Eur. Phys. J. D **14**, 141–145 (2001)

THE EUROPEAN PHYSICAL JOURNAL D

EDP Sciences
© Società Italiana di Fisica
Springer-Verlag 2001

Non-iterative optimal design of quantum controls

Z. Murtha and H. Rabitz[a]

Department of Chemistry, Princeton University, Princeton, NJ 08544, USA

Received 30 June 2000 and Received in final form 22 November 2000

Abstract. A non-iterative means for quantum control design is introduced with the aim of offering practical designs that can later be fine-tuned with laboratory closed-loop techniques. The procedure recognizes that Hamiltonians for realistic system control applications are rarely known accurately. The algorithm takes advantage of this fact by allowing for managed deviations in the equations of motion, thus removing the standard Lagrange multiplier. Suitable time-dependent cost functional weights are introduced that eliminate the traditional final time matching condition, thereby producing non-iterative design equations as an initial value problem. Removal of the final time condition also eliminates the demand that the target state be reached at any artificially imposed time. Tests on a simple molecular system indicate that the algorithm leads to well-behaved designs and that the weight functions are adequately estimated by order of magnitude analysis.

PACS. 02.30.Xx Calculus of variations – 02.30.Yy Control theory – 33.80.Wz Other multiphoton processes

1 Introduction

There is active interest in the control of quantum systems, especially stimulated by the recent successes of closed-loop learning control methods in the laboratory [1–4]. These techniques are capable of selectively breaking chemical bonds and controlling special atomic and molecular excitations by rapidly performing many experiments, fine-tuning the laser pulses, and gradually homing in on an optimal field for the objective. This learning procedure would benefit from a practical design method for narrowing the space of search possibilities, but optimal control theory [5–10] and tracking theory [11,12] are not always feasible because of several computational difficulties. An application of optimal control theory can require lengthy iterative integration before converging to an optimal field. Tracking theory has the advantage of requiring a single solution of Schrödinger's equation, but it has the disadvantage of attempting to guide the system over a specific path toward the target, which is a constraint that may be overly demanding [13]. Furthermore, both design procedures are based on calculations with Hamiltonians which are not known well for most systems of interest. A theoretical design method that minimizes both computation time and controller intuition on the subtle control complexities would be a useful tool. In formulating such a method, it would also be desirable to take advantage of the lack of precise knowledge of the Hamiltonian, rather than have it as a burden. Ideally, we would also like to combine the optimization features of optimal control with the single solution of the Schrödinger equation offered by tracking. In addition, we desire to lift the requirement of terminating the algorithm at an arbitrarily imposed final time τ. The algorithm proposed in this paper has these combined features to efficiently estimate control design for laboratory refinement.

Most current quantum optimal control design techniques seek to find the control field, $\varepsilon(t)$, that minimizes a cost functional, generally written as [6]

$$J = \int_0^\tau \mathrm{d}t\, \alpha(t) \left(\langle \mathcal{O}(t) \rangle - \tilde{\mathsf{O}} \right)^2 + \int_0^\tau \mathrm{d}t\, \beta(t) \varepsilon^2(t) + \int_0^\tau \mathrm{d}t\, \langle \lambda(t) | \Omega | \psi \rangle + \text{c.c.} \quad (1)$$

where c.c. denotes complex conjugate, $\tilde{\mathsf{O}}$ is the target value for $\langle \mathcal{O}(\tau) \rangle$, and $\Omega = \mathcal{H} - \mathrm{i}\hbar\partial/\partial t$. The Hamiltonian $\mathcal{H} = \mathcal{H}_\circ - \mu\varepsilon(t)$ has the standard form where $\mathcal{H}_\circ$ describes the free motion and μ is the system dipole. Here, $\alpha(t) \geq 0$, and $\beta(t) \geq 0$ weight the importance of nearing the target and minimizing the energy fluence of the field, respectively. The Lagrange multiplier state function $|\lambda(t)\rangle$ is introduced to guarantee that the solution propagates precisely according to the Hamiltonian $\mathcal{H}$ in Schrödinger's equation.

Minimizing equation (1) raises two vexing computational issues. First, the minimization leads directly to a final time condition for $|\lambda(\tau)\rangle$, which converts the design calculation into a burdensome boundary value problem in time. Second, the presence of the Lagrange multiplier forces the propagation to be exactly governed by

[a] e-mail: hrabitz@princeton.edu

the Hamiltonian, which is not known well for virtually all non-trivial applications. Thus, the third term of equation (1) artificially constrains the minimization with equations that are known to be inexact. The present paper seeks to address these design issues by introducing a new cost functional with the form

$$J = \int_0^{t_\mathrm{f}} \mathrm{d}t\, \alpha(t) \left(\langle \mathcal{O}(t) \rangle - \tilde{\mathsf{O}} \right)^2 + \int_0^{t_\mathrm{f}} \mathrm{d}t\, \beta(t) \varepsilon^2(t) + \int_0^{t_\mathrm{f}} \mathrm{d}t\, \langle \Omega\psi | \Gamma(t) | \Omega\psi \rangle \quad (2)$$

where a new positive semi-definite operator $\Gamma(t)$ is introduced to weight the deviation from exact satisfaction of the equations of motion. A suitable choice for $\Gamma(t)$ will allow for finding an optimal field without artificially restricting the design to exactly follow the motion of an imprecise reference Hamiltonian. The first and second terms of the new cost functional remain unchanged from conventional optimal design, but the inclusion of the $\Gamma(t)$ term in the equation to be minimized is unique to this non-iterative design method. Also, the upper integration limit in equation (2), t_f, is no longer a fixed final time τ, but rather a floating time which is allowed to grow until satisfactory control is achieved.

Imperfectly satisfied equations of motion have been treated previously through consideration of random disturbances or constraints that affect the robustness of the field [14]. A penalty algorithm [15] like equation (2) has been considered, but with a sequence of Γ weights in an iterative homotopy method to achieve the solution to minimizing equation (1). The present paper will seek to demonstrate that the cost functional in equation (2) can yield good design solutions with minimal computational effort. Section 2 will outline the nature of design with equation (2). Section 3 will illustrate the design procedure with a simple system. Some brief final comments are given in Section 4.

2 Features of the design cost functional

For equation (2) to offer any advantage over previous quantum design control methods, it must simplify the computational issues raised in Section 1. This approach will have merit if it removes excess restrictions and dependence on the input. Because closed-loop laboratory learning techniques are capable of fine-tuning control fields, a design method that inexpensively produces a reasonable control field estimate will provide an adequate operational framework.

The weight function $\Gamma(t)$ has an important role in the variational equation arising from equation (2). A carefully considered choice for $\Gamma(t)$ will allow a solution that works successfully within a suitably small window around the reference input Hamiltonian. This approach eliminates the previous control design problem of forcing the solution to satisfy the dynamics of an estimated Hamiltonian.

Another limitation of previous optimal design techniques was the requirement of specifying a final time, τ, which directly resulted in the burdensome process of iterative integration of the design equations. Because we are now free to set $\Gamma(t)$ at our discretion, we may impose the condition

$$\lim_{t \to \infty} \Gamma(t) \to 0. \quad (3)$$

In practice, the infinite time limit of $\Gamma(t)$ will never be taken, as we only require that $\Gamma(t)$ sufficiently diminish over a physically acceptable control period t_f in equation (2). Equation (3) will cause the final time condition to automatically vanish, thus eliminating explicit dependence of the solution on τ. Eliminating the final time matching criteria on the dynamical state converts the design process from a boundary value problem in time into an initial value problem. This conversion is critical since initial value problems do not require iterative solutions. Thus, we can gain the same non-iterated advantage of tracking control, but without the burden of actually imposing a specific path to the target.

The utility of equation (3) is more readily seen when the behavior of $\alpha(t)$ and $\Gamma(t)$ are considered together. Generally, the control objective is to achieve the target in finite time, without concern for the path taken. With that logic, it is reasonable to conclude that $\alpha(t)$ should be constructed as a smoothly increasing function to enhance the likelihood of achieving the target as the design progresses. On the other hand, $\Gamma(t)$ will be specified as a smoothly decreasing function of time to satisfy equation (3). The only other criterion on $\alpha(t)$, $\beta(t)$, and $\Gamma(t)$ calls for their magnitudes to roughly balance the terms in equation (2). These magnitudes may be simply estimated from the expected system dynamics. Beyond these loose physically imposed specifications on $\alpha(t)$, $\beta(t)$, and $\Gamma(t)$, there should be wide latitude in their choices. This expectation is verified in the illustration of Section 3.

These weight function considerations will result in the minimization of equation (2) producing associated dynamics that closely adhere to the equations of motion at short times, but become increasingly driven toward the target as the time t_f progresses. By construction, there will be a time $t_\mathrm{c} < t_\mathrm{f}$ around which there is a smooth transition from closely satisfying the equations of motion with the estimated reference Hamiltonian $\mathcal{H}$ to more attentively focusing on achieving the target. By these means, it is hoped that the control field design will achieve an acceptable signal in the target for an array of reasonable Hamiltonians in a neighborhood of the reference Hamiltonian. Specifically, if the true Hamiltonian lies within this neighborhood, then final laboratory refinement of the field should be readily attained. In the following section, we will apply these concepts to a system [5] that has been a useful testing ground for control design.

3 Illustration

The system chosen to illustrate these design concepts is a linear chain of coupled harmonic oscillators because it provides a simple example that is qualitatively easy to analyze. The example is also non-trivial for control as the objective will be a bond stretch at one end of the molecule with minimal disturbances to the other bonds, but energy will only enter through a single bond at the opposite end of the molecule. The chain has N bonds, and the equations of motion for the average positions, $\langle \mathbf{q} \rangle$, and momenta, $\langle \mathbf{p} \rangle$, are [5]

$$\dot{\mathbf{z}}(t) = \mathsf{M}\mathbf{z}(t) + \varepsilon(t)\mathbf{b}, \tag{4}$$

where

$$\mathbf{z}(t) = \begin{pmatrix} \langle \mathbf{q}(t) \rangle \\ \langle \mathbf{p}(t) \rangle \end{pmatrix} \tag{5}$$

$$\mathsf{M} = \begin{pmatrix} \mathsf{O} & \mathsf{G} \\ -\mathsf{F} & \mathsf{O} \end{pmatrix}. \tag{6}$$

Here, F is the diagonal force constant matrix where $\mathsf{F}_{ij} = \delta_{ij} k_i$, and G is the reduced mass matrix where $\mathsf{G}_{ij} = \delta_{ij}\left(\mu_{i+1} + \mu_i\right) - \delta_{i,(j+1)}\mu_i - \delta_{i,(j-1)}\mu_{i+1}$. The vector $\mathbf{b} = \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\mu} \end{pmatrix}$ contains the vector of constant transition dipoles, $\boldsymbol{\mu}$. The initial condition is chosen as the molecule at rest, $\mathbf{z}(0) = \mathbf{0}$. The cost functional in equation (2) now has the explicit form

$$\begin{aligned} J = & \int_0^{t_\mathrm{f}} \mathrm{d}t \left(\mathbf{z}(t) - \tilde{\mathbf{z}}\right)^{\mathrm{T}} \mathsf{A}(t) \left(\mathbf{z}(t) - \tilde{\mathbf{z}}\right) \\ & + \int_0^{t_\mathrm{f}} \mathrm{d}t \, \beta(t) \varepsilon^2(t) \\ & + \int_0^{t_\mathrm{f}} \mathrm{d}t \, \left[\mathsf{M}\mathbf{z}(t) + \varepsilon(t)\mathbf{b} - \dot{\mathbf{z}}(t)\right]^{\mathrm{T}} \\ & \times \Gamma(t) \left[\mathsf{M}\mathbf{z}(t) + \varepsilon(t)\mathbf{b} - \dot{\mathbf{z}}(t)\right] \end{aligned} \tag{7}$$

where A and Γ are symmetric positive semi-definite matrices of the form

$$\mathsf{A}(t) = \alpha(t) \begin{pmatrix} \mathsf{I} & \mathsf{O} \\ \mathsf{O} & c_1 \mathsf{I} \end{pmatrix} \tag{8}$$

$$\Gamma(t) = \gamma(t) \begin{pmatrix} \mathsf{I} & \mathsf{O} \\ \mathsf{O} & c_2 \mathsf{I} \end{pmatrix}. \tag{9}$$

Here, I is the $N \times N$ identity matrix, and the constants c_1 and c_2 are included to balance the relative magnitudes of $\langle \mathbf{q}(t) \rangle$ and $\langle \mathbf{p}(t) \rangle$ in the target and the equations of motion. The forms in equations (8, 9) are only guides, and the structures may be altered. Most critical is that reasonable physical intuition be satisfactory here, and the numerical results below support this key point.

To minimize equation (7) and find the optimal field, we separately set to zero the first order variations with respect to ε and $\mathbf{z}$. This leads to the control field

$$\varepsilon(t) = \frac{-\mathbf{b}^{\mathrm{T}} \Gamma(t) (\mathsf{M}\mathbf{z}(t) - \dot{\mathbf{z}}(t))}{\beta(t) + \mathbf{b}^{\mathrm{T}} \Gamma(t) \mathbf{b}} \tag{10}$$

and the equations of motion

$$\mathsf{A}(t)(\mathbf{z} - \tilde{\mathbf{z}}) + \left(\mathsf{M} + \frac{\mathrm{d}}{\mathrm{d}t}\right) \left[\Gamma(t) \left(\mathsf{M}\mathbf{z} + \varepsilon\mathbf{b} - \dot{\mathbf{z}}\right)\right] = \mathbf{0}. \tag{11}$$

Here we explicitly utilized equation (3), assuming that the equation will be propagated for a sufficiently long time t_f to validate using the asymptotic final condition. Equations (10, 11) are the operational design equations for this application. Equation (11) differs from equation (4) in having explicit dependence on the target through the term $\mathsf{A}(t)(\mathbf{z} - \tilde{\mathbf{z}})$, and thus, equation (11) tends to evolve toward $\tilde{\mathbf{z}}$.

Substituting equation (10) into equation (11) results in the second order linear differential equations

$$\begin{aligned} \ddot{\mathbf{z}} = & \left(\mathsf{M} - \mathsf{R}^{-1}\mathsf{M}^{\mathrm{T}}\mathsf{R} - \mathsf{R}^{-1}\frac{\mathrm{d}\mathsf{R}}{\mathrm{d}t}\right) \dot{\mathbf{z}} \\ & + \left(\mathsf{R}^{-1}\mathsf{A} + \mathsf{R}^{-1}\mathsf{M}^{\mathrm{T}}\mathsf{R}\mathsf{M} + \mathsf{R}^{-1}\frac{\mathrm{d}\mathsf{R}}{\mathrm{d}t}\mathsf{M}\right) \mathbf{z} \\ & - \mathsf{R}^{-1}\mathsf{A}\tilde{\mathbf{z}} \end{aligned} \tag{12}$$

where

$$\mathsf{R}(t) = \Gamma(t) - \left(\frac{\Gamma(t)\mathbf{b}\mathbf{b}^{\mathrm{T}}\Gamma(t)}{\beta + \mathbf{b}^{\mathrm{T}}\Gamma(t)\mathbf{b}}\right). \tag{13}$$

One may prove that R is positive definite, such that R^{-1} exists in equation (12).

Many computer codes can easily integrate equation (12) as two coupled first order equations. Since there is no longer a final condition on $\mathbf{z}(\tau)$, we are free to choose a physically consistent initial condition. To do so, we assume the reasonable condition that $\varepsilon(0) = 0$, which implies that $\dot{\mathbf{z}}(0) = \mathsf{M}\mathbf{z}(0)$ to be consistent with equation (10). Thus, the design effort reduces to a forward propagation of equation (12) resulting in $\mathbf{z}$ and $\dot{\mathbf{z}}$ which define the field in equation (10). There is no upper limit on the time integration, and it can proceed until the desired dynamical evolution is achieved.

As an illustration of the above concepts for this example, a chain of $N = 5$ bonds was considered. Each of the force constants was 1.15 a.u. and each mass was 2×10^4 a.u. The molecule had a single non-zero dipole at bond five, $\boldsymbol{\mu}_5 = 0.295$ a.u., and the goal was to effect a significant stretch at the opposite end, $\langle \mathbf{q}_1 \rangle$. Thus, $\tilde{\mathbf{z}}$ had a single non-zero term at $\tilde{\mathbf{z}}_1 = 0.1$ a.u., which, because of the linear nature of this problem in equation (4), merely acts as a scale factor for the field, and consequently $\langle \mathbf{q} \rangle$ and $\langle \mathbf{p} \rangle$. In the $\mathsf{A}(t)$ matrix, c_1 was set to zero, putting no cost on the momenta. The constant c_2 in the $\Gamma(t)$ matrix was 1×10^{-4}, based on an order of magnitude estimate

of the likely dynamics of the positions and momenta. The weight function $\beta(t)$ was set to the constant 0.01, and because of their smooth character, hyperbolic tangents were chosen to model the weight functions $\alpha(t)$ and $\gamma(t)$. The excitation is allowed to occur on whatever time-scale is appropriate for the system. Merely a rough estimate will suffice to appropriately specify the trade-off time t_c, taken as 4 000 a.u., between $\alpha(t)$ rising and $\gamma(t)$ falling. The mean values of $\alpha(t)$ and $\gamma(t)$ are 0.0025 and 10 000 respectively. It is not useful to compare the magnitude of the weight functions alone, but only the overall terms in equation (2) as they roughly balance each other.

A central feature of this non-iterative design procedure is the ability to advantageously work with inexact Hamiltonians. To assess this capability, the control field was tested on 100 random Hamiltonians taken from a neighborhood around the reference Hamiltonian used in the design process. This ensemble of Hamiltonians was generated by varying the force constants in the reference Hamiltonian within the window 0.15 a.u. (*i.e.*, ±8% variation around the nominal value).

At any time t_f each bond has a distribution of the maximum stretch achieved over the interval $0 \leq t \leq t_f$ due to the ensemble of Hamiltonians. Because the displacement waves will travel along the chain from the dipole bond to the target bond, it is useful to plot maximum displacement over the interval in this manner to remove the dependence of the statistics on the arbitrary times t_f that are chosen for analysis. If there is a time t_f at which the distribution for the target bond is near the desired target value and significantly higher than the distribution for the other bonds, then we conclude that sufficient control was achieved because at t_f, the target bond will have a stretch that was not equaled by any other bond over the time *interval* $t \in [0, t_f]$. In that case, the control field provides adequate control for any Hamiltonian near the reference making the design viable for laboratory refinement. In practical applications, one would want to avoid such statistical testing. Here we perform it to assess the methodology.

The results of applying the non-iterative design techniques successfully produced selective excitation of the target bond over the ensemble of Hamiltonians. A typical result is shown in Figure 1. The target bond distribution is plotted along with the dipole bond distribution. As the dipole bond is the site of energy input, it typically has the largest displacement of the non-target bonds. Selective excitation begins to appear at $t_f \simeq 8\,000$ a.u. when the mean target bond displacement is significantly larger than the mean dipole bond displacement. The statistics of the maximum dipole bond displacement reaches a local extension at $t_f \simeq 7\,500$ a.u., evidenced by the maximum displacement remaining constant in Figure 1 (*i.e.*, that actual dipole bond displacement decreased on the interval $7\,500\,\text{a.u.} \lesssim t \lesssim 8\,500$ a.u., thereby making the statistics of the *maximum* displacement nearly constant on the same interval for t_f). In contrast, the target bond begins to expand until it reaches a maximum at $\sim 8\,000$ a.u., when excitation has peaked. The magnitude of the excitation of the target bond is actually beyond the objective value of 0.1 a.u., and the degree of excitation over that of the non-target bonds is significant despite our having made no attempt to either minimize the momenta, or strictly demand reduced non-target displacements. A characteristic feature of these plots is that after $\gamma(t)$ has decreased to the point where control over the system has been achieved, the dynamics escape the command of the algorithm. In this case, this time occurs near 9 000 a.u., but this causes no problem in practice, as the field can be smoothly turned off prior to that. The design field itself has a broad band spectral structure (not shown here) with peaks corresponding to the second and third highest normal mode frequencies. The control field is plotted as a function of time in Figure 2. The structure after $\sim 8\,000$ a.u. may be smoothly damped without affecting the quality of the control results in Figure 1. The form of the field in Figure 2 is similar to those found by optimal control in the same class of models.

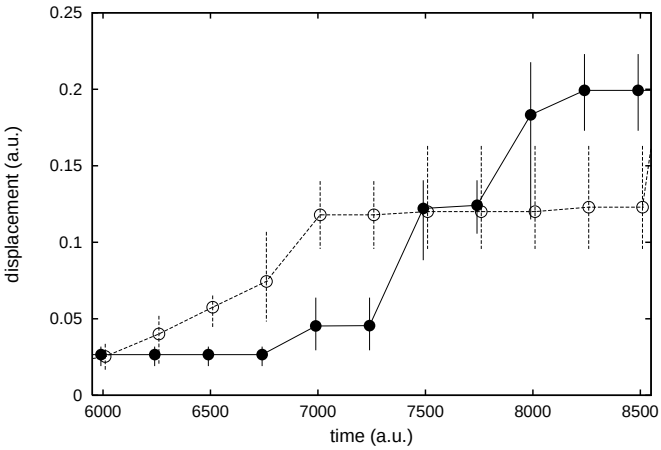

Fig. 1. Typical results for control of a coupled harmonic chain. At time intervals of 250 a.u. for t_f, the vertical dotted lines span the distribution of the maximum stretch achieved by the dipole bond over the interval $t \in [0, t_f]$, calculated by an ensemble of Hamiltonians and driven by the design field. The solid line represents the same data for the target bond. The open and closed circles mark the arithmetic means of the two distributions. Significant excitation is seen at 8 000 a.u. when the target bond achieves a displacement that is larger than the dipole bond has achieved to that point.

By running a large ensemble of functions α, β, and γ, we verified consistent physical behavior. Importantly, systematic changes in the weight functions led to intuitive responses in the system dynamics. As expected, the magnitude of α directly affects the ability to realize the objective, with a larger α being more effective at selective bond excitation. Relatively larger β functions tends to decrease the magnitude of the control field, and the function γ controls how closely the ensemble of Hamiltonians matched the design dynamics. These patterns emerged readily from an ensemble of tested weight functions. Furthermore, physically estimated orders of magnitude for these functions were found to be quite adequate. A broad

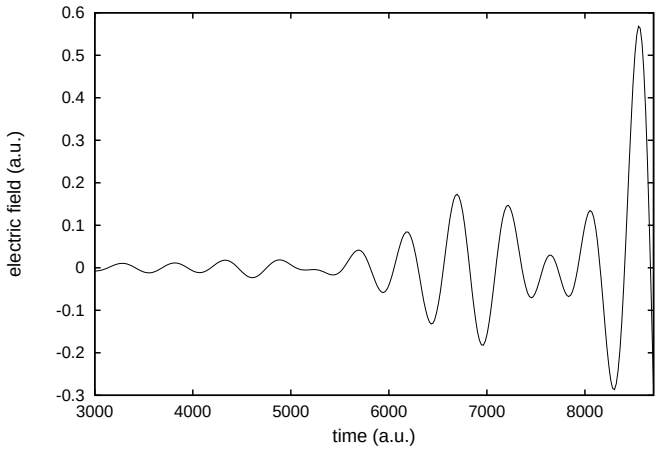

Fig. 2. The control field for Figure 1 as a function of time. In practice, the field can be smoothly turned off after the design objective has been suitably met at $\sim 8\,000$ a.u.

range of successful possibilities for α, β, and γ is evidenced by the robustness of the control results over a large sampling of the weight function space. This behavior is very encouraging for the efficacy of the design procedure. Once the three terms are in reasonable balance, a design can be executed, and it should then be possible to fine-tune the solution in the laboratory for better results.

4 Conclusion

The algorithm introduced in this paper may provide a powerful alternative to the classic Lagrange multiplier formulation. The success with the simple test system indicates that minimizing equation (2) leads to a well-behaved design procedure. The algorithm aims to provide good control designs while making significant reductions in the necessary computational effort, so it should be possible to consider systems that would have otherwise been too complex to solve with iterated optimization. The key to the success is through reducing the excess restrictions imposed by the Lagrange multiplier and the demand that the target be hit at a particular time. Further exploration needs to be done with more complex systems using the Schrödinger equation to verify that targets can still be reached with straightforward manipulation of the weighting functions. In the case of a system that propagates according to the Schrödinger equation, the design equations of motion will contain the operator $\Omega\Gamma\Omega$ acting on $|\psi\rangle$, but the system can again be rewritten as coupled first order equations. Careful testing of the algorithm under a variety of conditions is necessary to define its full utility.

The authors acknowledge support from the NSF and DOD.

References

1. R. Judson, H. Rabitz, Phys. Rev. Lett. **68**, 1500 (1992).
2. A. Assion, T. Baumert, M. Bergt, T. Brixner, B. Kiefer, V. Seyfried, M. Strehle, G. Gerber, Science **282**, 919 (1998).
3. C.J. Bardeen, V. Yakovlev, K. Wilson, S. Carpenter, P. Weber, W.S. Warren, Chem. Phys. Lett. **280**, 151 (1997).
4. T.C. Weinacht, J.L. White, P.H. Bucksbaum, J. Phys. Chem. A **103**, 10166 (1999).
5. S. Shi, A. Woody, H. Rabitz, J. Chem. Phys. **88**, 6870 (1988).
6. S. Shi, H. Rabitz, J. Chem. Phys. **92**, 2927 (1990).
7. S.A. Rice *et al.*, Adv. Chem. Phys. **101**, 213 (1997).
8. R. Kosloff, S.A. Rice, P. Gaspard, S. Tersigni, D.J. Tannor, Chem. Phys. **139**, 201 (1989).
9. J. Cao, K. Wilson, Phys. Rev. A **55**, 4477 (1997).
10. J. Manz, K. Sundermann, R. de Vivie-Riedle, Chem. Phys. Lett. **290**, 415 (1998).
11. P. Gross, H. Singh, H. Rabitz, K. Mease, G.M. Huang, Phys. Rev. A **47**, 4593 (1993).
12. Y. Chen, P. Gross, V. Ramakrishna, H. Rabitz, K. Mease, J. Chem. Phys. **102**, 8001 (1995).
13. W.S. Zhu, H. Rabitz, J. Chem. Phys. **110**, 7142 (1999).
14. J.G.B. Beumee, H. Rabitz, J. Chem. Phys. **97**, 1353 (1992).
15. H. Shen, J.P. Dussault, A.D. Bandrauk, Chem. Phys. Lett. **221**, 498 (1994).