementation of the algorithm follows.

(i) Using an initial guess field $\boldsymbol{\epsilon}^{(0)}$, the states φ_{fk} are propagated in reverse from $t=T$ to $t=0$ to determine $\hat{\mathbf{U}}(t,T;\boldsymbol{\epsilon}^{(0)})\hat{\mathbf{O}}|\varphi_{ik}\rangle$ on the time grid of states.

(ii) The new field is determined in the interleaved grid point $t=\Delta t/2$ using the approximation

$$\Delta\boldsymbol{\epsilon}(\Delta t/2)\approx-\frac{s(\Delta t/2)}{\lambda_0\hbar}\mathrm{Im}\left[\sum_{k=1}^{N} a_k(\boldsymbol{\epsilon}^{(0)})\langle\varphi_{ik}|\hat{\mathbf{O}}^{\dagger}\hat{\mathbf{U}}^{\dagger}(0,T;\boldsymbol{\epsilon}^{(0)})\times\hat{\boldsymbol{\mu}}\hat{\mathbf{U}}(0,0;\boldsymbol{\epsilon}^{(1)})|\varphi_{ik}\rangle\right]. \quad (47)$$

Note that $\hat{\mathbf{U}}(0,0;\boldsymbol{\epsilon}^{(1)})|\varphi_{ik}\rangle=|\varphi_{ik}\rangle$. Then the new field in the first field time grid point is obtained, $\boldsymbol{\epsilon}^{(1)}(\Delta t/2)=\boldsymbol{\epsilon}^{(0)}+\Delta\boldsymbol{\epsilon}(\Delta t/2)$, and used to propagate $|\varphi_{ik}(t=0)\rangle$ to the next state grid point $t=\Delta t$. The same process is used to obtain the new field $\boldsymbol{\epsilon}^{(1)}$ in the next field time grid point $t=\Delta t+\Delta t/2$, evaluating the correction with the already known states in the state grid point $t=\Delta t$. The process is repeated to obtain $\boldsymbol{\epsilon}^{(1)}$ in all the field time grid points.

(iii) The new field $\boldsymbol{\epsilon}^{(1)}$ is used as input to the new iteration

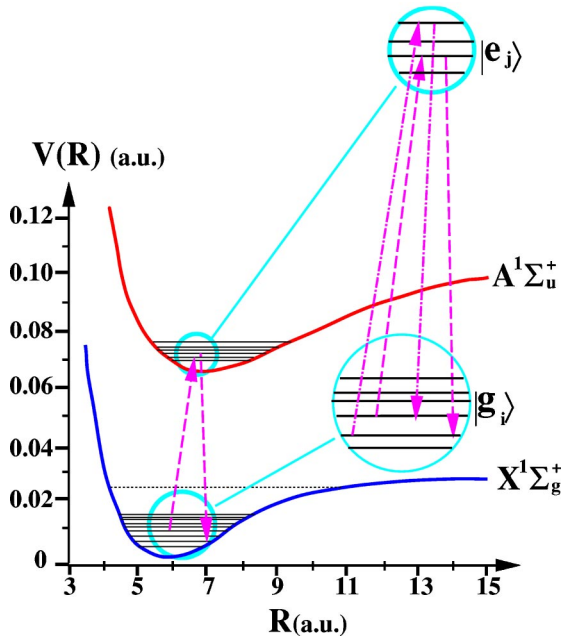

FIG. 1. Schematic representation of a molecular model based on the vibrational levels in the $X\,^1\Sigma_g^+$ (lower) and $A\,^1\Sigma_u^+$ (upper) electronic surfaces of the molecule Na_2. Atomic units are chosen $\hbar=1$. R denotes the internuclear distance. The arrows indicate two of the possible transitions induced by the driving field between arbitrary levels in the lower and upper surfaces. On the right is a magnified view of some of the energy levels involved and transitions between them.

$(\epsilon^{(0)}=\epsilon^{(1)})$ and the process is repeated until the required convergence is achieved.

More elaborate methods to deal with the implicit time dependence of Eq. (39) have been developed. For example, approximating the dynamics in between grid points by the free evolution with $\hat{\mathbf{H}}_0$ [22]. The simple procedure, which is able to keep the monotonic behavior of the optimization method, was found sufficient.

The present implementation is based on a forward time propagation. Using the same formalism, the optimization can be accomplished also by a backward time propagation. It is also possible to combine both cases, and to perform the optimization in the forward and backward propagations [22,26]. In the current studies, these other procedures were found to be inferior, slowing down the convergence rate.

IV. THE FOURIER TRANSFORM EXAMPLE IN A MOLECULAR MODEL

As an illustration the implementation of Q-qubit Fourier transform in a two-electronic-surfaces molecular model was studied. Figure 1 shows a schematic view of a model based on the electronic manifolds of Na_2.

The Hamiltonian of the system describes a ground and excited electronic potential-energy surface coupled by a transition dipole operator:

$$\hat{\mathbf{H}}=\hat{\mathbf{H}}_g\otimes|G\rangle\langle G|+\hat{\mathbf{H}}_e\otimes|E\rangle\langle E|-\hat{\boldsymbol{\mu}}\otimes(|G\rangle\langle E| +|E\rangle\langle G|)\epsilon(t), \tag{48}$$

where $|G\rangle$ and $|E\rangle$ are the ground and excited electronic states and $\hat{\mathbf{H}}_g$ and $\hat{\mathbf{H}}_e$ are the corresponding vibrational Hamiltonians. The electronic surfaces are coupled by the transition dipole operator $\hat{\boldsymbol{\mu}}$, controlled by the shaped field $\epsilon(t)$.

The present model is a simplification of the Na_2 Hilbert space restricting the number of vibrational levels. On the ground $X\,^1\Sigma_g^+$ electronic state the first 40 vibrational levels selected out from the 66 bound states are used. In the excited $A\,^1\Sigma_u^+$ state, the lowest 20 vibrational states are used out of the 210 bound levels. The vibrational Hamiltonians become therefore

$$\hat{\mathbf{H}}_g=\sum_{i=1}^{40} E_{gi}|g_i\rangle\langle g_i|; \quad \hat{\mathbf{H}}_e=\sum_{j=1}^{20} E_{ej}|e_j\rangle\langle e_j|. \tag{49}$$

For Na_2 the 00 transition frequency between the ground vibrational levels of each surface is $\Omega\equiv E_{e1}-E_{g1}\approx 0.066\,01$ a.u. (~ 1.8 eV). A transition dipole operator independent of the internuclear distance R was considered, $\hat{\boldsymbol{\mu}}=\mu_0(|G\rangle\langle E|+|E\rangle\langle G|)$. This model is sufficient for the illustrative purpose of demonstrating the execution of an algorithm in a molecular setting.

The $N=2^Q$ first levels of the ground electronic surface are chosen as the registers representing the Q qubits. The unitary transformation implemented is a Fourier transform [27] invoked on the N levels on the $X\,^1\Sigma_g^+$ electronic state representing the qubit(s). The unitary transformation is implemented through transitions between the two electronic manifolds, cf. Fig. 1.

An implementation of the iterative algorithm is chosen where the $|g_i\rangle\otimes|G\rangle$ and $|e_j\rangle\otimes|E\rangle$ eigenstates are used as the basis $\{|m\rangle\}$. The $N=2^Q$ first states in the lower surface are used as the basis $\{|n\rangle\}$ of the relevant subspace. The first $N-1$ energy levels plus the linear combination $\Sigma_{n=1}^N|n\rangle/\sqrt{N}$ are used as the set $|l\rangle$ for the state-to-state formulation. The wavefunction propagations were carried out by using a Newton polynomial integrator [28]. The final time for the implementation is $T=4.5\times10^4$ a.u. (≈ 1 ps). In all the cases a Gaussian shape function $s(t)=\exp\{-32(t/T-1/2)^2\}$ and a guess field $\epsilon_{guess}(t)=\epsilon_0 s(t)\cos(\Omega t)$ were chosen.

The implementation of the Fourier transform in two qubits ($N=4$) embedded in the set of 60 levels is used for comparing the performance of the methods. Figure 2 shows the change in the normalized functional, defined as $J_{norm}\equiv J/N$ for F_{re} and F_{ss}, and $J_{norm}\equiv J/N^2$ for F_{sm}, with the progression of the iterative algorithm. In all the cases the target value of the normalized functional is -1. A large reduction in the value of the functionals is accomplished in a small number of iterations. Note the behavior of the simultaneous state-to-state formulation F_{ss} with an insufficient choice of the states $|l\rangle$. The algorithm finds a minimum of the objective, but, as shown in Fig. 4, the fidelity saturates at a very low value meaning that this field does not generate the target unitary transformation.

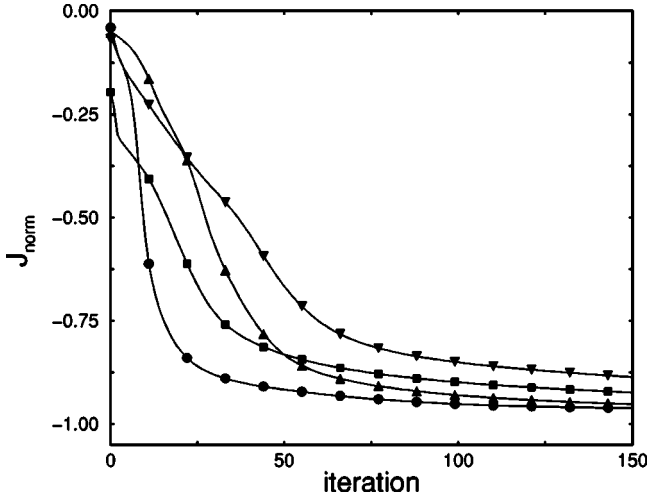

FIG. 2. Normalized functional J_{norm} vs the number of iteration: F_{re} (squares), F_{sm} (circles), F_{ss} (triangles up) for implementing a fast Fourier transform in four levels. The line with triangles pointing down corresponds to F_{ss} functional when $\{|l\rangle\}$ is chosen as the orthogonal basis $\{|n\rangle\}$. The objective is reached when $J_{norm}=-1$. $\lambda_0=10^3$ and $\epsilon_0=5\times10^{-3}$ a.u. in all the cases.

Figure 3 shows the value of τ for the field obtained in each iteration. The same initial guess was used in all the cases which constituted the starting point for all the iterative optimizations. However, the final results depend on the particular functional used. As discussed before the method based on F_{re} finds a solution with a phase factor $\exp(-i\phi)\approx+1$.

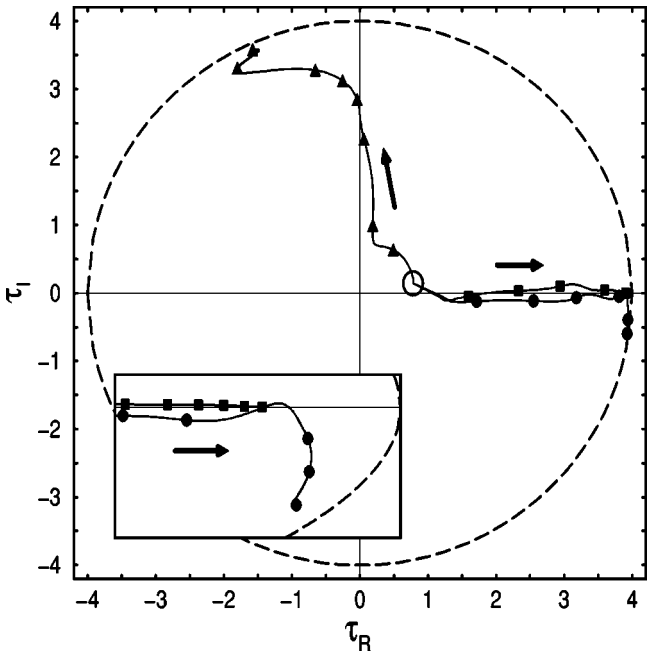

FIG. 3. Evolution of the optimization in the complex τ plane for the case in Fig. 2. The lines correspond to F_{sm} (circles), F_{re} (squares), and F_{ss} (triangles up). The open circle indicates the value of τ for the common guess field. The dashed black line is the circle $|\tau|=N$ indicating the target of the methods. The arrows mark the direction of convergence. The insert enlarges the region corresponding to the real axes close to the circumference.

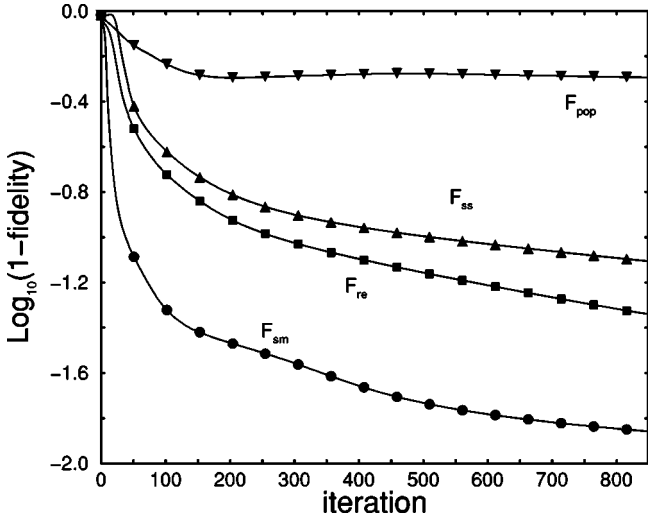

FIG. 4. $\mathrm{Log}_{10}[1-(\mathrm{fidelity})]$ of the implementation of the two-qubit Fourier transform vs the number of iterations N_{it} for the optimization in Fig. 2. The lines correspond to F_{re} (squares), F_{sm} (circles), and F_{ss} (triangles up). F_{pop} (triangles down) denotes the case when the set $\{|l\rangle\}$ is chosen as the orthogonal basis $\{|n\rangle\}$ for the functional F_{ss}.

For the purpose of quantum computing the target unitary transformation has to achieve high accuracy. The fidelity functional

$$(\mathrm{fidelity})=|\tau|^2/N^2 \tag{50}$$

is used to indicate the quality of the solution. Figure 4 shows the improvement of the fidelity versus the iteration. The square modulus functional F_{sm} [Eq. (6)] shows a faster convergence rate than the other two functionals.

In Fig. 5 the Fourier transform of the field for each of the

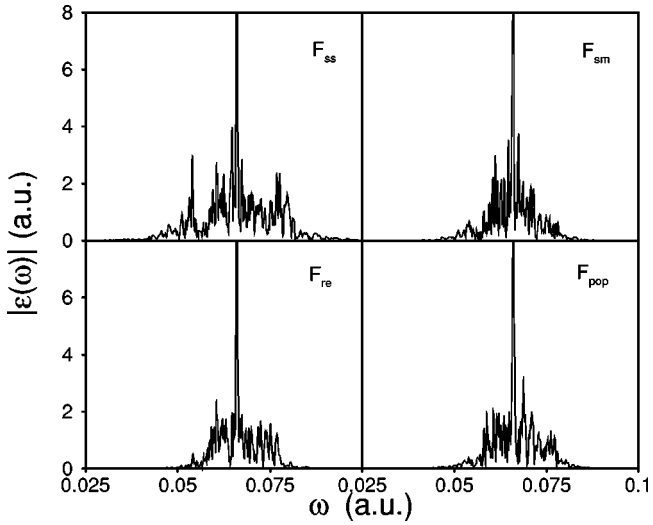

FIG. 5. Fourier transform of the optimal field result of the optimization in Fig. 4 for the functionals F_{re}, F_{sm}, and F_{ss}. Atomic units are chosen $\hbar=1$. The case F_{pop} is also shown.

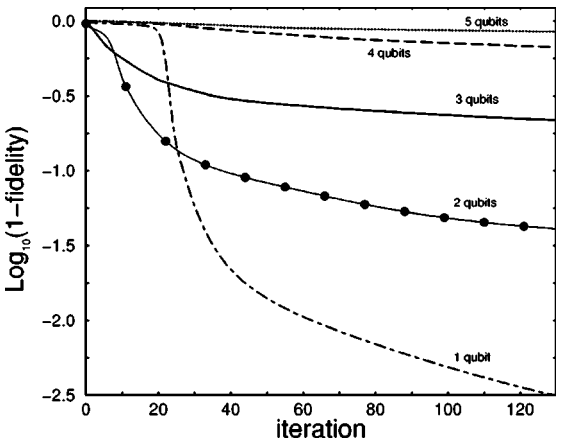

FIG. 6. $\mathrm{Log}_{10}[1-(\mathrm{fidelity})]$ vs the number of iterations for implementing a Fourier transform in 2 (dashed-dotted line), 4 (circles), 8 (solid line), 16 (dashed line), and 32 (dotted line) levels.

optimization procedures is shown. The large peak at the frequency Ω, seen in all cases, is the result of the choice of the guess field. Besides, a similar width in frequencies is found. However, the fidelity reached by the solution corresponding to the square modulus functional F_{sm} is significantly better than in the other cases for the same number of iterations.

The molecular model is also used to compare the convergence of the unitary transformation with the size N of the relevant subspace. Figure 6 shows the improvement in the fidelity versus the number of iterations for implementing a Fourier transform in 2, 4, 8, 16, and 32 levels (1, 2, 3, 4, and 5 qubits, respectively). The convergence characteristics in the initial iterations strongly depends on the initial guess and the parameter λ_0. For example the initial guess seems inappropriate for the one-qubit case which displays an initial very slow convergence until after 25 iterations the right track is found. After a large number of iterations the convergence characteristics settled meaning that each new iteration gave only a slight improvement on the previous one. As the iteration proceeds the rate of convergence decreases in all cases, scaling approximately as the inverse of the number of iterations. Comparing the rate of convergence for the different number of qubits after a large number of iterations the rate seems to be inversely proportional to the number of levels. High fidelity was obtained for one-, two-, three-qubit cases by continuing to 600 iterations. The results allow one to compare the integrated intensity of the optimal field:

$$\mathcal{I}=\int_0^T|\mu_0\epsilon(t)|dt. \tag{51}$$

The initial integrated intensity for all cases was identical. The optimization procedure changed $\mathcal{I}$ depending on the number of qubits. The converged results show a moderate increase of $\mathcal{I}$ with the number of levels starting from $\mathcal{I}=42$ for one qubit to $\mathcal{I}=54$ for two qubits and $\mathcal{I}=78$ for three qubits.

V. DISCUSSION

An implicit assumption in the optimization procedure is that the system is controllable. This means that a field $\epsilon(t)$ exists which implements the unitary transformation up to a prespecified tolerance. The problem of controllability has been the subject of several studies [29–33]. In the context of unitary transformations it has been shown [30] that if the commutators of the operators $\hat{\mathbf{H}}_0$ and $\hat{\boldsymbol{\mu}}$ generate the complete Lie group SU(N), the system is completely controllable. In more concrete terms addressing the Na_2 model, it is expected to be completely controllable. The reason is that the energy levels are nondegenerate and in addition each transition is distinct, characterized by a different Frank-Condon factor $\langle e_j|\hat{\boldsymbol{\mu}}|g_i\rangle$. This controllability property will be true in almost any nonsymmetric molecular system.

A far reaching conclusion is therefore that for any unitary transformation contained in the Hilbert space of the molecule, there is a driving field that implements the transformation in one step. In a molecular system this task could be achieved in a time scale of a picosecond. Since a field that executes such a unitary transformation exists, how difficult is it to find it? Does this optimal field have reasonable intensity and bandwidth?

The OCT scheme can be considered as a classical algorithm employed for the inverse problem of finding the field that generates a predefined unitary transformation. The difficulty of the inversion process is related to the scaling properties of the numerical effort with respect to the number of levels N. The best OCT algorithm based on the F_{sm} functional is then used for estimating the scaling.

Simulating the quantum evolution is the major numerical task of the algorithm implementing OCT. The basic step is a single vector matrix multiplication which represents the operation of the Hamiltonian on the wave function. This task scales as $\mathcal{O}(M^2)$ for direct vector-matrix multiplication or $\mathcal{O}(M\log M)$ for grid methods based on fast Fourier transform [34]. The time propagation requires N_t steps which scale as $\mathcal{O}(T\Delta E)$, where ΔE is the energy range of the problem.

The simulation of a unitary transformation in the relevant subspace turns out to be N times more costly. Summarizing, the numerical cost of the classical simulation of the quantum propagation scales as $(\mathrm{Cost})\sim\mathcal{O}(2^Q M^2 T\Delta E)$. This scaling relation is consistent with the fact that a classical simulation of a quantum unitary transformation scales exponentially with the number of qubits.

The numerical cost of the OCT iterative algorithm used for inversion can now be examined. The crucial question is how many iterations are required to obtain the field that implements the unitary transformation up to a specified fidelity. For this discussion it is convenient to introduce $f=\log_{10}\{1-(\mathrm{fidelity})\}$. The analysis of the results of Sec. IV show that only the initial iterative steps are very sensitive to the choice of the initial guess field. Eventually an asymptotic behavior is reached where the rate of change of f becomes inversely proportional to the number of the iterations steps. In addition, cf. Fig. 6, the rate of convergence is also inversely proportional to the number of levels. This relation

implies that the number of iterations N_{it} required to achieve the specified fidelity becomes

$$N_{it} \approx b e^{2^Q |f|/a}, \tag{52}$$

where the coefficients a and b are positive. The data confirm that the coefficient a is independent of the number of levels N. A consequence of Eq. (52) is that the numerical resources required on a classical computer in order to implement the proposed scheme scale exponentially with the number of levels N. Finding the field that implements in a single step a large unitary transformation is therefore prohibitively expensive. Thus fields that achieve high fidelity are only feasible for unitary transformations with a small relevant subspace. The limiting case would be the one-dimensional state-to-state optimization.

Quantum control is based on interferences between many distinct pathways [1]. State-to-state coherent control finds a constructive interference which leads exclusively to the final state. The controllability depends on having a sufficient amount of interference pathways. Implementing a unitary transformation by interferences is more complex. In this case the interference pathways from one state to another have to avoid other interference paths which connect other states. The possible number of interference pathways becomes the crucial resource that allows one to generate the transformation.

For weak fields, the number of pathways connecting two states in the subspace is linearly related to the number of auxiliary states on the excited surface. Practically the bandwidth of the pulse determines this number. This means that the bandwidth in a weak-field implementation of a unitary transformations has to increase exponentially when the number of levels N increases. The picture is completely altered when the intensity is allowed to increase. Rabi cycling increases the number of interference pathways exponentially. The number of Rabi cycles can be estimated from the integrated intensity $J_{Rabi} \sim \mathcal{I}/2\pi$, cf. Eq. (51), which leads to an estimation of the number of interference pathways $\mathcal{O}(M^{J_{Rabi}}) \sim \mathcal{O}(M^{\mathcal{I}/2\pi})$. This estimation is consistent with the results of Sec. IV where only a moderate increase in $\mathcal{I}$ was observed when the number of qubits in the transformation increased. The estimated number of Rabi cycles changed from $J_{Rabi} \sim 6$ for $Q=1$ to $J_{Rabi} \sim 8$ for $Q=2$ to $J_{Rabi} \sim 12$ for $Q=3$. This means that the increase in resources of implementing a unitary transformation with Q qubits in a molecular environment will scale with a low power of $T\Delta E$, where ΔE is the pulse energy.

In summary, we note the following points.

(1) A unified approach for obtaining the field that implements a unitary transformation has enabled the assessment of various formulations. In addition, a algorithm based on the square modulus of τ was developed. This scheme was found to have superior convergence properties with respect to the number of iterations.

(2) A unitary transformation could be implemented in a molecular environment in a time scale of picoseconds with reasonable bandwidth and intensity. For intense field conditions the physical resources scale moderately with the number of qubits in the transformation.

(3) The inversion problem of finding the field that induces a unitary transformation seems to be a hard numerical problem scaling unfavorably with the number of levels in the transformation.

ACKNOWLEDGMENTS

J.P.P. acknowledges financial support of the Gobierno de Canarias. This work was supported by Spanish MCT Grant No. BFM2001-3349, Gobierno de Canarias PI2002-009, and the Israel Science Foundation. The Fritz Haber Center is supported by the Minerva Gesellschaft für die Forschung, GmbH München, Germany. We thank Christiane Koch for her assistance and encouragement. Also we thank Zohar Amitay, David Tannor, Shlomo Sklarz, and Lajos Diosi for helpful discussions.

APPENDIX: THE VARIATIONAL METHOD

An alternative to the Krotov method of optimization is the variational method [1,21]. This method has been used previously in the simultaneous N state-to-state transitions formulation [12] and for the evolution operator formulation using the functional F_{re} [11]. In the last case the variational method was generalized in terms of the evolution equation for the unitary transformation. Unlike the Krotov method the variational method does not offer a direct algorithm to minimize F_{sm}. For simplicity only the optimization of the functional F_{ss} is discussed.

The variational method is based on the functional [12]

$$\begin{aligned} K(\{\psi_{il},\psi_{fl}\},\Delta\epsilon) &= \sum_{l=1}^{N} |\langle \psi_{il}(T)| \hat{\mathbf{O}} |l\rangle|^2 - \int_0^T \frac{\lambda_0}{s(t)} |\Delta\epsilon|^2 \, dt \\ &\quad - 2\,\mathrm{Re}\left[\sum_{l=1}^{N} \langle \psi_{il}(T)| \hat{\mathbf{O}} |l\rangle \right. \\ &\quad \left. \times \int_0^T \langle \psi_{fl}(t)| \left(\frac{d}{dt} + \frac{i}{\hbar} \hat{\mathbf{H}}(\tilde{\epsilon}+\Delta\epsilon) \right) |\psi_{il}(t)\rangle \right], \end{aligned} \tag{A1}$$

with the additional condition $|\psi_{il}(t=0)\rangle = |l\rangle$. The set of states $\{|l\rangle\}$ and the target unitary transformation $\hat{\mathbf{O}}$ were introduced in Sec. II. $\{|\psi_{il}(t)\}$ denotes the initial states driven by the field to the final states $\hat{\mathbf{O}}|l\rangle$. The terms $|\varphi_{fl}(t)\rangle$ are interpreted as Lagrange multipliers used as a constraint to impose the Schrödinger equation. The two first terms are equivalent to the functional (14) of the Krotov method. The parameter λ_0 is now interpreted as a Lagrange multiplier. The functional (A1) differs from the common formulation of OCT in the form of the field term $\tilde{\epsilon}+\Delta\epsilon$. $\tilde{\epsilon}$ is a reference field and $\Delta\epsilon$ must be interpreted as the correction used to converge to the optimal field that implements the target uni-

tary transformation. Setting $\tilde{\epsilon}=0$ and interpreting $\Delta\epsilon$ as the field the common form is re-attained.

By applying the calculus of variations, requiring $\delta K=0$, with respect to each element of the set $\{\psi_{il}(t)\}$, the evolution equations are reconstructed,

$$\frac{d}{dt}|\psi_{il}(t)\rangle=-\frac{i}{\hbar}\hat{\mathbf{H}}(\tilde{\epsilon}+\Delta\epsilon)|\psi_{il}(t)\rangle, \tag{A2}$$

with the condition $|\psi_{il}(t=0)\rangle=|l\rangle$ and formal solution $|\psi_{il}(t)\rangle=\hat{\mathbf{U}}(t,0;\tilde{\epsilon}+\Delta\epsilon)|l\rangle$. The variations with respect to the set $\{\psi_{fl}(t)\}$ gives

$$\frac{d}{dt}|\psi_{fl}(t)\rangle=-\frac{i}{\hbar}\hat{\mathbf{H}}(\tilde{\epsilon}+\Delta\epsilon)|\psi_{fl}(t)\rangle, \tag{A3}$$

with the condition $|\psi_{fl}(t=T)\rangle=\hat{\mathbf{O}}|l\rangle$. The formal solution is $|\psi_{fl}(t)\rangle=\hat{\mathbf{U}}(t,T;\tilde{\epsilon}+\Delta\epsilon)\hat{\mathbf{O}}|l\rangle$. Finally, variations with respect to $\Delta\epsilon$ lead to the correction to the field

$$\Delta\epsilon(t)=-\frac{s(t)}{\lambda_0\hbar}\mathrm{Im}\left[\sum_{l=1}^{N} b_l\langle l|\hat{\mathbf{O}}^{\dagger}\hat{\mathbf{U}}^{\dagger}(t,T;\tilde{\epsilon}+\Delta\epsilon)\times\hat{\boldsymbol{\mu}}\,\hat{\mathbf{U}}(t,0;\tilde{\epsilon}+\Delta\epsilon)|l\rangle\right], \tag{A4}$$

with

$$b_l=\langle l|\hat{\mathbf{U}}^{\dagger}(T,0;\tilde{\epsilon}+\Delta\epsilon)\hat{\mathbf{O}}|l\rangle. \tag{A5}$$

The correction to the field (A4) is the starting point of the iterative algorithm to find the optimal field. In such a case the correction to the field is implicit in the backward and forward propagation of the states in $\Delta\epsilon$. Several iterative methods have been proposed [22]. In the simplest approach, a guess field $\epsilon^{(0)}$ is used to evaluate $\Delta\epsilon$, which will be used to obtain the input field in the next iteration. Usually it does not converge. An alternative procedure [22] is to evaluate $\hat{\mathbf{U}}^{\dagger}$ in Eq. (A4) using the field in the previous iteration and then to simultaneously obtain the correction to the field and evaluate $\hat{\mathbf{U}}$ with the new field. This iterative algorithm is identical to the one obtained from the Krotov method in Sec. III. A study comparing different iterative algorithms based on the Krotov and variational methods for the problem of state-to-state optimization is described in Ref. [26].

[1] S.A. Rice and M. Zhao, *Optimal Control of Molecular Dynamics* (Wiley, New York, 2000).
[2] M. Shapiro and P. Brumer, *Principles of the Quantum Control of Molecular Processes* (Wiley, New Jersey, 2003).
[3] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[4] S. Lloyd, Phys. Rev. Lett. **75**, 346 (1995).
[5] S.G. Schirmer, A.D. Greentree, V. Ramakishna, and H. Rabitz, J. Phys. A **35**, 8315 (2002).
[6] L. Tian and S. Lloyd, Phys. Rev. A **62**, 050301 (2000).
[7] R. Zadoyan, D. Kohen, D.A. Lidar, and V.A. Apkarian, Chem. Phys. **266**, 323 (2001).
[8] Z. Amitay, R. Kosloff, and S.R. Leone, Chem. Phys. Lett. **359**, 8 (2002); J. Vala, Z. Amitay, B. Zhang, S.R. Leone, and R. Kosloff, Phys. Rev. A **66**, 062316 (2002).
[9] C.M. Tesch, L. Kurtz, and R. de Vivie-Riedle, Chem. Phys. Lett. **343**, 633 (2001).
[10] C. Rangan and P.H. Bucksbaum, Phys. Rev. A **64**, 033417 (2001).
[11] J.P. Palao and R. Kosloff, Phys. Rev. Lett. **89**, 188301 (2002).
[12] C.M. Tesch and R. de Vivie-Riedle, Phys. Rev. Lett. **89**, 157901 (2002).
[13] D. J. Tannor, V. Kazakov, and V. Orlov, in *Time Dependent Quantum Molecular Dynamics*, Vol. 299 of *NATO Advanced Study Institute, Series B: Physics* (Plenum, New York, 1992).
[14] A.M. Weiner, Rev. Sci. Instrum. **71**, 1929 (2000).
[15] T. Brixner and G. Gerber, Opt. Lett. **26**, 557 (2001).
[16] A. Bartana, R. Kosloff, and D.J. Tannor, Chem. Phys. **267**, 195 (2001).
[17] Y. Ohtsuki, J. Chem. Phys. **119**, 661 (2003).
[18] A. Bartana, R. Kosloff, and D.J. Tannor, J. Chem. Phys. **99**, 196 (1993).
[19] M. Sugawara, J. Chem. Phys. **118**, 6784 (2003).
[20] R. Kosloff, S.A. Rice, P. Gaspard, S. Tersigni, and D.J. Tannor, Chem. Phys. **139**, 201 (1989).
[21] A.P. Peirce, M.A. Dahleh, and H. Rabitz, Phys. Rev. A **37**, 4950 (1988).
[22] W. Zhu, J. Botina, and H. Rabitz, J. Chem. Phys. **108**, 1953 (1998).
[23] S.E. Sklarz and D.J. Tannor, Phys. Rev. A **66**, 053619 (2002).
[24] K. Sundermann and R. de Vivie-Riedle, J. Chem. Phys. **110**, 1896 (1999).
[25] T. Hornung, M. Motzkus, and R. de Vivie-Riedle, Phys. Rev. A **65**, 021403 (2002).
[26] Y. Maday and G. Turinici, J. Chem. Phys. **118**, 8191 (2003).
[27] Y.S. Weinstein, M.A. Pravia, E.M. Fortunato, S. Lloyd, and D.G. Cory, Phys. Rev. Lett. **86**, 1889 (2001).
[28] R. Kosloff, Annu. Rev. Phys. Chem. **45**, 145 (1994).
[29] V. Ramakrishna, K. Flores, H. Rabitz, and R.J. Ober, Phys. Rev. A **62**, 053409 (2000).
[30] V. Ramakrishna, M.V. Salapaka, M. Dahleh, H. Rabitz, and A. Peirce, Phys. Rev. A **51**, 960 (1995).
[31] G. Turinici and H. Rabitz, Chem. Phys. **267**, 1 (1995).
[32] S.G. Shirmer, A.I. solomon, and J.V. Leahy, J. Phys. A **35**, 8551 (2002).
[33] G.M. Huang, T.J. Tarn, and J.W. Clark, J. Math. Phys. **24**, 2608 (1983).
[34] Ronnie Kosloff, J. Phys.: Condens. Matter **92**, 2087 (1988).