A Fourier-Coefficient based Solution of an Optimal Control Problem in Quantum Chemistry*

Katharina Kormann,[†] Sverker Holmgren,[‡]
and Hans O. Karlsson[§]

Abstract

We consider an optimal control problem for the time-dependent Schrödinger equation modeling molecular dynamics. Given a molecule in its ground state, the interaction with a tuned laser pulse can result in an excitation to a state of interest. By these means, one can optimize the yield of chemical reactions. The problem of designing an optimal laser pulse can be posed as an optimal control problem. We reformulate the optimization problem by Fourier-transforming the electric field of the laser and narrow the frequency band. In this way, we reduce the dimensionality of the control variable. This allows for storing an approximate Hessian and, thereby, we can solve the optimization problem with a quasi-Newton method. Such an implementation provides superlinear convergence. We show computational results for a Raman-transition example and give numerical evidence that our algorithm can outperform the standard Krotov-like method which does not employ approximative second derivatives.

Key words   quantum optimal control · Schrödinger equation · BFGS

1 Introduction

Ever since the introduction of femtosecond lasers in the 1980s, chemists have had means at hand to watch the motion of nuclei in a molecule. Such experiments reveal processes on the quantum level. These techniques were even brought one step further when scientists started using ultra-fast laser pulses to manipulate the state of molecules. Given a molecule in its ground state, the

interaction with a designed laser pulse can yield an excitation to some target state. In this way, the quantum state of molecules can be controlled and specific chemical reactions initiated.

We consider the problems of how to optimally choose the electric field to get a maximal cross-correlation with a target state and how to make the wave packet follow a prescribed trajectory. This problem is commonly formulated as an optimal control problem with the pulse field as the control variable (cf. [22, Chap. 16.2]).

In quantum control, originally only targets for some final time were addressed. Shi et al. [18] solved the optimal control problem [11] by a gradient-type method. Somlói et al. [21] proposed to use the Krotov method [12] for which they obtained faster convergence. The idea of the Krotov scheme is to solve the first-order optimality condition for the control parameter and solve the resulting equation by fixed-point iteration. Zhu et al. [28, 29] refined this ansatz in a similar fashion as Gauss–Seidel iteration improves Jacobi's method. Maday and Turinici [13] generalized this further and formulated a framework of Krotov-like methods where the old and the new field are combined in various ways. Later, Serban et al. [17] also explained how to handle time-dependent targets using the Krotov method.

One main bottleneck in optimal control procedures used for femtosecond pulse shaping is the computer memory. The number of time-steps needed for the simulation is usually very large. Therefore, care has to be taken when the first-order optimality condition (which requires solution of a primal and a dual problem) is evaluated, so that the memory consumption is not prohibitive. Furthermore, it is not possible to store an (approximate) Hessian for large time intervals.

In this article, we propose to Fourier transform the pulse and to reformulate the problem such that the Fourier coefficients are optimized. This type of procedure imitates the way pulses are normally shaped in an experimental setup [27, 25] and reduces the dimensionality of the optimization problem since not all the frequencies that can be resolved in the discrete time domain (that is chosen for numerical optimization) are actually of interest. Moreover, we propose to use a memory efficient way of computing the gradient (cf. Ref. [26]). For solving the reformulated problem, we use a quasi-Newton method with the Broyden–Fletcher–Goldfarb–Shanno (BFGS) update formula. The BFGS algorithm was applied to quantum optimal control for Bose–Einstein condensate by von Winckel and Borzì in the context of their study of L_2- versus H_1-based minimization [24].

Note that all the algorithms mentioned above only find local optima. Finding a global minimizer is even more complicated since then these procedures would have to be combined with some global optimization strategy.

The outline of this article is as follows. In the next section, we will introduce the continuous optimization problem. A discretization of the problem is discussed in Sec. 3. We then describe our implementation of the optimal control

problem in Sec. 4, and a numerical study for the rubidium diatom demonstrates the potential of our approach in Sec. 5. Sec. 6 gives some concluding remarks.

2 Optimization Problem

The interaction of a molecule with an electric field is modeled by the time-dependent Schrödinger equation (TDSE),

$$i\hbar \frac{\partial}{\partial t} \psi = \left(\hat{H}_0 + \hat{\mu}\varepsilon(t) \right) \psi,$$
$$\psi(x,0) = \psi_0,$$
(1)

where $\hat{H}_0 = \hat{T}(\frac{\partial^2}{\partial x_i^2}) + \hat{V}(x)$ is the molecular Hamiltonian with kinetic ($\hat{T}$) and potential ($\hat{V}$) energy of the system, $\hat{\mu}$ is the molecular dipole moment vector, and $\varepsilon(t)$ is the electric field of the laser pulse [22, Chap. 13].

Our goal is to design a laser field that maximizes the transition from the given initial state to some target state after a certain time t_f. Furthermore, we sometimes want to include a time-dependent objective, e.g., penalization of the transition to some undesired state. Conversely, we can specify a density that should be maximized. Let us define the generic projection operator

$$\hat{O} = \hat{O}_1(t) + \delta(t - t_f) t_f \hat{O}_2,$$

and formulate the general minimization problem

$$\text{minimize}_{\varepsilon \in L_2([0,t_f])} \quad \mathscr{J}_1(\varepsilon, \psi) = \frac{1}{t_f} \int_0^{t_f} \langle \psi(\cdot, t) | \hat{O} | \psi(\cdot, t) \rangle \, dt$$

$$\text{subject to} \quad i\hbar \frac{\partial}{\partial t} \psi = \left(\hat{H}_0 + \hat{\mu}\varepsilon(t) \right) \psi,$$
$$\psi(x,0) = \psi_0.$$
(2)

The operator $\hat{O}_2$ projects onto the orthogonal complement of the target state and $\hat{O}_1(t)$ represents the time-dependent objective. For instance, we can set $\hat{O}_1 \equiv 0$ and let $\hat{O}_2$ be the projection to the orthogonal complement of a wavefunction ϕ, i.e.,

$$\mathscr{J}_1 = \langle \phi - \psi(\cdot, t_f) | \phi - \psi(\cdot, t_f) \rangle,$$
(3)

to formulate the important special case of finding a laser pulse that yields maximal overlap with ϕ at final time.

The optimization problem (2) is generally unbounded: An optimization procedure would tend to choose pulses with energy tending to infinity. Therefore, we have to make sure to optimize the quality, not the quantity of the field. We want to bound the strength $\int_0^{t_f} |\varepsilon(t)|^2 \, dt$ of the laser pulse, and accomplish this by adding a Tikhonov regularization term [23],

$$\mathscr{J}_2(\varepsilon) = \eta \int_0^{t_f} |\varepsilon(t)|^2 \, dt, \qquad \eta > 0,$$

3

to the objective function. The weighting of the total pulse energy by the parameter η influences the strength of the pulse as well as the properties of the optimization problem. The larger the value of η the weaker the optimal pulse will be and the more regular the problem becomes. To summarize, we consider the following optimization problem,

$$\text{minimize}_{\varepsilon \in L_2([0,t_f])} \quad \mathscr{J}(\varepsilon, \psi) = \frac{1}{t_f} \int_0^{t_f} \langle \psi(\cdot, t) | \widehat{O} | \psi(\cdot, t) \rangle \, dt$$

$$+ \eta \int_0^{t_f} |\varepsilon(t)|^2 \, dt \tag{4}$$

$$\text{subject to} \quad i\hbar \frac{\partial}{\partial t} \psi = \left(\hat{H}_0 + \hat{\mu} \varepsilon(t) \right) \psi,$$

$$\psi(x, 0) = \psi_0.$$

Peirce et al. [15] show existence of an optimal solution to the control problem (4) for the objective (3). Their proof also applies to more general objectives, as long as the objective function is weakly lower semicontinuous in ψ.

For the solution of the problem (4), we introduce the Lagrange function,

$$\mathscr{L}(\varepsilon, \psi, \chi) = \frac{1}{t_f} \int_0^{t_f} \langle \psi(\cdot, t) | \widehat{O} | \psi(\cdot, t) \rangle \, dt + \eta \int_0^{t_f} |\varepsilon(t)|^2 \, dt$$

$$- 2 \cdot \int_0^{t_f} \Re \langle \chi | \left(\frac{\partial}{\partial t} + \frac{i}{\hbar} \left(\hat{H}_0 + \hat{\mu} \varepsilon(t) \right) \right) | \psi \rangle \, dt.$$

Variation of $\mathscr{L}$ with respect to ψ yields the following adjoint equation for the Lagrange multiplier χ,

$$\frac{\partial}{\partial t} \chi = -\frac{i}{\hbar} \left(\hat{H}_0 + \hat{\mu} \varepsilon(t) \right) \chi - \frac{1}{t_f} \widehat{O}_1(t) \psi(r, t),$$

$$\chi(x, t_f) = \widehat{O}_2 \psi(x, t_f). \tag{5}$$

The TDSE is self-adjoint because the Hamiltonian is symmetric. Hence the dual equation is well-posed and can be solved by the same means as the primal problem, but backwards in time. The sensitivity is then given by

$$\frac{\delta \mathscr{L}}{\delta \varepsilon} = \frac{1}{\hbar} \Im \left(\langle \chi | \hat{\mu} | \psi \rangle \right) + \eta \varepsilon.$$

The optimality system for the optimal control problem (4) is thus given by the state equation (1), the adjoint equation (5), and the first order optimality condition,

$$\frac{1}{\hbar} \Im \left(\langle \chi | \hat{\mu} | \psi \rangle \right) + \eta \varepsilon = 0. \tag{6}$$

2.1 Optimization in Fourier Space

The control variable of the optimization problem is the pulse $\varepsilon(t)$. After discretization in time this is a vector with the value of the pulse at the grid points in time. The dimensionality of the optimization problem thus corresponds to the number of time steps. A fruitful complementary view is given by considering the Fourier-transformed pulse and the amplitudes corresponding to the different frequencies. We propose to use the Fourier coefficients as control variables and to narrow the frequency band, hereby focusing on relevant wave numbers.

A main motivation for examining the Fourier-transformed pulse is the experimental setup for femtosecond pulse shaping. Here, the laser field is split into different frequencies and each frequency is separately manipulated [27]. Technical constraints, such as the frequency band of the laser or the assembly of the shaper, restrict the shapes that can be produced in the laboratory. For this reason, an optimization loop without constrictions of the wave numbers may propose a pulse that cannot be realized.

Moreover, narrowing the frequency band results in a reduction of the dimensionality of the control variable, which makes the optimization problem both more regular and less memory-intensive. Physical reasoning can help to identify the most interesting frequencies when we are seeking to reduce the dimensionality of the control variables. According to Fermi's Golden Rule (cf. Chap. 14 in [22]), a pulse with an oscillatory frequency corresponding to the energy difference between initial and target state gives a high transition probability. This intuitive rule does not include phase information and interference effects of intermediate states. Therefore, the solution of the optimal control problem can usually not be reduced to only one Fourier coefficient. However, in practice, the interesting frequencies are indeed found around the one given by the Golden Rule.

Given the domain $\Omega \subset \mathbb{R}$ of interesting frequencies, the optimization problem in Fourier space is given as

$$\text{minimize}_{\hat{\varepsilon} \in L_2(\Omega)} \quad \mathcal{I}(\hat{\varepsilon}) = \langle \psi(\cdot, t_f) - \phi, \psi(\cdot, t_f) - \phi \rangle + \hat{\eta} \int_\Omega |\hat{\varepsilon}(\omega)|^2 \, d\omega$$

$$\text{subject to} \quad i\hbar \frac{\partial}{\partial t} \psi = \left(\hat{H}_0 + \hat{\mu} \frac{1}{2\pi} \int_\Omega \hat{\varepsilon}(\omega) e^{i\omega t} \, d\omega \right) \psi,$$

$$\psi(x, 0) = \psi_0. \tag{7}$$

Remark. In pulse shaping by feedback control [8], Fourier coefficients are often considered as control parameters [1, 19] because the experimental setup is usually based on grading of the laser pulse.

3 Discretization

To solve the optimal control problem formulated in the previous section, we have to repeatedly solve the Schrödinger equation and to evaluate the sensitiv-

ity. In order to do so numerically, we have to discretize the problem. We will first describe the discretization scheme and then discuss the discrete formulation of the optimal control problem.

3.1 Discretization of the Schrödinger Equation

We discretize the Schrödinger equation using the method of lines. A grid in the spatial dimensions is introduced and we compute the kinetic energy by a pseudospectral method with Fourier basis. After discretization in space, we end up with a system of ordinary differential equations of the form

$$\frac{d}{dt} v = \left(H_0 + \mu \varepsilon(t) \right) v,$$
$$v(0) = v_0,$$

where H_0 and μ are matrices representing the molecular Hamiltonian and the transition dipole moment, respectively, and v_0 gives the initial value at the grid points. Then, we also discretize in time and denote the step length by Δt and the number of grid points by N. For sufficiently small time steps, the evolution operator (propagating the solution from t to $t + \Delta t$) can be computed based on the Magnus expansion [7]. Truncating the expansion after the first term yields

$$v(t + \Delta t) = e^{-i\Delta t (H_0 + \mu \int_t^{t+\Delta t} \varepsilon(\tau) \, d\tau)} v(t).$$

The matrix exponential can efficiently be computed using Strang splitting [16] which is also known as split operator. Let $\mathbf{e} = (e_n)_{0 \le n \le N}$ be a discrete representation of the pulse. The integral over the pulse on the nth interval can be approximated by $\frac{\Delta t}{2}(e_n + e_{n+1})$. To sum up, this yields the following standard second-order time-propagation scheme for discretization of the Schrödinger equation:

$$u_{n+1} = \underbrace{e^{-i\frac{\Delta t}{2}T} e^{-i\Delta t (V + \frac{1}{2}\mu(e_n + e_{n+1}))} e^{-i\frac{\Delta t}{2}T}}_{:= \mathcal{M}_n(e_n, e_{n+1})} u_n.$$

There are two main reasons to choose the Strang splitting method. Firstly, the discrete evolution operator is unitary, i.e., $(\mathcal{M}_n)^* \mathcal{M}_n = I$. The numerical method thus retains the self-adjointness of the continuous problem which will be very important when formulating the discrete adjoint problem in Sec. 3.2. Moreover, the TDSE has to be solved hundreds of times during the optimization procedure, and the use of a highly accurate (and computationally more expensive) method is out of reach. As reported in [10], Strang splitting is very efficient for low-accuracy computations.

3.2 Discrete Formulation of the Optimal Control Problem

In Sec. 2, the continuous formulation of the optimal control problem resulted in expression (6) for the gradient. However, discretization and gradient computation do not commute (cf. Chap. 3.2 in [6] and Chap. 4 in [2]). Therefore, it is often

favorable to first discretize and then formulate the optimal control problem. By pursuing this approach, we are provided with the exact gradient information for the discrete optimization problem (to which we actually apply the numerical optimization algorithm). To derive this, let us first reconsider the optimization problem in time domain. We refer to Fisher and Jennings' work [3] for a detailed discussion on discrete-time optimal control problems.

As introduced in the previous section, the discrete primal problem is given by

$$u_0 = v_0,$$
$$u_{n+1} = \mathcal{M}_n(e_n, e_{n+1}) u_n, \qquad n = 0, \dots, N-1.$$

The discretized objective function then reads

$$J(\mathbf{e}) = \Delta x\, u_N^H O_2 u_N + \frac{\Delta x \cdot \Delta t}{t_f} \sum_{k=0}^{N-1} u_k^H O_1 u_k + \eta \cdot \Delta t \sum_{k=0}^{N} e_k^2.$$

The primal equation is enforced by using the Lagrangian

$$L(\mathbf{e}) = J(\mathbf{e}) - 2 \cdot \Delta x \cdot \Delta t \Re \left(\sum_{k=0}^{N-1} \lambda_{k+1}^H (\mathcal{M}_k(e_k, e_{k+1}) u_k - u_{k+1}) \right),$$

where

$$\lambda_N = -O_2 u_N \cdot \frac{1}{\Delta t},$$
$$\lambda_n = (\mathcal{M}_n(e_n, e_{n+1}))^H \lambda_{n+1} - \frac{1}{T} O_1 u_n, \qquad n = N-1, \dots, 1.$$

Exploiting the fact that the evolution operator $\mathcal{M}_k$ only depends on components k and $k+1$ of $\mathbf{e}$, we can compute the derivative of L with respect to e_n,

$$\frac{\partial L}{\partial e_n} = \Delta t \left(\eta \cdot e_n - \Re \left(\lambda_{n+1}^H \frac{\partial}{\partial e_n} \mathcal{M}_n u_n + \lambda_n^H \frac{\partial}{\partial e_n} \mathcal{M}_{n-1} u_{n-1}. \right) \cdot \Delta x \right).$$

For e_0 and e_N, we will only get the first or the second term within the real part, respectively. Note also that the evolution operator is symmetric in e_n and e_{n-1} which is why the first term in $\frac{\partial L}{\partial e_n}$ is the same as the second term in $\frac{\partial L}{\partial e_{n+1}}$.

We are however actually interested in is the optimization in frequency domain for a number of discrete frequencies Ω. To compute the gradient of $I(\hat{\mathbf{e}}) = J(\mathbf{e})$ with respect to the Fourier coefficients $\hat{e}_\omega$, $\omega \in \Omega$, we use the chain rule $\frac{\partial I}{\partial \hat{e}_i} = \frac{\partial J}{\partial e_i} \frac{\partial e_i}{\partial \hat{e}_i}$. The derivative $\frac{\partial e_i}{\partial \hat{e}_i}$ is the same in each iteration and can therefore be precomputed.

4 Optimization Algorithm and Implementation

In this section, we discuss the BFGS optimization method and compare it with a Krotov-like scheme, which is the standard algorithm for quantum optimal control.

4.1 BFGS

We use the so-called Broyden–Fletcher–Goldfarb–Shanno (BFGS) method [5] to solve the optimization problem (7). BFGS is a method for unconstrained minimization. The method is applicable since we have formulated the optimization problem such that explicit constraints are avoided. This includes the Schrödinger equation which is not treated as a constraint but as part of the objective function.

The BFGS method converges superlinearly. It uses the objective function value as well as its gradient and builds an approximate Hessian. This approximate Hessian makes the method superior compared to first order methods like conjugate gradient both with respect to convergence properties and sensitivity to scaling of the optimization parameters. The algorithm is most efficient when the approximate Hessian can be stored. However, there also exist limited-memory versions with linear memory consumption that are more suited for large-scale problems. In our experiments in Sec. 5, we use the implementation within the Matlab optimization toolbox (Matlab version 7.8).

When using the BFGS algorithm, we have to compute the value of the objective function as well as the gradient $\nabla_e L \cdot \nabla_{\hat{e}} e$ in each iteration. Especially when computing the gradient, care has to be taken not to use too much computer memory. This is the subject of the following subsection.

4.2 Memory Efficient Implementation

Computing the gradient involves solving the primal and the dual equation. Since the primal equation has to be solved forward and the dual backward in time, it is not possible to compute both problems simultaneously. A straightforward implementation would therefore be to first solve the primal problem saving all the values u_j. Then, the dual problem is solved backwards and the inner product $\langle \lambda_{j+1}, \frac{d}{de_j} \mathcal{M}_j u_j \rangle$ can be computed step by step. For high-dimensional problems, storing the forward solution for every time step would require far too much memory.

However, at the expense of one additional PDE-solving procedure, the gradient computations can be implemented in a very memory efficient way: First, we solve the primal problem without saving any intermediate step. This provides u_N as well as λ_N. Because of the time reversibility of the Schrödinger equation (and our discretization of it), we can now solve the dual — as well as the primal — equation backward in time. If we do this simultaneously, we only need the memory to compute both problems and one additional wave function (for the computation of the derivative with respect to the control variable). Note that Werschnik and Gross [26] discuss a similar procedure for the Krotov method. The algorithm can be summarized as follows:

1. Solve the primal Schrödinger equation for u_N

2. Compute λ_N from u_N

3. For j from $N-1$ to 0

    (a) Compute $u_j = \mathcal{M}_j^* u_{j+1}$

    (b) Compute $\frac{\partial}{\partial e_j} \mathcal{M}_j u_j$

    (c) Set $temp = 2\Re \langle \lambda_{j+1}, \frac{\partial}{\partial e_j} \mathcal{M}_j u_j \rangle$

    (d) Add $temp$ to $\frac{\partial L}{\partial e_j}$ and $\frac{\partial L}{\partial e_{j+1}}$

    (e) Compute $\lambda_j = \mathcal{M}_j^* \lambda_{j+1}$

4. Compute $\nabla_{\mathbf{e}} L \cdot \nabla_{\hat{\mathbf{e}}} e$, where $\nabla_{\hat{\mathbf{e}}} e$ can be precomputed.

Note that step 3(e) and 3(a) in the next cycle can be computed in parallel.

4.3 Comparison with the Krotov Method

We now compare our BFGS-based optimization scheme with the Krotov method. The idea of the latter algorithm is to reformulate the first order optimality condition (6) as

$$\varepsilon(t) = -\frac{1}{\eta \cdot \hbar} \Im\left(\langle \chi(t) | \hat{\mu} | \psi(t) \rangle \right). \tag{8}$$

This expression is a non-linear equation for the laser field. The Krotov method solves equation (8) as a fixed point iteration. One starts with some initial guess — often $\varepsilon \equiv 0$ — and then alternately solves for ψ and χ based on a pulse that is updated in each iteration. There are different strategies for updating the pulse, summarized in [13], based on two parameters σ and τ. The kth step of a Krotov-like method reads

$$i\hbar \frac{\partial}{\partial t} \psi^k = \left(\hat{H}_0 + \hat{\mu} \varepsilon^{k,0}(t) \right) \psi^k, \quad \psi^k(x,0) = \psi_0,$$

$$\varepsilon^{k,0}(t) = (1-\sigma)\varepsilon^{k-1,1}(t) - \frac{\sigma}{\eta \cdot \hbar} \Im\left(\langle \chi^{k-1}(t) | \hat{\mu} | \psi^k(t) \rangle \right),$$

$$i\hbar \frac{\partial}{\partial t} \chi^k = \left(\hat{H}_0 + \hat{\mu} \varepsilon^{k,1}(t) \right) \chi^k - \frac{1}{t_f} \hat{O}_1(t) \psi^k(r,t), \quad \chi^k(x,t_f) = \hat{O}_2 \psi(r,t_f),$$

$$\varepsilon^{k,1}(t) = (1-\tau)\varepsilon^{k,0}(t) - \frac{\tau}{\eta \cdot \hbar} \Im\left(\langle \chi^k(t) | \hat{\mu} | \psi^k(t) \rangle \right).$$

For the parameter choice $\sigma = 1$ and $\tau = 0$, the original Krotov method [21] is obtained and, for $\sigma = \tau = 1$, the algorithm corresponds to the one proposed by Zhu et al. [28, 29].

We propose to restrict the space of possible pulses by Fourier transforming and truncating the expansion. Additional constraints entail a global minimum that is larger or equal the one of the original problem. Naturally, the minimum

will be equal if all the frequencies that are included in the optimal pulse are part of the Fourier-coefficient based formulation. Therefore, it is important to carefully choose the wave numbers that should be included in the reformulated problem. However, for a good choice of the frequencies, the optimal solution of the new problem will not be considerably worse.

The memory efficient scheme for the gradient evaluation of Algorithm 4.1 can also be applied to the original Krotov [9] [20] method (cf. [26]).

We expect the optimization framework described in this article to converge faster than a Krotov-like method, since the BFGS algorithm converges superlinearly thanks to an approximation of the Hessian, whereas the Krotov method is entirely based on the first order optimality conditions. A numerical comparison is provided in Sec. 5.1.

Remark. Zhu et al. [28] claim quadratic convergence of their implementation of the Krotov method. However, the underlying definition of the convergence rate does not comply with the standard definition (cf., e.g., [5, Chap. 2.5]).

5 Numerical Results

In this section, we consider two different configurations of the rubidium dimer (Rb_2). In both cases, we want to design a pulse that maximizes some cross-correlation at final time. The first setting only considers a final-time objective and the second problem additionally includes a time-dependent penalization term.

5.1 Comparison with Krotov-like Method

We consider an example of the Rb_2 molecule with three bounded states. The data is taken from Ref. [4]. Fig. 1 illustrates the configuration of the problem.

In this case, we have the following form of the molecular Hamiltonian

$$H_0 = \begin{pmatrix} T + V_X & 0 & 0 \\ 0 & T + V_A & X_{Ab} \\ 0 & X_{Ab} & T + V_b \end{pmatrix},$$

with V_X, V_A, and V_b being the potential energy surfaces (PES) for the ground state $^1\Sigma_g^+$, the $^1\Sigma_u^+$ excited state, and the $^3\Pi_u$ excited state, respectively. The transition dipole moment for coupling the first and the second surface is given by

$$\mu = \begin{pmatrix} 0 & X_{XA} & 0 \\ X_{XA} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The goal is to design a femtosecond laser pulse which maximizes the cross-correlation with the lowest eigen state on the $^3\Pi_u$-surface by exciting the ground state to the $^1\Sigma_u^+$-state.

10
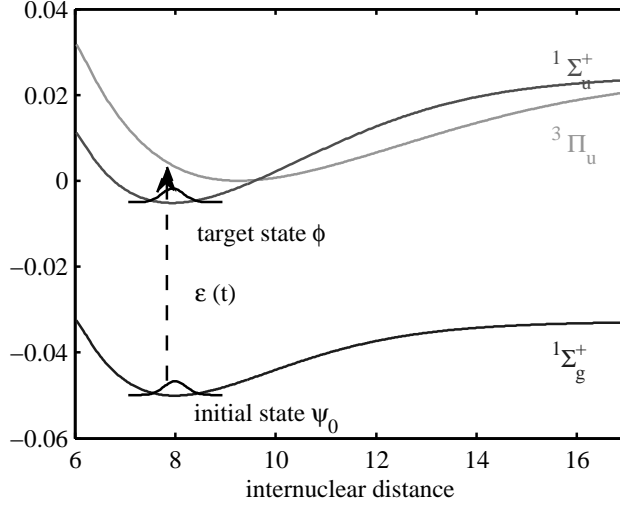
Figure 1: Wavepacket evolution control for Rb_2 in Sec. 5.1.

We design a pulse of length 100 fs using both the Krotov-like method by Zhu and Rabitz [29] and our Fourier-based BFGS method. We use a step size of $\Delta t = 0.01$ fs. Hence, we have 10,000 time steps but we only include the lowest 60 sine and cosine frequencies (i. e., we have 120 control parameters in the BFGS method versus 10,000 for the Krotov scheme). The penalty parameter η is set to 0.7 which gives a reasonable pulse strength. To compare the performance of the methods, we view the overlap with the target state as a function of the number of (gradient) evaluations in Fig. 2. The results show that our method clearly outperforms the Krotov-like method. The optimized pulses for both methods are shown in Figs. 4 and 3. In both cases, the pulse is intense in the beginning, after half the time, and in the end. However, the Krotov method gives a peak also after about one quarter of the time.

5.2   Results for a time-dependent target

In this section, we consider the Rb_2 molecule involving different electronic states. We want to achieve a Raman-like transition from the lowest to the next-lowest eigenstate of the first level via excitation to the $^1\Sigma_u^+$-state, while excitation to the $^1\Pi_g$-state should be avoided. The configuration is displayed in Fig. 5. In our computations, we dismiss interactions with additional electronic states. A similar problem was considered in Ref. [14] and tackled with a Krotov-like method.
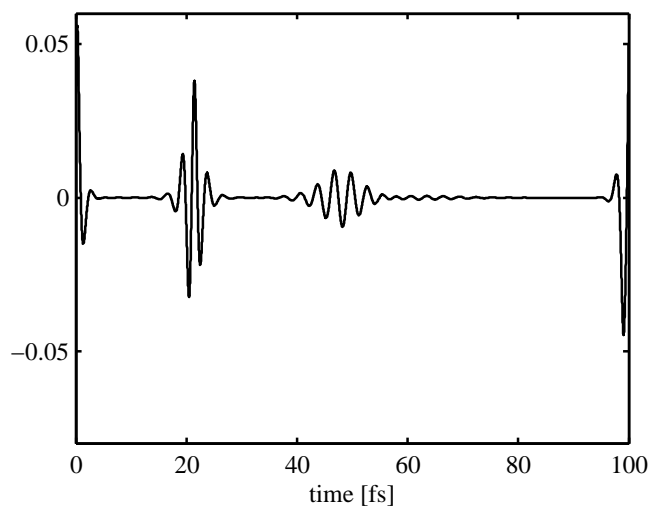
For our computations, we take a maximum time $t_f = 500$ fs and a time step size of $\Delta t = 0.01$ fs. On this time grid, we include the wave numbers 101–420 (sine and cosine) and optimize the corresponding Fourier-coefficients.
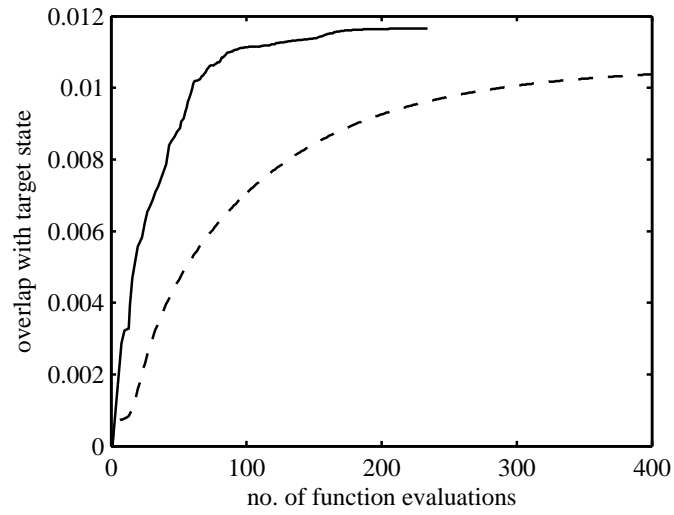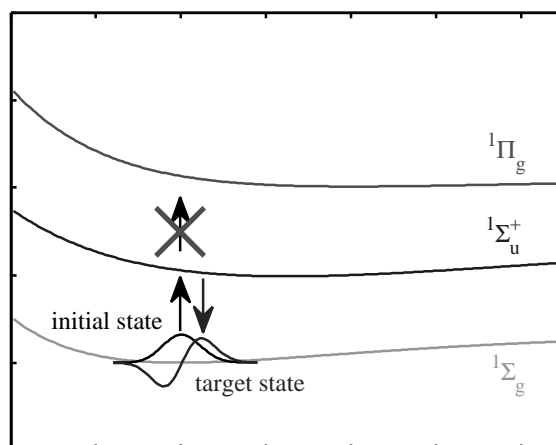
Figure 2: Comparison of the convergence properties of Krotov(– –) and BFGS(—).
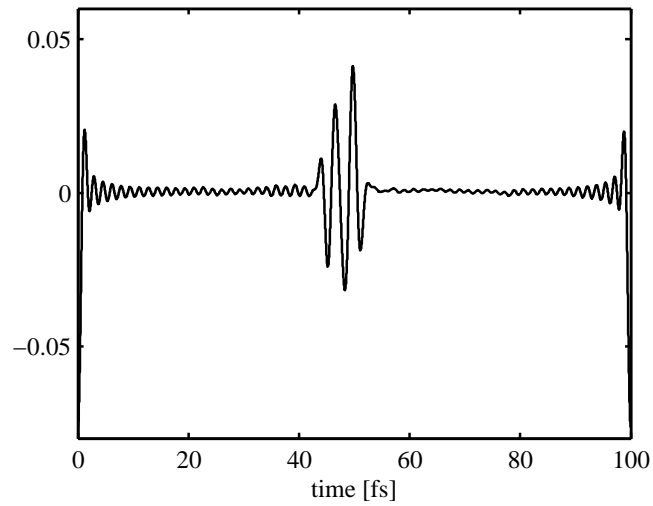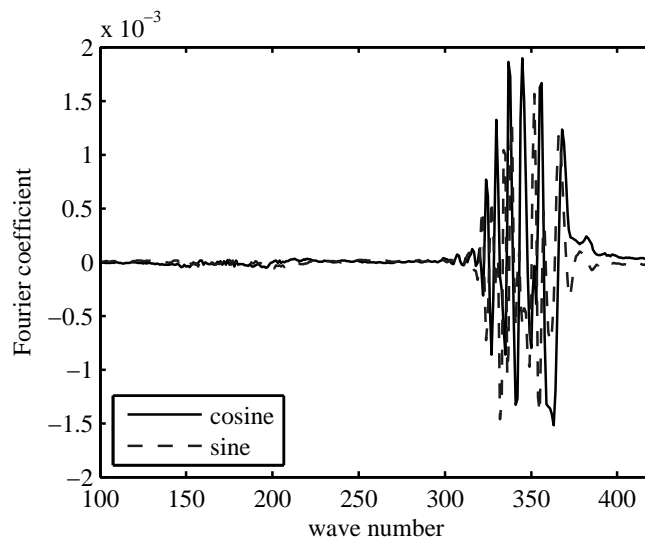

Figure 3: Pulse optimized with Fourier-coefficient based BFGS method.

Figure 4: Pulse optimized with Krotov-like method.

Figure 5: Configuration of optimization problem in Sec. 5.2.

13

Figure 6: Fourier coefficients of optimized pulse for time-dependent target.

The BFGS algorithm stops after 161 iterations (172 function and gradient evaluations). The resulting Fourier coefficients are shown in Fig. 6 and the corresponding pulse can be seen in Fig. 7. Employing this pulse the molecule is transfered to the target state with a probability of 99.3%. During the simulation time an average fraction of $3.3 \cdot 10^{-4}$ of the wave packet is on the forbidden state ($^1\Pi_g$).

In order to study the effect of the time-dependent penalty term, we perform the same experiment including only the final-time objective. In this case, the optimization procedure stops after 248 iterations (317 evaluations). We then get a transition probability of 99.5% and an average population of the $^1\Pi_g$ state of $2.2 \cdot 10^{-3}$. Thus, the penalty term indeed enforces a solution with less population on the unwanted state.

Comparing our results with the ones obtained for a much longer time span in Ref. [14], we see that we reach almost as high transfer probability. Comparing the number of iterations required in the algorithm, our results are obtained with less operations in the time-dependent but more in the time-independent target case. Note however that the authors of Ref. [14] selected the initial value as well as the weighting of the penalty term in a very specific way, employing advanced knowledge about the system whereas our algorithm starts off from a random initial guess. It is obvious that the number of iterations needed can be drastically reduced by choosing a sophisticated starting point. Therefore, it is difficult to compare the two studies. In our computations, we intentionally choose a random initial guess to demonstrate that good results can be obtained even with very little prior knowledge.
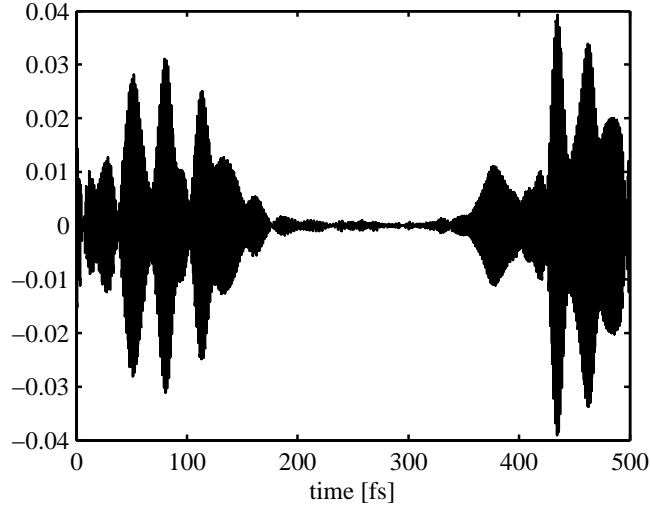
14

Figure 7: Optimized pulse for time-dependent target.

5.3 Automatic choice of frequencies

The computations discussed so far have been based on a certain choice of the Fourier coefficients. In this section, we will demonstrate an ad hoc algorithm for choosing these frequencies automatically.

Let us assume we are only able to include $n/2$ sine and cosine frequencies due to memory restrictions. Suppose further we want to search for the most important frequencies among $m > n/2$ different frequencies. For simplicity, let $a = 2m/n$ be an integer. Then, the following strategy can be used:

(i) For j from 1 to a
Solve the optimal control problem including the wave numbers $(j-1)n/2$ to $jn/2$. Record fitness and optimized Fourier coefficients. Use a low tolerance or a restriction to a low number of frequencies.

(ii) Identify the $n/2$ most important frequencies. To do so, take the frequencies with the (in modulus) highest coefficients from each interval. The higher the fitness value in each interval the more frequency values should be taken from the corresponding interval.

(iii) Solve the optimal control problem based on the frequencies identified in (ii).

In order to demonstrate the performance of this frequency search, we apply it to the Rb$_2$ example discussed in Sec. 5.2, but now with a pulse of length $t_f = 1$ ps. We assume that $n = 500$ and consider the wave number span 1–1000. With the procedure described above, 250 wave numbers within the range 501 to 894 are

15
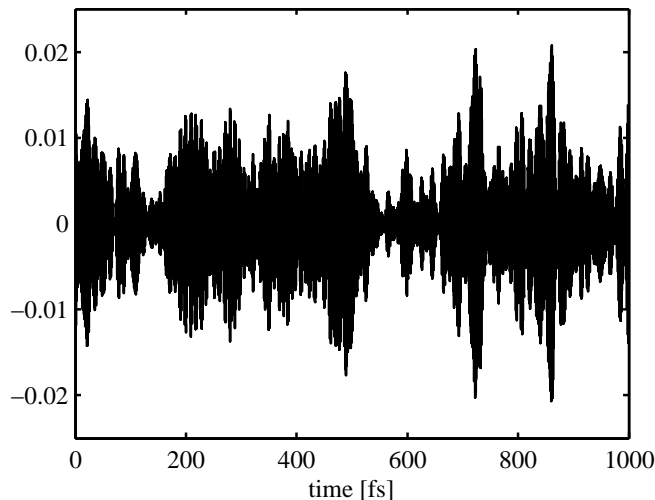
Figure 8: Optimized pulse with automatic frequency choice.

chosen. Based on these frequencies, the optimization routine stops for the pulse shown in Fig. 8 after 88 iterations (95 gradient evaluations). This pulse yields a transition to the prescribed state with a probability of 98.9 % and an average population of $9.1 \cdot 10^{-5}$ in the $^1\Pi_g$ state.

The strategy for frequency identification discussed above is not very elaborate. Nevertheless, it performs satisfactory in the numerical test. An important parameter is the number of iteration or the tolerance that is set in the preliminary step (i). One observation is that the algorithm quickly stops for frequency intervals that are unimportant, so that it is relatively easy to dismiss a whole frequency band. Still, it is not clear how fast more and less important frequencies can be separated within frequency bands that are of importance. Another shortcoming of the presented strategy is that effects that only appear when certain frequencies (that belong to different problems in (i)) interact, are excluded.

6 Conclusions

We have discussed a reformulation of the quantum optimal control problem to shape the Fourier coefficients of the pulse. In this way, we can narrow the frequency band to physically and technically relevant wave numbers. By this procedure, one can make sure that the optimized pulse can be realized in an experimental setup. Furthermore, a reduction of the dimensionality of the (discrete) control variable is achieved, and it becomes feasible to tackle the optimization problem using a quasi-Newton method, exploiting its favorable convergence properties.

16

Our results show that we can reach our target with a high probability even when starting with a random or zero initial guess. We have also seen that the algorithm clearly outperforms a Krotov-like method, which is the standard approach within the field. Furthermore, we demonstrate how to identify the most important frequencies numerically. The method described is not very elaborate but exhibits good result in our experiments.

Directions of future work include an automation of the choice of the initial value, possibly combined with a global optimization strategy. Furthermore, better insight in the way different frequencies interact could be valuable for improvements in our strategy to identify the most important frequencies.

References

[1] A. Assion, T. Baumert, M. Bergt, T. Brixner, B. Kiefer, V. Seyfried, M. Strehle, and G. Gerber. Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses. *Science*, 282:919 – 922, 1998.

[2] M. Berggren. Numerical solution of a flow-control problem: Vorticity reduction by dynamic boundary action. *SIAM J. Sci. Comput.*, 19:829 –860, 1998.

[3] M. E. Fisher and L. S. Jennings. Discrete-time optimal control problems with general constraints. *ACM Transactions on Mathematical Software*, 18:401 – 413, 1992.

[4] N. Gador, B. Zhang, H. O. Karlsson, and T. Hansson. Dynamical interference structures in fully coupled bound-bound state quantum wave-packet dynamics. *Phys. Rev. A*, 70:033418, 2004.

[5] I. Griva, S. G. Nash, and A. Sofer. *Linear and Nonlinear Optimization*. SIAM, Philadelphia, 2009.

[6] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*. Springer Netherlands, Dordrecht, 2009.

[7] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. *Acta Numerica*, 9:215 – 365, 2000.

[8] R. S. Judson and H. Rabitz. Teaching lasers to control molecules. *Phys. Rev. Lett.*, 68(10):1500–1503, Mar 1992.

[9] C. P. Koch, J. P. Palao, R. Kosloff, and F. Masnou-Seeuws. Stabilization of ultracold molecules using optimal control theory. *Phys. Rev. A*, 70:013402, 2004.

[10] K. Kormann, S. Holmgren, and H. O. Karlsson. Accurate time propagation for the Schrödinger equation with an explicitly time-dependent Hamiltonian. *J. Chem. Phys.*, 128:184101, 2008.

[11] V. F. Krotov. *Global methods in optimal control theory*. Marcel Dekker, New York, 1996.

[12] V. F. Krotov and I. N. Feldman. An iterative method for solving optimal control problems. *Eng. Cybern.*, 21:123–130, 1983.

[13] Y. Maday and G. Turinici. New formulations of monotonically convergent quantum control algorithms. *J. Chem. Phys.*, 118:8191–8196, 2003.

[14] J. P. Palao, R. Kosloff, and C. P. Koch. Protecting coherence in optimal control theory: State-dependent constraint approach. *Phys. Rev. A*, 77:063412, 2008.

[15] A. P. Peirce, M. A. Dahleh, and H. Rabitz. Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications. *Phys. Rev. A*, 37:4950 – 4964, 1988.

[16] P. Schwender, F. Seyl, and R. Schinke. Photodissociation of Ar_2^+ in strong laser fields. *Chem. Phys.*, 217:233 – 247, 1997.

[17] I. Serban, J. Werschnik, and E. K. U. Gross. Optimal control of time-dependent targets. *Phys. Rev. A*, 71:053810, 2005.

[18] S. Shi, A. Woody, and H. Rabitz. Optimal control of selective vibrational excitation in harmonic linear chain molecules. *J. Chem. Phys.*, 88:6870 – 6883, 1988.

[19] O. M. Shir, C. Siedschlag, T. Bäck, and M. J. J. Vrakking. Niching in evolution strategies and its application to laser pulse shaping. In E. Talbi, P. Liardet, P. Collet, E. Lutton, and M. Schoenauer, editors, *Artificial Evolution, LNCS 3871*, pages 85 – 96. Springer, Berlin, 2006.

[20] S. E. Sklarz and D. J. Tannor. Loading a Bose–Einstein condensate onto an optical lattice: An application of optimal control theory to the nonlinear schrödinger equation. *Phys. Rev. A*, 66:053619, 2002.

[21] J. Somlói, V. A. Kazakov, and D. J. Tannor. Controlled dissociation of I_2 via optical transitions between the X and B electronic states. *Chem. Phys.*, 172:85 – 98, 1993.

[22] D. J. Tannor. *Introduction to Quantum Mechanics: A Time-Dependent Perspective*. University Science Books, Sausalito, 2007.

[23] C. R. Vogel. *Computational Methods for Inverse Problems*. SIAM, Philadelphia, 2002.

[24] G. von Winckel and A. Borzì. Computational techniques for a quantum control problem with H_1-cost. *Inverse Problems*, 24:034007, 2008.

[25] A. M. Weiner. Femtosecond pulse shaping using spatial light modulators. *Review of Scientific Instruments*, 71:1929–1960, 2000.

[26] J. Werschnik and E. K. U. Gross. Quantum optimal control theory. *J. Phys. B: At. Mol. Opt. Phys.*, 40:R175 – R211, 2007.

[27] M. Wollenhaupt, A. Assion, and T. Baumert. Femtosecond laser pulses: Linear properties, manipulation, generation and measurement. In F. Träger, editor, *Springer Handbook of Lasers and Optics*, pages 937 – 983. Springer Science and Business Media, LLC New York, 2007.

[28] W. Zhu, J. Botina, and H. Rabitz. Rapidly convergent iteration methods for quantum optimal control of population. *J. Chem. Phys.*, 108:1953 – 1963, 1998.

[29] W. Zhu and H. Rabitz. A rapid monotonically convergent iteration algorithm for quantum optimal control over the expectation value of a positive definite operator. *J. Chem. Phys.*, 109:385 – 391, 1998.