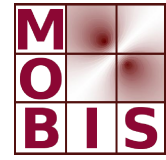




SpezialForschungsBereich F 32



Karl-Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz



**A GLOBALIZED NEWTON
METHOD FOR THE
ACCURATE SOLUTION OF A
DIPOLE QUANTUM
CONTROL PROBLEM**

Gregory von Winkel Alfio Borzi
Stefan Volkwein

SFB-Report No. 2009-004

March 2009

A-8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)



SFB sponsors:

- **Austrian Science Fund (FWF)**
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**



A GLOBALIZED NEWTON METHOD FOR THE ACCURATE SOLUTION OF A DIPOLE QUANTUM CONTROL PROBLEM*

G. VON WINCKEL[†], A. BORZI[‡], AND S. VOLKWEIN[§]

Abstract. A theoretical and computational framework is presented to obtain accurate controls for fast quantum state transitions that are needed in a host of applications such as nano electronic devices and quantum computing. This method is based on a reduced Hessian Krylov-Newton scheme applied to a norm-preserving discrete model of a dipole quantum control problem. The use of second-order numerical methods for solving the control problem is justified proving existence of optimal solutions and analyzing first- and second-order optimality conditions. Criteria for the discretization of the nonconvex optimization problem and for the formulation of the Hessian are given to ensure accurate gradients and a symmetric Hessian. Robustness of the Newton approach is obtained using a globalization strategy with a robust linesearch procedure. Results of numerical experiments demonstrate that the Newton approach presented in this paper is able to provide fast and accurate controls for high-energy state transitions.

Key words. Optimal control theory, Schrödinger equation, Newton method.

AMS subject classifications. 35Q40, 49K20, 65H10, 65M06, 65N06, 90C53.

1. Introduction. The control of quantum electronic states in physical systems has a host of applications such as quantum computers [5], control of photochemical processes [37], and semiconductor lasers [18, 31]. Quantum computers, in particular, have been the subject of much research interest since they hold the promise of performing complex calculations in polynomial time. In a quantum computer the classical logic states 0 and 1 are replaced by states of a quantum system. One such possibility is a two-level system where the occupancy of the ground state could be analogous to the logical 0 and the first excited state represents logic state 1.

In most of the envisioned applications, it is important to define fast control mechanisms that cannot be constructed based on perturbation theory strategies or on a-priori parameterized control fields. This fact motivates the increasing interest in the optimal control theory framework [24] within which many recent successful results for quantum control problems [7, 8, 9, 23, 26, 27, 28, 34, 35] have been obtained. A pioneering work in this field was done by Peirce, Dahleh, and Rabitz [33] who investigated the optimal dipole control of a diatomic molecule represented by a Morse potential. The focus in this and in other early papers (see [11, 13]) was to validate the ability of the optimal control framework to provide suitable quantum control mechanisms. However, computational difficulties due to the structure of the optimality system with bilinear control and the nonconvexity of the optimization problem led to research focused more on finite-dimensional Schrödinger equations; see, e.g., [7, 9, 28, 35]. In this case, the computational schemes of choice have been the monotonic iterative scheme [28, 35] and accelerated versions of the gradient scheme [7, 8]. These first-order schemes perform well for finite-level quantum systems and provide acceptable

*Supported in part by the Austrian Science Fund FWF project P18136-N13 “Quantum optimal control of semiconductor nanostructures”, and SFB F32 projects “Fast Multigrid Methods for Inverse Problems” and “Numerical Methods for Optimal Control with PDE Constraints”.

[†]Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität, Heinrichstraße 36, 8010 Graz, Austria (gregory.von-winckel@uni-graz.at).

[‡]Università degli Studi del Sannio, Dipartimento e Facoltà di Ingegneria, Palazzo Dell’Aquila Bosco Lucarelli, Corso Garibaldi 107, 82100 Benevento, Italy (alfio.borzi@unisannio.it).

[§]Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität, Heinrichstraße 36, 8010 Graz, Austria (stefan.volkwein@uni-graz.at)

results [27, 41] when applied to infinite-dimensional systems. However, they cannot provide second-order convergence typical of the Newton method and their convergence behavior may not be robust with respect to changes of values of discretization and optimization parameters.

It is the purpose of this work to develop an accurate, efficient, and robust Newton scheme for infinite-dimensional quantum systems in the most representative case of a dipole control structure. We remark that, although the application of the Newton scheme to solve constrained optimization problems is well known, its use for solving quantum control problems has been less successful. The reason for this fact is manifold and will be illustrated in this paper. In particular, we discuss the accurate construction of the gradient and the setting of the Hessian of the reduced optimization problem. The former requires: 1) an appropriate discretization scheme that is norm-preserving and second-order accurate also in the case of time-varying potential; 2) a discretize-before-optimize approach to avoid any inconsistency between the optimality condition and its discrete approximation; 3) a gradient which is defined in the same functional space where the control function is sought. This last point is automatically fulfilled in a full Newton scheme. Point 2) results are necessary as in our experience bilinear control problems have very flat minima and therefore are prone to gradient inconsistency when using a optimize-before-discretize approach.

The other important aspect for a successful development of the Newton scheme for solving our problem is the construction (in the sense of application) of the Hessian. For this purpose, we notice that a formal derivation of this operator, which does not take into consideration the fact that quantum control problems are defined in complex Hilbert spaces, results in nonsymmetric Hessians and thus in nonconverging schemes; see a related discussion in [4]. For this purpose, we introduce a real-valued matrix representation of complex variables and obtain the Hessian within this formalism. Together with this fact, we provide a new theoretical analysis of first- and second-order optimality conditions giving criteria such that the Hessian becomes positive definite in a neighborhood of the optimal solution. Because the theoretical analysis is quite involved, for ease of reading we present this analysis in the Appendix while we recall the main results where needed.

The Newton scheme which results taking into consideration all issues mentioned above still may lack of robustness because of nonconvexity of the optimization problem. This fact appears less clearly with large values of the regularization parameter and when considering large time intervals for the control. However, in application we need controls that are fast and accurate, i.e. with less regularization. For this purpose, we augment the Newton scheme with a robust linesearch algorithm which exploits a priori estimates and uses continuation techniques to solve our problem with small regularization. Results of numerical experiments demonstrate that the resulting globalized Newton scheme is able to compute fast controls for high-energy transitions with typical second-order convergence.

In the next section, we introduce a dipole quantum control problem and discuss modeling issues concerning the objective and the governing equation. We employ the optimal control framework by formulating a cost to minimize with the equality constraint that the particle dynamics satisfy the time-dependent Schrödinger equation (TDSE). The cost functional is designed in such a way to avoid specification of the phase of the target function. Further, we review some properties of the TDSE, define the optimality system to be solved, and show that the constrained optimization problem can not have a unique minimizer.

In Section 3, a second-order accurate discrete approximation to the Schrödinger equation with dipole control potential is introduced which shares the unitarity for the continuous TDSE. We use a real-valued matrix representation of complex variables for the discretization of the governing equation and of the objective. Hence we derive the discrete optimality system. From this we obtain the reduced objective and correspondingly define the gradient and the reduced Hessian. With this setting, in Section 4 we present all details of our Newton method that includes initialization and globalization issues. In particular, we discuss a robust linesearch where the unitarity of the TDSE is used to compute a maximum feasible steplength for the line search which also provides a useful criteria indicating when a unit steplength is feasible.

In Section 5, we investigate the ability of the optimal control framework to provide fast and accurate controls for high-energy state transitions and validate our Newton scheme. Results of numerical experiments show that with our approach we are able to obtain controls for very short time intervals and to steer high-energy transition while solving the optimality system to high accuracy, that is, to very small values of the norm of the reduced gradient. In addition, we present a comparison of the performance of the Newton scheme with that of a steepest descent algorithm and of a nonlinear conjugate gradient method, showing that our Newton approach outperforms gradient based schemes.

In the Appendix, we define our functional setting and discuss existence and uniqueness of solution to the TDSE problem. We use these results to define the reduced optimization problem and illustrate the optimality conditions. Then we prove existence of optimal solutions and provide criteria such that the second-order sufficient optimality condition holds.

A section of conclusion completes the exposition of our work.

2. Model problem and optimal control framework. We illustrate the model of a charged quantum particle subject to a stationary confining potential and a time-dependent electric control field [3]. A dipole approximation of the electric field is considered and we formulate the control problem of steering transitions of the particle among stationary states. In this section, we focus on the physical properties of the model and on the formulation of the control problem, including the main theoretical statements on the optimality system, while in the Appendix we collect details of our functional analytic setting and present our proofs of the existence of an optimal control solution including necessary and sufficient optimality conditions.

The quantum state of a particle is described by a wavefunction $\psi : Q \rightarrow \mathbb{C}$ that is governed by the time-dependent Schrödinger equation (TDSE) as follows

$$i\partial_t \psi(x, t) = \{-\partial_x^2 + V(x, t)\} \psi(x, t), \quad (x, t) \in Q = \Omega \times (0, T), \quad (2.1)$$

where Ω is the spatial domain and $(0, T)$ is the time interval, and we choose the scaling of the Planck constant $\hbar = 1$ and the mass $m = 1/2$. The potential $V(x, t)$ consists of a stationary part $V_0(x)$ and a time varying control part.

In the quantum mechanical framework, a dynamically stable system like an atom or a molecule exists in the presence of a stationary confining potential, that is, a potential with a 'well' envelope [14]. In this case one considers the following eigenproblem

$$\{-\partial_x^2 + V_0(x) - \lambda_j\} \phi_j(x) = 0, \quad j = 1, 2, \dots, \quad (2.2)$$

whose eigenfunctions represent the stationary states and the eigenvalues λ_j represent the energy of the corresponding states. The time-evolution of these states is given

by $\psi_j(x, t) = \phi_j(x) \exp(-i\lambda_j t)$. A representative stationary potential with various applications in semiconductor nanostructures [20], define on $\Omega = (0, \ell)$, is the infinite barrier well potential where $V_0(x) = 0$ for $x \in (0, \ell)$ and $V_0(0) = +\infty$ and $V_0(\ell) = +\infty$. The infinite barrier condition is equivalent to homogeneous Dirichlet boundary conditions for the wavefunction and thus we have $\lambda_j = \frac{j^2\pi^2}{\ell^2}$ and $\phi_j(x) = \sin(j\pi x/\ell)$.

Although the methodology in this work is not limited to the infinite quantum well potential, in application it is important to consider this case to determine a control function $V(x, t)$ which allows transitions of a charged particle from one stationary state to another of a quantum well over a short time interval. A physically meaningful control mechanism is an electric control field modeling a laser pulse. Using the dipole approximation the total potential results in the form $V(x, t) = V_0(x) + u(t)x$, where $u : (0, T) \rightarrow \mathbb{R}$ is the modulating control amplitude. Our approach generalizes to n dimensions considering a vector-valued control $\mathbf{u}(t) \cdot \mathbf{x}$ with $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{u} : (0, T) \rightarrow \mathbb{R}^n$.

Next, we discuss some important properties of the solution to the TDSE and report a few results from perturbation theory. We write $\Re(z)$ and $\Im(z)$ for the real and imaginary part of a complex $z \in \mathbb{C}$. Moreover, z^* stands for the complex conjugate of z and $|z|_{\mathbb{C}} = \sqrt{z^*z}$ for its absolute value. For our discussion, we define $\mathcal{H} = L^2(\Omega; \mathbb{C})$, the Hilbert space endowed with the inner product

$$(\varphi, \psi)_{\mathcal{H}} = \int_{\Omega} \varphi(x)^* \psi(x) dx \quad \text{for } \varphi, \psi \in \mathcal{H},$$

and the induced norm $\|\varphi\|_{\mathcal{H}}$ for $\varphi \in \mathcal{H}$. The Hilbert space $\mathcal{V} = H_0^1(\Omega; \mathbb{C})$ is given by

$$\mathcal{V} = \left\{ \varphi \in \mathcal{H} \mid \|\varphi\|_{\mathcal{V}} = \left(\int_{\Omega} |\varphi'(x)|_{\mathbb{C}}^2 dx \right)^{1/2} < \infty, \quad \varphi = 0 \text{ on } \partial\Omega \right\},$$

supplied with the inner product $(\varphi, \psi)_{\mathcal{V}} = (\varphi', \psi')_{\mathcal{H}}$ for $\varphi, \psi \in \mathcal{V}$ and the induced norm $\|\varphi\|_{\mathcal{V}}$. We also need the Hilbert space

$$\mathcal{W} = L^2(0, T; H_0^1(\Omega; \mathbb{C}) \cap H^2(\Omega; \mathbb{C})) \cap H^1(0, T; L^2(\Omega; \mathbb{C}))$$

For more details on the above Lebesgue and Sobolev spaces and more weaker spaces see the Appendix and, e.g., [1, 17].

Now, consider the TDSE with an initial state of the quantum system given by $\psi_0(x) \in \mathcal{V}$ at $t = 0$. One recognizes that the Schrödinger evolution operator is time-reversible (non-dissipative) and therefore ψ cannot have better regularity than ψ_0 [10]. We also see that with a time-varying potential there is no energy conservation. In fact we want to change the energy of the system. However, we have mass conservation as stated by the following

PROPOSITION 2.1. *Let $V(x, t) = V_0(x) + u(t)x$ and $\|\psi_0(\cdot)\|_{\mathcal{H}} = 1$, then we have $\|\psi(\cdot, t)\|_{\mathcal{H}} = 1$ for all $t \in [0, T]$.*

Proof. The time rate of change of the total probability is

$$\partial_t \|\psi\|_{\mathcal{H}}^2 = (\psi, \psi_t)_{\mathcal{H}} + (\psi_t, \psi)_{\mathcal{H}} = 2\Re(\psi, \psi_t)_{\mathcal{H}} = 2\Re(\psi, i\psi_{xx} - iu(t)x\psi)_{\mathcal{H}} \quad (2.3)$$

where

$$(\psi, i\psi_{xx} - iu(t)x\psi)_{\mathcal{H}} = \int_{\Omega} \psi^* (i\psi_{xx} - iu(t)x\psi) dx = -i\|\psi_x\|_{\mathcal{H}}^2 - iu(t) \int_{\Omega} \psi^* x\psi dx.$$

Since x is a self-adjoint operator (in the sense that $A : \mathcal{H} \rightarrow \mathcal{H}$ defined by $(A\psi)(x) = x\psi(x)$ is a self-adjoint operator) with respect to the inner product, it means that we are taking the real part of a purely imaginary quantity in (2.3). It follows that $\partial_t \|\psi\|_{\mathcal{H}}^2 = 0$ for all time and so $\|\psi(\cdot, t)\|_{\mathcal{H}}^2 = \|\psi_0\|_{\mathcal{H}}^2$. \square

We now show that a spatially symmetric potential results in symmetric or anti-symmetric eigenfunctions. Consider $\Omega = (0, \ell)$ and let $\ell_m = \ell/2$.

PROPOSITION 2.2. *If the stationary Schrödinger equation has a symmetric potential, $V_0(\ell_m + x) = V_0(\ell_m - x)$, then the eigenfunctions must have even or odd parity with respect to ℓ_m .*

Proof. Let \mathcal{P} be the parity operator defined $\mathcal{P}f(\ell_m + x) = f(\ell_m - x)$ and $H_0 = -\partial_x^2 + V_0$ be the stationary Hamiltonian with $H_0\phi_j = \lambda_j\phi_j$. We have

$$\mathcal{P}H_0\phi_j(\ell_m + x) = \lambda_j\phi_j(\ell_m - x) = H_0\mathcal{P}\phi_j(\ell_m + x) \Rightarrow [\mathcal{P}, H_0] = 0. \quad (2.4)$$

Since the operators commute, ϕ_j must be an eigenfunction of the parity operator $\mathcal{P}\phi_j = \mu_j\phi_j$. Given that $\mathcal{P}^2\phi_j = \phi_j$, the parity eigenvalue must be $\mu_j = \pm 1$. Therefore, all eigenfunctions must be either symmetric $\phi_j(\ell_m + x) = \phi_j(\ell_m - x)$ or antisymmetric $\phi_j(\ell_m + x) = -\phi_j(\ell_m - x)$. \square

An important result of perturbation theory is that in a long time horizon a time-harmonic control $u(t)$ is able to induce transition between two eigenstates if its frequency equals the difference of energy of the two states [14]. Therefore, it is relatively easy to control state transitions for long time intervals as we show in the example below. However, the problem becomes very difficult if short time intervals are considered. Here short means that $T \approx 2\pi/\omega$ and in this case the perturbation theory is inapplicable and the control $u(t)$ deviates greatly from an harmonic function. In this case we use an optimal control approach.

To formulate the optimal control problem, we have to decide in which functional space the control is sought. From the previous discussion, it appears that the control space $\mathcal{U} = H_0^1(0, T; \mathbb{R})$ is the most appropriate for dipole controls as it means that the laser pulse cannot change instantaneously and it accommodates sinusoidal functions for long time controls. In the optimal control framework, this choice means that the objective of the optimization has a regularization term of the form $\|u\|_{\mathcal{U}}^2$, where this norm is induced by the following inner product

$$(u, v)_{\mathcal{U}} = \int_0^T (u(t)v(t) + \alpha \dot{u}(t)\dot{v}(t)) dt \quad \text{for } u, v \in \mathcal{U},$$

with $0 < \alpha \ll 1$. Notice that with this norm the control is continuous since $H_0^1(0, T)$ is compactly embedded in $C^0([0, T])$ in one dimension. Use of smaller values of α allows for controls with larger rates of change. We have that the control is zero at the beginning and at the end of the time interval which is the maximum time-window for the laser pulse.

Our control problem requires to finding a control $u \in \mathcal{U}$ such that a quantum system initially in the state ψ_0 evolves with (2.1) to a state $\psi(\cdot, T)$ that is as close as possible to a desired target configuration $\tilde{\psi}$. This aim is formulated by requiring to minimize the objective given by the following cost functional

$$\min_{\psi \in \mathcal{W}, u \in \mathcal{U}} J(\psi, u) := \frac{1}{2} \left(1 - \|P\psi(\cdot, T)\|_{\mathcal{H}}^2 \right) + \frac{\gamma}{2} \|u\|_{\mathcal{U}}^2, \quad (2.5)$$

under the constraint given by the TDSE, including the initial condition, denoted as follows

$$e(\psi, u) := \{\partial_t - iH(u)\} \psi = 0, \quad (2.6)$$

where $H(u) = -\partial_x^2 + V_0(x) + u(t)x$ and we introduce the projector $P\psi = (\tilde{\psi}, \psi)_{\mathcal{H}} \tilde{\psi}$. The goal of the first term of the cost functional is to track the given terminal state $\tilde{\psi}$ up to a global phase $e^{i\varphi T}$ which cannot be specified. In the Appendix, we prove existence of a unique solution to $e(\psi, u) = 0$ for a given $u \in \mathcal{U}$ and discuss the differentiability properties of the operator e and of the objective $J(\psi, u)$ as required for the optimal control formulation.

Before we discuss the solution of this constrained optimization problem, we consider the case of a control of sinusoidal type designed to drive our quantum model from the first eigenstate $\phi_1(x) = \sin(\pi x/2)$ to the second one $\phi_2(x) = \sin(2\pi x/2)$ where we choose $\ell = 2$. We take a control of the form

$$u(t) = u_0 \left[\sin(\omega t) - \sin(\omega T) \frac{t}{T} \right], \quad \omega = \lambda_2 - \lambda_1, \quad (2.7)$$

with $\lambda_2 - \lambda_1 = 3\pi^2/4$ and the linear term is such that $u(0) = 0$ and $u(T) = 0$. With this control in (2.6) and $\psi_0 = \phi_1$, we solve the forward problem and obtain $\psi(\cdot, T)$ which is used in (2.5) together with $\tilde{\psi} = \phi_2$ to determine $J(\psi, u)$. Results with this setting are reported in the two pictures of Figure 2.1. The left picture shows that

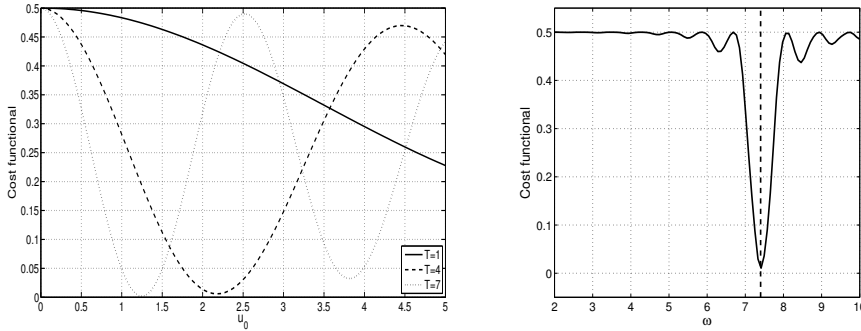


FIG. 2.1. Left: Values of the objective with nearly time-harmonic control with fixed frequency $\omega = 3\pi^2/4$ and varying amplitude u_0 for three different time horizons T . Right: Values of the objective for fixed amplitude $u_0 = 1$ and $T = 8$ for variable frequency.

choosing $\omega = 3\pi^2/4$, smaller amplitudes are required in correspondence to larger time horizons in order to attain the minimum of the objective. However, with $T = 1$ we have $\omega T \approx 2\pi$ and the objective does not change considerably as we increase the amplitude. On the other hand, in the right picture we see that, choosing T sufficiently large and a fixed amplitude, a clear minimum is obtained for $\omega = 3\pi^2/4$.

We show that a control suitable for fast state transition can be obtained by in the optimal control formulation given by (2.5) and (2.6). To characterize the solution to this problem, we introduce the following Lagrangian

$$L(\psi, u, p) = J(\psi, u) + \Re \int_0^T \int_{\Omega} p^*(x, t) e(\psi, u)(x, t) dx dt, \quad (2.8)$$

where p is the Lagrange multiplier. We prove in the Appendix that any minima of (2.5) and (2.6) corresponds to an extremal point of the Lagrangian; see also [24, 25]. Therefore, by taking the Frechét derivatives of $L(\psi, u, p)$ with respect to the optimization variables gives the following first-order optimality system that characterizes the optimal solution, we have

$$\begin{aligned} \{i\partial_t + \partial_x^2 - V_0(x) - u(t)x\} \psi(x, t) &= 0, \\ \{i\partial_t + \partial_x^2 - V_0(x) - u(t)x\} p(x, t) &= 0, \\ -\gamma u + \gamma \alpha \ddot{u} + \Re \int_{\Omega} p^*(x, t) x \psi(x, t) dx &= 0 \end{aligned} \quad (2.9)$$

This system consists of the state equation, the adjoint equation, and the optimality condition, respectively, with homogeneous Dirichlet boundary conditions, and initial and terminal conditions given by

$$\begin{aligned} \psi(x, 0) &= \psi_0(x), \\ p(x, T) &= i(\tilde{\psi}(\cdot), \psi(\cdot, T))_{\mathcal{H}} \tilde{\psi}(x), \\ u(0) &= 0, \quad u(T) = 0. \end{aligned} \quad (2.10)$$

In the Appendix, we prove that there exists at least one solution to (2.9)-(2.10). We also prove that if γ is sufficiently large and the projection $\|P\psi(T)\|_{\mathcal{H}}^2$ is sufficiently small, then the second-order sufficient optimality condition holds and our optimization problem is locally strictly convex.

Notice that the control is a function of time only, and the state and adjoint variables can be seen as implicit functions of the control. Therefore the dimensionality of the optimization problem can be reduced significantly introducing a reduced cost functional $\tilde{J}(u) = J(\psi(u), u)$. In the Appendix we show that the corresponding gradient is given by

$$(\nabla \tilde{J}(u))(t) = \gamma u(t) - \gamma \alpha \ddot{u}(t) - \Re \int_{\Omega} p^*(x, t) x \psi(x, t) dx. \quad (2.11)$$

Therefore we have that $\nabla \tilde{J}(u) \in H^{-1}(0, T; \mathbb{R})$ which is problematic with a gradient-based approach because the gradient is not in the same space of the solution and thus neither it provides an update to the control along the descent direction. It has been shown [41], that this problem can be solved by using the Riesz representation of the gradient in the $H_0^1(0, T)$ space as a means of Sobolev smoothing. Although we will ultimately be working with a discrete optimality system, the idea of formulating the gradient in a weighted ℓ^2 space follows analogously.

A main difficulty in the analysis and solution of our quantum control problem is that it may admit multiple solutions (as most bilinear control problems). We prove that this is the case in the following.

PROPOSITION 2.3. *Let the initial and target states be eigenfunctions and the stationary potential be symmetric, then the reduced cost functional does not have a unique minimizer. In particular, if $u^*(t)$ is a minimizer, then so is $-u^*(t)$ and consequently $\nabla \tilde{J}(u)$ is nonconvex independently of regularization.*

Proof. Let $\psi(x, t)$ be a solution to the TDSE with symmetric stationary potential

$$[i\partial_t + \partial_x^2 - V_0(x) - u(t)x] \psi(x, t) = 0$$

and

$$[i\partial_t + \partial_x^2 - V_0(-x) - (-u(t))(-x)]\psi(x, t) = 0.$$

This implies that if a control $u(t)$ yields a solution $\psi(x, t)$, then $-u(t)$ yields a solution which is spatially reversed $\psi(-x, t)$. The projection of the final state onto the target has the value $(\psi(\cdot), \psi(\cdot, T))_{\mathcal{H}}$. Since $\tilde{\psi}(x)$ is an eigenfunction, it has either even or odd parity. The final value of the wavefunction $\psi(x, T)$ is given by the control $u(t)$ and $\mathcal{P}\psi(x, T)$ is given by $-u(t)$. The parity operator is self-adjoint on L^2 ($\mathcal{P} = \mathcal{P}^*$), by virtue of commuting with the self-adjoint Hamiltonian. This gives the relationship

$$\left| \int_{\Omega} \mathcal{P}\psi^*(x, T)\tilde{\psi}(x)dx \right|^2 = \left| \int_{\Omega} \psi^*(x, T)\mathcal{P}\tilde{\psi}(x)dx \right|^2 = \left| \int_{\Omega} \psi^*(x, T)[\pm\tilde{\psi}(x)]dx \right|^2$$

The cost functional depends only on the magnitude of this projection, therefore $\tilde{J}(u) = \tilde{J}(-u)$. The consequence of the above nonuniqueness property is that the cost functional cannot be globally convex, independently of how large the regularization parameter may be. \square

3. Formulation of the discrete optimal control problem. In a PDE-based optimization problem, there are two possible discretization procedures. One is the optimize-before-discretize approach in which we discretize the optimality system (2.9) choosing appropriate discretization schemes for the forward equation, for the adjoint equation, and for the optimality equation. The drawback with this approach is the possible inconsistency between the discretized objective and the reduced gradient given by the discrete optimality condition; see [12]. This means discrepancy between the directional derivative $(\nabla\tilde{J}(u), \phi)_{\mathcal{H}}$ and its approximation $\frac{\tilde{J}(u+\epsilon\phi) - \tilde{J}(u-\epsilon\phi)}{2\epsilon}$, which, however, can be controlled at the cost of increasing accuracy of discretization by using finer meshes. The other drawback of the optimize-before-discretize approach is that the Hessian may not be symmetric.

In our experience, gradient inconsistency is usually not negligible in the case of hyperbolic- and Schrödinger-type equations with bilinear controls. For this reason, we pursue the approach of discretize-before-optimize where the consistency between the reduced objective and its gradient is guaranteed. The first step is to discretize the state equation and the Lagrangian and then take derivatives to obtain first- and second-order optimality conditions. The difficulty of this approach is that the approximation scheme resulting for the adjoint equation may be numerically disadvantageous.

In any case, the discretization of the Schrödinger equation must be norm-preserving to ensure a discrete $\partial_t\|\psi(\cdot, t)\|_{\mathcal{H}}^2 = 0$. This is essential, otherwise it happens that we may compute a control which attains a minimum of the functional by violating the underlying physical constraint. Moreover, we want a scheme that is second-order accurate also with time-varying potentials.

We know that the classical Crank-Nicolson scheme is norm-preserving and unconditionally stable when solving the Schrödinger equation with stationary potential [33]. However, in quantum control problems, the Hamiltonian between different time steps is different, $H(t_k) \neq H(t_{k-1})$, and in this case the Crank-Nicolson scheme is not norm-preserving. For this reason, we define a modified Crank-Nicolson (MCN) method by first integrating numerically the semigroup operator and then using a Padé approximant.

Consider the semigroup operator $\psi(\delta t) = \exp(\Theta(\delta t))\psi(0)$, where $\Theta(t)$ is given by the Magnus expansion [22]. This is an infinite series representation for a time-dependent Hamiltonian. In the case of a dipole control with $H(t) = -\partial_x^2 + V_0(x) + u(t)x$, the Magnus expansion terminates after three terms.

$$\Theta(t) = -i \int_0^t H(\tau) d\tau + \frac{1}{2} \int_0^t \left[\int_0^\tau H(\sigma) d\sigma, H(\tau) \right] d\tau + \dots \quad (3.1)$$

We choose to approximate $\Theta(t)$ by the first term in the expansion. The analysis of a similar second-order truncation with a midpoint rule is carried out by Hochbruck and Lubich in [21]; see also [39]. In our approach, we approximate the integral in the first term of (3.1) with the trapezoidal rule. The choice of the trapezoidal rule over the midpoint rule results in control and state variables that are defined on the same time step. Since the Magnus formula and the integral are approximated to second order, the overall scheme is second-order accurate also in the case when the potential is time-dependent. In addition, we can prove that our scheme is unconditionally stable.

Now, let N_t be the number of time steps of size $\delta t = \frac{T}{N_t}$ and N_x be the number of intervals of the Ω discretization. The TDSE discretized by our MCN scheme results in the following

$$\psi_k - \psi_{k-1} = -\frac{i\delta t}{4} [H(t_k) + H(t_{k-1})][\psi_k + \psi_{k-1}]. \quad (3.2)$$

Spatial discretization of the Hamiltonian $H(t_k)$ is carried out using linear finite elements on a uniform grid which results in a matrix \mathbf{H}_k . We have that $\mathbf{H}_k = \mathbf{H}_k^\top$, which is important for preserving unitarity of the time-stepping method. Let $\mathbf{A}_k = \frac{\delta t}{4} [\mathbf{H}_k + \mathbf{H}_{k-1}]$. This gives the desirable property that $\mathbf{A}_k = \mathbf{A}_k^\top$, which is consistent with the infinite-dimensional Hamiltonian operator being Hermitian.

We notice an important fact: While we have no difficulty with the complex-valued variable representation in gradient-based optimization approaches [41], the real-valued matrix representation of complex variables is necessary for constructing the Hessian; see [40]. Suppose $z_1, z_2 \in \mathbb{C}$ with $z_\ell = x_\ell + iy_\ell$ for $\ell = 1, 2$. These can be represented as vectors in \mathbb{R}^2 with $Z_\ell = \begin{pmatrix} x_\ell & y_\ell \end{pmatrix}^\top$. Notice that, considering Z_1 and Z_2 as vectors in \mathbb{R}^2 , it is not possible to use vector multiplication between them which results equivalent to multiplication between the corresponding complex variables. For this purpose, a matrix representation of one of the two vectors is employed as follows.

$$Z_1 Z_2 = \begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 x_2 - y_1 y_2 \\ x_1 y_2 + x_2 y_1 \end{pmatrix} \quad (3.3)$$

In this representation, complex conjugacy is performed via the transpose operation. This representation is also valid for matrices and vectors and in such cases leads to block systems. The spatially-discrete form of Equation (3.2) contains a term $-i\mathbf{A}_k$, for which therefore a real-valued representation is needed. Including the left-hand side terms, we obtain a block matrix \mathbf{B}_k as follows

$$\mathbf{B}_k = \begin{pmatrix} \mathbf{I} & \mathbf{A}_k \\ -\mathbf{A}_k & \mathbf{I} \end{pmatrix}. \quad (3.4)$$

This gives the following representation of the equality constraint

$$\mathbf{e}_k(\mathbf{y}, \mathbf{u}) = \mathbf{B}_k \mathbf{y}_k - \mathbf{B}_k^\top \mathbf{y}_{k-1}, \quad \mathbf{y}_k = \begin{pmatrix} \Re[\psi_k] \\ \Im[\psi_k] \end{pmatrix}, \quad (3.5)$$

where \mathbf{y} is a compact notation for the set of state vectors at each time step $\mathbf{y}_1, \dots, \mathbf{y}_{N_t}$ and similarly for \mathbf{u} .

Given a total number N_x of grid points, the wavefunction at each time step will be a vector in \mathbb{C}^{N_x-2} , therefore the state vector at each step will be in \mathbb{R}^{2N_x-4} . We introduce the matrix operator \mathbf{S} which corresponds, in matrix representation form, to multiplication by i . We have that

$$\mathbf{S} = \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}, \quad (\tilde{\psi}, \psi)_{\mathcal{H}} \text{ corresponds to } \begin{pmatrix} \tilde{\mathbf{y}}^\top \\ \tilde{\mathbf{y}}^\top \mathbf{S} \end{pmatrix} \mathbf{y} \quad (3.6)$$

In this representation, we can rewrite the original cost functional in the form

$$J(\mathbf{y}, \mathbf{u}) = \frac{1}{2} \left[1 - \mathbf{y}_{N_t}^\top \begin{pmatrix} \tilde{\mathbf{y}} & -\mathbf{S}\tilde{\mathbf{y}} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{y}}^\top \\ \tilde{\mathbf{y}}^\top \mathbf{S} \end{pmatrix} \mathbf{y}_{N_t} \right] + \frac{\gamma}{2} \mathbf{u}^\top \mathbf{K} \mathbf{u} \quad (3.7)$$

The matrix \mathbf{K} is symmetric and positive definite and is a linear finite-element discretization of the Helmholtz operator $I - \alpha \partial_x^2$. This means that the weighted inner product $\mathbf{u}^\top \mathbf{K} \mathbf{v}$ is a second-order approximation to $(u, v)_U$.

We consider the Lagrangian

$$\mathbf{L}(\mathbf{y}, \mathbf{u}, \mathbf{p}) = J(\mathbf{y}, \mathbf{u}) + \sum_{k=1}^{N_t} \mathbf{p}_k^\top \mathbf{e}_k(\mathbf{y}, \mathbf{u}) \quad (3.8)$$

Differentiating this Lagrangian with respect to its arguments and setting the derivatives to zero gives the discrete first-order optimality system

$$\mathbf{B}_k \mathbf{y}_k = \mathbf{B}_k^\top \mathbf{y}_{k-1}, \quad (3.9)$$

$$\mathbf{B}_k^\top \mathbf{p}_k = \mathbf{B}_{k+1} \mathbf{p}_{k+1}, \quad (3.10)$$

$$\gamma \mathbf{K} \mathbf{u} = \mathbf{f}. \quad (3.11)$$

The control vector $\mathbf{u} = (u_1, \dots, u_{N_t-1})$ has $N_t - 1$ elements as the control is set to zero at the initial and final times. The elements of the vector $\mathbf{f} = (f_1, \dots, f_{N_t-1})$ are given as

$$f_j = \mathbf{p}_j^\top \mathbf{V}(\mathbf{y}_j + \mathbf{y}_{j-1}) + \mathbf{p}_{j+1}^\top \mathbf{V}(\mathbf{y}_{j+1} + \mathbf{y}_j), \quad \mathbf{V} = \frac{\delta t}{4} \begin{pmatrix} \mathbf{0} & -\mathbf{X} \\ \mathbf{X} & \mathbf{0} \end{pmatrix}, \quad (3.12)$$

where \mathbf{V} is the discrete approximation of $\frac{i\delta t}{4} x$ and $\mathbf{X} = \text{diag}(x_1, \dots, x_{N_x-1})$. The initial and terminal conditions are as follows

$$\mathbf{y}_0 = \begin{pmatrix} \Re e[\psi_0] \\ \Im m[\psi_0] \end{pmatrix} \quad (3.13)$$

$$\mathbf{B}_{N_t}^\top \mathbf{p}_{N_t} = \begin{pmatrix} \tilde{\mathbf{y}} & -\mathbf{S}\tilde{\mathbf{y}} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{y}}^\top \mathbf{y}_{N_t} \\ \tilde{\mathbf{y}}^\top \mathbf{S} \mathbf{y}_{N_t} \end{pmatrix} \quad (3.14)$$

Notice that our modified Crank-Nicolson scheme results in a unitary time-stepping, while the resulting scheme for the adjoint equation is non-unitary. This situation is reversed in using the standard Crank-Nicolson scheme for the forward problem. We emphasize that because the cost functional depends on the state variable, it is essential to preserve the norm of the state variable.

Now, consider the reduced cost functional $\tilde{J}(\mathbf{u}) = J(\mathbf{y}(\mathbf{u}), \mathbf{u})$. The gradient of this objective function is

$$\nabla \tilde{J}(\mathbf{u}) = \gamma \mathbf{K} \mathbf{u} - \mathbf{f} \quad (3.15)$$

Since the control itself has only $N_t - 1$ degrees of freedom and the state variable has $(N_t - 1)(N_x - 2)$ degrees of freedom, the matrix of the optimality system is very large with $[(2N_x - 3)(N_t - 1)]^2$ elements, albeit sparse. In contrast, the reduced Hessian will be a full matrix with $(N_t - 1)^2$ entries. Because of the structure of the optimal control problem, applying the Hessian to a vector is not significantly more expensive than computing the gradient, so the reduced Hessian approach is suitable for use with a Krylov solver. The application of the Hessian is presented in 2. The reduced Hessian is formulated following [6].

The equality constraint can be written as a block bidiagonal system where $\mathbf{e}(\mathbf{y}, \mathbf{u})$ is a matrix and \mathbf{y} is a column-stacked vector. We have

$$\nabla^2 \tilde{J}(\mathbf{u}) = \mathbf{L}_{\mathbf{u}\mathbf{u}} + \mathbf{e}_{\mathbf{u}}^* \mathbf{e}_{\mathbf{y}}^{-\top} \mathbf{L}_{\mathbf{y}\mathbf{y}} \mathbf{e}_{\mathbf{y}}^{-1} \mathbf{e}_{\mathbf{u}} - \mathbf{e}_{\mathbf{u}}^{\top} \mathbf{e}_{\mathbf{y}}^{-\top} \mathbf{L}_{\mathbf{y}\mathbf{u}} - \mathbf{L}_{\mathbf{y}\mathbf{u}}^{\top} \mathbf{e}_{\mathbf{y}}^{-1} \mathbf{e}_{\mathbf{u}}. \quad (3.16)$$

Linearizing the system about a control \mathbf{u} , the change in the state variable \mathbf{y} due to a change $\delta \mathbf{u}$ is $\delta \mathbf{y} = -\mathbf{e}_{\mathbf{y}}^{-1} \mathbf{e}_{\mathbf{u}} \delta \mathbf{u}$. We have that

$$[\nabla^2 J(\mathbf{u})] \delta \mathbf{u} = \mathbf{L}_{\mathbf{u}\mathbf{u}} \delta \mathbf{u} + \mathbf{e}_{\mathbf{u}}^{\top} \delta \mathbf{p} + \mathbf{L}_{\mathbf{y}\mathbf{u}}^{\top} \delta \mathbf{y}, \quad (3.17)$$

where $\delta \mathbf{y} = -\mathbf{e}_{\mathbf{y}}^{-1} \mathbf{e}_{\mathbf{u}} \delta \mathbf{u}$ and $\delta \mathbf{p} = -\mathbf{e}_{\mathbf{y}}^{-\top} [\mathbf{L}_{\mathbf{y}\mathbf{y}} \delta \mathbf{y} + \mathbf{L}_{\mathbf{y}\mathbf{u}} \delta \mathbf{u}]$. We can write the linearized equality constraint with the following

$$\mathbf{e}_{\mathbf{y}} = \begin{pmatrix} \mathbf{B}_1 & & & & \\ -\mathbf{B}_1^{\top} & \mathbf{B}_2 & & & \\ & \ddots & \ddots & & \\ & & & -\mathbf{B}_{N_t-1}^{\top} & \mathbf{B}_{N_t} \end{pmatrix}, \quad \mathbf{e}_{\mathbf{u}} = \begin{pmatrix} \mathbf{c}_1 & & & & \\ \mathbf{c}_2 & \mathbf{c}_2 & & & \\ & \ddots & \ddots & & \\ & & & \mathbf{c}_{N_t-1} & \mathbf{c}_{N_t-1} \\ & & & & \mathbf{c}_{N_t} \end{pmatrix} \quad (3.18)$$

where $\mathbf{c}_k = \mathbf{V}(\mathbf{y}_k + \mathbf{y}_{k-1})$. The application of $\mathbf{e}_{\mathbf{y}}^{-1}$ and $\mathbf{e}_{\mathbf{y}}^{-\top}$ follows the two-term recursion form of the forward and adjoint equations given in (3.9) and (3.10). The matrix $\mathbf{L}_{\mathbf{u}\mathbf{u}}$ is given by $\gamma \mathbf{K}$ and $\mathbf{L}_{\mathbf{y}\mathbf{u}}$ has the form

$$\mathbf{L}_{\mathbf{y}\mathbf{u}} = \begin{pmatrix} \mathbf{V}^{\top}(\mathbf{p}_1 + \mathbf{p}_2) & \mathbf{V}^{\top} \mathbf{p}_2 & & & \\ \mathbf{V}^{\top} \mathbf{p}_2 & \mathbf{V}^{\top}(\mathbf{p}_2 + \mathbf{p}_3) & \mathbf{V}^{\top} \mathbf{p}_3 & & \\ & \ddots & \ddots & & \\ & & & \mathbf{V}^{\top} \mathbf{p}_{N_t-1} & \mathbf{V}^{\top}(\mathbf{p}_{N_t-1} + \mathbf{p}_{N_t}) \\ & & & & \mathbf{V}^{\top} \mathbf{p}_{N_t} \end{pmatrix} \quad (3.19)$$

It is the $\mathbf{L}_{\mathbf{y}\mathbf{y}}$ block that necessitates the real variable formulation, because the Lagrangian is not analytic in ψ . $\mathbf{L}_{\mathbf{y}\mathbf{y}}$ is a block matrix of $N_t - 1 \times N_t - 1$ blocks, each of size $N_x - 2 \times N_x - 2$. However, since the cost depends only on final-time observation, these blocks are all zero except for the lowest right-most block. This block contains the projection matrix of at least rank 2, i.e. $\mathbf{P} = \begin{pmatrix} \tilde{\mathbf{y}} & -\mathbf{S}\tilde{\mathbf{y}} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{y}} & -\mathbf{S}\tilde{\mathbf{y}} \end{pmatrix}^{\top}$. Writing this explicitly in real and imaginary parts, results in the matrix

$$\mathbf{P} = \begin{pmatrix} \psi_r \psi_r^{\top} + \psi_i \psi_i^{\top} & 0 \\ 0 & \psi_r \psi_r^{\top} + \psi_i \psi_i^{\top} \end{pmatrix} \quad (3.20)$$

There is no way to form this operator utilizing the dyadic product of a complex-valued vector with itself. Indeed, the fact that a rank-2 matrix is needed instead of a rank-1 matrix indicates the lack of an equivalent complex representation.

We have completed the formulation of the discrete optimal control problem that has been tailored to guarantee a consistent gradient and an appropriate construction of the Hessian.

4. A globalized Newton method. In this section, we discuss our globalized Newton approach focusing on two issues. First, we discuss the problem of starting the Newton procedure which arises due to the fact that, in a quantum control setting, the starting configuration and the target are usually orthogonal functions. In particular, it is meaningful to require to reach the target state exactly, which is possible if the weight of the regularization terms goes to zero. However, in this case it is difficult to argue positive definiteness of the Hessian. The second issue arises from the lack of convexity of our optimization problem and the need of a robust and efficient linesearch procedure.

Our numerical experience shows that the initialization of the minimizing iteration is a delicate step. Guessing an initial control to be nearly harmonic as in equation (2.7) can be a viable strategy when $\omega T \gg 1$. However, for short time intervals, a more robust initialization is required. Clearly, an initial guess of a zero control does not work. As the initial and target states are both eigenfunctions, they are orthogonal and the projection onto the target will be zero in the absence of a non-zero control function. Then the particle remains in the initial state and, $P\psi = (\tilde{\psi}, \psi)_{\mathcal{H}} \tilde{\psi} = 0$ which makes the adjoint variable and consequently the gradient identically zero. We have that the cost functional has a local maximum. If one were to modify the target state to be some $\hat{\psi} \in \text{span}(\psi_0, \tilde{\psi})$, then the projection onto the modified target will be non-zero, which in turn gives a non-zero gradient.

An additional difficulty arises from the parity of the eigenstates. In particular, if ψ_0 and $\tilde{\psi}$ are either both even or both odd, then the integral term in (2.11) will vanish due to an antisymmetric integrand. For this reason, we recommend modifying the target state to $\hat{\psi} \in \text{span}(\psi_0, \tilde{\psi}, x\tilde{\psi})$. Once a control is computed for the transition $\psi_0 \rightarrow \hat{\psi}$, it may be used as a starting guess for the transition $\psi_0 \rightarrow \tilde{\psi}$. This approach mimics an homotopy method. For very fast controls and large energy transitions, we use a sequence of intermediate targets $\hat{\psi}_1, \hat{\psi}_2, \dots$ which approach $\tilde{\psi}$ in L^2 .

The choice of regularization parameters plays an important role in the initialization as well. It is desirable to start having a large regularization, that gives faster convergence in computing the Newton direction with `symmlq`. At the same time, a large weight of the cost will compromise the goal of attaining the target as close as possible. Therefore we employ successive reductions of the regularization parameter. We have found that halving γ whenever $\|\nabla \tilde{J}(u)\|_{\mathcal{U}}$ reaches a given tolerance provides robust convergence to optimal controls corresponding to very small regularization.

We start the discussion of our globalized Newton scheme giving the workflow of the main Algorithm 1, followed by the illustration of the linesearch procedure given by Algorithm 3. To improve readability, we use the notation of the continuous formulation. Their discrete counterpart is discussed in the previous section.

Notice that in Algorithm 3 we anticipate the possible lack of positive definiteness, while still exploiting the symmetry of the Hessian, by using the Krylov-type symmetric LQ method [32]. In our experience, `symmlq` consistently computes search directions in less time than other Krylov methods, such as GMRES or BiCG. If the Hessian has negative eigenvalues, `symmlq` may compute an ascent direction. Whether the

```

Data: Given  $\psi_0, \tilde{\psi}, \gamma, \alpha, T, u = 0$ 
Choose  $\hat{\psi}$ ;
while  $\|\nabla \tilde{J}(\mathbf{u})\| > tol$  do
    Compute search direction  $\delta u$  with Algorithm 3;
    Compute  $a^*$  with Algorithm 4;
     $\mathbf{u} \leftarrow \mathbf{u} + a^* \delta \mathbf{u}$ ;
end
 $\hat{\psi} \leftarrow \tilde{\psi}$ ;
 $iter \leftarrow 0$ ;
while  $iter < maxit$  do
    while  $\|\nabla \tilde{J}(\mathbf{u})\| > tol$  do
        Compute search direction  $\delta u$  with Algorithm 3 ;
        Compute  $a^*$  with Algorithm 4;
         $\mathbf{u} \leftarrow \mathbf{u} + a^* \delta \mathbf{u}$ ;
    end
     $\gamma \leftarrow \gamma/2$ ;
     $iter \leftarrow iter + 1$ ;
end
    
```

Algorithm 1: Complete minimization program

```

Data: Given a control  $\mathbf{u}$ , state variable  $\mathbf{y}$  and adjoint variable  $\mathbf{p}$ 
Solve  $\mathbf{e}_y \delta \mathbf{y} = -\mathbf{e}_u \delta \mathbf{u}$  ;
Solve  $\mathbf{e}_y^\top \delta \mathbf{p} = -[\mathbf{L}_{yy} \delta \mathbf{y} + \mathbf{L}_{yu} \delta \mathbf{u}]$  ;
 $[\nabla^2 \tilde{J}(\mathbf{u})] \delta \mathbf{u} = \mathbf{L}_{uu} \delta \mathbf{u} + \mathbf{e}_u^\top \delta \mathbf{p} + \mathbf{L}_{yu}^\top \delta \mathbf{y}$ 
    
```

Algorithm 2: Applying the reduced Hessian to a test vector $\delta \mathbf{u}$

direction is an ascent or descent can be determined from the sign of its projection onto the gradient. In the cases where $\delta \mathbf{u}$ is an ascent direction, $-\delta \mathbf{u}$ is a descent direction and thus it is used instead.

We have shown in Proposition 2.3 that the objective is nonconvex and therefore a linesearch is required to globalize the Newton method. Once a search direction \mathbf{d} is computed by the Krylov-Newton solver, the aim is to compute a steplength a such that it satisfies the strong Wolfe conditions (SWC) given by

$$\tilde{J}(\mathbf{u} + a \mathbf{d}) \leq \tilde{J}(\mathbf{u}) + c_1 a \mathbf{d}^\top \nabla \tilde{J}(\mathbf{u}), \quad 0 < c_1 \ll 1, \quad (4.1)$$

$$|\mathbf{d}^\top \nabla \tilde{J}(\mathbf{u} + a \mathbf{d})| \leq c_2 |\mathbf{d}^\top \nabla \tilde{J}(\mathbf{u})|, \quad c_1 < c_2 < 1. \quad (4.2)$$

When using a linesearch with a Newton-type method, it is often recommended that one begins with $a = 1$; see, e.g., [30]. This is an excellent choice when the functional is locally quadratic. However, in our optimization problem, the desired steplength can be orders of magnitude smaller which happens when the higher order derivatives of $J(\mathbf{u} + a \mathbf{d})$ with respect to a are large compared to the first and second-order derivatives.

We use our knowledge of the model to define the initial steplength. In our case, the tracking part of the cost functional is bounded between 0 and 1 for all controls by virtue of the unitarity of the state equation; see Proposition 2.1. On the other hand, the cost functional is bounded from below by the regularization term. This amounts

Data: Given a control \mathbf{u} , make an initial guess of search direction
 $\delta\mathbf{u} = -\mathbf{L}_{\mathbf{u}\mathbf{u}}^{-1}\nabla\tilde{J}(\mathbf{u})$
 Use Algorithm 2 to apply the Hessian, iteratively solve $\nabla^2\tilde{J}(\mathbf{u})\delta\mathbf{u} = -\nabla\tilde{J}(\mathbf{u})$
 with `symmlq`;
if $\delta\mathbf{u}^\top\nabla\tilde{J}(\mathbf{u}) > 0$ **then**
 $\delta\mathbf{u} \leftarrow -\delta\mathbf{u}$;
end

Algorithm 3: Computing the search direction

to a quadratic polynomial and we can write

$$\tilde{J}(\mathbf{u} + a\mathbf{d}) \geq m_2a^2 + m_1a + m_0,$$

where the coefficients are $m_0 = \frac{\gamma}{2}\mathbf{u}^\top\mathbf{K}\mathbf{u} - J(\mathbf{u}) \leq 0$, $m_1 = \gamma\mathbf{u}^\top\mathbf{K}\mathbf{d}$, and $m_2 = \frac{\gamma}{2}\mathbf{d}^\top\mathbf{K}\mathbf{d}$. Since $m_0 \leq 0$, the equation $m_2a^2 + m_1a + m_0 = 0$ admits real roots and we can establish an upper bound on the maximum feasible steplength as follows

$$a_{\max} = \frac{\sqrt{m_1^2 - 4m_0m_2} - m_1}{2m_2}. \quad (4.3)$$

In addition, this upper bound provides a valuable metric in the sense that if $a_{\max} \ll 1$ occurs, it means that a locally quadratic model is not valid. We have the following new result:

PROPOSITION 4.1. *For sufficiently small $c_1 > 0$, there exists at least one steplength a^* which satisfies the SWC condition in the interval $(0, a_{\max}]$*

Proof. Because $\tilde{J}(\mathbf{u} + a_{\max}\mathbf{d}) \geq \tilde{J}(\mathbf{u})$ and $\mathbf{d}^\top\nabla\tilde{J}(\mathbf{u}) < 0$, by the intermediate value theorem, there must be at least one value of $a^* \in (0, a_{\max}]$ such that $\tilde{J}(\mathbf{u} + a^*\mathbf{d}) < 0$ and $\mathbf{d}^\top\nabla\tilde{J}(\mathbf{u} + a^*\mathbf{d}) = 0$. \square

In order to have a robust and efficient linesearch scheme we combine two methods. First, we apply the Algorithm 4 that, for a given a descent direction \mathbf{d} , computes the maximum feasible steplength. If $a = 1$ satisfies the SWC condition, it is accepted. Otherwise, it is assumed that the functional is locally approximated by a cubic polynomial and for this purpose we construct a cubic Hermite interpolant on $[0, 1]$. The minimum a_m of this polynomial function is tested for the SWC condition. If this condition is not satisfied, we apply a more robust scheme represented by the bisection method of Algorithm 5 given below. At the end a minimizer will be bracketed by $[0, a_{\max}]$.

We have implemented a modified bisection method to compute a minimum of a C^2 function. For this purpose, we write a set of criteria which ensure that a twice continuously differentiable function $f(x)$ must have at least one local minimizer $x^* \in (x_l, x_r)$. We have the following

PROPOSITION 4.2. *Suppose that f is continuously differentiable. If $f'(x_l) < 0$ and $f'(x_r) > 0$, then there must be at least one point in $\tilde{x} \in (x_l, x_r)$ such that \tilde{x} is a local minimum.*

PROPOSITION 4.3. *Suppose that f is continuously differentiable. If $f'(x_l) < 0$ and $f(x_r) > f(x_l)$ ($f'(x_l) > 0$, $f(x_r) < f(x_l)$), then there must be at least one point in $\tilde{x} \in (x_l, x_r)$ such that \tilde{x} is a local minimum.*


```

Data: Given a descent direction  $p$  and the function  $f(a) = \tilde{J}(\mathbf{u} + a \mathbf{d})$  and
 $f'(a) = \mathbf{d}^\top \nabla \tilde{J}(\mathbf{u} + a \mathbf{d})$ 
Compute  $a_{max}$  based on equation 4.3
if  $a_{max} > 2$  then
    Evaluate  $f(1)$  and  $f'(1)$  ;
    if  $a = 1$  satisfies SWC then
        |  $a^* \leftarrow 1$ ;
    else
        Construct cubic model on  $[0, 1]$  and compute its minimum  $a_m$ ;
        Evaluate  $f(a_m)$  and  $f'(a_m)$ 
        if  $a = a_m$  satisfies SWC then
            |  $a^* \leftarrow a_m$ ;
        else
            if  $[0, a_m]$  brackets a minimum then
                |  $a_r \leftarrow a_m$ ;

                else if  $[0, 1]$  brackets a minimum then
                    |  $a_r \leftarrow 1$ ;

                else
                    |  $a_r \leftarrow a_{max}$ ;
                end
                 $a^* \leftarrow \text{bisect}(0, a_r)$ .
            end
        end
    end
else
    |  $a^* \leftarrow \text{bisect}(0, a_{max})$ 
end
    
```

Algorithm 4: Linesearch algorithm

These two propositions give the guideline for the formulation of the following minimization algorithm. If for any reason, Algorithm 5 returns a critical point a^* which does not satisfy the SWC condition, then we can use the fact that for sufficiently small value of c_1 there exists a point that does satisfy the SWC and lies between 0 and any critical point which does not. We can then use a^* as an upper bound and apply the bisection scheme on the interval $(0, a^*)$.

5. Numerical Results. In this section, we are concerned with the evaluation of the optimal control formulation for dipole control of a quantum system and with the investigation of the computational performance of the proposed Newton scheme. We consider the cost for state transitions for different choices of the regularization parameters and we compare the convergence behavior of the Krylov-Newton scheme with that of gradient-type schemes. We work in the discrete H^1 space; see [41]. We focus on fast state transitions and we compare results with T ranging from very small to moderate values. Unless otherwise stated, we take a spatial domain $\Omega = [0, 2]$ with $N_x = 100$ grid points and consider the time interval $(0, T)$ subdivided in $N_t = 100$ time steps.

In Figure 5.1, we report the values of the objective at the optimum of state transition $1 \rightarrow 2$ for different values of the regularization parameters. We notice that

```

Data:  $a_l$  and  $a_r$  which bracket a minimum point.  $L = a_r - a_l$ .
 $f(a) = \tilde{J}(\mathbf{u} + a \mathbf{d})$  and  $f'(a) = \mathbf{d}^\top \nabla \tilde{J}(\mathbf{u} + a \mathbf{d})$ 
while  $L > tol$  do
  Compute the midpoint  $a_m = \frac{1}{2}(a_l + a_r)$  and evaluate  $f(a_m)$  and  $f'(a_m)$ 
  if  $a_m$  satisfies SWC then
    |  $a^* \leftarrow a_m$  ;
  end
  if  $f'(a_l) < 0$  and either  $f'(a_r) > 0$  or  $f(a_r) > f(a_l)$  then
    |  $a_r \leftarrow a_m$  ;
  else if  $f'(a_l) > 0$  and  $f'(a_r) < 0$  or  $f(a_r) < f(a_l)$  then
    |  $a_r \leftarrow a_m$  ;
  else
    |  $a_l \leftarrow a_m$  ;
  end
   $L \leftarrow (a_r - a_l)$ 
end

```

Algorithm 5: Bisection minimizer

for a given γ , the objective value increases while increasing the value of α , that is, penalizing highly varying controls. We see that our scheme can explore a large range of values of γ such that good tracking of the target function can be guaranteed. In Figure 5.1, we also plot the optimal controls corresponding to a given α for a range of γ values. We notice that, as the weight of the cost reduces, the control acquires more structure and we obtain an improved tracking.

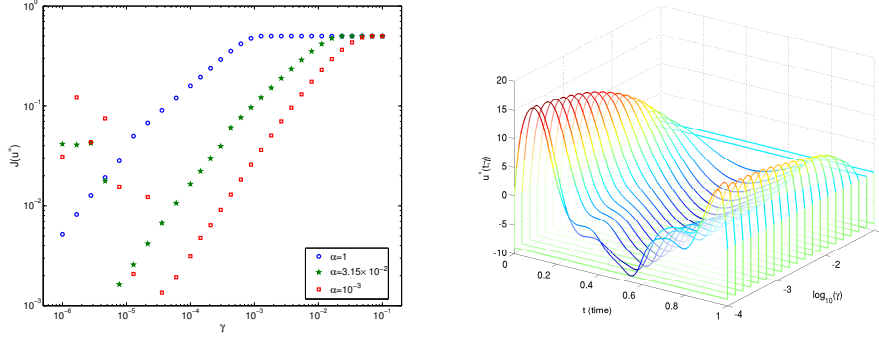


FIG. 5.1. Left: The cost for the state transition $1 \rightarrow 2$ given an optimal control u^* for different choices of the regularization parameters, γ and α . Right: The optimal control for the state transition $1 \rightarrow 2$ as a function of time with $\alpha = 10^{-3}$ for a range of fixed values of γ .

The results shown in Figure 5.1, suggest that the resulting optimal control are not simple harmonic functions. This fact appears more clearly in the plot of Figure 5.2 where optimal controls for different state transitions are depicted.

In correspondence to the control of the transition $1 \rightarrow 2$, we provide in Figure 5.3 a picture of the corresponding time evolution of $|\psi(x, t)|^2$. We can clearly see the transition occurring for the state.

We remark that large differences are obtained between optimal controls corresponding to different time horizons. In Figure 5.4, we see that as the terminal time T increases from the very small value $T = 0.75$ to the moderate value $T = 4$, the ampli-

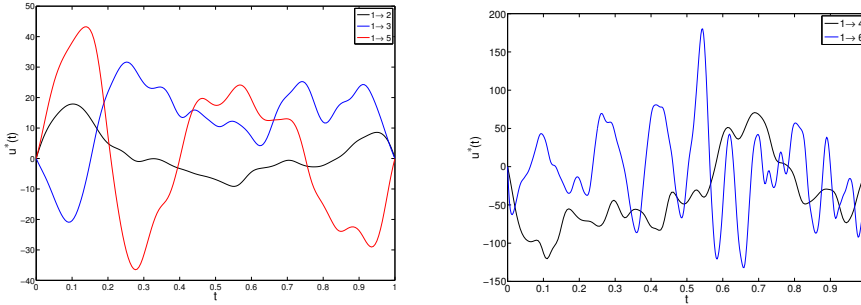


FIG. 5.2. Left: Optimal controls for transitions from the first state to the second, the third, and the fifth states. Right: Optimal controls for transitions from the first state to the fourth and the sixth states.

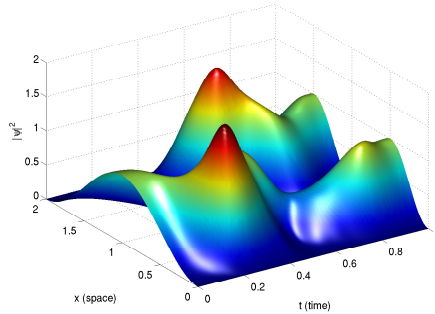


FIG. 5.3. Optimal control transition of $|\psi(x, t)|^2$ from first to second eigenstate.

tude of the control decreases of two orders of magnitude and becomes less oscillating. This result points out the difficulty of solving for short terminal times.

Next, we investigate the computational performance of our optimal control solver. We start with Table 5.1, where we compare the convergence behavior of the Krylov-Newton (KN) scheme, of the steepest descent (SD) scheme, and of the nonlinear conjugate gradient (NCG) scheme. We use the NCG method proposed by Hager and Zhang [19] which in our case appears to be the most competitive among NCG schemes.

For these experiments, we choose the initial state $\psi_0(x) = \sin(\pi x)$ and the target state is $\tilde{\psi}(x) = \frac{1}{\sqrt{2}}[\sin(\pi x) + \sin(2\pi x)]$. We take $T = 1$. The regularization parameters are $\gamma = 10^{-1}$ and $\alpha = 10^{-3}$. Results with this setting are reported in Table 5.1 that provides the convergence history of the iterative schemes along nine iterations. We can see that the Krylov-Newton scheme does not descend as rapidly in the first steps followed by accelerate convergence in the subsequent steps thus overperforming the SD scheme and the NCG scheme.

A useful metric of computational performance of the schemes discussed here is a comparison of CPU times to reduce the value of the norm $\|\nabla \tilde{J}(u)\|$, representing the error in solving the optimality condition. We consider the initial state as $\phi_1(x)$ and the target as $\psi = \frac{1}{\sqrt{2}}(\phi_1 + \phi_2)$, $T = 1$, and fixed regularization parameter $\alpha = 10^{-2}$. We compare for two different values $\gamma = 10^{-1}$ and $\gamma = 10^{-4}$ to show the influence of regularization on the convergence of the control problem. In Figure 5.5, we plot

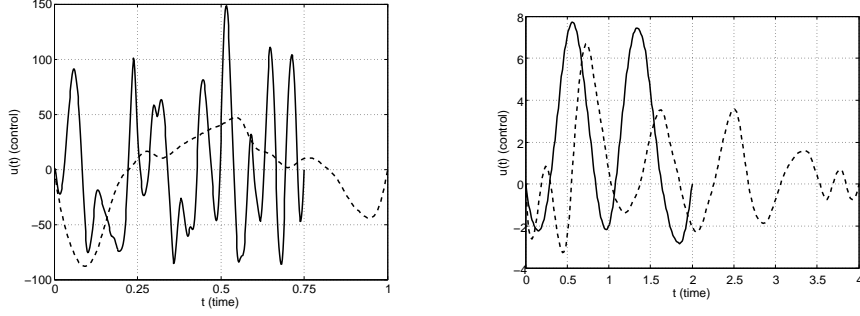


FIG. 5.4. Left: The optimal control u for the state 1-to-2 transition with $T = 0.75$ in solid line. The dashed line is the optimal control, multiplied by factor of 5, for $T = 1$ so that it can be shown on the same scale. Right: The optimal controls for $T = 2$ and $T = 4$ with solid and dashed lines respectively.

TABLE 5.1

Convergence of the SD scheme, the NCG scheme, and the KN scheme to reach the optimal cost $\tilde{J}^* = \tilde{J}(u^*)$.

Iteration	$\tilde{J}_{SD} - \tilde{J}^*$	$\tilde{J}_{NCG} - \tilde{J}^*$	$\tilde{J}_{KN} - \tilde{J}^*$
1	2.4969×10^{-1}	2.4969×10^{-1}	2.4969×10^{-1}
2	1.3070×10^{-2}	1.3070×10^{-2}	1.5346×10^{-2}
3	6.4184×10^{-3}	6.4184×10^{-3}	5.1099×10^{-3}
4	5.5337×10^{-3}	5.3438×10^{-3}	2.2381×10^{-4}
5	4.8170×10^{-3}	3.1011×10^{-3}	1.8383×10^{-4}
6	4.2081×10^{-3}	2.3384×10^{-3}	1.6253×10^{-5}
7	3.6768×10^{-3}	1.2475×10^{-3}	2.7534×10^{-6}
8	3.2177×10^{-3}	9.1869×10^{-5}	3.3921×10^{-7}
9	2.8141×10^{-3}	5.9258×10^{-5}	4.7022×10^{-9}

the values of $\|\nabla \tilde{J}(u_k)\|$ for a sequence of iterations $k = 1, 2, \dots$, as a function of CPU time to illustrate the relative performance of the KN method compared to the NCG scheme. In both cases, the Matlab Krylov solver, `symmlq` is used with a tolerance of 10^{-5} and the \mathbf{L}_{uu} block is used as a preconditioner. The plots in Figure 5.5 show that the NCG scheme provides comparable computational performance to the KN scheme when choosing larger values of γ . However, as γ is taken smaller, the KN method converges significantly faster to the optimal control solution.

6. Conclusion. We discussed an optimal control approach to the construction of an accurate dipole control mechanism which is able to induce transition between

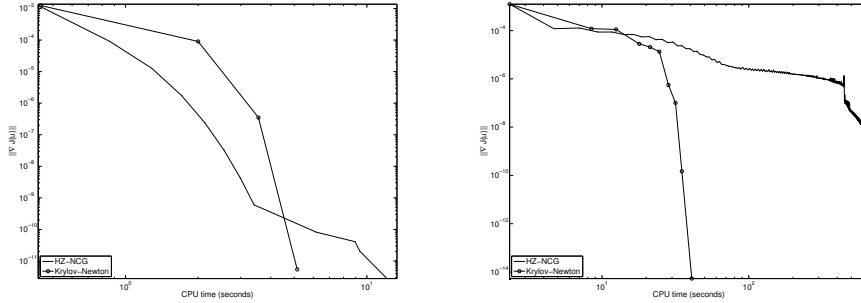


FIG. 5.5. Convergence results for $\gamma = 10^{-1}$ (left) and $\gamma = 10^{-4}$ (right) with the KN scheme and the NCG scheme; $\alpha = 10^{-2}$.

quantum states within a short time horizon where the classical perturbation approach is no longer valid. For this purpose we introduced appropriate discretization schemes and a globalized Newton method with a properly designed linesearch for computing the optimal dipole control. Results of numerical experiments were reported to show the ability of the resulting scheme to provide fast optimal controls that induce transition with high energy differences.

Appendix A. The dipole quantum control problem.

We present all details of our functional analytic setting and present our results on the existence of an optimal control solution including necessary and sufficient optimality conditions. For ease of presentation, in this appendix we repeat a few definitions concerning functional spaces. To the best of our knowledge, this is the first systematic and complete analysis of first- and second-order necessary and sufficient optimality conditions for the dipole quantum control problem.

Let $\Omega = (0, \ell)$ be an open interval with $\ell > 0$. Recall that for $q \in [1, \infty)$ the Lebesgue space $L^q(\Omega; \mathbb{C})$ is defined as

$$L^q(\Omega; \mathbb{C}) = \left\{ \varphi : \Omega \rightarrow \mathbb{C} \mid \varphi \text{ measurable and } \|\varphi\|_{L^q(\Omega; \mathbb{C})} = \left(\int_{\Omega} |\varphi(x)|_{\mathbb{C}}^q dx \right)^{\frac{1}{q}} < \infty \right\}.$$

In particular, we set $\mathcal{H} = L^2(\Omega; \mathbb{C})$, which is a Hilbert space endowed with the inner product

$$(\varphi, \psi)_{\mathcal{H}} = \int_{\Omega} \varphi(x)^* \psi(x) dx \quad \text{for } \varphi, \psi \in \mathcal{H}$$

and the induced norm $\|\varphi\|_{\mathcal{H}}$ for $\varphi \in \mathcal{H}$. The Hilbert space $\mathcal{V} = H_0^1(\Omega; \mathbb{C})$ is given by

$$\mathcal{V} = \left\{ \varphi \in \mathcal{H} \mid \|\varphi\|_{\mathcal{V}} = \left(\int_{\Omega} |\varphi'(x)|_{\mathbb{C}}^2 dx \right)^{1/2} < \infty \right\}$$

supplied by the inner product

$$(\varphi, \psi)_{\mathcal{V}} = (\varphi', \psi')_{\mathcal{H}} \quad \text{for } \varphi, \psi \in \mathcal{V}$$

and the induced norm $\|\varphi\|_{\mathcal{V}} = \sqrt{(\varphi, \varphi)_{\mathcal{V}}}$ for $\varphi \in \mathcal{V}$. For more details on Lebesgue and Sobolev spaces we refer to [1, 17], for instance.

For $T > 0$ let $Q = \Omega \times (0, T)$. If Y is a Banach space, we define the Bochner space $L^2(0, T; Y)$ as the space of all measurable functions $\varphi : [0, T] \rightarrow Y$ satisfying

$$\|\varphi\|_{L^2(0, T; Y)} = \left(\int_0^T \|\varphi(t)\|_Y^2 dt \right)^{1/2},$$

where we write $\varphi(t)$ for the function on Y only; see, e.g., [1, 17]. Analogously, the spaces $H_0^1(0, T; Y)$ and $H^1(0, T; Y)$ are defined. In particular, we write $L^2(0, T)$, $H_0^1(0, T)$ or $H^1(0, T)$ if $Y = \mathbb{R}$.

Let us introduce the Hilbert space

$$W(0, T) = \left\{ \varphi \in L^2(0, T; \mathcal{V}) \mid \frac{d\varphi}{dt} \in L^2(0, T; \mathcal{V}') \right\},$$

where $\mathcal{V}' = H^{-1}(\Omega; \mathbb{C})$ denotes the dual space of $\mathcal{V} = H_0^1(\Omega; \mathbb{C})$. The space $W(0, T)$ is endowed with the inner product

$$(\varphi, \psi)_{W(0, T)} = \int_0^T (\varphi_t(t), \psi_t(t))_{\mathcal{V}'} + (\varphi(t), \psi(t))_{\mathcal{V}} dt \quad \text{for } \varphi, \psi \in W(0, T)$$

and the associated induced norm.

A.1. The state equation. Let $u : (0, T) \rightarrow \mathbb{R}$ be a given control input function. We suppose that our potential V has the form

$$V(x, t) = V_0(x) + u(t)x \quad \text{for } (x, t) \in Q \quad (\text{A.1})$$

where $V_0 : \Omega \rightarrow \mathbb{R}$ is fixed. Then, the wave function $\psi : Q \rightarrow \mathbb{C}$ is governed by the time-dependent Schrödinger equation:

$$i \psi_t(x, t) = -\psi_{xx}(x, t) + V(x, t)\psi(x, t) \quad \text{for } (x, t) \in Q, \quad (\text{A.2a})$$

$$\psi(0, t) = \psi(L, t) = 0 \quad \text{for } t \in (0, T), \quad (\text{A.2b})$$

$$\psi(x, 0) = \psi_0(x) \quad \text{for } x \in \Omega, \quad (\text{A.2c})$$

where i is the imaginary unit and $\psi_0 : \Omega \rightarrow \mathbb{C}$ is a given initial wave function distribution. We say that ψ is a weak solution to (A.2) if $\psi \in W(0, T)$ holds with $\psi(0) = \psi_0$ in \mathcal{H} and

$$\int_0^T (i \psi_t, \varphi)_{\mathcal{V}', \mathcal{V}} dt = \int_0^T \int_{\Omega} \psi_x \varphi_x^* + V \psi \varphi^* dx dt \quad \text{for all } \varphi \in L^2(0, T; \mathcal{V}) \quad (\text{A.3})$$

is satisfied, where $(\cdot, \cdot)_{\mathcal{V}', \mathcal{V}}$ denotes the dual pairing between \mathcal{V} and \mathcal{V}' .

PROPOSITION A.1. *Let $\psi_0 \in \mathcal{H}$, $V_0 \in L^2(\Omega)$ and $V(x, t) = V_0(x) + u(t)x$ for $(x, t) \in Q$. Then, there exists a unique weak solution ψ to (A.1) satisfying $\|\psi(t)\|_{\mathcal{H}} = \|\psi_0\|_{\mathcal{H}}$ for all $t \in [0, T]$. If, in addition, $\psi \in L^\infty(\Omega)$ holds, then*

$$\|\psi\|_{W(0, T)} \leq C \left(1 + \|u\|_{L^\infty(0, T)} \right). \quad (\text{A.4})$$

Proof. For the existence of a weak solution to (A.1) we refer the reader to [42]. From (A.3) we infer that

$$(i \psi_t(t), \psi(t))_{\mathcal{V}', \mathcal{V}} = \int_{\Omega} |\psi_x(t)|_{\mathbb{C}}^2 + V(\cdot, t) |\psi(t)|_{\mathbb{C}}^2 dx \quad \text{for almost all (f.a.a.) } t \in (0, T).$$

Recall that for $\psi \in W(0, T)$ we have

$$(i\psi_t(t), \psi(t))_{\mathcal{V}', \mathcal{V}} = \frac{i}{2} \frac{d}{dt} \|\psi(t)\|_{\mathcal{H}}^2 \quad \text{f.a.a. } t \in (0, T).$$

see, e.g., [17]. By assumption, $V(x, t) \in \mathbb{R}$ f.a.a. $(x, t) \in Q$. Thus,

$$\frac{d}{dt} \|\psi(t)\|_{\mathcal{H}}^2 = -2i \int_{\Omega} |\psi_x(t)|_{\mathbb{C}}^2 + V(\cdot, t) |\psi(t)|_{\mathbb{C}}^2 dx \quad \text{f.a.a. } t \in (0, T). \quad (\text{A.5})$$

The left-hand side of A.5 is purely real, whereas the right-hand side of A.5 is purely imaginary. Thus, (A.5) can only be satisfied both sides are zero. Consequently, $\|\psi(t)\|_{\mathcal{H}}^2$ is constant, i.e.,

$$\|\psi(t)\|_{\mathcal{H}} = \|\psi_0\|_{\mathcal{H}} \quad \text{f.a.a. } t \in [0, T]. \quad (\text{A.6})$$

Moreover,

$$\begin{aligned} \|\psi\|_{L^2(0, T; \mathcal{V})}^2 &= - \int_0^T \int_{\Omega} (V_0(x) + u(t)x) |\psi(x, t)|_{\mathbb{C}}^2 dx dt \\ &\leq (\|V_0\|_{L^\infty(\Omega)} + \|u\|_{L^\infty(0, T)}) \|\psi\|_{L^2(0, T; \mathcal{H})}^2 \\ &\leq (\|V_0\|_{L^\infty(\Omega)} + \|u\|_{L^\infty(0, T)}) \|\psi_0\|_{\mathcal{H}}^2 \end{aligned}$$

and (A.4) imply (A.2a). \square

Next we introduce the control space $\mathcal{U} = H_0^1(0, T)$ supplied with the inner product

$$(u, v)_{\mathcal{U}} = \int_0^T u(t)v(t) + \alpha \dot{u}(t)\dot{v}(t) dt \quad \text{for } u, v \in \mathcal{U}$$

with $0 < \alpha \ll 1$ and the induced norm $\|u\|_{\mathcal{U}} = \sqrt{(u, u)_{\mathcal{U}}}$ for $u \in \mathcal{U}$. Use of small values of α allows for controls with larger rate of changes. Since the aim of the optimal control problem will be to drive a particle from one eigenstate to another, it is required that the control is zero at the beginning and the end of the time interval.

Let us define the Hilbert spaces

$$\mathcal{Z} = W(0, T) \times \mathcal{U} \quad \text{and} \quad \mathcal{Y} = L^2(0, T; \mathcal{V}) \times \mathcal{H}$$

endowed with their natural product topology. We identify the dual \mathcal{Y}' of \mathcal{Y} with $L^2(0, T; \mathcal{V}') \times \mathcal{H}$.

To formulate the Schrödinger equation as an abstract operator equation we define the nonlinear operator $e = (e_1, e_2) : \mathcal{Z} \rightarrow \mathcal{Y}'$ by

$$(e(z), \varphi)_{\mathcal{Y}', \mathcal{Y}} = \int_0^T \left((i\psi_t, p)_{\mathcal{V}', \mathcal{V}} - \int_{\Omega} \psi_x p_x^* - V \psi p^* dx \right) dt + (\psi(0) - \psi_0, p_0)_{\mathcal{H}}$$

for $z = (\psi, u) \in \mathcal{Z}$ and $\varphi = (p, p_0) \in \mathcal{Y}$. Recall that the potential V given by (A.1) depends on the control u . For given $u \in \mathcal{U}$ the function ψ is a weak solution to (A.2) if and only if $e(\psi, u) = 0$ in \mathcal{Y}' .

LEMMA A.2. *The operator $e : \mathcal{Z} \rightarrow \mathcal{Y}'$ is twice Fréchet-differentiable and its Fréchet derivatives at $\bar{z} = (\bar{\psi}, \bar{u}) \in \mathcal{Z}$ are given by*

$$\begin{aligned} (e'(\bar{z})z, \varphi)_{\mathcal{Y}', \mathcal{Y}} &= \int_0^T \left((i\psi_t, p)_{\mathcal{V}', \mathcal{V}} - \int_{\Omega} \psi_x p_x^* - (u(t)x\bar{\psi} + (V_0 + \bar{u}(t)x)\psi)p^* dx \right) dt \\ &\quad + (\psi(0), p_0)_{\mathcal{H}}, \\ (e''(\bar{z})(z, \tilde{z}), \varphi)_{\mathcal{Y}', \mathcal{Y}} &= \int_0^T \int_{\Omega} (u(t)x\tilde{\psi} + \tilde{u}(t)x\psi)p^* dx dt \end{aligned}$$

for any directions $z = (\psi, u)$, $\tilde{z} = (\tilde{\psi}, \tilde{u}) \in \mathcal{Z}$ and for $\varphi = (p, p_0) \in \mathcal{Y}$. In particular, the second Fréchet derivative is Lipschitz-continuous on \mathcal{Z} .

Proof. Let $\bar{z} = (\bar{\psi}, \bar{u}) \in \mathcal{Z}$ be arbitrary. Then, by the Sobolev embedding theorem [1, 17] we have $\bar{u} \in C([0, T])$ so that the claim follows by using standard arguments. \square

To ensure existence of Lagrange multiplier we will make use of the next result.

PROPOSITION A.3. *The linear operator $e_{\psi}(\bar{z}) : \mathcal{Z} \rightarrow \mathcal{Y}'$ is bijective for every $\bar{x} \in \mathcal{Z}$, where e_{ψ} denotes the partial Fréchet derivative of e with respect to ψ .*

Proof. Let $\bar{z} = (\bar{\psi}, \bar{u}) \in \mathcal{Z}$ be chosen arbitrary. The operator $e_{\psi}(\bar{z})$ is bijective, if the equation $e_{\psi}(\bar{z})\psi = (F, \phi_0)$ possesses a unique solution $\psi \in \mathcal{Z}$ for any $F \in L^2(0, T; \mathcal{V}')$ and $\phi_0 \in \mathcal{H}$. Thus, ψ solves

$$(i\psi_t(t), \varphi)_{\mathcal{V}', \mathcal{V}} + a(t; \psi, \varphi) = (F(t), \varphi)_{\mathcal{V}', \mathcal{V}} \quad \text{for all } \varphi \in \mathcal{V}, \quad (\text{A.7a})$$

$$\psi(\cdot, 0) = \phi_0 \quad \text{in } \Omega, \quad (\text{A.7b})$$

where the time-dependent bilinear form $a(t; \cdot, \cdot) : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{C}$ is defined as

$$a(t; \phi, \varphi) = \int_{\Omega} \phi'(x)\varphi'(x)^* + (V_0(x) + \bar{u}(t)x)\phi(x)\varphi(x)^* dx \quad \text{for } \phi, \varphi \in \mathcal{V} \text{ and } t \in [0, T].$$

Since $\bar{u} \in H^1(0, T)$ holds and $H^1(0, T)$ is continuously embedded in $C([0, T])$, there exists a constant $c_1 > 0$ such that

$$|\bar{u}(t)| \leq c_1 \|u\|_{\mathcal{U}} \quad \text{for all } t \in [0, T]. \quad (\text{A.8})$$

Moreover, \mathcal{V} is embedded into $L^4(\Omega; \mathbb{C})$. Thus, there exists a constant $c_2 > 0$ satisfying $\|\varphi\|_{L^4(\Omega; \mathbb{C})} \leq c_2 \|\varphi\|_{\mathcal{V}}$ for all $\varphi \in \mathcal{V}$. Hence,

$$\begin{aligned} |a(t; \phi, \varphi)| &\leq \|\phi\|_{\mathcal{V}} \|\varphi\|_{\mathcal{V}} + (\|V_0\|_{L^2(\Omega)} + c_1 \|u\|_{\mathcal{U}}) \|\phi\|_{L^4(\Omega; \mathbb{C})} \|\varphi\|_{L^4(\Omega; \mathbb{C})} \\ &\leq \left(1 + c_2 (\|V_0\|_{L^2(\Omega)} + c_1 \|u\|_{\mathcal{U}}) \right) \|\phi\|_{\mathcal{V}} \|\varphi\|_{\mathcal{V}} \quad \text{for } \phi, \varphi \in \mathcal{V}. \end{aligned}$$

so that the bilinear form a is bounded. Using Agmon's inequality [38]

$$\|\varphi\|_{L^\infty(\Omega; \mathbb{C})} \leq c_3 \|\varphi\|_{\mathcal{H}}^{1/2} \|\varphi\|_{\mathcal{V}}^{1/2} \quad \text{for } \varphi \in \mathcal{V}$$

and Young's inequality [2]

$$ab \leq \frac{1}{\varepsilon^p} \frac{a^p}{p} + \varepsilon^q \frac{b^q}{q}$$

with $a = c_3(\|V_0\|_{L^2(\Omega)} + c_1 \|u\|_{\mathcal{U}})\|\varphi\|_{\mathcal{H}}$, $b = \|\varphi\|_{\mathcal{V}}$, $\varepsilon = 1$, $p = 4/3$, $q = 4$ we find

$$\begin{aligned} a(t; \varphi, \varphi) &\geq \|\varphi\|_{\mathcal{V}}^2 - (\|V_0\|_{L^2(\Omega)} + c_1 \|u\|_{\mathcal{U}})\|\varphi\|_{\mathcal{H}}\|\varphi\|_{L^\infty(\Omega; \mathbb{C})} \\ &\geq \|\varphi\|_{\mathcal{V}}^2 - c_3(\|V_0\|_{L^2(\Omega)} + c_1 \|u\|_{\mathcal{U}})\|\varphi\|_{\mathcal{H}}^{3/2}\|\varphi\|_{\mathcal{V}}^{1/2} \geq c_4 \|\varphi\|_{\mathcal{V}}^2 - c_5 \|\varphi\|_{\mathcal{H}}^2, \end{aligned}$$

where $c_4 = 3/4$ and $c_5 = 3(c_3(\|V_0\|_{L^2(\Omega)} + c_1 \|u\|_{\mathcal{U}}))^{4/3}/4$. Now the claim follows from a complex variant of classical results for linear evolution problems; see, e.g., [15]. \square

REMARK A.4. It follows from Proposition A.3 that the operator $e'(\bar{z})$ is surjective for every $\bar{z} = (\bar{\psi}, \bar{u}) \in \mathcal{Z}$. This implies that every point \bar{z} is a regular point. \diamond

Let $\bar{z} = (\bar{\psi}, \bar{u}) \in \mathcal{Z}$ be arbitrary. By $\ker e'(\bar{z}) \subset \mathcal{Z}$ we denote the null space of the operator $e'(\bar{z})$. Suppose that $z = (\psi, u) \in \ker e'(\bar{z})$ is given. Then, ψ satisfies

$$\begin{aligned} i \psi_t(x, t) &= -\psi_{xx}(x, t) + (V_0(x) + \bar{u}(t)x)\psi(x, t) \\ &\quad + u(t)x\bar{\psi}(x, t) \end{aligned} \quad \text{f.a.a. } (x, t) \in Q, \quad (\text{A.9a})$$

$$\psi(0, t) = \psi(L, t) = 0 \quad \text{f.a.a. } t \in (0, T), \quad (\text{A.9b})$$

$$\psi(x, 0) = 0 \quad \text{f.a.a. } x \in \Omega. \quad (\text{A.9c})$$

Multiplying (A.9a) by $\psi(x, t)^*$ and integrating over Ω imply that

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\psi(t)\|_{\mathcal{H}}^2 + i \|\psi(t)\|_{\mathcal{V}}^2 \\ = i \int_{\Omega} (V_0(x) + \bar{u}(t)x) |\psi(x, t)|^2 + u(t)x\bar{\psi}(x, t)\psi(x, t)^* dx. \end{aligned} \quad (\text{A.10})$$

The real part of (A.10) reads

$$\frac{1}{2} \frac{d}{dt} \|\psi(t)\|_{\mathcal{H}}^2 = -u(t) \int_{\Omega} x \Im(\bar{\psi}(x, t)\psi(x, t)^*) dx. \quad (\text{A.11})$$

where we have used that $V_0(x)$ as well as $u(t)$ are real-valued. We infer from (A.9) and Young's inequality [2] that

$$\frac{1}{2} \frac{d}{dt} \|\psi(t)\|_{\mathcal{H}}^2 \leq |u(t)| \|\bar{\psi}(t)\|_{\mathcal{H}} \|\psi(t)\|_{\mathcal{H}} \leq \frac{1}{2} \left(|u(t)|^2 \|\bar{\psi}(t)\|_{\mathcal{H}}^2 + \|\psi(t)\|_{\mathcal{H}}^2 \right).$$

Using Gronwall's inequality we find that

$$\|\psi(t)\|_{\mathcal{H}}^2 \leq e^{t/2} \left(\|\psi(0)\|_{\mathcal{H}}^2 + \frac{1}{2} \int_0^t |u(s)|^2 \|\bar{\psi}(s)\|_{\mathcal{H}}^2 ds \right) \quad (\text{A.12})$$

Combining (A.8), (A.9c), and (A.12) we obtain

$$\|\psi(t)\|_{\mathcal{H}}^2 \leq c_2 \|u\|_{\mathcal{U}}^2 \quad \text{f.a.a. } t \in [0, T]. \quad (\text{A.13})$$

with the constant $c_2 = c_1 e^{T/2} \|\bar{\psi}\|_{L^2(0, T; \mathcal{H})}^2 / 2$.

Now we consider the imaginary part of (A.10). Using (A.13) Hölder's and Young's inequality we find

$$\begin{aligned} \|\psi(t)\|_{\mathcal{V}}^2 &= \int_{\Omega} (V_0(x) + \bar{u}(t)x) |\psi(x, t)|^2 + u(t)x \Re(\bar{\psi}(x, t)\psi(x, t)^*) dx \\ &\leq (\|V_0\|_{L^\infty(\Omega)} + |\bar{u}(t)|) \|\psi(t)\|_{\mathcal{H}}^2 + \frac{1}{2} \left(|u(t)|^2 + \|\bar{\psi}(t)\|_{\mathcal{H}}^2 \|\psi(t)\|_{\mathcal{H}}^2 \right) \\ &\leq c_2 (\|V_0\|_{L^\infty(\Omega)} + |\bar{u}(t)|) \|u\|_{\mathcal{U}}^2 + \frac{1}{2} \left(|u(t)|^2 + c_2 \|\bar{\psi}(t)\|_{\mathcal{H}}^2 \|u\|_{\mathcal{U}}^2 \right). \end{aligned} \quad (\text{A.14})$$

Estimate (A.14) implies

$$\|\psi\|_{L^2(0,T;\mathcal{V})}^2 \leq \left(c_2(T\|V_0\|_{L^\infty(\Omega)} + \|\bar{u}\|_{L^1(0,T)}) + \frac{1}{2} + \frac{c_2}{2} \|\bar{\psi}\|_{L^2(0,T;\mathcal{H})}^2 \right) \|u\|_{\mathcal{U}}^2. \quad (\text{A.15})$$

Setting

$$c_3 = \sqrt{\max \left(c_2 T, c_2(T\|V_0\|_{L^\infty(\Omega)} + \|\bar{u}\|_{L^1(0,T)}) + \frac{1}{2} + \frac{c_2}{2} \|\bar{\psi}\|_{L^2(0,T;\mathcal{H})}^2 \right)}$$

we conclude from (A.13) and (A.15) that

$$\|\psi\|_{L^2(0,T;\mathcal{H})} + \|\psi\|_{L^2(0,T;\mathcal{V})} \leq c_3 \|u\|_{\mathcal{U}}. \quad (\text{A.16})$$

PROPOSITION A.5. *Suppose that $\bar{z} \in Z$ and that $V_0 \in L^\infty(\Omega)$. Then,*

$$\|\psi\|_{W(0,T)} \leq C_{ker} \|u\|_{\mathcal{U}} \quad \text{for all } z = (\psi, u) \in \ker e'(\bar{z})$$

for a constant $C_{ker} > 0$.

Proof. We have already derived an estimate in the $L^\infty(0, T; \mathcal{H})$ and the $L^2(0, T; \mathcal{V})$ -norms; see (A.16). Recall that

$$\|\psi_t\|_{L^2(0,T;\mathcal{V}')} = \sup_{\|\varphi\|_{L^2(0,T;\mathcal{V})}=1} \int_0^T (\psi_t(t), \varphi(t))_{\mathcal{V}', \mathcal{V}} dt.$$

Now, the proof follows directly from (A.9a) and (A.16). \square

A.2. The minimization problem. Given an initial wave function ψ_0 , we wish to find a control $u \in \mathcal{U}$ such that $\psi(T) \in \mathcal{H}$ is in some sense close to a given desired target wave function $\tilde{\psi} \in \mathcal{H}$. This is to say that the aim of the optimal control problem is to minimize the cost functional $J : \mathcal{Z} \rightarrow \mathbb{R}$ given by

$$J(z) = \frac{1}{2} (1 - \|P\psi(T)\|_{\mathcal{H}}^2) + \frac{\gamma}{2} \|u\|_{\mathcal{U}}^2, \quad z = (\psi, u) \in \mathcal{Z}, \quad (\text{A.17})$$

where the linear and bounded projection operator $P : \mathcal{H} \rightarrow \mathcal{H}$ is defined as

$$P\phi = (\tilde{\psi}, \phi)_{\mathcal{H}} \tilde{\psi} \quad \text{for } \phi \in \mathcal{H}$$

and $\gamma > 0$ is a regularization parameter.

LEMMA A.6. *The cost functional J is twice Fréchet-differentiable. Its Fréchet derivatives at $\bar{z} = (\bar{\psi}, \bar{u}) \in \mathcal{Z}$ are*

$$\begin{aligned} J'(\bar{z})z &= -(P\bar{\psi}(T), P\psi(T))_{\mathcal{H}} + \gamma(\bar{u}, u)_{\mathcal{U}}, \\ J''(\bar{z})(z, \tilde{z}) &= -(P\tilde{\psi}(T), P\psi(T))_{\mathcal{H}} + \gamma(\tilde{u}, u)_{\mathcal{U}} \end{aligned}$$

for any directions $z = (\psi, u)$, $\tilde{z} = (\tilde{\psi}, \tilde{u}) \in \mathcal{Z}$ and for $\varphi = (p, p_0) \in \mathcal{Y}$. In particular, the second Fréchet derivative is Lipschitz-continuous on \mathcal{Z} .

Proof. The claim follows by standard arguments so that the proof is omitted here.

\square

The set of admissible solutions of the optimal control problem that will be introduced now is given by

$$\mathcal{F}(\mathbf{P}) = \{z \in \mathcal{Z} \mid e(z) = 0 \text{ in } \mathcal{Y}'\}.$$

Then, the optimal control problems reads

$$\min J(z) \quad \text{subject to (s.t.) } z \in \mathcal{F}(\mathbf{P}). \quad (\mathbf{P})$$

THEOREM A.7. *There exists at least one optimal solution $z^\circ = (\psi^\circ, u^\circ)$ to (\mathbf{P}) .*

Proof. By Proposition A.1 the set $\mathcal{F}(\mathbf{P})$ is nonempty. Let $\{z^n\}_{n \in \mathbb{N}}$, $z^n = (\psi^n, u^n)$, be a minimizing sequence in $\mathcal{F}(\mathbf{P})$. Then, $\|u^n\|_{\mathcal{U}}$ is bounded. Recall that \mathcal{U} is (compactly) embedded into $L^\infty(\Omega)$; see [1, p. 144]. By (A.4) the sequence $\{\psi^n\}_{n \in \mathbb{N}}$ is bounded in $W(0, T)$. Thus, there exist a subsequence $\{z^{n_k}\}_{k \in \mathbb{N}}$, $z^{n_k} = (\psi^{n_k}, u^{n_k})$, and an element $z^\circ = (\psi^\circ, u^\circ) \in \mathcal{Z}$ such that

$$\psi^{n_k} \rightharpoonup \psi^\circ \quad \text{as } k \rightarrow \infty \quad \text{in } W(0, T), \quad (\text{A.18a})$$

$$u^{n_k} \rightharpoonup u^\circ \quad \text{as } k \rightarrow \infty \quad \text{in } \mathcal{U}. \quad (\text{A.18b})$$

Since \mathcal{U} is compactly embedded into $C([0, T])$ we conclude from (A.18b) that

$$u^{n_k} \rightarrow u^\circ \text{ as } k \rightarrow \infty \text{ in } C([0, T]). \quad (\text{A.19})$$

Combining $z^{n_k} \in \mathcal{F}(\mathbf{P})$ for all k , (A.18a) and (A.19) we have

$$0 = \lim_{k \rightarrow \infty} (e(z^{n_k}), \varphi)_{\mathcal{Y}', \mathcal{Y}} = (e(x^\circ), \varphi)_{\mathcal{Y}', \mathcal{Y}} \quad \text{for all } \varphi \in \mathcal{Y}.$$

Since the cost functional is weakly lower semicontinuous [36, p. 377] the claim follows directly. \square

A.3. The reduced problem. Let $u \in \mathcal{U}$ be given. Then, by Proposition A.1 there exists a unique weak solution to (A.1). Thus, the solution operator $\mathcal{S} : \mathcal{U} \rightarrow W(0, T)$ is well-defined. Boundedness of \mathcal{S} follows from (A.4). We introduce the so-called reduced cost functional

$$\tilde{J}(u) = J(\mathcal{S}(u), u) \quad \text{for } u \in \mathcal{U}$$

and the reduced problem

$$\min \tilde{J}(u) \quad \text{s.t. } u \in \mathcal{U} \quad (\tilde{\mathbf{P}})$$

which is, in contrast to (\mathbf{P}) , an unconstrained optimal control problem. If $u^\circ \in \mathcal{U}$ is a solution to $(\tilde{\mathbf{P}})$, then $z^\circ = (\mathcal{S}(u^\circ), u^\circ)$ solves (\mathbf{P}) .

From Lemmas A.2 and A.6 it follows that \tilde{J} is twice continuously Fréchet-differentiable. In particular, we have at $\bar{u} \in \mathcal{U}$ and

$$(\tilde{J}'(\bar{u}), u)_{\mathcal{U}} = (J_\psi(\mathcal{S}(\bar{u}), \bar{u}), \mathcal{S}'(\bar{u})u)_{W(0, T)} + (J_u(\mathcal{S}(\bar{u}), \bar{u}), u)_{\mathcal{U}} \quad (\text{A.20})$$

for a direction $u \in \mathcal{U}$. We derive from $e(\mathcal{S}(\bar{u}), \bar{u}) = 0$ that

$$0 = e_\psi(\mathcal{S}(\bar{u}), \bar{u})\mathcal{S}'(\bar{u})u + e_u(\mathcal{S}(\bar{u}), \bar{u})u$$

for a direction $u \in \mathcal{U}$. By Proposition A.3 the operator $e_\psi(\mathcal{S}(\bar{u}), \bar{u})$ is invertible. Thus, setting $\bar{z} = (\mathcal{S}(\bar{u}), \bar{u}) \in \mathcal{Z}$ and $\psi = \mathcal{S}'(\bar{u})u \in W(0, T)$ we derive

$$\psi = -e_\psi(\bar{z})^{-1}e_u(\bar{z})u. \quad (\text{A.21})$$

Inserting (A.21) into (A.20) we obtain

$$\begin{aligned} (\tilde{J}'(\bar{u}), u)_{\mathcal{U}} &= -(J_\psi(\bar{z}), e_\psi(\bar{z})^{-1}e_u(\bar{z})u)_{W(0, T)} + (J_u(\bar{z}), u)_{\mathcal{U}} \\ &= (J_u(\bar{z}), u)_{\mathcal{U}} - (e_\psi(\bar{z})^{-*}J_\psi(\bar{z}), e_u(\bar{z})u)_{\mathcal{Y}'}, \\ &= (J_u(\bar{z}) - e_u(\bar{z})^*e_\psi(\bar{z})^{-*}J_\psi(\bar{z}), u)_{\mathcal{U}} \end{aligned} \quad (\text{A.22})$$

for a direction $u \in \mathcal{U}$, where $e_\psi(\bar{z})^{-*} : W(0, T) \rightarrow \mathcal{Y}'$ and $e_u(\bar{z})^* : \mathcal{Y}' \rightarrow \mathcal{U}$ denote the adjoint operators of $e_\psi(\bar{z})^{-1} : \mathcal{Y}' \rightarrow W(0, T)$ and $e_u(\bar{z}) : \mathcal{U} \rightarrow \mathcal{Y}'$, respectively, satisfying

$$\begin{aligned} (e_\psi(\bar{z})^{-*}\varphi, G)_{\mathcal{Y}'} &= (\varphi, e_\psi(\bar{z})^{-1}G)_{W(0, T)} && \text{for all } (\varphi, G) \in W(0, T) \times \mathcal{Y}', \\ (e_u(\bar{z})^*F, v)_{\mathcal{Y}'} &= (F, e_u(\bar{z})v)_{\mathcal{Y}'} && \text{for all } (F, v) \in \mathcal{Y}' \times \mathcal{U}. \end{aligned}$$

It follows from (A.22) that the gradient of the reduced cost functional is given by

$$\tilde{J}'(\bar{u}) = J_u(\bar{z}) - e_u(\bar{z})^*e_\psi(\bar{z})^{-*}J_\psi(\bar{z}) \quad \text{in } \mathcal{U}.$$

A.4. Optimality conditions. Let $u^\circ \in \mathcal{U}$ be a solution to $(\tilde{\mathbf{P}})$. We write $\psi^\circ = \mathcal{S}(u^\circ)$ and $x^\circ = (\psi^\circ, u^\circ)$. To investigate first-order necessary optimality conditions for $(\tilde{\mathbf{P}})$ or (\mathbf{P}) we introduce the Lagrange functional $L : \mathcal{Z} \times \mathcal{Y} \rightarrow \mathbb{R}$ by

$$L(z, \lambda) = J(z) + \Re e(e(z), \lambda)_{\mathcal{Y}', \mathcal{Y}} \quad \text{for } z = (\psi, u) \in \mathcal{Z} \text{ and } \lambda = (p, p_0) \in \mathcal{Y}.$$

It follows from Proposition A.3, Remark A.4 that there exists a unique Lagrange multiplier $\lambda^\circ = (p^\circ, p_0^\circ)$ satisfying together with the optimal solution z° the system

$$L_\psi(z^\circ, \lambda^\circ)z = 0 \quad \text{for all } z \in \mathcal{Z}, \quad (\text{A.23a})$$

$$L_u(z^\circ, \lambda^\circ)u = 0 \quad \text{for all } u \in \mathcal{U}, \quad (\text{A.23b})$$

$$e(z^\circ) = 0 \quad \text{in } \mathcal{Y}'; \quad (\text{A.23c})$$

see, e.g., [25, 29]. We derive from (A.23a) that p° is a weak solution to

$$i p_t^\circ(x, t) = -p_{xx}^\circ(z, t) + (V_0(x) + u^\circ(t)x)p^\circ(x, t) \quad \text{f.a.a. } (x, t) \in Q, \quad (\text{A.24a})$$

$$p^\circ(0, t) = p^\circ(L, t) = 0 \quad \text{f.a.a. } t \in (0, T), \quad (\text{A.24b})$$

$$p^\circ(x, T) = i(P\psi^\circ(T))(x) \quad \text{f.a.a. } x \in \Omega, \quad (\text{A.24c})$$

in particular, p° lies in $W(0, T)$ and $p_0^\circ = p^\circ(T) \in \mathcal{H}$. Using (A.23b) we obtain that $u^\circ \in \mathcal{U}$ is a weak solution to

$$-\alpha\gamma\ddot{u}^\circ(t) + \gamma u^\circ(t) = -\Re e\left(\int_{\Omega} x\psi^\circ(x, t)p^\circ(x, t)^* dx\right) \quad \text{f.a.a. } t \in (0, T), \quad (\text{A.25a})$$

$$u^\circ(0) = u^\circ(T) = 0. \quad (\text{A.25b})$$

Finally, (A.23c) implies that ψ° satisfies the state equation (A.2) with control input $u = u^\circ$.

REMARK A.8. Analogously to the proof of Proposition A.1 we obtain that $\|p^\circ(t)\|_{\mathcal{H}} = \|p^\circ(T)\|_{\mathcal{H}}$ f.a.a. $t \in [0, T]$. Thus, using (A.24c) we have $\|p^\circ\|_{L^2(0,T;\mathcal{H})} = \sqrt{T} \|P\psi^\circ(T)\|_{\mathcal{H}}$. \diamond

Notice that the gradient of the reduced cost functional is given by

$$\tilde{J}'(u^\circ) = L_u(z^\circ, \lambda^\circ) \quad \text{in } \mathcal{U}.$$

To investigate second-order sufficient optimality conditions we derive from Lemmas A.2 and A.6 that the second Fréchet derivative of the Lagrangian at an optimal solution $z^\circ = (\psi^\circ, u^\circ) \in \mathcal{F}(\mathbf{P})$ and at its associated dual variable $\lambda^\circ = (p^\circ, p_0^\circ) \in \mathcal{Y}$ satisfies

$$L_{zz}(z^\circ, \lambda^\circ)(z, z) = \Re \int_0^T \int_{\Omega} 2u(t)x\psi(x, t)p^\circ(x, t)^* dxdt - \|P\psi(T)\|_{\mathcal{H}}^2 + \|u\|_{\mathcal{U}}^2 \quad (\text{A.26})$$

for any direction $z = (\psi, u) \in \mathcal{Z}$.

THEOREM A.9. *Suppose that $z^\circ = (\psi^\circ, u^\circ) \in \mathcal{F}(\mathbf{P})$ is an optimal solution to (\mathbf{P}) and that $\lambda^\circ = (p^\circ, p_0^\circ) \in \mathcal{Y}$ is the associated dual variable. If γ is sufficiently large, if $\|p^\circ\|_{L^2(0,T;\mathcal{H})}$ is sufficiently small and if*

$$\|P\psi(T)\|_{\mathcal{H}}^2 \leq \frac{\gamma}{4} \|u\|_{\mathcal{U}}^2 \quad \text{for all } z \in \ker e'(z^\circ), \quad (\text{A.27})$$

then the second-order sufficient optimality condition holds, i.e., there exists a constant $\kappa > 0$ such that

$$L_{zz}(z^\circ, \lambda^\circ)(z, z) \geq \kappa \|z\|_{\mathcal{Z}}^2 \quad \text{for all } z \in \ker e'(z^\circ). \quad (\text{A.28})$$

Proof. We derive from (A.8), (A.16), Proposition A.5, and (A.26) that

$$\begin{aligned} L_{zz}(z^\circ, \lambda^\circ)(z, z) &\geq -\|u\|_{C([0,T])} \|\psi\|_{L^2(0,T;\mathcal{H})} \|p^\circ\|_{L^2(0,T;\mathcal{H})} - \|P\psi(T)\|_{\mathcal{H}}^2 + \gamma \|u\|_{\mathcal{U}}^2 \\ &\geq -c_1 \|u\|_{\mathcal{U}} \|\psi\|_{L^2(0,T;\mathcal{H})} \|p^\circ\|_{L^2(0,T;\mathcal{H})} - \|P\psi(T)\|_{\mathcal{H}}^2 \\ &\quad + \frac{\gamma}{4} \|u\|_{\mathcal{U}}^2 + \frac{\gamma}{2C_{ker}} \|\psi\|_{W(0,T)}^2 \\ &\geq \left(\frac{\gamma}{2} - \frac{c_1}{2} - \frac{c_1 c_3}{2} \|p^\circ\|_{L^2(0,T;\mathcal{H})} \right) \|u\|_{\mathcal{U}}^2 + \frac{\gamma}{2C_{ker}} \|\psi\|_{W(0,T)}^2. \end{aligned}$$

If

$$\kappa = \min \left(\frac{\gamma}{2C_{ker}}, \frac{\gamma}{2} - \frac{c_1}{2} - \frac{c_1 c_3}{2} \|p^\circ\|_{L^2(0,T;\mathcal{H})} \right) > 0 \quad (\text{A.29})$$

holds, then we obtain (A.28). \square

REMARK A.10. Utilizing Remark A.8 we have the following sufficient condition for Theorem A.9: If

$$\|P\psi^\circ(T)\|_{\mathcal{H}} < \frac{\gamma}{2c_1 c_3 \sqrt{T}} \quad \text{nd} \quad \gamma > 2c_1$$

hold, condition (A.29) is satisfied. \diamond

REFERENCES

- [1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York-London, 1975.
- [2] H. W. Alt, *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung*, Springer-Verlag, Berlin, 1992.
- [3] P.W. Atkins and R.S. Friedman, *Molecular Quantum Mechanics*, Oxford University Press.
- [4] R. Becker and B. Vexler, *Optimal control of the convection-diffusion equation using stabilized finite element methods*, *Numerische Mathematik*, 106 (2007), pp. 349–367.
- [5] C.H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, *Strengths and weakness of quantum computing*, *SIAM J. Computing*, 26 (1997), pp. 1510–1523.
- [6] A. Borzi and U. Hohenester, *Multigrid optimization schemes for solving Bose-Einstein condensate control problems*, *SIAM J. Scientific Computing*, 30 (2008), pp. 441–462.
- [7] A. Borzi, J. Salomon, and S. Volkwein, *Formulation and numerical solution of finite-level quantum optimal control problems*, *J. Comput. Appl. Math.*, 216 (2008), pp. 170–197.
- [8] A. Borzi, G. Stadler, and U. Hohenester, *Optimal quantum control in nanostructures: Theory and application to a generic three-level system*, *Phys. Rev. A* 66, (2002) 053811.
- [9] U. Boscain, G. Charlot, J.-P. Gauthier, S. Guerin, H.-R. Jauslin, *Optimal control in laser-induced population transfer for two- and three-level quantum systems*, *Journal of Mathematical Physics*, 43 (2002), pp. 2107–2132.
- [10] J. Bourgain, *Global Solutions of Nonlinear Schrödinger Equations*, American Mathematical Society Colloquium Publications, 46. American Mathematical Society, Providence, RI (1999).
- [11] P.W. Brumer and M. Shapiro, *Principles of the Quantum Control of Molecular Processes*, Wiley-VCH, Berlin, 2003.
- [12] J. Burkardt, M. Gunzburger, and J. Peterson, *Insensitive functionals, inconsistent gradients, spurious minima, and regularized functionals in flow optimization problems*, *Int. J. Comput. Fluid Dyn.*, 16 (2002), pp. 171–185.
- [13] A.G. Butkovskiy and Yu.I. Samoilenko, *Control of quantum-mechanical processes and systems*, Kluwer, 1990.
- [14] A.Z. Capri, *Nonrelativistic Quantum Mechanics*, World Scientific Publishing Co. Inc., River Edge, NJ, 2002.
- [15] R. Dautray, J.-P. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution Problems*, Springer-Verlag, Berlin, 1992.
- [16] P. Ditz and A. Borzi, *A cascadic monotonic time-discretized algorithm for finite-level quantum control computation*, *Computer Physics Communications*, 178 (2008), 393–399.
- [17] L.C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, Rhode Island, 2002.
- [18] D. Gaveaux, *Semiconductor lasers: Quantum wells meet nanowires*, Vol. 2, *Nature*, October 2008.
- [19] W. W. Hager and H. Zhang, *A new conjugate gradient method with guaranteed descent and an efficient line search*, *SIAM J. Optim.*, 16 (2005), pp. 170–192
- [20] P. Harrison, *Quantum Wells, Wires and Dots: Theoretical and Computational Physics of Semiconductor Nanostructures*, Wiley, 2005.
- [21] M. Hochbruck and C. Lubich, *On Magnus integrators for time-dependent Schrödinger equations*, *SIAM J. Numer. Anal.*, 41 (2003), pp. 945–963.
- [22] A. Iserles, H.Z. Munthe-Kaas, S. Nørsett, and A. Zanna, *Lie-group methods*, *Acta Numerica*, Cambridge Univ. Press, 9 (2000), pp. 215–365.
- [23] R. Kosloff, S.A. Rice, P. Gaspard, S. Tersigni, and D. Tannor, *Wavepacket Dancing: Achieving Chemical Selectivity by Shaping Light Pulses*, *Chem. Phys.*, 139 (1989), pp. 201–220.
- [24] J.L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.
- [25] D.G. Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, Inc., New York, 1969.
- [26] Y. Maday, J. Salomon and G. Turinici, *Monotonic time-discretized schemes in quantum control*, *Num. Math.*, 103 (2006), pp. 323–338.
- [27] Y. Maday, J. Salomon, and G. Turinici, *Parareal in time control for quantum systems*, *SIAM J. Num. Anal.*, 45 (2007), pp. 2468–2482.
- [28] Y. Maday and G. Turinici, *New formulations of monotonically convergent quantum control algorithms*, *J. Chem. Phys.*, 118 (2003), pp. 8191–8196.
- [29] H. Maurer and J. Zowe, *First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, *Mathematical Programming*, 16 (1979), pp. 98–110.
- [30] J. Nocedal and S.J. Wright, *Numerical optimization, Springer Series in Operations Research*

- and *Financial Engineering*, Springer, New York, second ed., 2006.
- [31] T. Numai, *Fundamentals of Semiconductor Lasers*, Springer-Verlag, New York, 2004
 - [32] C.C. Paige and M.A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617-629.
 - [33] A.P. Peirce, M.A. Dahleh, and H. Rabitz, *Optimal control of quantum-mechanical systems: existence, numerical approximation, and applications*, Physical Review A, 37 (1988), pp. 4950-4964.
 - [34] H. Rabitz, M. Hsieh, and C. Rosenthal, *Quantum optimally controlled transition landscapes*, Science, 303 (2004), p. 998.
 - [35] H. Rabitz, G. Turinici, and E. Brown, *Control of Molecular Motion: Concepts, Procedures, and Future Prospects*, Ch. 14 in Handbook of Numerical Analysis, Volume X, P. Ciarlet and J. Lions, Eds., Elsevier, Amsterdam, 2003.
 - [36] M. Reed and B. Simon, *Methods of Modern Mathematical Physics. Volume 1: Functional Analysis*, Academic Press, Inc., Boston, 1980.
 - [37] B.J. Sussman, D. Townsend, M.Y. Ivanov, and A. Stolow, *Dynamic Stark control of photochemical processes*, Science, 13 (2006), pp. 278 - 281.
 - [38] R. Temam, *Infinite-Dimensional Systems in Mechanics and Physics*, Springer-Verlag, New York, 1988.
 - [39] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag, Berlin, 2006.
 - [40] A. van den Bos, *Complex Gradient and Hessian*, IEEE Proc.-Vis. Image Signal Processing, 141(6) (1994), pp. 380-82.
 - [41] G. von Winckel and A. Borzì, Computational techniques for a quantum control problem with H^1 -cost, Inverse Problems, 24 (2008), 034007.
 - [42] K. Yajima, *Existence of Solutions for Schrödinger Evolution Equations*, Communications in Mathematical Physics, 110 (1987), pp. 415-26.