New formulations of monotonically convergent quantum control algorithms

Cite as: J. Chem. Phys. **118**, 8191 (2003); https://doi.org/10.1063/1.1564043
Submitted: 16 October 2002 . Accepted: 06 February 2003 . Published Online: 23 April 2003

Yvon Maday, and Gabriel Turinici

New formulations of monotonically convergent quantum control algorithms

Yvon Maday[a)]
Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie, Boîte Courrier 187, 75252 Paris Cedex 05, France

Gabriel Turinici[b)]
INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt B.P. 105, 78153 Le Chesnay Cedex, France

(Received 16 October 2002; accepted 6 February 2003)

Most of the numerical simulation in quantum (bilinear) control have used one of the monotonically convergent algorithms of Krotov (introduced by Tannor *et al.*) or of Zhu and Rabitz. However, until now no explicit relationship has been revealed between the two algorithms in order to understand their common properties. Within this framework, we propose in this paper a unified formulation that comprises both algorithms and that extends to a new class of monotonically convergent algorithms. Numerical results show that the newly derived algorithms behave as well as (and sometimes better than) the well-known algorithms cited above. © *2003 American Institute of Physics.*
[DOI: 10.1063/1.1564043]

I. INTRODUCTION

Laser control of complex molecular and solid-state systems is becoming feasible, especially since the introduction[1] of closed loop laboratory learning techniques and their successful implementation.[2–7] However, at the level of the numerical simulations, much work is still to be done in order to bring the size of the systems that can be treated accurately to practical dimensions. Of course, using a good algorithm to solve the quantum control critical point equations is very important to the reduction of the overall cost; one of the most used algorithms is the Zhu and Rabitz[8] that extends an algorithm due to Krotov.[9] Both formulations share a common property: the algorithms are **guaranteed** to improve at each step a cost functional that describes the required control objectives. The introduction of this class of algorithms impulsionned much the early simulations on quantum control that were initially dependent on the nonlinear conjugate algorithm that displayed poor convergence for this highly nonlinear nonconvex cost functional. However, no general analysis to explain in depth the numerical properties of those two algorithms is available to date. In an attempt to fill this gap, the present paper presents a unified framework for the monotonically convergent algorithms that contains as particular cases the two classical methods cited above. The balance of the paper is as follows: The necessary background and definitions of the quantum control settings are given in the Sec. II; the unified formulation of the monotonically convergent algorithms is presented in Sec. III followed by relevant numerical results in Sec. IV. A discussion and some concluding remarks are presented in Sec. V. Finally, we present in the Appendix a method that takes even more advantage of the new algorithm formulations introduced in this paper.

a)Electronic mail: maday@ann.jussieu.fr
b)CERMICS-ENPC, Champs sur Marne, 77455 Marne la Vallée Cedex, France; Electronic mail: Gabriel.Turinici@inria.fr,

II. QUANTUM CONTROL SETTING AND OPTIMAL CONTROL EQUATIONS

Consider a quantum system with internal Hamiltonian H_0 prepared in the initial state $\Psi_0(x)$, where x denotes the relevant spatial coordinates; the state $\Psi(x,t)$ at time t satisfies the time-dependent Schrödinger equation (we set $\hbar=1$),

$$i\frac{\partial}{\partial t}\Psi(x,t)=H_0\Psi(x,t), \tag{1}$$

$$\Psi(x,t=0)=\Psi_0(x).$$

In the presence of an external interaction taken here as an electric field modeled by a coupling operator with amplitude $\epsilon(t)\in\mathbb{R}$ and a time independent dipole moment operator μ, the new Hamiltonian $H=H_0-\epsilon(t)\mu$ gives rise to the dynamical equations to be controlled,

$$\begin{cases} i\dfrac{\partial}{\partial t}\Psi(x,t)=(H_0-\epsilon(t)\mu)\Psi(x,t), \\ \Psi(x,t=0)=\Psi_0(x). \end{cases} \tag{2}$$

The **optimal control** approach allows us to assess the fitness of the final state $\Psi(T)$ to a prescribed goal. This is achieved through the introduction of a cost functional J to be maximized; this cost functional includes on the one hand terms that describe how well the objectives have been met and on the other hand terms that penalize undesired effects. One simple example of a cost functional is

$$J(\epsilon)=\langle\Psi(T)|O|\Psi(T)\rangle-\alpha\int_0^T\epsilon^2(t)dt, \tag{3}$$

where $\alpha>0$ is a parameter (it may also depend of time cf. Ref. 10) and O is the observable operator that encodes the

goal; the larger the value $\langle\Psi(T)|O|\Psi(T)\rangle$ the better the control objectives have been met; note that, in general, attaining the maximal possible value of $\langle\Psi(T)|O|\Psi(T)\rangle$ is at the price of a large laser fluence $\int_0^T \epsilon^2(t)dt$. The optimum evolution will therefore strike a balance between using a not too expensive laser fluence while simultaneously ensuring the desired observable has an acceptable (large) value.

The maximization of the cost functional $J(\epsilon)$ is realized by solving the Euler–Lagrange critical point equations; a standard way to write these equations is to introduce an *adjoint state* $\chi(x,t)$ (used as a Lagrange multiplier). The following critical point equations are thus obtained:[8]

$$\begin{cases} i\dfrac{\partial}{\partial t}\Psi(x,t)=(H_0-\epsilon(t)\mu)\Psi(x,t), \\ \Psi(x,t=0)=\Psi_0(x), \end{cases} \tag{4}$$

$$\begin{cases} i\dfrac{\partial}{\partial t}\chi(x,t)=(H_0-\epsilon(t)\mu)\chi(x,t), \\ \chi(x,t=T)=O\Psi(x,T), \end{cases} \tag{5}$$

$$\alpha\epsilon(t)=-\mathrm{Im}\langle\chi|\mu|\Psi\rangle(t). \tag{6}$$

III. UNIFIED FORMULATION OF MONOTONICALLY CONVERGENT ALGORITHMS

Efficient choices for solving in practice the critical point Eqs. (4)–(6) are given by the monotonically convergent algorithms[8,9] that are guaranteed to improve the cost functional J at each iteration. In the formulation of Krotov used by Tannor *et al.*[9] the optimization of J is achieved by iteratively solving at the step k,

$$\begin{cases} i\dfrac{\partial}{\partial t}\Psi^k(x,t)=(H_0-\epsilon^k(t)\mu)\Psi^k(x,t), \\ \Psi^k(x,t=0)=\Psi_0(x), \end{cases} \tag{7}$$

$$\epsilon^k(t)=-\frac{1}{\alpha}\,\mathrm{Im}\langle\chi^{k-1}|\mu|\Psi^k\rangle(t), \tag{8}$$

$$\begin{cases} i\dfrac{\partial}{\partial t}\chi^k(x,t)=(H_0-\epsilon^k(t)\mu)\chi^k(x,t), \\ \chi^k(x,t=T)=O\Psi^k(x,T). \end{cases} \tag{9}$$

Solving in practice Eqs. (7) and (8) is realized by propagating the nonlinear equation,

$$\begin{cases} i\dfrac{\partial}{\partial t}\Psi^k(x,t)=\left(H_0+\dfrac{1}{\alpha}\,\mathrm{Im}\langle\chi^{k-1}|\mu|\Psi^k\rangle(t)\mu\right)\Psi^k(x,t), \\ \Psi^k(x,t=0)=\Psi_0(x). \end{cases} \tag{10}$$

Designed to improve the convergence of the above algorithm, the formulation Zhu and Rabitz[8] introduces the nonlinear propagation also in the adjoint equation; the overall algorithm can be written as

$$\begin{cases} i\dfrac{\partial}{\partial t}\Psi^k(x,t)=(H_0-\epsilon^k(t)\mu)\Psi^k(x,t), \\ \Psi^k(x,t=0)=\Psi_0(x), \end{cases} \tag{11}$$

$$\epsilon^k(t)=-\frac{1}{\alpha}\,\mathrm{Im}\langle\chi^{k-1}|\mu|\Psi^k\rangle(t), \tag{12}$$

$$\begin{cases} i\dfrac{\partial}{\partial t}\chi^k(x,t)=(H_0-\tilde{\epsilon}^k(t)\mu)\chi^k(x,t), \\ \chi^k(x,t=T)=O\Psi^k(x,T), \end{cases} \tag{13}$$

$$\tilde{\epsilon}^k(t)=-\frac{1}{\alpha}\,\mathrm{Im}\langle\chi^k|\mu|\Psi^k\rangle(t). \tag{14}$$

Remark 1: Finding the precise mathematical setting that allow to give a rigorous proof of the convergence of ϵ^k and Ψ^k to their respective limit values is still an open problem.

Despite the apparent similarity between the two algorithms no studies were pursued to better understand the relationships between the two formulations. However some relationships exist and the formulation below, the main topic of this paper, will make this explicit. Consider the algorithm described by the resolution of the following equations at step k:

$$\begin{cases} i\dfrac{\partial}{\partial t}\Psi^k(x,t)=(H_0-\epsilon^k(t)\mu)\Psi^k(x,t), \\ \Psi^k(x,t=0)=\Psi_0(x), \end{cases} \tag{15}$$

$$\epsilon^k(t)=(1-\delta)\tilde{\epsilon}^{k-1}(t)-\frac{\delta}{\alpha}\,\mathrm{Im}\langle\chi^{k-1}|\mu|\Psi^k\rangle(t), \tag{16}$$

$$\begin{cases} i\dfrac{\partial}{\partial t}\chi^k(x,t)=(H_0-\tilde{\epsilon}^k(t)\mu)\chi^k(x,t), \\ \chi^k(x,t=T)=O\Psi^k(x,T), \end{cases} \tag{17}$$

$$\tilde{\epsilon}^k(t)=(1-\eta)\epsilon^k(t)-\frac{\eta}{\alpha}\,\mathrm{Im}\langle\chi^k|\mu|\Psi^k\rangle(t). \tag{18}$$

We recognize the Zhu and Rabitz algorithm for $\delta=1$ and $\eta=1$ and the Krotov (as in Tannor *et al.*[9]) formulation for $\delta=1$ and $\eta=0$. Of course, this general formulation remains only a mathematical artifice if monotonic convergence properties fails to hold for the above algorithm. The proof that monotonicity is still valid for this formulation is given in the following:

Theorem 1: *Suppose O is a self-adjoint positive semidefinite observable. Then, for any $\eta,\delta\in[0,2]$ the algorithm given in Eqs. (15)–(18) converges monotonically in the sense that $J(\epsilon^{k+1})\geqslant J(\epsilon^k)$.*

Proof: Let us evaluate the difference between the values of the cost functional between two successive iterations. Suppose that $\eta\neq 0$, $\delta\neq 0$. Then,

$$J(\epsilon^{k+1})-J(\epsilon^k)=\langle\Psi^{k+1}(T)|O|\Psi^{k+1}(T)\rangle-\alpha\int_0^T\epsilon^{k+1}(t)^2dt-\langle\Psi^k(T)|O|\Psi^k(T)\rangle+\alpha\int_0^T\epsilon^k(t)^2dt$$

$$=\langle\Psi^{k+1}(T)-\Psi^k(T)|O|\Psi^{k+1}(T)-\Psi^k(T)\rangle+2\,\mathrm{Re}\langle\Psi^{k+1}(T)-\Psi^k(T)|O|\Psi^k(T)\rangle+\alpha\int_0^T\epsilon^k(t)^2dt$$

$$-\alpha\int_0^T\epsilon^{k+1}(t)^2dt. \tag{19}$$

Since we also have

$$2\,\mathrm{Re}\langle\Psi^{k+1}(T)-\Psi^k(T)|O|\Psi^k(T)\rangle=2\,\mathrm{Re}\langle\Psi^{k+1}(T)-\Psi^k(T),O\Psi^k(T)\rangle$$

$$=2\,\mathrm{Re}\langle\Psi^{k+1}(T)-\Psi^k(T),\chi^k(T)\rangle$$

$$=2\,\mathrm{Re}\int_0^T\left\langle\frac{\partial(\Psi^{k+1}(t)-\Psi^k(t))}{\partial t},\chi^k(t)\right\rangle+\left\langle\Psi^{k+1}(t)-\Psi^k(t),\frac{\partial\chi^k(t)}{\partial t}\right\rangle dt$$

$$=2\,\mathrm{Re}\int_0^T\left\langle\frac{H_0-\mu\epsilon^{k+1}}{i}\Psi^{k+1}(t)-\frac{H_0-\mu\epsilon^k}{i}\Psi^k(t),\chi^k(t)\right\rangle$$

$$+\left\langle\Psi^{k+1}(t)-\Psi^k(t),\frac{H_0-\mu\tilde{\epsilon}^k}{i}\chi^k(t)\right\rangle dt$$

$$=2\,\mathrm{Re}\int_0^T\epsilon^{k+1}\left\langle\frac{-\mu}{i}\Psi^{k+1}(t),\chi^k(t)\right\rangle-\epsilon^k\left\langle\frac{-\mu}{i}\Psi^k(t),\chi^k(t)\right\rangle$$

$$+\tilde{\epsilon}^k\left\langle\Psi^{k+1}(t)-\Psi^k(t),\frac{-\mu}{i}\chi^k(t)\right\rangle dt$$

$$=2\int_0^T\epsilon^{k+1}\cdot\frac{\alpha(\epsilon^{k+1}-(1-\delta)\tilde{\epsilon}^k)}{\delta}-\epsilon^k\cdot\frac{\alpha(\tilde{\epsilon}^k-(1-\eta)\epsilon^k)}{\eta}-\tilde{\epsilon}^k\cdot\frac{\alpha(\epsilon^{k+1}-(1-\delta)\tilde{\epsilon}^k)}{\delta}$$

$$+\tilde{\epsilon}^k\cdot\frac{\alpha(\tilde{\epsilon}^k-(1-\eta)\epsilon^k)}{\eta}dt, \tag{20}$$

we thus obtain

$$J(\epsilon^{k+1})-J(\epsilon^k)=\langle\Psi^{k+1}(T)-\Psi^k(T)|O|\Psi^{k+1}(T)-\Psi^k(T)\rangle$$

$$+\alpha\int_0^T\left(\frac{2}{\delta}-1\right)(\epsilon^{k+1}-\tilde{\epsilon}^k)^2+\left(\frac{2}{\eta}-1\right)(\tilde{\epsilon}^k-\epsilon^k)^2, \tag{21}$$

which is positive under the hypothesis of the theorem. Note that the case $\delta=0$ yields $\epsilon^{k+1}=\tilde{\epsilon}^k$ and $\eta=0$ yields $\tilde{\epsilon}^k=\epsilon^k$ and the conclusion is reached after straightforward term elimination.

Each step of this algorithm will therefore result in an increase of the value of the cost functional; this increase is expected to be important for initial steps where the critical point equations are not fulfilled and the difference between successive fields ϵ^k and ϵ^{k+1} will be important.

Remark 2: Although initially seen as a relaxation procedure, Eqs. (16)–(18) proved to generate monotonic algorithms even for values of δ and η larger than 1 which is not an intuitive fact; as it will be seen in the next section this leads to new numerical behavior of the algorithms.

Although not conceptually different, an algorithm similar to Eqs. (15)–(18) can be devised,

$$\begin{cases} i\dfrac{\partial}{\partial t}\Psi^k(x,t)=(H_0-\epsilon^k(t)\mu)\Psi^k(x,t), \\ \Psi^k(x,t=0)=\Psi_0(x), \end{cases} \tag{22}$$

$$\epsilon^k(t)=(1-\delta)\epsilon^{k-1}(t)-\frac{\delta}{\alpha}\,\mathrm{Im}\langle\chi^{k-1}|\mu|\Psi^k\rangle(t), \tag{23}$$

$$\begin{cases} i\dfrac{\partial}{\partial t}\chi^k(x,t)=(H_0-\tilde{\epsilon}^k(t)\mu)\chi^k(x,t), \\ \chi^k(x,t=T)=O\Psi^k(x,T), \end{cases} \tag{24}$$

$$\tilde{\epsilon}^k(t)=(1-\eta)\epsilon^k(t)-\frac{\eta}{\alpha}\,\mathrm{Im}\langle\chi^k|\mu|\Psi^k\rangle(t), \tag{25}$$

for which the following convergence theorem holds:

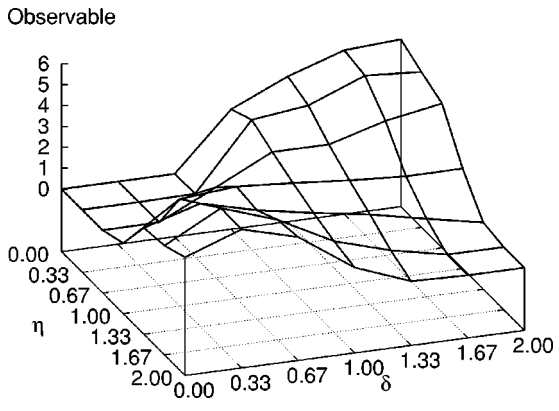

FIG. 1. First test case: dependence of the obtained observable $\langle\Psi^5|O|\Psi^5\rangle$ of the parameters δ and η. Best choices of parameters are $(\delta=2,\eta=0)$, $(\delta=5/3,\eta=0)$, $(\delta=4/3,\eta=0)$,$(\delta=1,\eta=0)$, $(\delta=5/3,\eta=1/3)$, and $(\delta=2,\eta=1/3)$.

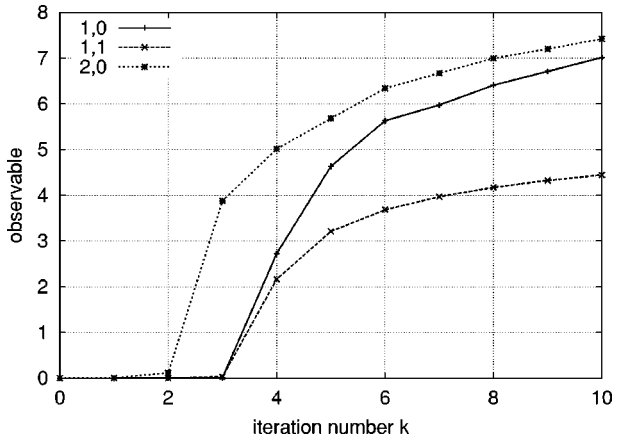

FIG. 2. Evolution of the observable for the first test case and different parameter values (δ,η).

Theorem 2: *Suppose O is a self-adjoint positive semidefinite observable. Then for any $\delta\in[0,2]$, $0\leqslant\eta\leqslant\delta(2-\delta)$ the algorithm given in Eqs. (22)–(25) converges monotonically in the sense that $J(\epsilon^{k+1})\geqslant J(\epsilon^k)$.*

Proof: The proof of this result being similar to the proof of Theorem 1 is left as an exercise to the reader.

IV. NUMERICAL RESULTS

In order to test the performances of the algorithm, two cases already treated in the literature were considered.[8] The system under consideration is the $O-H$ bond that vibrates in a Morse type potential. We refer the reader to Ref. 8 for the numerical details concerning this system. The goal is to localize the wave packet at a given location x'; this is expressed via the observable $O(x)=(\gamma_0/\sqrt{\pi})e^{-\gamma_0^2(x-x')^2}$ through the requirement that $\langle\Psi|O|\Psi\rangle$ is maximized. The two cases analyzed differ in the initial state used to control the system: in the first example this is the ground state while in the second case it is the first excited state, and also in the value of the parameter x': $x'=2.5$ and $x'=1.821$, respectively.

Numerical results are presented in Figs. 1, 2, and 3. In all the figures the plotted quantity is the observable (and not the cost functional). In order to compare possible choices for the parameters δ and η we plot in Fig. 1 the surfaces of the observables values $\langle\Psi^k|O|\Psi^k\rangle$ for the iteration $k=5$ for different choices of the parameters δ and η. It is seen that the choice $(\delta=2,\eta=0)$ is very good and that it outperforms (in this particular case) the algorithms of Zhu and Rabitz $(\delta=1,\eta=1)$ and Krotov $(\delta=1,\eta=0)$ which still display good performance.

Remark 3: The numerical procedure used to propagate the state and the electric field was chosen to be first order for the field update as in Ref. 11, Eq. (46) and second order split-operator for the wave function update [the same results are also obtained with the second order scheme as in Ref. 11, Eq. (47)]; all cases share the same numerical scheme.

A detailed plot of the evolution of the observable $\langle\Psi^k|O|\Psi^k\rangle$ during the iterations k is presented in Figs. 2 and 3 for different choices of the parameters δ and η. In both cases the choice $(\delta=2,\eta=0)$ is giving the fastest algorithm. Note that, in practice, the numerical implementation of the schemes requires that, for large values of δ and η, very small time step be used (especially after several iterations at about $k>5$). This diminution of the time step is related, in our opinion, to the **nonlinearity** induced by these parameters in the corresponding evolution Eqs. (15) and (17); larger δ and η are, more nonlinear the evolution Eqs. (15) and (17) become. Of course, even if from the numerical point of view a nonlinear equation requires small time steps and it is therefore not convenient to treat, in this control setting it happened to have the merit of leading very fast to fields $\epsilon(t)$ that allows to control the system. For instance, for the first case, which is a difficult setting because the initial state has almost no overlap with the target, it is seen from Fig. 2 that the scheme with $(\delta=2,\eta=0)$ is able to find at iteration $k=3$ an acceptable laser field, one iteration in advance with respect to other schemes that at $k=3$ still have negligible overlap with the target. It is not obvious *a priori* to determine the best values for δ and η; in addition these optimal values may vary from one case to the other. To address this, we propose

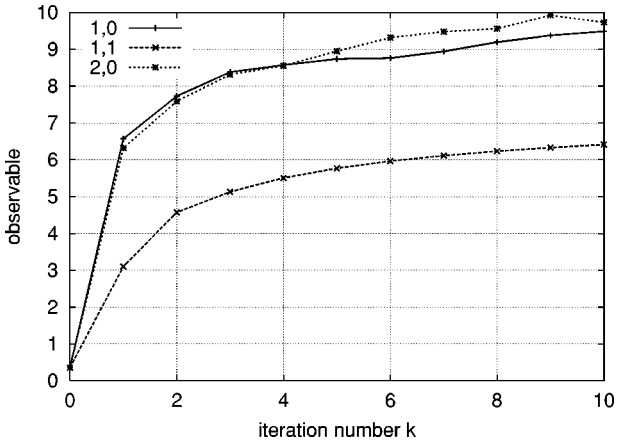

FIG. 3. Evolution of the observable for the second test case and different parameter values (δ,η).

in the Appendix a method that automatically search these values. It will be extended in future works.[12]

Let us mention that small values of the parameters δ and η may also prove interesting as they allow for almost linear evolution equations which may stabilize sensible numerical schemes.[13]

Finally, let us remark that, although derived for constant parameter α, nothing prevents us from using this scheme in the framework of a time-dependent $\alpha(t)$ as in Ref. 10.

V. DISCUSSION AND CONCLUSIONS

In an attempt to extend the class of monotonically convergent algorithms used for the numerical resolution of quantum control problems, a new class of algorithms has been introduced in this paper; its convergence properties were proven in Theorems 1 and 2 and their numerical properties were presented in Sec. IV.

The existence of this new procedure gives the flexibility to choose between different alternatives according to the difficulty of the problem to be treated; the earlier schemes $\delta=1,\eta=1$ or $\delta=1,\eta=0$ may thus be complemented for instance with the scheme $\delta=2,\eta=0$ that may improve the initial convergence to the target (eventually switching to smaller δ values after the initial iterations if numerical instabilities due to the large time steps are experienced); the suggested implementation is the automatic search for optimal parameters as in the Appendix. On the contrary, if for some reason a less nonlinear evolution is desired we suggest using from the beginning small values of the parameters δ and η.

APPENDIX A: DERIVATION OF OPTIMALITY EQUATIONS FOR PARAMETER δ WHEN $\eta=0$

In a search of an optimal setting of the δ and η parameters one can consider an automatic update of δ,η during the iterations in k in Eqs. (15)–(18). We will present here the necessary theoretical computations in a simplified case where $\eta=0$. The goal of this method is to update the parameter δ after each interation in $k+1$ with the value that maximizes the quantity $J(\epsilon^{k+1})-J(\epsilon^k)$. The computation of the gradient of $J(\epsilon^{k+1})-J(\epsilon^k)$ with respect to δ is then required. Suppose that in Eqs. (15)–(18) Ψ^k is computed with the field ϵ^k and Ψ^{k+1} is computed with a given δ. Then it follows from Eq. (21) that

$$\begin{aligned}\frac{\partial}{\partial\delta}(J(\epsilon^{k+1})-J(\epsilon^k))&=2\,\mathrm{Re}\left\langle\frac{\partial}{\partial\delta}\Psi^{k+1}(T)\middle|O\middle|\Psi^{k+1}(T)-\Psi^k(T)\right\rangle\\&\quad+\frac{\partial}{\partial\delta}\left\{\alpha\int_0^T\left(\frac{2}{\delta}-1\right)\right.\\&\quad\left.\times\delta^2\left(\epsilon^k+\frac{1}{\alpha}\,\mathrm{Im}\langle\chi^k|\mu|\Psi^{k+1}\rangle(t)\right)^2dt\right\}.\end{aligned}\tag{A1}$$

Indeed fact that $\eta=0$ imply

$$\frac{\epsilon^{k+1}-\epsilon^k}{\delta}=-\epsilon^k-\frac{1}{\alpha}\,\mathrm{Im}\langle\chi^k|\mu|\Psi^{k+1}\rangle(t).$$

We can write from Eq. (A1),

$$\begin{aligned}&\frac{\partial}{\partial\delta}(J(\epsilon^{k+1})-J(\epsilon^k))\\&\quad=2\,\mathrm{Re}\left\langle\frac{\partial\Psi^{k+1}(T)}{\partial\delta}\middle|O\middle|\Psi^{k+1}(T)-\Psi^k(T)\right\rangle\qquad\text{(A2)}\\&\quad+2\alpha(1-\delta)\int_0^T\left(\frac{\epsilon^{k+1}-\epsilon^k}{\delta}\right)^2dt+2(2-\delta)\\&\quad\times\int_0^T(\epsilon^k-\epsilon^{k+1})\,\mathrm{Im}\left\langle\chi^k\middle|\mu\middle|\frac{\partial\Psi^{k+1}}{\partial\delta}\right\rangle(t)dt.\qquad\text{(A3)}\end{aligned}$$

Note that

$$\begin{aligned}\frac{\partial\epsilon^{k+1}}{\partial\delta}&=-\epsilon^k-\frac{1}{\alpha}\,\mathrm{Im}\langle\chi^k|\mu|\Psi^{k+1}\rangle(t)\\&\quad-\frac{\delta}{\alpha}\,\mathrm{Im}\left\langle\chi^k\middle|\mu\middle|\frac{\partial\Psi^{k+1}}{\partial\delta}\right\rangle\\&=\frac{\epsilon^{k+1}-\epsilon^k}{\delta}-\frac{\delta}{\alpha}\,\mathrm{Im}\left\langle\chi^k\middle|\mu\middle|\frac{\partial\Psi^{k+1}}{\partial\delta}\right\rangle.\end{aligned}$$

We can **formally** derive from Eq. (15) the evolution equation for $\partial\Psi^{k+1}/\partial\delta$,

$$\begin{aligned}i\frac{\partial}{\partial t}\left(\frac{\partial\Psi^{k+1}}{\partial\delta}\right)&=(H_0-\epsilon^{k+1}(t)\mu)\left(\frac{\partial\Psi^{k+1}}{\partial\delta}\right)-\left(\frac{\epsilon^{k+1}-\epsilon^k}{\delta}\right.\\&\quad\left.-\frac{\delta}{\alpha}\,\mathrm{Im}\left\langle\chi^k\middle|\mu\middle|\frac{\partial\Psi^{k+1}}{\partial\delta}\right\rangle\right)\mu\Psi^{k+1},\qquad\text{(A4)}\end{aligned}$$

$$\frac{\partial\Psi^{k+1}}{\partial\delta}(t=0)=0.\tag{A5}$$

Let us now consider an additional function $\Lambda(t)$ defined at time T by $\Lambda(T)=O(\Psi^{k+1}-\Psi^k)(T)$ and at all other times $t\leqslant T$ by an evolution equation to be made precise later. It follows that

$$\begin{aligned}&2\,\mathrm{Re}\left\langle\frac{\partial\Psi^{k+1}(T)}{\partial\delta}\middle|O\middle|\Psi^{k+1}(T)-\Psi^k(T)\right\rangle\\&\quad=2Re\left\langle\frac{\partial\Psi^{k+1}(T)}{\partial\delta},\Lambda(T)\right\rangle\\&\quad=2\,\mathrm{Re}\int_0^T\left\langle\frac{\partial}{\partial t}\left(\frac{\partial\Psi^{k+1}}{\partial\delta}\right)(t),\Lambda(t)\right\rangle\\&\quad+\left\langle\frac{\partial\Psi^{k+1}(t)}{\partial\delta},\frac{\partial}{\partial t}\Lambda(t)\right\rangle dt.\end{aligned}\tag{A6}$$

Replacing Eq. (A4) in Eq. (A6) we obtain

$$2\,\mathrm{Re}\left\langle \frac{\partial \Psi^{k+1}(T)}{\partial \delta}\middle| O \middle| \Psi^{k+1}(T)-\Psi^{k}(T)\right\rangle$$
$$=2\,\mathrm{Re}\int_0^T \left\langle i\mu\frac{\epsilon^{k+1}-\epsilon^{k}}{\delta}\Psi^{k+1},\Lambda\right\rangle \tag{A7}$$
$$+\left\langle \frac{\partial \Psi^{k+1}(t)}{\partial \delta}, i(H_0-\epsilon^{k+1}(t)\mu)\Lambda+\frac{\partial \Lambda}{\partial t}\right.$$
$$\left.-i\left(\frac{\delta}{\alpha}\mathrm{Re}\langle i\mu\Psi^{k+1},\Lambda\rangle\right)\mu\chi^{k}\right\rangle dt. \tag{A8}$$

It follows then that

$$\frac{\partial (J(\epsilon^{k+1})-J(\epsilon^{k}))}{\partial \delta}$$
$$=2\alpha(1-\delta)\int_0^T\left(\frac{\epsilon^{k+1}-\epsilon^{k}}{\delta}\right)^2 dt$$
$$+2\,\mathrm{Re}\int_0^T\left\langle i\mu\frac{\epsilon^{k+1}-\epsilon^{k}}{\delta}\Psi^{k+1},\Lambda\right\rangle \tag{A9}$$
$$+2\,\mathrm{Re}\int_0^T\left\langle \frac{\partial \Psi^{k+1}(t)}{\partial \delta}, i(H_0-\epsilon^{k+1}\mu)\Lambda+\frac{\partial \Lambda}{\partial t}\right.$$
$$\left.-i\left(\frac{\delta}{\alpha}\,\mathrm{Re}\langle i\mu\Psi^{k+1},\Lambda\rangle+(2-\delta)(\epsilon^{k+1}-\epsilon^{k})\right)\mu\chi^{k}\right\rangle dt. \tag{A10}$$

If suffices now to set the evolution equation for Λ as

$$i\frac{\partial \Lambda}{\partial t}=(H_0-\epsilon^{k+1}\mu)\Lambda-\left(\frac{\delta}{\alpha}\,\mathrm{Re}\langle i\mu\Psi^{k+1},\Lambda\rangle\right.$$
$$\left.+(2-\delta)(\epsilon^{k+1}-\epsilon^{k})\right)\mu\chi^{k}, \tag{A11}$$
$$\Lambda(T)=O(\Psi^{k+1}-\Psi^{k})(T), \tag{A12}$$

to get

$$\frac{\partial (J(\epsilon^{k+1})-J(\epsilon^{k}))}{\partial \delta}=2\alpha(1-\delta)\int_0^T\left(\frac{\epsilon^{k+1}-\epsilon^{k}}{\delta}\right)^2 dt$$
$$+2\,\mathrm{Re}\int_0^T\left\langle i\mu\frac{\epsilon^{k+1}-\epsilon^{k}}{\delta}\Psi^{k+1},\Lambda\right\rangle. \tag{A13}$$

Formulas (A11)–(A13) allow us to compute the gradient of the cost functional improvement $J(\epsilon^{k+1})-J(\epsilon^{k})$ with respect to the parameter δ. Simple methods can be imagined that update at each iteration k the parameter δ so that it approaches the optimal value that maximizes the cost functional increase. The following algorithm can therefore be proposed: At each iteration k solve the Eqs. (15)–(18) with $\delta=\delta_k$ and $\eta=0$, together with Eqs. (A11)–(A13); set then for the next $k+1$th iteration,

$$\delta_{k+1}=1+\frac{2\,\mathrm{Re}\int_0^T\left\langle i\mu\frac{\epsilon^{k+1}-\epsilon^{k}}{\delta_k}\Psi^{k+1},\Lambda\right\rangle dt}{2\alpha\int_0^T\left(\frac{\epsilon^{k+1}-\epsilon^{k}}{\delta_k}\right)^2 dt}, \tag{A14}$$

when this quantity is between 0 and 2. When δ_{k+1} computed by Eq. (A14) is negative set it to zero and when is larger than 2set it to 2.

Remark 4: The resolution of Eqs. (A11)–(A13) only takes half the effort required to solve Eqs. (15)–(18) and therefore the automatic optimization procedure proposed above is 50% more expensive than fixed parameter computations (in addition parallel processing can also be envisioned). In practice we have noticed that the speed-up achieved by these improvements including also the choice $\eta\neq 0$ compensates for the increased cost and in addition gives rise to more robust algorithms. As yet, the results have only been verified on a few number of cases and further investigations, necessary in order to quantitatively assess these improvements, will be reported elsewhere.

[1] R. S. Judson and H. Rabitz, Phys. Rev. Lett. **68**, 1500 (1992).
[2] R. J. Levis, G. Menkir, and H. Rabitz, Science **292**, 709 (2001).
[3] A. Assion, T. Baumert, M. Bergt, T. Brixner, B. Kiefer, V. Seyfried, M. Strehle, and G. Gerber, Science **282**, 919 (1998).
[4] M. Bergt, T. Brixner, B. Kiefer, M. Strehle, and G. Gerber, J. Phys. Chem. A **103**, 10381 (1999).
[5] T. Weinacht, J. Ahn, and P. Bucksbaum, Nature (London) **397**, 233 (1999).
[6] C. Bardeen, V. V. Yakovlev, K. R. Wilson, S. D. Carpenter, P. M. Weber, and W. S. Warren, Chem. Phys. Lett. **280**, 151 (1997).
[7] C. J. Bardeen, V. V. Yakovlev, J. A. Squier, and K. R. Wilson, J. Am. Chem. Soc. **120**, 13023 (1998).
[8] W. Zhu and H. Rabitz, J. Chem. Phys. **109**, 385 (1998).
[9] D. Tannor, V. Kazakov, and V. Orlov, in *Time Dependent Quantum Molecular Dynamics*, edited by J. Broeckhove and L. Lathouwers (Plenum, New York, 1992), pp. 347–360.
[10] T. Hornung, M. Motzkus, and R. de Vivie-Riedle, J. Chem. Phys. **115**, 3105 (2001).
[11] W. Zhu, J. Botina, and H. Rabitz, J. Chem. Phys. **108**, 1953 (1998).
[12] J. Salomon, Ph.D. thesis, Paris VI University (unpublished).
[13] Y. Maday and G. Turinici, I.J.Q.C. (unpublished).