SIAM J. SCI. COMPUT.
Vol. 30, No. 1, pp. 441–462

MULTIGRID OPTIMIZATION SCHEMES FOR SOLVING BOSE–EINSTEIN CONDENSATE CONTROL PROBLEMS*

A. BORZÌ† AND U. HOHENESTER‡

Abstract. The control of the transport of Bose–Einstein condensates in magnetic microtraps is formulated within the framework of optimal control theory and solved by multigrid optimization schemes. The time evolution of the wave function of Bose–Einstein condensates is governed by the Gross–Pitaevskii equation and can be manipulated through variation of a controllable magnetic confinement potential. In order to define an optimal control strategy, an appropriate cost functional is introduced that must be minimized under the constraint given by the dynamic equation. The resulting optimality system consists of two nonlinear Schrödinger-type equations with opposite time orientation coupled with an elliptic equation for the control function. These equations are approximated by using a time-splitting pseudospectral method and finite differences. To solve the resulting problem a cascadic nonlinear conjugate gradient scheme and a multigrid optimization scheme are considered. The convergence properties of these two schemes are investigated theoretically, and their computational performance is discussed based on results of numerical experiments. It appears that the multigrid optimization scheme provides a robust optimization strategy.

Key words. Bose–Einstein condensates, Gross–Pitaevskii equation, optimal control theory, optimality conditions, cascadic acceleration, nonlinear conjugate gradient method, multigrid method

AMS subject classifications. 35Q40, 49K10, 65M06, 81V80

DOI. 10.1137/070686135

1. Introduction.

A Bose–Einstein condensate (BEC) is a state of matter formed by bosons (e.g., helium-4, rubidium) cooled to temperatures very near to absolute zero. Under such conditions, the atoms with magnetic spin collapse into the lowest quantum state sharing the same wave function, and quantum effects become apparent on a macroscopic scale. Coherent manipulation of this wave function is one of the ultimate goals of modern atom optics, because trapping and manipulating cold neutral atoms in microtraps near surfaces of atomic chips is a promising approach towards full control of matter waves on small scales; see [16, 17, 19, 20, 40], and references therein. The possibility to manipulate a single quantum system with extremely high precision has boosted research in various fields ranging from atom interferometry [19, 36], over quantum gates [13] and resonant condensate transport [30], to microscopic magnetic-field imaging [39]. In the vast majority of these schemes the wave function of Bose–Einstein condensates, trapped in the vicinity of an atom chip, is manipulated through variation of the magnetic confinement potential, which is achieved by changing the currents through gate wires mounted on the chip or modifying the strength of additional radio-frequency fields [17, 24].

An important situation is the case where one is aiming at an efficient wave function transfer from a given initial to a final desired state. This problem was first considered by Hänsel et al. [19] for a trapped-atom interferometer setup where a dilute condensate should be split through variation of the confinement potential from

*Received by the editors March 23, 2007; accepted for publication (in revised form) September 25, 2007; published electronically February 1, 2008. This work was supported in part by the Austrian Science Fund FWF project P18136-N13 "Quantum optimal control of semiconductor nanostructures."
http://www.siam.org/journals/sisc/30-1/68613.html

†Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstr. 36, 8010 Graz, Austria (alfio.borzi@uni-graz.at).

‡Institut für Physik, Karl–Franzens–Universität Graz, Universitätsplatz 5, 8010 Graz, Austria (ulrich.hohenester@uni-graz.at).

a single- to a double-well ground state by using a scheme that optimizes adiabatic transfer by minimizing transitions to excited states.

Recently, we proposed [23] quantum optimal control of transport of Bose–Einstein condensates in magnetic microtraps. In [23] the focus was on modeling issues related to the optimal control formulation, its physical interpretation, and demonstration by results of simulation of the advantages provided by the optimal control approach. In this paper, we give a detailed discussion of the optimality conditions, illustrate some important implementation aspects concerning discretization of the optimality system, introduce two multigrid optimization schemes for solving the Bose–Einstein condensate control problem, and investigate the convergence properties of these schemes from a theoretical and scientific computing point of view showing the effectiveness of the proposed techniques.

From a mathematical viewpoint, the purpose of this paper goes beyond the numerical solution of the Bose–Einstein condensate control problem. In fact, we consider a challenging problem of bilinear control of an evolution model that describes nonlinear wave function evolution in complex Hilbert spaces. Bilinear control represents a class of nonlinear control strategies with the aim to obtain better system response than possible with linear control. The resulting bilinear systems pose new theoretical and computational problems which are open or have been only partially addressed. We contribute to this field of research with the formulation of effective multigrid solution strategies that can accommodate the presence of nonlinearities and bilinear control structure in wavelike evolution problems.

Our focus on multigrid techniques is motivated by recent advances in computational optimization [25, 27] based on reinterpreting multigrid ideas from an optimization point of view that leads to new concepts in the definition of the various multigrid components. In particular, we have that, under appropriate conditions, the coarse-grid correction provides a descent direction for the functional to be minimized. Augmenting the coarse-grid correction step with a line-search strategy and using appropriate one-grid optimization schemes as "smoothers" will result in fast and robust multigrid optimization procedures. In this framework the coarse-grid correction does not merely accelerate the one-grid optimization scheme that takes the place of the smoother. More than this, it provides effective search directions, and it gives the possibility to accommodate robust globalization strategies. The enhancement of robustness obtained with the coarse-grid correction turns out to be essential in the solution of our quantum control problem.

In the formulation of quantum optimal control of transport of Bose–Einstein condensates, the objective of the control is quantified through a cost functional depending on the final state as well as the physically motivated constraint of smooth varying magnetic control fields. This functional is minimized subject to the condition that the time dynamics of the condensate is governed by the Gross–Pitaevskii equation [15]. We show that the quantum optimal control approach [1, 9, 11, 12, 22, 26, 29, 31, 32] provides a versatile tool for determining efficient control strategies with realistic confinement potentials and nonlinearities in the condensate dynamics. In principle, with our approach we are able to explore a larger control space than the adiabatic scheme described above, where only transitions within a finite number of excited states are considered.

The formulation of the optimal control of the transport of Bose–Einstein condensates results in a coupled system of two nonlinear Schrödinger equations, with opposite time orientation, and an elliptic equation for the control. These are the

so-called state and adjoint equations and the optimality condition and represent the first-order necessary optimality conditions. The solution of this optimality system characterizes a solution of the optimal control problem. The optimality conditions for the Bose–Einstein condensate control problem are discussed in the next section.

To solve the nonlinear optimality system we proceed in section 3 by introducing a time-splitting spectral discretization [4, 3] adapted to the resulting state and adjoint equations and use finite differences to discretize the optimality condition. Thus we obtain second-order accuracy in time discretization and spectral accuracy in space. For a discussion on the importance of spectral accuracy for accurate computation of quantum evolution, see, e.g., [7].

In section 4, we introduce and discuss two multigrid optimization strategies: the cascadic nonlinear conjugate gradient scheme and the multigrid optimization scheme. The former results from combining an appropriate extension of a newly proposed nonlinear conjugate gradient method [14] with a cascadic acceleration scheme. Convergence of the proposed nonlinear conjugate gradient method is discussed, and it is pointed out that under appropriate conditions on the line search a sufficient descent condition is obtained. The multigrid optimization scheme [27] is a recent extension of the nonlinear multigrid strategy to optimization problems. This method appears to be an effective multigrid strategy for optimization problems having a minimum of requirements on the structure of the minimization problem. We give a detailed discussion of the convergence properties of this scheme in section 4.3. Both the cascadic nonlinear conjugate gradient scheme and the multigrid optimization scheme considered in this paper are new in the context of quantum optimal control computation.

In section 5, results of numerical experiments are reported to compare the computational performance of the two solution strategies considered in this paper. We obtain that the multigrid optimization scheme provides a more robust solution for faster and slower control setting, independently of the discretization, the optimization parameters, and the strength of the nonlinearity. We also report results of optimized control to show the advantage of the optimal control formulation.

A conclusion section completes the exposition of our work.

2. Formulation of the control problem. The mean-field dynamics of a coherent Bose–Einstein condensate is described by the Gross–Pitaevskii equation [15]

$$i\frac{\partial}{\partial t}\,\psi(x,t) = \left(-\frac{1}{2}\nabla^2 + V(x,\lambda(t)) + g\,|\psi(x,t)|^2\right)\psi(x,t), \tag{1}$$

where $x \in \Omega$ and $t \in [0,T]$, with g a coupling constant related to the scattering length of the atoms, density, and transversal confinement. Notice that, in principle, the evolution of the Bose–Einstein condensate is defined in unbounded spaces, while the fact that V is a confinement potential results in wave functions ψ whose support is localized in a bounded region. Therefore, with Ω we represent a spatial domain that is large enough to represent ψ during evolution.

We set $\hbar = 1$ and measure mass in units of the atom mass and length in units of micrometers. For a condensate of ^{87}Rb (rubidium) atoms, the time and energy scales are then given by 1.37 milliseconds and 5.58 nanokelvins, respectively. We assume that the wave function $\psi(x,t)$ is normalized to one $\int_\Omega |\psi(x,t)|^2 dx = 1$, $t \geq 0$, and therefore g in (1) incorporates the number of atoms N_A. Alternatively as in, e.g., [15], we could choose to normalize the wave function to N_A in the condensates, and then g represents the strength of interatomic interaction.

We consider confining potentials $V_\lambda(x,t) = V(x,\lambda(t))$ that are produced by magnetic microtraps whose variation is described by a control function $\lambda(t)$. We assume that $\lambda(t)$ is real and single-valued, although different situations, e.g., microtraps controlled by several parameters, could be treated equally well. In the following we treat the case typical in experimental settings where the initial and final potential configurations are given, and therefore we require that $\lambda(t)$ takes initial and final values of zero and one, respectively. These two extremal values correspond to the case where the potential $V_\lambda(x,t)$ is convex and to the case where it has a double-well structure. Furthermore it is physically reasonable to require that V_λ be spatially symmetric with respect to the origin of coordinates $V_\lambda(x,t) = V_\lambda(-x,t)$ and $V_\lambda(x,\cdot) \in C^2(\Omega)$ and to assume that V_λ is twice continuously differentiable in λ.

Suppose that initially the system is in the ground state ψ_0 for the potential $V(x,\lambda)|_{\lambda=0}$. Upon varying $\lambda(t)$ in the time interval $t \in [0,T]$ from zero to one, the system will pass through a sequence of states and will end up in the final state $\psi(T)$. Our purpose is to determine an optimal time evolution of $\lambda(t)$ that allows us to channel the system from the initial state ψ_0 at time zero to a desired state ψ_d at final time T. In accordance with [19], we assume ψ_d to be the ground state for the potential $V(x,\lambda)|_{\lambda=1}$ at time T. The ground state for a given potential $V(x,\lambda)$ is defined as the stationary state $\phi(x)$, with $\int_\Omega |\phi(x)|^2 dx = 1$, that minimizes the energy [2]

$$E_\lambda(\phi) = \int_\Omega \left(\frac{1}{2} |\nabla \phi(x)|^2 + V_\lambda(x)\, |\phi(x)|^2 + \frac{g}{2}\, |\phi(x)|^4 \right) dx.$$

For our purpose, in order to define a well-defined control problem, we introduce the following cost functional:

$$J(\psi,\lambda) = \frac{1}{2}\big(1 - \big|\langle \psi_d, \psi(T)\rangle\big|^2\big) + \frac{\gamma}{2} \int_0^T \left(\dot\lambda(t)\right)^2 dt, \tag{2}$$

with $\langle u, v\rangle = \int_\Omega u(x)^* v(x)\, dx$ the usual inner product in complex spaces, where the $*$ denotes complex conjugate, and $\|u\| = \langle u, u\rangle^{1/2}$.

The goal of the first term of the cost functional is to track the state ψ to a given terminal state ψ_d at $t = T$. This choice is different than the one given by $\frac{1}{2}\|\psi(T) - \psi_d\|^2$ as used in, e.g., [9]. In fact, in (2), the final wave function is required to match the desired target function only up to a global phase $e^{i\phi T}$ which cannot be specified. That is, we are free to specify the target function up to a phase shift.

The second term in the cost functional is from one hand for the regularization of the problem so that existence of at least one optimal control is ensured. On the other hand, the particular choice of regularization in (2) aims at penalizing fast varying confinement potentials that are more difficult to realize in real experiments. The regularization parameter $\gamma > 0$ allows one to vary the relative importance of the objectives represented by the two terms.

The control problem under consideration is therefore to minimize $J(\psi,\lambda)$ subject to the condition that $\psi(x,t)$ fulfills the Gross–Pitaevskii equation (1) with given initial conditions.

To solve this problem we introduce the Lagrange function

$$L(\psi,p,\lambda) = J(\psi,\lambda) + \Re e \left(p \,\Big|\, i\dot\psi - \left(-\frac{1}{2}\nabla^2 + V_\lambda + g|\psi|^2 \right) \psi \right), \tag{3}$$

with $(u|v) = \int_0^T \int_\Omega u(x,t)^*\, v(x,t) dx\, dt$, and $p(x,t)$ is the Lagrange multiplier.

By formally equating to zero the Frechét derivatives of L with respect to the triple (ψ, p, λ), we obtain the following optimality system characterizing the solution to our optimal control problem. We have

$$i\frac{\partial \psi}{\partial t} = \left(-\frac{1}{2}\nabla^2 + V_\lambda + g|\psi|^2\right)\psi, \tag{4}$$

$$i\frac{\partial p}{\partial t} = \left(-\frac{1}{2}\nabla^2 + V_\lambda + 2g|\psi|^2\right)p + g\,\psi^2\,p^*, \tag{5}$$

$$\gamma\,\frac{d^2\lambda}{dt^2} = -\Re e\left\langle \psi, \frac{\partial V_\lambda}{\partial \lambda}\,p\right\rangle, \tag{6}$$

which has to be solved together with the initial and terminal conditions

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

$$ip(T) = -\langle \psi_d, \psi(T)\rangle\,\psi_d, \tag{8}$$

$$\lambda(0) = 0\,, \quad \lambda(T) = 1. \tag{9}$$

Equation (8) follows from computing time integration by parts to determine the Frechét derivative with respect to the variable ψ. Notice that, while the state equation (4) with initial condition $\psi(0) = \psi_0$ evolves forward in time, the adjoint equation (5) with terminal condition (8) is marching backwards. The control equation (6) provides the optimality condition that determines the optimal control. Notice that because of H^1 regularization we have a natural setting to impose the required Dirichlet boundary conditions on the control function: $\lambda(0) = 0$ and $\lambda(T) = 1$.

We have that (4) is uniquely solvable for every $\lambda \in H^1(0, T; \mathbb{R})$ such that V_λ is a symmetric double-well potential; see [33]. Thus, it is meaningful to introduce the so-called reduced cost functional $\hat{J} : H^1(0, T; \mathbb{R}) \to \mathbb{R}$ given by

$$\hat{J}(\lambda) = J(\psi(\lambda), \lambda), \tag{10}$$

where $\psi(\lambda)$ denotes the unique solution to (4) for given λ. One can show that the gradient of $\hat{J}$ with respect to λ is given by

$$\nabla \hat{J}(\lambda) = -\gamma\,\frac{d^2\lambda}{dt^2} - \Re e\left\langle \psi, \frac{\partial V_\lambda}{\partial \lambda}\,p\right\rangle, \tag{11}$$

where ψ and p solve the state and the adjoint equations, respectively, with given λ.

2.1. Second-order optimality conditions. Quantum optimal control problems are usually nonconvex optimization problems so that different local minima are possible. Iterative numerical methods will converge to a local minimum close to the given starting point. The solutions to the first-order conditions (4)–(9) are necessary for optimality and to characterize extremal points. Local minima are obtained if second-order optimality conditions are also satisfied. In the following, we show that second-order optimality conditions are (locally) satisfied if a sufficient tracking of the target function is obtained.

To discuss second-order optimality conditions, assume that (ψ, p, λ) satisfy the optimality system (4)–(9). Let us denote the nonlinear Schrödinger equation (4) by $c(\psi, \lambda) = 0$. Consider that J and c are twice continuously differentiable and the second-order sufficient conditions for a minimum are given by the optimality system and the following:

$$L_{zz}(\psi, p, \lambda)(v, v) \geq c_1\,||v||^2, \quad c_1 > 0, \text{ for all } v \in \mathcal{N}(c'(\psi, \lambda)), \tag{12}$$

where $z = (\psi, \lambda)$ and c' represents the linearized state equation. We assume that the null space $\mathcal{N}(c'(\psi,\lambda))$ can be represented by $\mathcal{N}(c'(\psi,\lambda)) = T(\psi,\lambda)\,\Lambda$, where Λ is the space where the control is defined and

$$T(\psi,\lambda) = \begin{bmatrix} -c_\psi^{-1}\, c_\lambda \\ I_\lambda \end{bmatrix},$$

and c_ψ, c_λ are evaluated at $(\psi(\lambda),\lambda)$. Therefore condition (12) becomes

$$\nabla^2 \hat{J}(\lambda)(w,w) \geq c_2\, ||w||^2, \quad c_2 > 0, \tag{13}$$

for all w in the control space. The operator $\nabla^2\hat{J}$ is the reduced Hessian defined by

$$\nabla^2 \hat{J} = T^*\, L_{zz}\, T.$$

That is, $\nabla^2\hat{J}(\psi(\lambda),\lambda)$ is given by

$$\nabla^2 \hat{J} = L_{\lambda\lambda} + W^*\, L_{\psi\psi}\, W - L_{\lambda\psi}\, W - W^*\, L_{\psi\lambda},$$

where $W = W(\psi(\lambda),\lambda) = c_\psi(\psi(\lambda),\lambda)^{-1}\, c_\lambda(\psi(\lambda),\lambda)$.

Notice that $\nabla^2\hat{J}$ is symmetric. Therefore condition (13) requires that, in order to have a minimum, all eigenvalues of the reduced Hessian must be positive.

The purpose of this discussion is to show that, by guaranteeing sufficient tracking, that is, sufficiently small values of $\frac{1}{2}\big(1 - \big|\langle\psi_d, \psi(T)\rangle\big|^2\big)$, second-order optimality conditions are satisfied, and therefore a local minimum is obtained. To show this fact, we consider a simplified model with properties similar to that of our quantum control problem. Consider the following optimal control problem with time-dependent variables:

$$\begin{cases} \min_{\lambda\in\Lambda} J(\psi,\lambda) & := & \frac{1}{2}|\psi(T) - \psi_d|^2 + \frac{\gamma}{2}\|\dot{\lambda}\|^2, \\ i\dot{\psi} & = & a\,\psi + \lambda\,\psi, \end{cases} \tag{14}$$

where $\dot{\psi} = d\psi/dt$ and $c(\psi,\lambda) = i\dot{\psi} - a\,\psi - \lambda\,\psi$, $a \in \mathbb{R}$. Define $(u,v) = \int_0^T u\, v^* dt$ and $\|u\| = (u,u)^{1/2}$.

Define the Lagrange function

$$L(\psi,p,\lambda) = J(\psi,\lambda) + \Re e(p, i\dot{\psi} - a\,\psi - \lambda\,\psi).$$

Equating to zero the Frechét derivatives of L with respect to the triple (ψ,p,λ), we obtain the following optimality system:

$$\begin{aligned} i\dot{\psi} &= a\,\psi + \lambda\,\psi, & \psi(0) &= \psi_0, \\ i\dot{p} &= a\,p + \lambda\,p, & ip(T) &= \psi(T) - \psi_d, \\ \gamma\,\ddot{\lambda} &= -\Re e(p\,\psi^*), & \lambda(0) &= 0,\ \lambda(T) = 1, \end{aligned}$$

where we require $|\psi_0| = 1$. Differentiating twice gives

$$(L_{\psi\psi}\delta u, \delta v) = \delta u(T)^*\, \delta v(T), \qquad (L_{\lambda\lambda}\delta u, \delta v) = \gamma\,(\dot{\delta u}, \dot{\delta v}),$$

$$(L_{\lambda\psi}\delta u, \delta v) = (L_{\psi\lambda}\delta u, \delta v) = (-p\,\delta u, \delta v).$$

Therefore, we obtain

$$(\nabla^2\hat{J}\,\delta u, \delta u) = (W\delta u)(W\delta u)^* + 2\Re e(p\,\delta u, W\delta u) + \gamma(\dot{\delta u}, \dot{\delta u}).$$

From the result above we can state positive definiteness of the reduced Hessian if $\Re e(p\,\delta u, W\delta u)$ is sufficiently small. Now notice that, because the evolution of $p(t)$ is unitary, we have $|p(t)| = |p(T)| = |\psi(T) - \psi_d|$. By using this result we obtain that $|\Re e(p\,\delta u, W\delta u)| \le C\,|\psi(T) - \psi_d|\,\|\delta u\|^2$ for some $C > 0$ depending on λ. Therefore we conclude that for sufficiently small values of $|\psi(T) - \psi_d|$ positive definiteness of the reduced Hessian is obtained.

As discussed below, positive definiteness of the reduced Hessian guarantees that the coarse-to-fine update direction obtained in the multigrid optimization scheme is a descent direction.

3. Discretization of the optimality system. In this section, we discuss the discretization of (4) and (5) using an unconditionally stable explicit second-order norm-preserving time-splitting spectral scheme (TSSP) [4, 3]. These properties make the TSSP scheme particularly suitable for computation on a hierarchy of grids. To introduce the time-splitting technique and the corresponding notation, we discuss the discretization of the forward equation. For the backward adjoint equation, additional work is required to implement the time-splitting method appropriately. This is due to the presence of the term $g\,\psi^2\,p^*$ in the adjoint equation. This problem is discussed in detail below.

For ease of notation we take $\Omega = (-L/2, L/2) \subset \mathbb{R}$, where L is large enough so that the support of the state and adjoint variables is well within the domain. Assume that the interval $(-L/2, L/2)$ is divided into N subintervals of size $h = L/N$. Subinterval end points are denoted by $x_j = (j-1)\,h - L/2$, $j = 1, \ldots, N$, where we take N to be an integer power of 2. The point x_{N+1} corresponds to $x = L/2$. Further we assume that the time interval $(0, T)$ is divided in M subintervals, and thus the time step size is given by $\delta t = T/M$. The approximation to $\psi(x,t)$ at x_j for the time $t_m = m\,\delta t$ is denoted with ψ_j^m. We set $\mu_k = \frac{2\pi}{L}k$. We assume that the function ψ is periodic in $(-L/2, L/2)$ in the sense that $\psi(-L/2^+) = \psi(L/2^-)$.

For a given continuous periodic function ψ, consider the polynomial

$$I_N\psi(x) = \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \tilde{\psi}_k\, e^{i\mu_k x}, \tag{15}$$

where

$$\tilde{\psi}_k = \frac{1}{N}\sum_{j=1}^{N} \psi(x_j)\, e^{-i\,\mu_k x_j}, \qquad \text{with} \quad x_j = (j-1)\,h - L/2. \tag{16}$$

The function $I_N\psi(x)$ is the $\frac{N}{2}$-degree trigonometric interpolant of ψ at the nodes x_j, i.e.,

$$I_N\psi(x_j) = \psi(x_j), \qquad j = 1, \ldots, N. \tag{17}$$

This polynomial is the discrete Fourier series of ψ.

The Fourier pseudospectral derivative of ψ is defined by $D_N\psi = (I_N\psi)'$. That is,

$$D_N\psi(x) = \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \tilde{\psi}'_k\, e^{i\,\mu_k x}, \qquad \text{where} \qquad \tilde{\psi}'_k = i\mu_k\tilde{\psi}_k.$$

Further, the Laplacian is given by

$$D_N^2\psi(x) = \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \tilde{\psi}_k''\, e^{i\mu_k x}, \qquad \text{where} \qquad \tilde{\psi}_k'' = -\mu_k^2\tilde{\psi}_k.$$

Now let $H = H_0 + V$, where $H_0 = -\frac{1}{2}\nabla^2$ is the free Hamiltonian and $V = V_\lambda + g|\psi|^2$ represents the effective potential including the magnetic confinement potential. Next we illustrate the time-splitting spectral scheme for the Gross–Pitaevskii equation written as follows; see [3] for more details:

$$i\frac{\partial\psi}{\partial t} = (H_0 + V)\,\psi.$$

The time-splitting scheme for this equation can formally be written in the following form:

$$\psi^{m+1} = e^{-i\frac{\delta t}{2}V^{m+1}} e^{-i\delta t\, H_0} e^{-i\frac{\delta t}{2}V^m}\,\psi^m. \tag{18}$$

From time t_m to time t_{m+1} we have three steps. For a $\delta t/2$ time step we first solve

$$i\frac{\partial\psi}{\partial t} = V\,\psi.$$

For $j = 1, \ldots, N$, we obtain

$$\psi_j^+ = \exp\left(-i(V_\lambda(x_j, t_m) + g|\psi_j^m|^2)\,\delta t/2\right)\,\psi_j^m. \tag{19}$$

With this value, we compute a full time step for the equation $i\frac{\partial\psi}{\partial t} = H_0\,\psi$, obtaining

$$\psi_j^{++} = \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \tilde{\psi}_k^+\, e^{\left(-i\mu_k^2\,\delta t/2\right)}\, e^{i\mu_k x_j}, \qquad j = 1, \ldots, N, \tag{20}$$

where

$$\tilde{\psi}_k^+ = \frac{1}{N}\sum_{j=1}^{N} \psi_j^+\, e^{-i\,\mu_k x_j}.$$

The final step consists in another $\delta t/2$ time step of the evolution governed by V with ψ^{++} as initial condition. Hence we have

$$\psi_j^{m+1} = \exp\left(-i(V_\lambda(x_j, t_{m+1}) + g|\psi_j^{++}|^2)\,\delta t/2\right)\,\psi_j^{++}, \qquad j = 1, \ldots, N. \tag{21}$$

Additional work is required to implement the time-splitting scheme for the adjoint equation (5) because of the presence of the term $g\,\psi^2\,p^*$ that prevents a straightforward application of an exponential solution formula. We proceed as follows. The adoint equation can be written in the following form:

$$i\frac{\partial}{\partial t}(p_r + i\,p_i) = \left(-\frac{1}{2}\nabla^2 + A\right)\,(p_r + i\,p_i) + (a_r + i\,a_i)\,(p_r - i\,p_i),$$

where $p = p_r + i\,p_i$, $A = V_\lambda + 2g|\psi|^2$, and $g\,\psi^2 = a_r + i\,a_i$. We focus on the first and third steps of the time-splitting formula, and therefore we consider

$$\frac{d}{dt}(p_r + i\,p_i) = -iA\,(p_r + i\,p_i) - i(a_r + i\,a_i)\,(p_r - i\,p_i).$$

This equation is equivalent to the following system of differential equations:

$$\dot{p}_r = Ap_i + (a_i p_r - a_r p_i), \tag{22}$$

$$\dot{p}_i = -Ap_r - (a_r p_r + a_i p_i). \tag{23}$$

We now introduce the Pauli matrices to write this system in a form suitable for exponential representation. The Pauli matrices are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and we write $\bar{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$.

Therefore the system (22)–(23) can be written as follows:

$$\frac{d}{dt}\begin{pmatrix} p_r \\ p_i \end{pmatrix} = i(\bar{u}\cdot\bar{\sigma})\begin{pmatrix} p_r \\ p_i \end{pmatrix},$$

where $\bar{u} = (ia_r, A, -ia_i)$. We can now use the exponential representation to obtain

$$\begin{pmatrix} p_r \\ p_i \end{pmatrix}(t+\delta t) = \exp\left(i\,\bar{u}\cdot\bar{\sigma}\,\delta t\right)\begin{pmatrix} p_r \\ p_i \end{pmatrix}(t). \tag{24}$$

Further, we use the fact that

$$\begin{aligned}\exp\left(i\,\bar{u}\cdot\bar{\sigma}\,\delta t\right) &= \cos\left(|\bar{u}|\,\delta t\right)\,I + i\,\sin\left(|\bar{u}|\,\delta t\right)\,\frac{\bar{u}}{|\bar{u}|}\cdot\bar{\sigma} \\ &= \cos\left(|\bar{u}|\,\delta t\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin\left(|\bar{u}|\,\delta t\right)\frac{1}{|\bar{u}|}\begin{pmatrix} a_i & A - a_r \\ -A - a_r & -a_i \end{pmatrix},\end{aligned}$$

where $|\bar{u}| = \sqrt{a_r^2 + a_i^2 + A^2}$. We can therefore use this formula in (24) and replace δt with $-\delta t/2$ to obtain the (backward) time evolution of the adjoint variable for the first and third steps of the time-splitting procedure. Notice that in this case because of the terms A and a_r, a_i, the values of the wave function ψ at intermediate half time steps are also required. For this purpose, the ψ function is also computed in the backward evolution by using (19)–(21) with δt replaced by $-\delta t$. This is possible since the time-splitting scheme is time-reversible.

Evaluation of the gradient of the reduced cost functional is given by the following:

$$\nabla\hat{J}(\lambda)^m = -\gamma\,\frac{\lambda^{m+1} - 2\lambda^m + \lambda^{m-1}}{\delta t^2} - \Re e\sum_{j=1}^{N} h\,(p_j^m)^*\,\frac{\partial V_\lambda}{\partial\lambda}|_{\lambda=\lambda^m}\,\psi_j^m.$$

We conclude this section by illustrating how to determine the initial-state and the target-state wave functions. In our setting the initial state ψ_0 and the target state ψ_d are the ground state wave functions of the Gross–Pitaevskii equation with single- ($\lambda = 0$) and double-well ($\lambda = 1$) potential, respectively. To determine these states we use the normalized gradient flow scheme [2]. In this approach, one considers

the evolution of the Gross–Pitaevskii equation with δt replaced by $-i\delta t$, and at each step of this evolution the resulting wave function is projected back to the unit sphere by normalization. A sketch of the normalized gradient flow scheme is given below. Consider a time step sequence $\{t_m\}_{m\geq 0}$ with time step size δt. Let $\phi(x, t_n)$ be the ground state wave function approximation after n iterations. We have the following.

ALGORITHM 1 (Normalized gradient flow method).

Step 1. *Given* $\phi(x, t_0)$, $t_0 = 0$, δt, $m = 0$.

Step 2. *Set* $t_{m+1} = t_m + \delta t$, *and solve*

$$\frac{\partial \phi}{\partial t}(x,t) = \left(\frac{1}{2}\nabla^2 - V_\lambda(x) - g\,|\phi(x,t)|^2\right)\phi(x,t), \qquad t_m < t < t_{m+1},$$

with $\phi(x, t_m)$ *as an initial condition.*

Step 3. *Normalize*

$$\phi(x, t_{m+1}) := \frac{\phi(x, t_{m+1})}{\|\phi(\cdot, t_{m+1})\|}.$$

If $\|\phi(\cdot, t_{m+1}) - \phi(\cdot, t_m)\| < tol$, *then stop.*

Step 4. *Set* $m = m + 1$; *goto Step* 2.

As an initial condition for this evolution we take a Gaussian, and the time-stepping process is iterated until the difference between two consecutive ground state approximations is less then $tol = 10^{-12}$.

4. Multigrid optimization strategies. Starting with this section, we focus on the iterative solution of the optimality system (4)–(9) approximated by using the discretization scheme previously discussed. The most popular and easy to implement method used to solve such problems is the gradient method with line search. This method uses a forward sweep of the state equation followed by a backward sweep of the adjoint equation, and then $-\nabla \hat{J}(\lambda)$ is determined and taken as descent direction in a line-search scheme to update the control field. This process is repeated until a suitable convergence criteria is satisfied, such as $\|\nabla \hat{J}(\lambda)\| < tol$. Numerical experience shows that the gradient method just described is inefficient because, after a few iterations, slowdown occurs.

A large improvement of this method is obtained by introducing the notion of conjugate directions. This leads to conjugate gradient algorithms that under suitable conditions have satisfactory convergence properties. The most general extension of these schemes is represented by nonlinear conjugate gradient (NCG) methods that can be applied to nonconvex problems; see, e.g., [18, 35]. The use of the NCG scheme to solve quantum control problems, apart from a less recent work [37], is new and appears to be competitive with state-of-the-art methods for quantum control computation [8]. In the control of linear finite-level quantum systems (finite-dimensional case), the NCG scheme proves to be efficient and robust [8], especially when it is used in combination with cascadic (nested) iteration techniques.

In this paper, we continue the development started in [8] and consider a cascadic NCG scheme to solve the problem of control of the nonlinear extended quantum Bose–Einstein condensate system (infinite-dimensional case). In addition we continue the development of multigrid optimization schemes for quantum control computation and present a MGOPT (multigrid optimization) scheme where the NCG iteration plays the role of the one-grid optimization scheme. The choice of the NCG scheme as the basic optimization method is suggested by the robustness of this scheme, the

fact that it has a relatively large convergence basin, and the easy of implementation. The MGOPT approach uses a refinement strategy similar to that of the cascadic approach in combination with a coarsening strategy similar to that of classical multigrid methods. In this framework, the coarse-grid correction provides a new effective descent direction for the functional to be minimized. The coarse-to-fine update direction is augmented with a line-search strategy and results in a fast and robust multigrid optimization procedure.

4.1. The nonlinear conjugate gradient method. We discuss the minimization of the real-valued differentiable function $\hat{J}(\lambda)$, representing the reduced cost functional $J(\psi(\lambda), \lambda)$, where $\psi(\lambda)$ denotes the solution to the state equation (4). Denote that $g(\lambda) = \nabla\hat{J}(\lambda)$. In the following, we formulate the NCG scheme in a continuous setting in the space of real-valued functions since the control function is real. We have $(u, v) = (u, v)_{L^2(0,T;\mathbb{R})} = \int_0^T u(t)v(t)dt$ and $\|u\| = \sqrt{(u, u)}$.

In the common NCG variants, the basic idea is to avoid matrix operations and express the search directions recursively as

$$d_{k+1} = -g_{k+1} + \beta_k\, d_k \tag{25}$$

for $k = 1, 2, \ldots$, with $d_1 = -g_1$. In general, for convergence it is required that d_k is a descent direction for any k; i.e., $(g_k, d_k) < 0$ holds.

The iterates for a minimum point are given by

$$\lambda_{k+1} = \lambda_k + \tau_k\, d_k, \tag{26}$$

where $\tau_k > 0$ is a step length. The parameter β_k is chosen so that (25)–(26) reduces to the linear CG scheme if $\hat{J}$ is a strictly convex quadratic function and τ_k is the exact one-dimensional minimizer of $\hat{J}$ along d_k.

We focus on the NCG scheme of Dai and Yuan [14] based on the formula

$$\beta_k = \beta_k^{DY} := \frac{(g_{k+1}, g_{k+1})}{(d_k, y_k)}, \tag{27}$$

where $y_k = g_{k+1} - g_k$. In [14], convergence of the proposed NCG scheme is established by requiring that the step length τ_k, determined by backtracking, satisfies the Armijo condition of sufficient decrease of $\hat{J}$'s value given by

$$\hat{J}(\lambda_k + \tau_k d_k) \le \hat{J}(\lambda_k) + \delta\, \tau_k\, (g_k, d_k) \tag{28}$$

together with the Wolfe condition

$$(g(\lambda_k + \tau_k d_k), d_k) > \sigma\, (g_k, d_k), \tag{29}$$

where $0 < \delta < \sigma < 1/2$; see [28]. The last condition means that the graph of $\hat{J}$ should not increase too rapidly beyond the minimum.

We implement the following NCG scheme.

ALGORITHM 2 (NCG method).

Step 1. *Given* $k = 1$, λ_1, $d_1 = -g_1$, *if* $\|g_1\| < tol$, *then stop.*

Step 2. *Compute* $\tau_k > 0$ *satisfying* (28)–(29).

Step 3. *Let* $\lambda_{k+1} = \lambda_k + \tau_k\, d_k$.

Step 4. *Compute* $g_{k+1} = \nabla\hat{J}(\lambda_{k+1})$.

If $\|g_{k+1}\| < tol_{abs}$ *or* $\|g_{k+1}\| < tol_{rel}\, \|g_1\|$ *or* $k = k_{\max}$, *then stop.*

Step 5. *Compute* β_k *by* (27).
Step 6. *Let* $d_{k+1} = -g_{k+1} + \beta_k\, d_k$.
Step 7. *Set* $k = k + 1$*; goto Step* 2.

Convergence of the NCG scheme defined above is discussed in [14]. In [8] the extension and convergence analysis of the NCG scheme for the case of complex quantum control fields is given. While we do not report the convergence proof, we recall the essential features that are required for the convergence analysis.

ASSUMPTION 3.

(1) *$\hat{J}$ is bounded from below and is continuously real differentiable in a neighborhood $\mathcal{N}$ of the level set $\mathcal{L} = \{\lambda \in H^1(0,T;\mathbb{R}) : \hat{J}(\lambda) \le \hat{J}(\lambda_1)\}$.*
(2) *$\nabla\hat{J}$ is Lipschitz continuous in $\mathcal{N}$; i.e., there exists a constant $M > 0$ such that*

$$\|\nabla\hat{J}(\lambda_1) - \nabla\hat{J}(\lambda_2)\| \le M\,\|\lambda_1 - \lambda_2\| \quad \textit{for all } \lambda_1, \lambda_2 \in \mathcal{N}.$$

Notice that in our setting Assumption 3(1) is satisfied. Moreover, we can prove that $\lambda \mapsto \hat{J}(\lambda)$ is twice continuously Fréchet differentiable so that Assumption 3(2) holds at least locally. Convergence of the NCG is stated in the next theorem, proved in [14]; see also [8].

THEOREM 4. *Suppose that λ_1 is a starting point for which Assumption* 3 *holds. Let the sequence $\{\lambda_k\}_{k\ge1}$ be generated by the NCG Algorithm* 2*. Then this algorithm either terminates at a stationary point or converges in the sense that*

$$\liminf_{k\to\infty} \|g_k\| = 0.$$

If, in addition, the sequence $\{\|d_k\|/\|g_k\|\}$ is bounded, then $\lim_{k\to\infty}\|g_k\| = 0$.

In the proof of this theorem, it is required that the standard Wolfe condition (29) is satisfied. By replacing (29) with the strong Wolfe condition

$$\big|(g(\lambda_k + \tau_k d_k), d_k)\big| \le -\sigma\,(g_k, d_k),$$

we have that the following sufficient descent condition is guaranteed:

$$(g_k, d_k) \le -c\,\|g_k\|^2,$$

where $c = 1/(1+\sigma)$ and for all $k \ge 1$. This condition is used in the convergence analysis of the MGOPT scheme, and it is implemented in the corresponding algorithm.

4.2. Cascadic NCG scheme. The motivation for combining the cascadic technique with the NCG scheme comes from our previous computational experience [8] and results given in [5, 34], where a cascadic conjugate gradient method is discussed and optimal computational complexity for elliptic problems is proved. While we are not able to extend these convergence results to the present case of nonconvex optimization, we are still able to obtain a considerable computational improvement.

To illustrate the cascadic NCG (CNCG) scheme, consider a hierarchy of nested grids with index $k = k_0, \dots, k_f$. The idea is to start from a coarse grid, with index k_0, where the size of the problem is small, and therefore the problem can be solved by the NCG iterative scheme with a reasonable computational effort. Let us denote with λ_{k_0} the solution obtained by this process with, e.g., "zero" initialization $\lambda^*_{k_0}$. The step that follows is to interpolate this solution to the next finer grid by using an interpolation operator I_k^{k+1}. Therefore we obtain an initialization of the NCG iterative process on the finer grid that is given by

$$\lambda^*_{k+1} = I_k^{k+1}\lambda_k,$$

where $k = k_0$. With this initialization, after a sufficient number of NCG iterations we obtain the solution λ_{k+1}. This process is repeated until the finest grid is reached and the desired solution is obtained.

An algorithm of this method is given by the following. Denote with $\lambda_k = NCG_k(\lambda_k^*)$ the result of the iteration, with λ_k^* as initialization, that is applied until a given stopping criteria is satisfied. We have the following.

ALGORITHM 5 (CNCG method).

Step 1. *Given* $k = k_0$, $\lambda_{k_0}^*$.
Step 2. *Compute* $\lambda_k = NCG_k(\lambda_k^*)$.
Step 3. *If* $k = k_f$, *then stop.*
Step 4. *Else if* $k < k_f$, *then interpolate* $\lambda_{k+1}^* = I_k^{k+1}\lambda_k$.
Step 5. *Set* $k = k + 1$; *goto Step* 2.

4.3. The MGOPT method. In [27, 25] the so-called MGOPT multigrid approach to optimization problems has been proposed that closely resembles the well-known full approximation storage scheme [10] and is similar to the damped nonlinear multigrid method discussed in [21]. Our interest in the MGOPT strategy is that it allows one to interpret multigrid schemes from an optimization point of view. This fact allows one to extend the multigrid methods to solve optimization problems, and it opens new perspectives for investigating convergence of nonlinear multigrid schemes using optimization theory; see [6]. In this framework, interpreting the action of the multigrid solution process in terms of low- and high-frequency components may not be possible. On the other hand, we have that with the MGOPT scheme there is the possibility to explore different minimizing directions provided by the one-grid optimization scheme and by the coarse-grid correction step. The convergence analysis shows that the latter provides a direction of descent whenever the Hessian of the optimization problem is positive definite. It appears that this new descent direction results in a more robust optimization algorithm.

The MGOPT method can be applied to optimization problems that can be formulated in a hierarchy of unconstrained optimization problems with varying resolution. We use the parameter $k = 1, 2, \ldots, K$ to index the various levels of resolution (or discretization). Here K denotes the finest resolution. For differential problems that can be defined on different grids, we have the usual multigrid notion of hierarchy of grids with k being the grid level.

We discuss the MGOPT solution to our optimization problem

$$\min_{\lambda_k} \hat{J}_k(\lambda_k), \tag{30}$$

where $\lambda_k = (\lambda_k^m)_{m=0,M_k}$ and $\hat{J}_k(\cdot)$ is the reduced cost functional with the state and adjoint equations defined by

$$Q_k(h_k, \delta t_k) = \{(x_j, t_m) : x_j = (j-1)h_k,\ 1 \le j \le N_k,\ \text{and}\ t_m = m\delta t_k,\ 0 \le m \le M_k\},$$

where $M_k = M/2^{K-k}$ and $N_k = N/2^{K-k}$, where $k = 1, 2, \ldots, K$.

In this framework $\lambda_k \in V_k$ is the (unconstrained) optimization variable in the space V_k of real $M_k + 1$-dimensional vectors. For variables defined on V_k we introduce the inner product $(\cdot,\cdot)_k$, which is the discrete version of $(u, v)_{L^2(0,T;\mathbb{R})}$. We define with $\|x\|_k = (x, x)_k^{1/2}$.

Among spaces V_k, restriction operators $I_k^{k-1} : V_k \to V_{k-1}$ and prolongation operators $I_{k-1}^k : V_{k-1} \to V_k$ are defined. We require that $(I_k^{k-1}u, v)_{k-1} = (u, I_{k-1}^k v)_k$

for all $u \in V_k$ and $v \in V_{k-1}$. This property is satisfied by the standard full-weighting intergrid transfer operators; see, e.g., [38].

Now denote with $\lambda_k^{(l)} = O_k(\lambda_k^{(l-1)})$ one step of an appropriate optimization scheme O_k, that provides a minimizing sequence $\{\lambda_k^{(l)}\}$ for $\hat{J}(\lambda_k)$. Later, O_k represents the action of the NCG scheme described in the previous section. Given an approximation $\lambda_k^{(l)}$ to the solution of (30), we require that the application of O_k results in sufficient reduction expressed by $\hat{J}_k(O_k(\lambda_k^{(l)})) < \hat{J}_k(\lambda_k^{(l)}) - \eta\|\nabla\hat{J}_k(\lambda_k^{(l)})\|^2$ for some $\eta \in (0,1)$.

We apply m_1 times the optimization iteration to (30) starting with the current approximation $\lambda_k^{(0)}$ to obtain the approximate solution $\tilde{\lambda}_k = \lambda_k^{(m_1)}$. Now the desired correction e_k to $\tilde{\lambda}_k$ for the minimum is defined by $\nabla\hat{J}_k(\tilde{\lambda}_k + e_k) = 0$. This can be equivalently written as

$$\nabla\hat{J}_k(\tilde{\lambda}_k + e_k) - \nabla\hat{J}_k(\tilde{\lambda}_k) = -\nabla\hat{J}_k(\tilde{\lambda}_k). \tag{31}$$

Next, we discuss the representation of the problem (31) on the coarser level $k-1$. To represent $\tilde{\lambda}_k + e_k$ on the coarse grid we write

$$\lambda_{k-1} = I_k^{k-1}\tilde{\lambda}_k + e_{k-1}. \tag{32}$$

Notice that this approach makes sense if we can assume that $e_k \in V_k$ is such that it can be approximated by $e_{k-1} \in V_{k-1}$.

To formulate (31) on the coarse space, we replace $\nabla\hat{J}_k(\cdot)$ by $\nabla\hat{J}_{k-1}(\cdot)$ on the left-hand side of this equation, $\tilde{\lambda}_k$ by $I_k^{k-1}\tilde{\lambda}_k$, and $\nabla\hat{J}_k(\tilde{\lambda}_k)$ on the right-hand side of (31) by $I_k^{k-1}\nabla\hat{J}_k(\tilde{\lambda}_k)$.

We get the following MGOPT equation:

$$\nabla\hat{J}_{k-1}(\lambda_{k-1}) = \nabla\hat{J}_{k-1}(I_k^{k-1}\tilde{\lambda}_k) - I_k^{k-1}\nabla\hat{J}_k(\tilde{\lambda}_k). \tag{33}$$

In agreement with standard multigrid terminology, we define the fine-to-coarse gradient correction given by

$$\phi_{k-1} = \nabla\hat{J}_{k-1}(I_k^{k-1}\tilde{\lambda}_k) - I_k^{k-1}\nabla\hat{J}_k(\tilde{\lambda}_k).$$

Therefore (33) becomes $\nabla\hat{J}_{k-1}(\lambda_{k-1}) - \phi_{k-1} = 0$ or, better,

$$\nabla\left(\hat{J}_{k-1}(\lambda_{k-1}) - \phi_{k-1}\lambda_{k-1}\right) = 0,$$

where $\phi_{k-1}\lambda_{k-1}$ means the vector product in V_{k-1} between ϕ_{k-1} and λ_{k-1}.

The following remark explains why we call ϕ_{k-1} the fine-to-coarse gradient correction.

Remark 6. We have that

$$\nabla\left(\hat{J}_{k-1}(\lambda_{k-1}) - \phi_{k-1}\lambda_{k-1}\right)\Big|_{\lambda_{k-1}=I_k^{k-1}\tilde{\lambda}_k} = I_k^{k-1}\nabla\hat{J}_k(\tilde{\lambda}_k);$$

that is, the gradient of the coarse-grid functional

$$\tilde{J}_{k-1}(\lambda_{k-1}) = \hat{J}_{k-1}(\lambda_{k-1}) - \phi_{k-1}\lambda_{k-1}$$

at $I_k^{k-1}\tilde{\lambda}_k$ equals the restriction of the gradient of the fine-grid functional at $\tilde{\lambda}_k$.

As a consequence, in correspondence to a minimum λ_k^* we have $\nabla \hat{J}_k(\lambda_k^*) = 0$ and $\nabla \tilde{J}_{k-1}(I_k^{k-1}\lambda_k^*) = 0$. This means that at convergence $\lambda_{k-1}^* = I_k^{k-1}\lambda_k^*$.

From an optimization point of view, solving (33) corresponds to the problem of minimizing $\tilde{J}_{k-1}(\lambda_{k-1})$. By applying $m = 1$ (V-cycle) or $m = 2$ (W-cycle) MGOPT steps to this problem, we obtain $\tilde{\lambda}_{k-1}$ such that $\tilde{J}_{k-1}(\tilde{\lambda}_{k-1}) < \tilde{J}_{k-1}(I_k^{k-1}\tilde{\lambda}_k)$. We are now ready to define the coarse-to-fine minimization step given by

$$\lambda_k = \tilde{\lambda}_k + \alpha\, I_{k-1}^k(\tilde{\lambda}_{k-1} - I_k^{k-1}\tilde{\lambda}_k). \tag{34}$$

Here α is a step length of minimization in the direction given by $I_{k-1}^k(\tilde{\lambda}_{k-1} - I_k^{k-1}\tilde{\lambda}_k)$. Under appropriate conditions discussed below, this is a direction of descent. The step length α is determined so to satisfy an Armijo-type condition of sufficient decrease.

The update step given by (34) is followed by m_2 times optimization steps by O_k. This completes the description of the iterative MGOPT scheme. One MGOPT iterate is described in the following. Let $\lambda_k^{(0)}$ be the starting approximation.

ALGORITHM 7 (MGOPT method).

Step 1. *If $k = 1$, solve $\min_{\lambda_k} \hat{J}_k(\lambda_k)$ directly.*

Step 2. *Preoptimization. $\lambda_k^{(l)} = O_k\,(\lambda_k^{(l-1)})$, $l = 1, \ldots, m_1$.*

Step 3. *Computation of the fine-to-coarse gradient correction*

$$\phi_{k-1} = \nabla \hat{J}_{k-1}(I_k^{k-1}\lambda_k^{(m_1)}) - I_k^{k-1}\nabla \hat{J}_k(\lambda_k^{(m_1)}).$$

Step 4. *Call m times MGOPT-(m_1, m_2) to solve $\min_{\lambda_{k-1}} \tilde{J}_{k-1}(\lambda_{k-1})$, where*

$$\tilde{J}_{k-1}(\lambda_{k-1}) = \hat{J}_{k-1}(\lambda_{k-1}) - \phi_{k-1}\lambda_{k-1}.$$

Step 5. *Coarse-to-fine minimization step with line search given by*

$$\lambda_k^{(m_1+1)} = \lambda_k^{(m_1)} + \alpha\, I_{k-1}^k(\lambda_{k-1} - I_k^{k-1}\lambda_k^{(m_1)}).$$

Step 6. *Postoptimization. $\lambda_k^{(l)} = O_k\,(\lambda_k^{(l-1)})$, $l = m_1 + 2, \ldots, m_1 + m_2 + 1$.*

To discuss the convergence properties of the MGOPT scheme applied to the control of the Bose–Einstein condensate transport problem, we now make some assumptions that can be proved by using techniques developed in [6]. We assume that the optimal solution (ψ^*, p^*, λ^*) to (4)–(9) under appropriate regularity assumptions and conditions on the optimization parameters satisfies second-order sufficient optimality conditions. Let us denote with $(\psi_k^*, p_k^*, \lambda_k^*)$ the optimal solution of the discretized optimal control problem. We assume that, for each k, $\hat{J}_k$ is twice Frechét differentiable and $\nabla^2 \hat{J}_k$ is (locally) positive definite and satisfies the condition $(\nabla^2 \hat{J}_k(\lambda)y, y)_k \geq \beta \|y\|_k^2$ together with $\|\nabla^2 \hat{J}_k(\lambda) - \nabla^2 \hat{J}_k(y)\| \leq \rho\, \|\lambda - y\|_k$ uniformly for some positive constants β and ρ. We use the expansion

$$\hat{J}_k(\lambda + z) = \hat{J}_k(\lambda) + (\nabla \hat{J}_k(\lambda), z)_k + \frac{1}{2}\int_0^1 (\nabla^2 \hat{J}_k(\lambda + tz)z, z)_k\, dt. \tag{35}$$

The discussion that follows is based on the following lemma [21].

LEMMA 8. *For $v, x, y \in V_k$ assume that $(\nabla \hat{J}_k(\lambda), y)_k \leq 0$, and let γ be such that*

$$0 \leq \gamma \leq -2\delta(\nabla \hat{J}_k(\lambda), y)_k \left[\int_0^1 (\nabla^2 \hat{J}_k(\lambda + t\gamma y)y, y)_k dt\right]^{-1} \qquad \text{for some } \delta \in [0, 1].$$

Then

$$-(1-\delta)\gamma(\nabla \hat{J}_k(\lambda), y)_k \le \hat{J}_k(\lambda) - \hat{J}_k(\lambda+\gamma y) \le -\gamma(\nabla \hat{J}_k(\lambda), y)_k. \tag{36}$$

Proof. Set $z = \gamma\, y$ in (35). The first inequality follows from the restriction to γ. The second inequality follows from the positivity of $\nabla^2 \hat{J}_k$. ∎

The first purpose of the present discussion on the convergence of the MGOPT method is to prove that we can find $0 < \alpha \le 2$ in Step 5 of Algorithm 7 such that an Armijo-type condition of sufficient decrease is satisfied. This is proved in the following lemma.

LEMMA 9. *For $v, x, y \in V_k$ assume that $(\nabla \hat{J}_k(\lambda), y)_k \le 0$, and let*

$$\alpha(\lambda, y) = \min\left\{2, \frac{-(\nabla \hat{J}_k(\lambda), y)_k}{(\nabla^2 \hat{J}_k(\lambda) y, y)_k + \rho\, \|y\|_k^3}\right\}. \tag{37}$$

Then

$$\hat{J}_k(\lambda + \alpha(\lambda, y) y) \le \hat{J}_k(\lambda) + \frac{1}{2}\alpha(\lambda, y)(\nabla \hat{J}_k(\lambda), y)_k. \tag{38}$$

Proof. For the proof it is enough to verify that Lemma 8 may be applied with $\gamma = \alpha(\lambda, y)$ and $\delta = 1/2$. Notice that

$$\int_0^1 (\nabla^2 \hat{J}_k(\lambda + t\alpha\, y) y, y)_k\, dt \le (\nabla^2 \hat{J}_k(\lambda) y, y)_k + \rho\, \|y\|_k^3.$$

Therefore we have

$$\alpha(\lambda, y) \le \frac{-(\nabla \hat{J}_k(\lambda), y)_k}{(\nabla^2 \hat{J}_k(\lambda) y, y)_k + \rho\, \|y\|_k^3} \le \frac{-(\nabla \hat{J}_k(\lambda), y)_k}{\int_0^1 (\nabla^2 \hat{J}_k(\lambda + t\alpha\, y) y, y)_k dt}.$$

Hence α satisfies the condition of Lemma 8 with $\delta = 1/2$. ∎

The following lemma states that the coarse-to-fine minimization step with step length α given by Lemma 9 is a minimizing step (without requiring exact solution of the coarse minimization problem).

LEMMA 10. *Take $\lambda_k \in V_k$. Denote that $\tilde{J}_{k-1}(\lambda_{k-1}) = \hat{J}_{k-1}(\lambda_{k-1}) - \phi_{k-1}\, \lambda_{k-1}$, where $\phi_{k-1} = \nabla \hat{J}_{k-1}(I_k^{k-1}\lambda_k) - I_k^{k-1} \nabla \hat{J}_k(\lambda_k)$. Let $\tilde{\lambda}_{k-1} \in V_{k-1}$ be such that $\tilde{J}_{k-1}(\tilde{\lambda}_{k-1}) \le \tilde{J}_{k-1}(I_k^{k-1}\lambda_k)$, and define $y_k = I_{k-1}^k(\tilde{\lambda}_{k-1} - I_k^{k-1}\lambda_k)$. Then*

$$\hat{J}_k(\lambda_k + \alpha(\lambda_k, y_k) y_k) \le \hat{J}_k(\lambda_k) + \frac{1}{2}\alpha(\lambda_k, y_k)(\nabla \hat{J}_k(\lambda_k), y_k)_k, \tag{39}$$

where $\alpha(\lambda_k, y_k)$ is defined in Lemma 9 (strict inequality holds if $\tilde{J}_{k-1}(\tilde{\lambda}_{k-1}) < \tilde{J}_{k-1}(I_k^{k-1}\lambda_k)$).

Proof. The proof follows from Lemma 9 after showing that $(\nabla \hat{J}_k(\lambda_k), y_k)_k \le 0$. From (35) we obtain that

$$(\nabla \tilde{J}_{k-1}(I_k^{k-1}\lambda_k), \tilde{\lambda}_{k-1} - I_k^{k-1}\lambda_k)_k \le \tilde{J}_{k-1}(\tilde{\lambda}_{k-1}) - \tilde{J}_{k-1}(I_k^{k-1}\lambda_k) \le 0.$$

Now we have

$$\begin{aligned}(\nabla \hat{J}_k(\lambda_k), y_k)_k &= (\nabla \hat{J}_k(\lambda_k), I_{k-1}^k(\tilde{\lambda}_{k-1} - I_k^{k-1}\lambda_k))_k \\ &= (I_k^{k-1}\nabla \hat{J}_k(\lambda_k), \tilde{\lambda}_{k-1} - I_k^{k-1}\lambda_k)_{k-1} \\ &= (\nabla \tilde{J}_{k-1}(I_k^{k-1}\lambda_k), \tilde{\lambda}_{k-1} - I_k^{k-1}\lambda_k)_{k-1} \le 0.\end{aligned}$$

For the last equality recall Remark 6. ∎

The following theorem states convergence of the MGOPT method.

THEOREM 11. *For each k, let λ_k^* be the solution to* (30). *Further, let $\hat{J}_k$ be twice Frechét differentiable, and let $\nabla^2 \hat{J}_k$ be locally Lipschitz continuous and satisfy $(\nabla^2 \hat{J}_k(\lambda_k^*)y, y)_k \geq \beta \|y\|_k^2$ together with $\|\nabla^2 \hat{J}_k(\lambda) - \nabla^2 \hat{J}_k(y)\| \leq \rho \|\lambda - y\|_k$ uniformly for some positive constants β and ρ in a neighborhood V_k^ϵ of λ_k^*. Then the MGOPT scheme given by Algorithm* 7 *provides a minimizing step.*

Proof. Let $\lambda_k^{(0)} \in V_k^\epsilon$. Then $A = \{\lambda \in V_k : \hat{J}_k(\lambda) \leq \hat{J}_k(\lambda_k^{(0)})\}$ is a compact set.

For $k = 2$, let λ_k be the result of the MGOPT step. We have $\tilde{\lambda}_{k-1} = \operatorname{argmin}_{\lambda \in V_{k-1}} \hat{J}_{k-1}(\lambda)$, and from Lemma 10 it follows that

$$\begin{aligned}\hat{J}_k(\lambda_k) = \hat{J}_k(O_k^{m_2}(\lambda_k^{m_1+1})) &\leq \hat{J}_k(\lambda_k^{(m_1)} + \alpha\, I_{k-1}^k(\lambda_{k-1} - I_k^{k-1}\lambda_k^{(m_1)})) \\ &\leq \hat{J}_k(O_k^{m_1}(\lambda_k^0)) \leq \hat{J}_k(\lambda_k^0),\end{aligned}$$

where strict inequality occurs in all steps whenever $\nabla \hat{J}_k$ is nonzero. For $k > 2$, due to the induction hypothesis and because of Lemma 10, the theorem holds. ∎

5. Numerical experiments. The purpose of this section is to validate the optimal control formulation and compare the CNCG method with the MGOPT method. The cascadic NCG scheme is a modern approach to optimization problems which is expected to perform well with quantum optimal control problems; see [8]. The MGOPT scheme is developed and applied for the first time to these problems and is expected to result in an improvement with respect to the CNCG scheme because of the additional coarse-grid optimization strategy. In fact, the MGOPT scheme improves robustness while the computational effort may depend on a problem's settings.

We remark that BEC experiments are carried out in three dimensions. Nevertheless trapped Bose–Einstein condensates are let move in a two-dimensional ($2D$) space (e.g., near the surface of atom chips), and the magnetic confinement potential used to manipulate the condensate usually acts along some preferred direction. Therefore we consider a one-dimensional space problem along the direction determined by the magnetic field. This is in agreement with the results presented in [23], where it is found that optimal control functions computed within a $1D$- and a $2D$-space framework are essentially identical.

We start by choosing a specific potential that is proposed in [24] to create condensates of trapped atoms coupled with a radio-frequency field. These can be realized by using simple and highly integrated wire geometries on atom chips. We have that V_λ for a given λ is a double-well potential of the following form:

$$V(x, \lambda) = -\frac{\lambda^2 d^2}{8c} x^2 + \frac{1}{c} x^4, \tag{40}$$

where $c = 40$ and d is a parameter corresponding to twice the distance of the two minima in the double-well potential.

In Figure 1, a visualization of the wave function evolution governed by the Gross–Pitaevskii equation with the potential (40) and $\lambda(t) = t/T$ is given. One can see the oscillatory character of the wave function. In this figure, the potential $V_\lambda(x)$ for $\lambda = 0$ and $\lambda = 1$ is also given. The initial condition is given by the ground state corresponding to $\lambda = 0$.

In the experiments, we consider the following setting. For the line search in the NCG scheme, we take $\delta = 10^{-4}$ and $\sigma = 0.4$ and use the strong Wolfe condition. This setting is used for both CNCG and MGOPT calculation. We choose $tol = 10^{-3}$

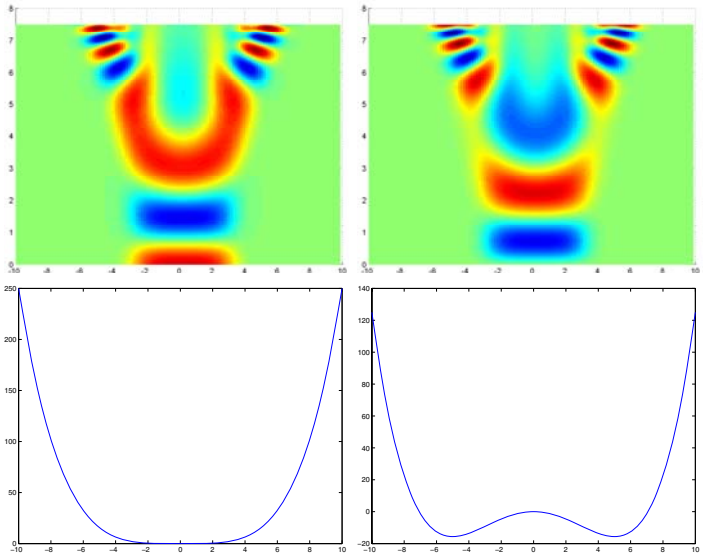

FIG. 1. *On top, the real (left) and imaginary parts of the solution* $\psi(x,t)$ *on the space-time domain (top) for* $\lambda(t) = t/T$ *and initial condition* ψ_0. *On the bottom, the potential* $V_\lambda(x)$ *for* $\lambda = 0$ *(left) and* $\lambda = 1$.

and $tol_{rel} = 10^{-2}$. For the line search for the coarse-to-fine minimization, we use the Armijo condition (28) also with $\delta = 10^{-4}$. The coarsest grid is 64×625 in space and time. Finer meshes are obtained with refinement by halving the mesh size. In the cascadic NCG algorithm, the solution process is transferred to the next finer grid whenever tolerance is reached or when the conditions of sufficient decrease of the strong Wolfe conditions cannot be fulfilled. In the MGOPT scheme, we use $m_1 = m_2 = 3$ and $m = 1$, and the same stopping criteria as in the CNCG scheme apply for this scheme. This choice is to some extent arbitrary; however, it is driven by the attempt to be as close as possible to the standard choice made in multigrid-based computing. The choice $m = 1$ corresponds to the standard V-cycle. With $m_1 = m_2 = 3$ we have that at least two steps of the NCG scheme are not steepest descent steps.

Our computational space domain is $\Omega = (-L/2, L/2)$ and T ranges between 5 and 10; the former corresponds to a faster control. We use $L = 20$ and $d = 10$ in (40). In our optimal quantum control calculations, we consider $\gamma \in (10^{-6}, 10^{-2})$ and $g = 10$. In all calculations we set a starting approximation given by $\lambda(t) = t/T + \sin(2\pi t/T)/10 + \sin(5\pi t/T)/10$. This choice is arbitrary, and it appears that other choices do not influence the final result.

We now assess the computational performance of the CNCG and MGOPT schemes. The results for this discussion are reported in Tables 1 and 2. We notice that usually the values of CPU time required by the MGOPT scheme are larger than the one related to CNCG and that better values of tracking are obtained with the MGOPT calculation. In fact, the CNCG scheme is usually not faster than the MGOPT scheme; instead, it comes faster to stagnation and stops. We remark that in most cases, also with various modifications of the CNCG's settings, it was not possible for the CNCG scheme to run longer and achieve the accuracy of the MGOPT solution.

In Table 1 we see that the MGOPT scheme provides a robust solution for faster and slower control settings. In Table 2, we obtain with MGOPT a robust optimal

TABLE 1
Computational performance of the CNCG and MGOPT schemes. Mesh 256×2500.

		CNCG		MGOPT	
γ	T	$\frac{1}{2}\left(1-\left\|\langle\psi_d,\psi(T)\rangle\right\|^2\right)$	CPU	$\frac{1}{2}\left(1-\left\|\langle\psi_d,\psi(T)\rangle\right\|^2\right)$	CPU
10^{-1}	5	$1.49\cdot10^{-1}$	112	$4.23\cdot10^{-2}$	941
10^{-3}	5	$1.40\cdot10^{-2}$	825	$2.97\cdot10^{-3}$	515
10^{-5}	5	$1.29\cdot10^{-2}$	205	$4.56\cdot10^{-3}$	213
10^{-1}	10	$3.23\cdot10^{-3}$	473	$4.38\cdot10^{-4}$	625
10^{-3}	10	$1.39\cdot10^{-3}$	239	$1.19\cdot10^{-4}$	930
10^{-5}	10	$3.63\cdot10^{-3}$	65	$2.27\cdot10^{-4}$	425

TABLE 2
Computational performance of the CNCG and MGOPT schemes; $T = 7.5$.

		CNCG		MGOPT	
γ	*Mesh*	$\frac{1}{2}\left(1-\left\|\langle\psi_d,\psi(T)\rangle\right\|^2\right)$	CPU	$\frac{1}{2}\left(1-\left\|\langle\psi_d,\psi(T)\rangle\right\|^2\right)$	CPU
10^{-2}	512×5000	$2.01\cdot10^{-2}$	2847	$5.47\cdot10^{-4}$	3091
10^{-4}	512×5000	$1.05\cdot10^{-3}$	524	$3.57\cdot10^{-4}$	2095
10^{-6}	512×5000	$3.55\cdot10^{-4}$	1077	$5.98\cdot10^{-4}$	772
10^{-2}	256×2500	$1.26\cdot10^{-2}$	580	$6.70\cdot10^{-4}$	695
10^{-4}	256×2500	$5.13\cdot10^{-4}$	90	$5.47\cdot10^{-4}$	299
10^{-6}	256×2500	$6.47\cdot10^{-4}$	77	$4.54\cdot10^{-4}$	758
10^{-2}	128×1250	$2.23\cdot10^{-2}$	17	$9.69\cdot10^{-4}$	116
10^{-4}	128×1250	$4.54\cdot10^{-4}$	202	$6.01\cdot10^{-4}$	82
10^{-6}	128×1250	$1.38\cdot10^{-2}$	14	$8.78\cdot10^{-4}$	78

TABLE 3
Computational performance of the CNCG and MGOPT schemes for different values of g; $T = 7.5$, $\gamma = 10^{-4}$.

		CNCG		MGOPT	
g	*Mesh*	$\frac{1}{2}\left(1-\left\|\langle\psi_d,\psi(T)\rangle\right\|^2\right)$	CPU	$\frac{1}{2}\left(1-\left\|\langle\psi_d,\psi(T)\rangle\right\|^2\right)$	CPU
25	128×1250	$3.89\cdot10^{-4}$	53	$7.08\cdot10^{-4}$	149
50	128×1250	$2.35\cdot10^{-3}$	80	$9.84\cdot10^{-3}$	76
75	128×1250	$5.54\cdot10^{-3}$	90	$1.85\cdot10^{-3}$	163
100	128×1250	$4.93\cdot10^{-1}$	13	$2.47\cdot10^{-1}$	27
100	256×2500	$4.94\cdot10^{-1}$	50	$5.44\cdot10^{-3}$	257

control solution independently of mesh size and of the optimization parameter. We should mention that simple gradient computation results in tracking values that are one to two orders of magnitude larger than the ones obtained with the nonlinear CG scheme.

We obtain that the tracking to the target state represented by ψ_d is not very sensitive to changes on the regularization parameter, although better tracking is usually obtained for smaller γ. By considering the CPU times we notice that similar effort is required in the case of fast control $T = 5$ as opposed to slower control $T = 10$.

In Table 3 we compare the computational performance of the CNCG and MGOPT schemes for different values of the strength of the nonlinearity. We obtain that, as g becomes larger, the tracking of the final target function is less satisfactory. This is in part due to the fact that larger g values require finer meshes for accurate solution. However, we can see that for large values of g and corresponding finer meshes the

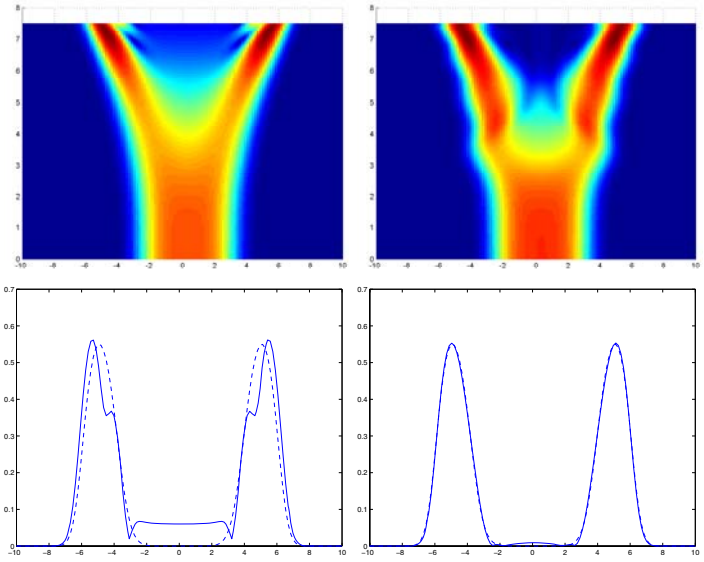

FIG. 2. *The function $|\psi(x,t)|$ on the space-time domain (top) for the linear (left) and the optimized (right) control. The corresponding profiles at $t = T$ (bottom, continuous line) compared to the desired state (dashed line). Mesh* 128×1250*; $\gamma = 10^{-4}$.*

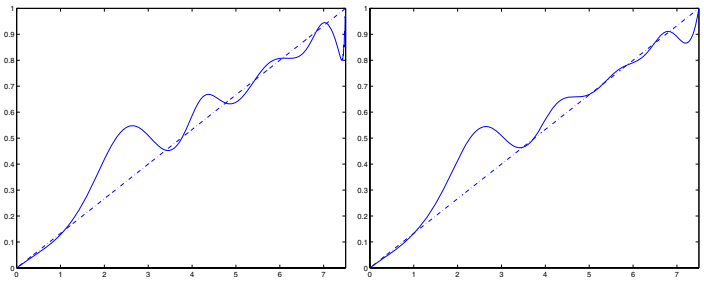

FIG. 3. *The optimal λ and its linear approximation. Result for $\gamma = 10^{-4}$ (left) and for $\gamma = 10^{-2}$.*

CNCG scheme may not converge, while the MGOPT scheme provides a convergent solution; compare the last two lines in Table 3.

We now compare the solution corresponding to an optimal λ with the one resulting from a linearly varying control function.

Figure 2 shows results for a wave function splitting for $T = 7.5$ for an optimized and a linear $\lambda(t)$ control function. One can see that the wave function becomes split along the time evolution and is transported in the respective minima. Excited vibrational states that originate during the initial splitting process can be seen by plotting the absolute value of the wave function profiles at final time. These are evident in the case of a linearly varying λ, where the split is not complete and part of the population remains localized between the two wells.

Correspondingly, the overlap with the desired ground state ψ_d is smaller than in the optimized case. In fact, for $\lambda(t) = t/T$ we obtain $\frac{1}{2}l(1-l|\langle\psi_d, \psi(T)\rangle|^2) = 6.26\,10^{-2}$, whereas for the optimal computed λ, we obtain $\frac{1}{2}l(1 - l|\langle\psi_d, \psi(T)\rangle|^2) = 6.01\,10^{-4}$. This tracking is obtained with the optimal control function shown in Figure 3. We

see that it substantially differs from a linear behavior. Notice that close to final time a fast varying control is obtained. A more smooth control is obtained with $\gamma = 10^{-2}$ that results in $\frac{1}{2}\left(1 - \left|\langle \psi_d, \psi(T)\rangle\right|^2\right) = 8.70\,10^{-4}$.

6. Conclusion. We investigated optimal control of transport of Bose–Einstein condensates in magnetic microtraps, giving a detailed formulation of the optimal control problem, including a discussion on optimality conditions. The discretization of the resulting optimization problem was illustrated, also addressing the special structure of the adjoint equation. The cascadic NCG and MGOPT multigrid optimization schemes were proposed and analyzed in detail.

The results presented in this paper suggest that the MGOPT scheme is one of the most robust optimization strategies suitable for quantum optimal control problems.

Acknowledgments. A. B. thanks Omar Ghattas and the J. Tinsley Oden Research Fellowship Program at the Institute for Computational Engineering and Sciences (ICES), University of Texas at Austin, for the nice hospitality and generous support. The authors thank P.-K. Rekdal and J. Schmiedmayer for helpful discussions. Special thanks also to Georg Stadler and Greg von Winckel and to the anonymous referees for helpful comments.

REFERENCES

[1] A. D. Bandrauk, M. C. Delfour, and C. Le Bris, eds., *Quantum Control: Mathematical and Numerical Challenges*, CRM Proc. Lecture Notes 33, American Mathematical Society, Providence, RI, 2004.

[2] W. Bao and Q. Du, *Computing the ground state solution of Bose–Einstein condensates by a normalized gradient flow*, SIAM J. Sci. Comput., 25 (2004), pp. 1674–1697.

[3] W. Bao, D. Jaksch, and P. A. Markowich, *Numerical solution of the Gross–Pitaevskii Equation for Bose–Einstein condensation*, J. Comput. Phys., 187 (2003), pp. 318–342.

[4] W. Bao, S. Jin, and P. A. Markowich, *On time-splitting spectral approximation for the Schrödinger equation in the semiclassical regime*, J. Comput. Phys., 175 (2002), pp. 487–524.

[5] F. A. Bornemann and P. Deuflhard, *The cascadic multigrid method for elliptic problems*, Numer. Math., 75 (1996), pp. 135–152.

[6] A. Borzì, *On the convergence of the MG/OPT method*, PAMM, 5 (2005), pp. 735–736.

[7] A. Borzì and E. Decker, *Analysis of a leap-frog pseudospectral scheme for the Schroedinger equation*, J. Comput. Appl. Math., 193 (2006), pp. 65–88.

[8] A. Borzì, J. Salomon, and S. Volkwein, *Formulation and numerical solution of finite-level quantum optimal control problems*, J. Comput. Appl. Math., to appear.

[9] A. Borzì, 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.

[10] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.

[11] P. W. Brumer and M. Shapiro, *Principles of the Quantum Control of Molecular Processes*, Wiley-VCH, Berlin, 2003.

[12] A. G. Butkovskiy and Yu. I. Samoilenko, *Control of Quantum-Mechanical Processes and Systems*, Kluwer, Norwell, MA, 1990.

[13] E. Charron, M. Cirone, A. Negretti, J. Schmiedmayer, and T. Calarco, *Theoretical analysis of the implementation of a quantum phase gate with neutral atoms on atom chips*, Phys. Rev. A, 74 (2006), 012308.

[14] Y. H. Dai and Y. Yuan, *A nonlinear conjugate gradient method with a strong global convergence property*, SIAM J. Optim., 10 (1999), pp. 177–182.

[15] F. Dalfovo, G. Stefano, L. P. Pitaevskii, and S. Stringari, *Theory of Bose–Einstein condensation in trapped gases*, Rev. Modern Phys., 71 (1999), pp. 463–512.

[16] R. Folman, P. Krüger, D. Cassettari, B. Hessmo, T. Maier, and J. Schmiedmayer, *Controlling cold atoms using nanofabricated surfaces: Atom chips*, Phys. Rev. Lett., 84 (2000), 4749.

[17] R. Folman, P. Krüger, D. Cassettari, J. Schmiedmayer, J. Denschlag, and C. Henkel, *Microscopic atom optics: From wires to an atom chip*, Adv. in Atom. Mol. and Opt. Phys., 48 (2002), 263.
[18] J. C. Gilbert and J. Nocedal, *Global convergence properties of conjugate gradient methods for optimization*, SIAM J. Optim., 2 (1992), pp. 21–42.
[19] W. Hänsel, J. Reichel, P. Hommelhoff, and T. W. Hänsch, *Trapped-atom interferometer in a magnetic microtrap*, Phys. Rev. A, 64 (2001), 063607.
[20] W. Hänsel, P. Hommelhoff, T. W. Hänsch, and J. Reichel, Nature, 413 (2001), pp. 498–501.
[21] W. Hackbusch and A. Reusken, *Analysis of a damped nonlinear multilevel method*, Numer. Math., 55 (1989), pp. 225–246.
[22] U. Hohenester, *Optical properties of semiconductor nanostructures: Decoherence versus quantum control*, in Handbook of Theoretical and Computational Nanotechnology, M. Rieth and W. Schommers, eds., American Scientific Publishers, Valencia, CA, 2005.
[23] U. Hohenester, P.-K. Rekdal, A. Borzì, and J. Schmiedmayer, *Optimal quantum control of Bose Einstein condensates in magnetic microtraps*, Phys. Rev. A, 75 (2007), 023602.
[24] I. Lesanovsky, T. Schumm, S. Hofferberth, L. M. Andersson, P. Krüger, and J. Schmiedmayer, *Adiabatic radio-frequency potentials for the coherent manipulation of matter waves*, Phys. Rev. A, 73 (2006), 033619.
[25] R. M. Lewis and S. G. Nash, *Model problems for the multigrid optimization of systems governed by differential equations*, SIAM J. Sci. Comput., 26 (2005), pp. 1811–1837.
[26] Y. Maday, J. Salomon, and G. Turinici, *Monotonic time-discretized schemes in quantum control*, Numer. Math., 103 (2006), pp. 323–338.
[27] S. G. Nash, *A multigrid approach to discretized optimization problems*, Optim. Methods Softw., 14 (2000), pp. 99–116.
[28] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, New York, 1999.
[29] J. P. Palao and R. Kosloff, *Quantum computing by an optimal control algorithm for unitary transformations*, Phys. Rev. Lett., 89 (2002), 188301.
[30] T. Paul, K. Richter, and P. Schlagheck, *Nonlinear resonant transport of Bose–Einstein condensates*, Phys. Rev. Lett., 94 (2005), 020404.
[31] A. P. Peirce, M. A. Dahleh, and H. Rabitz, *Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications*, Phys. Rev. A, 37 (1988), pp. 4950–4964.
[32] H. Rabitz, G. Turinici, and E. Brown, *Control of molecular motion: Concepts, procedures, and future prospects*, in Handb. Numer. Anal. 10, P. Ciarlet and J. Lions, eds., Elsevier, Amsterdam, 2003.
[33] A. Sacchetti, *Nonlinear time-dependent one-dimensional Schrödinger equation with double-well potential*, SIAM J. Math. Anal., 35 (2004), pp. 1160–1176.
[34] V. V. Shaidurov, *Some estimates of the rate of convergence of the cascadic conjugate gradient method*, Comput. Math. Appl., 31 (1996), pp. 161–171.
[35] D. F. Shanno, *Conjugate gradient methods with inexact searches*, Math. Oper. Res., 3 (1978), pp. 244–256.
[36] T. Schumm, S. Hofferberth, L.M. Andersson, S. Wildermuth, S. Groth, I. Bar–Joseph, J. Schmiedmayer, and P. Krüger, *Matter–wave interferometry in a double well on an atom chip*, Nature Phys., 1 (2005), pp. 57–62.
[37] S. H. Tersigni, P. Gaspard, and S. A. Rice, *On using shaped light pulses to control the selectivity of product formation in a chemical reaction: An application to a multiple level system*, J. Chem. Phys., 93 (1990), pp. 1670–1680.
[38] U. Trottenberg, C. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, London, 2001.
[39] S. Wildermuth, S. Hofferberth, I. Lesanovsky, E. Haller, L. Mauritz–Andersson, S. Groth, I. Bar–Joseph, P. Krüger, and J. Schmiedmayer, *Microscopic magnetic-field imaging*, Nature, 435 (2005), p. 440.
[40] J. Fortagh and C. Zimmermann, *Magnetic microtraps for ultracold atoms*, Rev. Modern Phys., 79 (2007), pp. 235–289.