Loading a Bose-Einstein condensate onto an optical lattice: An application of optimal control theory to the nonlinear Schrödinger equation

Shlomo E. Sklarz and David J. Tannor*

Department of Chemical Physics, Weizmann Institute of Science, 76100 Rehovot, Israel
(Received 15 May 2002; published 26 November 2002)

Using a set of general methods developed by Krotov [A. I. Konnov and V. A. Krotov, Automation and Remote Control **60**, 1427 (1999)], we extend the capabilities of optimal control theory to the nonlinear Schrödinger equation (NLSE). The paper begins with a general review of the Krotov approach to optimization. Although the linearized version of the method is sufficient for the linear Schrödinger equation, the full flexibility of the general method is required for treatment of the nonlinear Schrödinger equation. Formal equations for the optimization of the NLSE, as well as a concrete algorithm are presented. As an illustration, we consider a Bose-Einstein condensate (BEC) initially at rest in a harmonic trap. A phase develops across the BEC when an optical lattice potential is turned on. The goal is to counter this effect and keep the phase flat by adjusting the trap strength. The problem is formulated in the language of optimal control theory (OCT) and solved using the above methodology. To our knowledge, this is the first rigorous application of OCT to the nonlinear Schrödinger equation, a capability that is bound to have numerous other applications.

I. INTRODUCTION

In recent years much progress has been made in the use of optimal control theory (OCT), to coherently control quantum-mechanical systems governed by the Schrödinger equation. Such systems include controlled manipulations of molecular wave packets [1–3], high harmonic generation [4], and realization of quantum computing algorithms [5].

In this paper, the application of OCT is extended in a systematic way to systems governed by the nonlinear Schrödinger equation (NLSE), such as solitons in fiber optics and Bose-Einstein condensates (BEC's) in atomic physics. We begin with a general description of the Krotov iterative method [6]. We describe first its application to quantum systems governed by the linear Schrödinger equation, and then show how a generalized version of this method [7] can be used to treat nonlinear problems.

Finally, we consider a concrete problem governed by the NLSE, namely a BEC evolving under the Gross-Pitaevskii equation. The use of a BEC as a realization of quantum computing is widely being considered, as this is a macroscopic entity which nevertheless behaves quantum mechanically. The fact that a BEC carries a definite phase that can be manipulated and controlled is a striking manifestation of this quality. For many computation applications, it is desirable to split the BEC up into localized pieces each of which can then be viewed as a quantum bit and manipulated as such. This is achieved by the switching on of an optical lattice potential; however, the switching on of the optical lattice causes a phase to accumulate across the BEC, which is undesirable for use in computing applications. It is the cancellation of this effect which is the goal of this model problem.

The outline of the paper is as follows: In Sec. II, the Krotov method is reviewed. Its application to linear problems, in general, and the Schrödinger equation, in particular, are discussed. Section III deals with the application of OCT to the NLSE problems and demonstrates this by solving the BEC problem mentioned above. Finally, in Sec. IV, we conclude and suggest further applications.

II. KROTOV METHOD OF OPTIMIZATION

A. Description of problem

Consider a state of some system which can be defined by a vector of variables ψ and which is controlled by a set of variables u, through the state equations of motion

$$\dot{\psi} = f(\psi, u). \tag{1}$$

The initial value of ψ, $\psi(0) = \psi_0$, is fixed but evolves over time to some final value $\psi(T) = \psi_T$. The history of evolving state vectors is called the state trajectory, and the history of control input is termed the control history or just the control [8].

Given the trajectory $\psi(t)$ and the control $u(t)$, we define a "process" $w = (\psi(t), u(t))$ as a pair of histories $\psi(t)$, $u(t)$ satisfying Eq. (1). We can now define the objective functional on the process w:

$$J[w] = F(\psi(T)) + \int_0^T f^0(t, \psi(t), u(t)) dt, \tag{2}$$

where $F(\psi(T))$ and $f^0(t, \psi(t), u(t))$ are general functions that represent the dependence of J on the terminal and intermediate time values of ψ, respectively. It is required to find a process w for which the objective obtains its smallest value.

*Fax: 972-8-9344123

B. Utility constructs and definitions

For a continuously differentiable scalar function $\phi(t,\psi)$, we define the following functional:

$$L[w;\phi] = G(\psi_T) - \int_0^T R(t,\psi(t),u(t))dt - \phi(0,\psi_0), \quad (3)$$

where

$$G(\psi_T) = F(\psi_T) + \phi(T,\psi_T), \quad (4)$$

$$R(t,\psi,u) = \frac{\partial \phi}{\partial \psi} f(t,\psi,u) - f^0(t,\psi,u) + \frac{\partial}{\partial t}\phi(t,\psi). \quad (5)$$

The functions R and G are designed to separate out the dependence of $L[w;\phi]$ on the final time and intermediate time, respectively.

It can be shown that for any function ϕ and process $w = (\psi(t),u(t))$, $L[w;\phi] = J[w]$. The derivation goes as follows:

$$L[w;\phi] = G(\psi_T) - \int_0^T R(t,\psi(t),u(t))dt - \phi(0,\psi_0) = F(\psi_T) + \phi(T,\psi_T) - \int_0^T \left\{ \frac{\partial \phi}{\partial \psi} f(t,\psi,u) - f^0(t,\psi,u) + \frac{\partial \phi}{\partial t} \right\} dt - \phi(0,\psi_0)$$

$$= F(\psi_T) + \phi(T,\psi_T) - \int_0^T \left\{ \frac{\partial \phi}{\partial \psi}\frac{\partial \psi}{\partial t} + \frac{\partial \phi}{\partial t} - f^0(t,\psi,u) \right\} dt - \phi(0,\psi_0)$$

$$= F(\psi_T) + \phi(T,\psi_T) - \int_0^T \frac{d\phi}{dt} dt + \int_0^T f^0(t,\psi,u)dt - \phi(0,\psi_0) = F(\psi_T) + \int_0^T f^0(t,\psi,u)dt = J[w]. \quad (6)$$

Obviously therefore, minimizing $L[w;\phi]$ for any ϕ minimizes $J[w]$, and minimizing $L[w;\phi]$ can be achieved by separately minimizing $G(\psi_T)$ and maximizing $R(t,\psi,u)$.

It is convenient for later reference to define the function H through the following relation:

$$R(t,\psi,u) \equiv H\left(t,\psi,u,\frac{\partial \phi(t,\psi)}{\partial \psi}\right) + \frac{\partial}{\partial t}\phi(t,\psi), \quad (7)$$

where

$$H(t,\psi,u,p) = pf(t,\psi,u) - f^0(t,\psi,u). \quad (8)$$

Note the extra parameter in H denoted by p, which emphasizes that ψ and $\partial \phi / \partial \psi$ should be treated as independent variables, with respect to H.

C. An iterative algorithm

We now return to our main goal and describe the Krotov iterative method for finding a sequence of processes $\{w_s\}$, which monotonically decrease the value of the objective $J[w]$ [6,9]. The central idea is that as we have complete freedom in choosing the potential $\phi(\psi,t)$, we can construct ϕ such that our current estimate of the state history will *maximize* $L[w;\phi]$, and so become the worst of all possible histories. We are then free to find a new estimate for the control $u(t)$, which will minimize $L[w;\phi]$ with respect to its explicit dependence on $u(t)$, without worrying about the effect of $u(t)$ on $L[w;\phi]$ through the change of $\psi(t)$, as that can only be improved.

We begin by taking an arbitrary control history $u^0(t)$ and the corresponding state trajectory $\psi^0(t)$, which constitute together a process w^0.

(1) We first construct a function $\phi(t,\psi)$ such that $L[w;\phi]$ is a maximum with respect to $\psi(t)$ at the point w^0. This is equivalent to the following two conditions:

$$R(t,\psi^0(t),u^0(t)) = \min_{\psi} R(t,\psi(t),u^0(t)) \quad (9)$$

$$G(\psi_T^0) = \max_{\psi} G(\psi_T), \quad (10)$$

where the functions R and G are calculated using the new $\phi(t,\psi)$. In other words, we choose $\phi(t,\psi)$ such that our current $\psi^0(t)$ will be the worst of all possible $\psi(t)$'s in minimizing the objective $L[w;\phi] = J[w]$ (maximizing R, minimizing G). Any change in ψ brought about by a new choice of $u(t)$ will now only improve the minimization of $J[w]$ [see Fig. 1].

(2) For $\phi(t,\psi(t))$, we find a control $u(t)$ that maximizes $H(t,\psi,u,(\partial \phi/\partial \psi))$ and denote it by
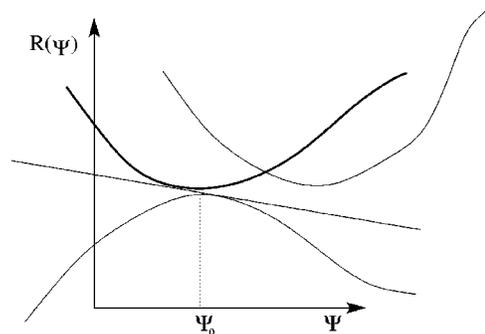
FIG. 1. Sketch of various $R(\psi)$ determined by different choices of $\phi(t,\psi)$. The construction of $\phi(t,\psi)$ is the one that minimizes $R(\psi)$ at $\psi = \psi_0$, i.e., is the worst R at the current ψ_0.

$$\tilde{u}(t,\psi)=\arg\max_u H\left(t,\psi,u,\frac{\partial\phi}{\partial\psi}\right)$$

$$=\arg\max_u R(t,\psi,u). \tag{11}$$

Note that the control $\tilde{u}(t,\psi)$ is still a function of ψ. This freedom will be removed in the next step.

(3) We require that $\tilde{u}(t,\psi)$ and $\psi(t)$ be consistent with each other through the equations of motion. The equation of motion (1) (with its initial conditions) together with the equation for the control $u=\tilde{u}(t,\psi)$ (11), provide two equations for the two unknowns u and ψ. These equations may be solved self-consistently for u and $\psi(t)$, obtaining the new process $w=(u,\psi)$.

(4) It is now guaranteed that minimization of the objective has been improved so that $J[w]<J[w^0]$; this completes the current iteration. The new w becomes a starting point for the next iteration, $w\rightarrow w^0$, and operations 1–3 can now be repeated to achieve further decrease in the objective.

We proceed to prove that indeed the new $J[w]\leqslant J[w^0]$. First, we assert using Eq. (6) that

$$J[w^0]-J[w]=L[w^0;\phi]-L[w;\phi]. \tag{12}$$

Also,

$$L[w^0;\phi]-L[w;\phi]=G(\psi_T^0)-G(\psi_T)+\int_0^T\{R(t,\psi(t),u(t))$$

$$-R(t,\psi^0(t),u^0(t))\}dt$$

$$=\Delta_1+\Delta_2+\Delta_3, \tag{13}$$

where

$$\Delta_1=G(\psi_T^0)-G(\psi_T), \tag{14}$$

$$\Delta_2=\int_0^T\{R(t,\psi(t),u(t))-R(t,\psi(t),u^0(t))\}dt, \tag{15}$$

$$\Delta_3=\int_0^T\{R(t,\psi(t),u^0(t))-R(t,\psi^0(t),u^0(t))\}dt. \tag{16}$$

The non-negativeness of Δ_3 and Δ_1 follow from conditions (9) and (10), respectively, and by Eq. (11) the choice of a new control ensures the non-negativeness of Δ_2. This completes the proof.

D. Construction of ϕ to first order in ψ

In implementing the above iterative method the main difficulty lies in step (1). Here it is necessary to determine a function $\phi(t,\psi)$ that, by conditions (9) and (10), will ensure the absolute maximum and minimum of the functions R and G, respectively, on the trajectory $\psi_0(t)$, i.e., to choose $\phi(t,\psi)$ to give the worst possible $L[\psi_0;\phi]$. A necessary condition for an extremum of R and G at $w^0=(\psi^0,u^0)$ is the existence of a stationary point there, but in order to make the conditions sufficient, it is necessary to add conditions of positivity and negativity on the second derivatives of R and G, respectively.

We leave the additional requirements on the second derivatives for a later section, and restrict ourselves in this section solely to determining the conditions for a stationary point in R and G, which are as follows: [10]

$$\frac{\partial}{\partial\psi}R(t,\psi^0,u^0)=\frac{\partial^2\phi(t,\psi^0)}{\partial\psi^2}f(t,\psi^0,u^0)+\frac{\partial\phi}{\partial\psi}\frac{\partial}{\partial\psi}f(t,\psi^0,u^0)-\frac{\partial}{\partial\psi}f^0(t,\psi^0,u^0)+\frac{\partial}{\partial t}\frac{\partial\phi(t,\psi^0)}{\partial\psi}$$

$$=\frac{\partial}{\partial\psi}H(t,\psi^0,u^0,\chi)+\frac{\partial^2\phi(t,\psi^0)}{\partial\psi^2}f(t,\psi^0,u^0)+\frac{\partial}{\partial t}\frac{\partial\phi(t,\psi^0)}{\partial\psi}$$

$$=\frac{\partial}{\partial\psi}H(t,\psi^0,u^0,\chi)+\left(\frac{\partial\psi}{\partial t}\frac{\partial}{\partial\psi}+\frac{\partial}{\partial t}\right)\frac{\partial\phi(t,\psi^0)}{\partial\psi}$$

$$=\frac{\partial}{\partial\psi}H(t,\psi^0,u^0,\chi)+\frac{d\chi}{dt}=0, \tag{17}$$

$$\frac{\partial G(t,\psi_T^0,u^0)}{\partial\psi_T}=\frac{\partial F(\psi_T^0,u^0)}{\partial\psi_T}+\frac{\partial\phi(T,\psi_T^0)}{\partial\psi_T}=\frac{\partial F(\psi_T^0,u^0)}{\partial\psi_T}+\chi_T=0, \tag{18}$$

where [11]

$$\chi(t)=\frac{\partial}{\partial\psi}\phi(t,\psi^0(t)). \tag{19}$$

Equations (17) and (18) can be restated as an equation of motion for the vector χ, and the definition of $H(t,\psi,u,p)$, (8), can be used to rewrite the equation of motion for ψ, (1) in the following compact form:

$$\dot{\chi} = -\frac{\partial}{\partial \psi} H(t, \psi^0, u^0, \chi)$$

with boundary conditions $\chi(T) = -\dfrac{\partial F(\psi_T^0, u^0)}{\partial \psi_T}$,

$$\text{(20)}$$

$$\dot{\psi} = \frac{\partial}{\partial \chi} H(t, \psi, u^0, \chi) \quad \text{with boundary conditions} \quad \psi(0) = \psi_0.$$

$$\text{(21)}$$

These equations constitute a so-called Hamiltonian system with a Hamiltonian $H(t, \psi, u, \chi)$, and the variables ψ and χ are said to be conjugate. Equation (19) shows that the variable χ represents the function $\phi(t, \psi)$ to first order in ψ. Equations (20) and (21) give the prescription for constructing $\chi(t)$.

Creating the conjugate variable χ to fulfill the above requirements enforces a stationary point in R and G with respect to ψ. As explained above it is also necessary, in the general case, that the stationary point be an extremum and therefore that the second derivatives of G and R, with respect to ψ, are negative and positive, respectively. However, for problems linear in ψ it so happens that a stationary point is sufficient. This will be illustrated by a concrete example in the following section, after which we return to our main line of discussion completing the conditions for obtaining $\phi(t, \psi)$ in the general nonlinear case.

E. A linear problem and application to the linear Schrödinger equation optimization

Consider a problem where the equations of motion are linear in the state variable,

$$\dot{\psi} = a(u) \psi, \quad F(\psi_T) = b \psi_T, \quad f^0 = f^0(u). \quad \text{(22)}$$

We proceed to show that it is sufficient to choose $\phi(t, \psi) = \chi(t) \psi$ to achieve monotonic increase in the objective at each iteration. The Hamiltonian for this problem by the definition (8) is $H(t, \psi, u, \chi) = \chi a(u) \psi - f^0(u)$, so we get by applying Eq. (20)

$$\dot{\chi} = -a(u^0) \chi \quad \text{with boundary conditions} \quad \chi(T) = -b. \quad \text{(23)}$$

Using the above we find that

$$R(t, \psi, u^0) = \chi a(u^0) \psi - f^0 + \frac{\partial \chi}{\partial t} \psi = \left(\chi a(u^0) + \frac{\partial \chi}{\partial t} \right) \psi - f^0$$

$$= -f^0(u^0), \quad \text{(24)}$$

$$G(\psi_T) = b \psi_T + \chi_T \psi_T = (b + \chi_T) \psi_T = 0, \quad \text{(25)}$$

which are independent of ψ. This implies that both Δ_1 and Δ_3 in Eqs. (14) and (16) vanish. By maximizing Δ_2 [Eq. (15)], the objective is guaranteed to decrease at each iteration. Note that in the linear case there is no need to check the second derivatives of G and R since, as R is linear in ψ and

we set $(\partial R / \partial \psi)|_{\psi^0} = 0$, R must be independent of ψ. (see Fig. 1). Therefore the control u can be made to maximize R without the resulting change in $\psi(t)$ having any effect on the objective.

The above example encompasses the problem of optimization of a quantum-mechanical wave function governed by the linear Schrödinger equation

$$|\dot{\psi}\rangle = -i\hat{H}(u)|\psi\rangle. \quad \text{(26)}$$

Some care is necessary, however, if the objective takes the form $F(\psi_T) = -\langle \psi_T | P | \psi_T \rangle$, which is not strictly linear in the state vector as in Eq. (22). Another complication arises from the fact that the state vector ψ is an element of a complex Hilbert space. We therefore work this problem out in full, and show that nevertheless the above choice of ϕ is sufficient in these problems just as in the linear example, due to the fact that the target projection operator P is positive definite [12,13].

As above, we set $\phi(t, \psi) = \langle \chi | \psi \rangle + \langle \psi | \chi \rangle$ (where we have included the complex conjugate as an extra independent state variable), and thus get for the Hamiltonian of the problem

$$H(t, \psi, u, \chi) = -i\langle \chi | \hat{H} | \psi \rangle + i\langle \psi | \hat{H}^\dagger | \chi \rangle - \lambda f^0(u)$$

$$= 2\operatorname{Im}\langle \chi | \hat{H}(u) | \psi \rangle - \lambda f^0(u), \quad \text{(27)}$$

which using Eq. (20) yields for the conjugate vector

$$|\dot{\chi}\rangle = -i\hat{H}^\dagger(u^0)|\chi\rangle$$

with boundary conditions $|\chi_T\rangle = P|\psi_T^0\rangle. \quad \text{(28)}$

Reinserting this equation and its complex conjugate into the formulas for R and G, we have

$$R(t, \psi, u^0) = -i\langle \chi | \hat{H}(u^0) | \psi \rangle + i\langle \psi | \hat{H}^\dagger(u^0) | \chi \rangle - \lambda f^0(u^0)$$

$$+ \langle \dot{\chi} | \psi \rangle + \langle \psi | \dot{\chi} \rangle$$

$$= -\lambda f^0(u^0), \quad \text{(29)}$$

$$G(\psi_T) = -\langle \psi_T | P | \psi_T \rangle + \langle \chi_T | \psi_T \rangle + \langle \psi_T | \chi_T \rangle$$

$$= -\langle \psi_T | P | \psi_T \rangle + \langle \psi_T^0 | P | \psi_T \rangle + \langle \psi_T | P | \psi_T^0 \rangle. \quad \text{(30)}$$

R is independent of ψ as above, which guarantees that Δ_3 of Eq. (16) vanishes. G is dependent on ψ_T; however (denoting $\Delta\psi = \psi - \psi^0$) the positiveness of $\Delta_1 = \langle \Delta\psi_T | P | \Delta\psi_T \rangle$ [Eq. (14)] is always guaranteed due to the positiveness of P. Alternatively, note that the second derivative of G, $\partial^2 G / \partial \psi_T \partial \psi_T^* = -P < 0$ is always negative due to the positiveness of P, assuring that the condition for a maximum of $G(\psi_0)$ is automatically met [14].

Another intricate point regarding these problems, which is often missed, is the following. In many problems in quantum mechanics it so happens that the equations of motion are linear in the control variable u, namely the field. This means that strictly speaking there is no proper maximum in the Hamiltonian of the control system $H(t, \psi, u, \chi)$ with respect to u, and stage (2) of the algorithm [Eq. (11)] cannot be properly fulfilled. This problem is often overcome by adding a penalty function $\lambda f^0(u)$ quadratic in u to the objective as

implicitly indicated above. The physical interpretation of this construction is that placing a penalty on the fluence of the field constrains the algorithm to search out the optimal *direction* of u rather than minimizing the objective by varying its magnitude. The price paid by this solution, however, is that the algorithm often exerts much effort into minimizing the superficial penalty part of the objective at the expense of the really required terminal part.

An alternative way to overcome this problem is by noticing that the algorithm does not really require that the Hamiltonian be maximized by u at each iteration. All that is really required is that the Hamiltonian be increased by the new choice of u, which is enough to ensure that Δ_2 [Eq. (10)] be non-negative. The penalty function can therefore be dropped and at each iteration u should be increased $u \rightarrow u\lambda^{-1}\,\partial H/\partial u$, where λ^{-1} is some *macroscopic* constant which can be chosen arbitrarily [15]. This is not to be confused, despite the formal similarity, with the gradient methods where the correction to u must always be small such that its effect on any change in ψ will remain in the linear regime.

F. Construction of ϕ to second order in ψ

As noted above, for an equation of motion nonlinear in the state variable, it is necessary to fulfill conditions on the second derivatives of R and G and therefore ϕ must be chosen to contain higher orders in ψ. We therefore take

$$\phi(t,\psi)=\chi_i\psi_i+\frac{1}{2}\Delta\psi_i^*\sigma_{ij}(t)\Delta\psi_j, \qquad (31)$$

where $\Delta\psi_i=\psi_i-\psi_i^0$, and the functions σ_{ij} are to be determined such as to obtain the required extrema in R and G. The conditions supplementary to Eqs. (17)–(18), necessary for fulfilling Eqs. (9)–(10), are the following system of differential inequalities:

$$d^2R=\Delta\psi_i^*\frac{\partial^2 R(t,\psi^0,u^0)}{\partial\psi_i^*\partial\psi_j}\Delta\psi_j, \quad d^2R\geqslant 0, \qquad (32)$$

$$d^2G=\Delta\psi_{Ti}^*\frac{\partial^2 G(\psi_T^0)}{\partial\psi_{Ti}^*\partial\psi_{Tj}}\Delta\psi_{Tj}, \quad d^2G\leqslant 0. \qquad (33)$$

For the positivity and negativity of the quadratic forms d^2R and d^2G, respectively, it suffices to set

$$\frac{\partial^2}{\partial\psi_i^*\partial\psi_j}R(t,\psi^0,u^0)=\begin{cases}0 & \text{for any } i\neq j \\ \delta_i\geqslant 0 & \text{for all } i=j\end{cases}, \qquad (34)$$

$$\frac{\partial^2}{\partial\psi_{Ti}^*\partial\psi_{Tj}}G(\psi_T^0)=\begin{cases}0 & \text{for any } i\neq j \\ -\alpha_i\leqslant 0 & \text{for all } i=j,z,\end{cases} \qquad (35)$$

where α_i and δ_i are some non-negative functions. Inserting the full form of R and G into the above equations yields a set of $n(n+1)/2$ equations of motion for the functions σ_{ij}, where n is the dimension of the state vector. This means that the dimension of the system grows proportionally to the

square of n, and for large n the expense of solving this system normally becomes too high.

In Ref. [7], it is proved that for certain classes of functionals conditions (9),(10) can be fulfilled by taking ϕ according to Eq. (31) with

$$\sigma_{ij}(t)=\begin{cases}\alpha(e^{\gamma(T-t)}-1)+\beta & \text{for } i=j \\ 0 & \text{for } i\neq j,\end{cases} \qquad (36)$$

where $\alpha,\beta<0$ and $\gamma>0$. Taking $\beta\leqslant-\partial^2 F(\psi_T^0)/\partial\psi_T^*\partial\psi_T$ always fulfills condition (35) and it can be shown that as $\gamma,|\alpha|\rightarrow\infty$, for these classes of functionals, condition (34) is also fulfilled. The strategy then is to begin with $\sigma=0$ and if the objective does not decrease, take increasingly larger $\gamma,|\alpha|,|\beta|$ until a decrease in the objective is achieved.

III. APPLICATION TO THE NLSE

A. General formulation

We wish to apply this algorithm to optimizing a quantum system governed by the nonlinear Schrödinger equation

$$|\dot\psi\rangle=-i\hat H_{NL}(|\psi|^2)|\psi\rangle=-i(\hat H+\mu|\psi|^2)|\psi\rangle, \qquad (37)$$

where $\hat H=\hat K+\hat V$ is the usual linear Hamiltonian operator consisting of kinetic and potential parts, and μ is the coefficient of the additional nonlinear term. The objective is defined as for the linear case, as minimizing

$$J=-\langle\psi_T|\hat P|\psi_T\rangle+\lambda\int dt f^0(u). \qquad (38)$$

In realization of step 1 of the iterative method, we choose $\phi(t,\psi)=\langle\chi|\psi\rangle+\langle\psi|\chi\rangle+\frac{1}{2}\langle\Delta\psi|\sigma|\Delta\psi\rangle$ and find the Hamiltonian of the system to be

$$H(t,\psi,u,\chi)=-i\langle\chi|\hat H_{NL}|\psi\rangle+i\langle\psi|\hat H_{NL}^\dagger|\chi\rangle-\lambda f^0(u)$$

$$=2\text{Im}\langle\chi|\hat H_{NL}|\psi\rangle-\lambda f^0(u), \qquad (39)$$

just as in the linear case. However it must be remembered that here $\hat H_{NL}$ depends on ψ and ψ^*, and therefore using Eq. (20) yields for the conjugate vector

$$|\dot\chi\rangle=-\frac{\partial}{\partial\psi^*}H(t,\psi^0,u^0,\chi)$$

$$=i\mu(\psi^0)^2|\chi^*\rangle-i(\hat H^\dagger+2\mu|\psi^0|^2)|\chi\rangle. \qquad (40)$$

Note that this equation differs in two respects from its linear counterpart [Eq. 28]. First, it involves $|\psi^0\rangle$ which does not cause special problems except that the vector $|\psi^0\rangle$ must be stored and used in propagating $|\chi\rangle$. The more cumbersome

difficulty arises from the fact that the equation obtained is no longer linear in $|\chi\rangle$, but rather contains an extra term linear in $|\chi^*\rangle (=\langle\chi|)$. Nevertheless the coupled equations

$$\frac{\partial}{\partial t}\begin{pmatrix} |\chi\rangle \\ |\chi^*\rangle \end{pmatrix} = -i\left\{ \begin{pmatrix} \hat{K} & 0 \\ 0 & -\hat{K} \end{pmatrix} + \begin{pmatrix} a & -b \\ b^* & -a \end{pmatrix} \right\} \begin{pmatrix} |\chi\rangle \\ |\chi^*\rangle \end{pmatrix},$$

(41)

where $\hat{K} = -(\hbar^2/2m)(\partial^2/\partial x^2)$, $a = \hat{V}(x) + 2\mu|\psi^0|^2$, and $b = \mu(\psi^0)^2$, or in matrix form

$$\tilde{\vec{\chi}} = -i(K+V)\vec{\chi}$$

(42)

can be solved using the split operator method for separating the treatment of the kinetic and potential parts of the Hamiltonian: $\vec{\chi}(t+\delta t) = e^{-iK\delta t/2} e^{-iV\delta t} e^{-iK\delta t/2} \vec{\chi}(t)$. The kinetic evolution can be computed using the usual Fourier-transform methods and the potential part can be evaluated by diagonalization to give

$$e^{-iV\delta t} = \frac{1}{D}\begin{pmatrix} \dfrac{|b|^2}{a-D}\cos(D\delta t) - ae^{iD\delta t} & ib\sin(D\delta t) \\[2ex] -ib^*\sin(D\delta t) & \dfrac{|b|^2}{a-D}\cos(D\delta t) - ae^{-iD\delta t} \end{pmatrix},$$

(43)

with $D = \sqrt{a^2 - |b|^2}$. Finally, introducing the complex-conjugation operator $\hat{C}$ and the Fourier-transform operator $\hat{Z}$, this procedure yields the following formula for the numerical propagation step:

$$|\chi(t+\delta t)\rangle = \hat{Z}^\dagger e^{i(\hbar^2 k^2/4m)\delta t}\hat{Z}\left\{\frac{1}{D}\left(\frac{|b|^2}{a-D}\cos(D\delta t) - a\exp(iD\delta t)\right) + i\frac{b}{D}\sin(D\delta t)\hat{C}\right\}\hat{Z}^\dagger e^{i(\hbar^2 k^2/4m)\delta t}\hat{Z}|\chi(t)\rangle.$$

(44)

Having obtained $|\chi(t)\rangle$ for all t, we now proceed to realize step (2) of the algorithm and, according to Eq. (11), find for each point in time a field that maximizes

$$H\left(t,\psi,u,\frac{\partial\phi}{\partial\psi}\right) = 2\text{Im}\langle\chi + \tfrac{1}{2}\sigma\Delta\psi.\hat{H}(u)|\psi\rangle - \lambda f^0(u).$$

Or mathematically,

$$\tilde{u}(t,\psi) = \arg\max_u H\left(t,\psi,u,\frac{\partial\phi}{\partial\psi}\right).$$

(45)

Step (3) of the algorithm is fulfilled simultaneously with step (2), by simultaneously propagating u and ψ such that each new $u(t)$ is used directly in propagating $\psi(t) \rightarrow \psi(t+\delta t)$. For σ, we use formula (36), and according to the algorithm described in the preceding section we begin with $\sigma = 0$ and increase the parameters $\gamma, |\alpha|, |\beta|$ until we achieve decrease of objective.

B. Application to a concrete problem

We consider a one-dimensional (1D) BEC confined by a harmonic trap and governed by the Gross-Pitaevskii equation

$$|\dot{\psi}\rangle = -i(\hat{K} + \hat{V} + NU_0|\psi|^2)|\psi\rangle,$$

(46)

where $\hat{K}$, $\hat{V}$ are as above, and NU_0 is the nonlinear atom-atom interaction strength, N being the number of atoms. The BEC is initially in the ground state of the trap potential, and is therefore stationary. An optical lattice is then switched on, having the effect of separating the BEC wave packet into a series of localized pieces. The potential-energy operator therefore takes the form $\hat{V} = Kx^2 + S(t)V_0\cos^2(kx)$, where K is the trap constant, k is the laser field wave number, V_0 is the lattice intensity, and the switching-on function of the field is denoted $S(t)$. In applications to quantum computing, these localized wave packets are to represent quantum bits. However, due to the nonlinearity of the equations, the condensate develops a phase that varies from one lattice site to the other (see Fig. 2), which is undesirable for quantum computing, since these algorithms assume that there is zero relative phase among the various single quantum bits. The problem is therefore to eliminate this phase profile by adjusting the trap strength during the switching on of the laser field. From the OCT perspective, the trap constant $K(t)$ is taken as the control and the objective is to minimize the variance of the phase of the wave packet, $\theta(x)$, at some final time T. The phase being a multivalued function poses problems; we therefore consider, instead, minimizing the variance of $\cos(\theta) = \tfrac{1}{2}[(\psi + \psi^*)/|\psi|]$ such that the objective becomes

$$J = \langle\cos^2(\theta_T)\rangle - \langle\cos(\theta_T)\rangle^2$$
$$= \langle\psi|\cos^2(\theta_T)|\psi\rangle - \langle\psi|\cos(\theta_T)|\psi\rangle^2.$$

(47)

Using the first part of Eq. (20), we get the equations of motion for χ as in Eq. (40); and taking a derivative of J with respect to ψ_T^*, we get [according to Eq. (20)], the boundary conditions

$$\chi_T = -\frac{\partial J}{\partial\psi_T^*} = -\text{Re}[\psi_T] + \frac{1}{2}|\psi_T|\langle\cos(\theta_T)\rangle\left(\frac{\psi_T}{\psi_T^*} + 3\right).$$

(48)
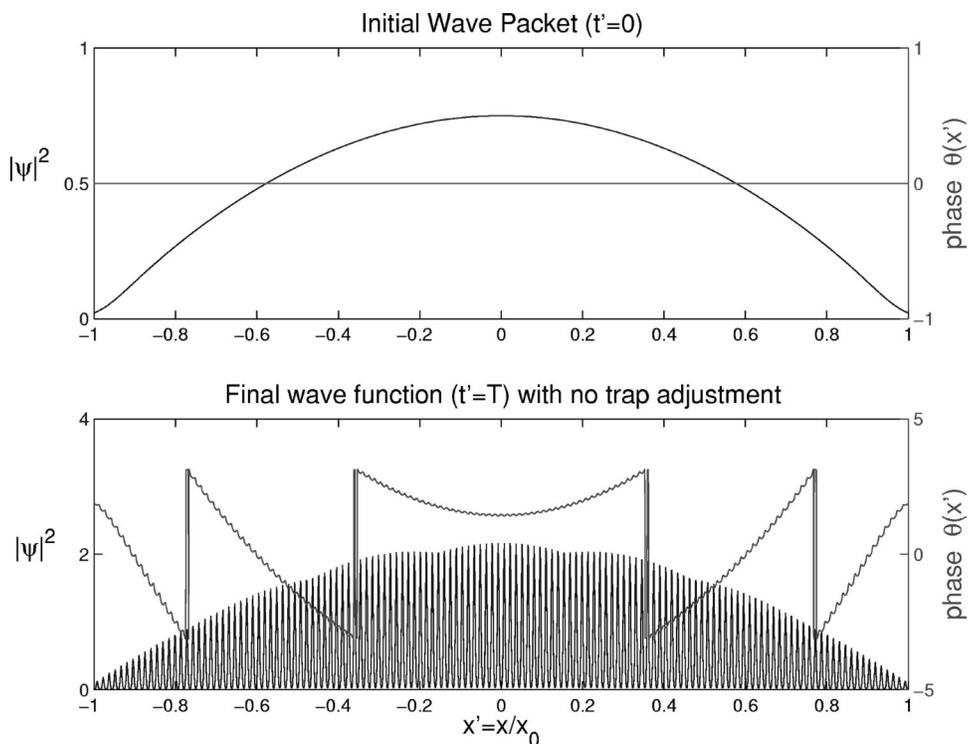
FIG. 2. Wave packet at $t=0$, and $t=T$ with no trap adjustments.
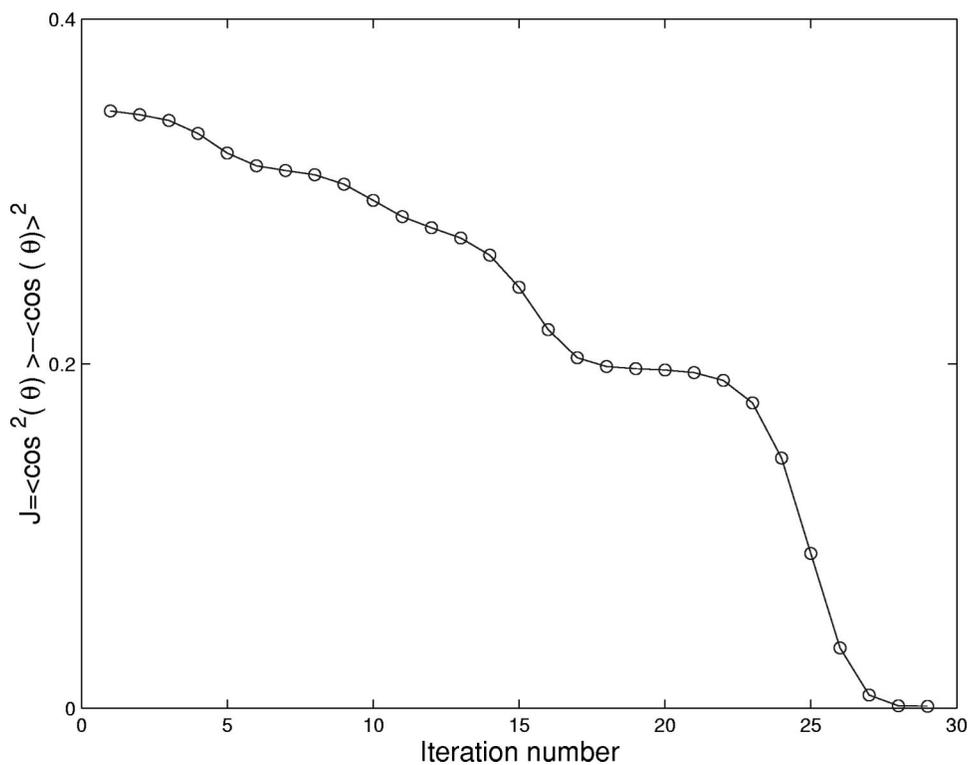
FIG. 3. Objective decrease as a function of iteration.
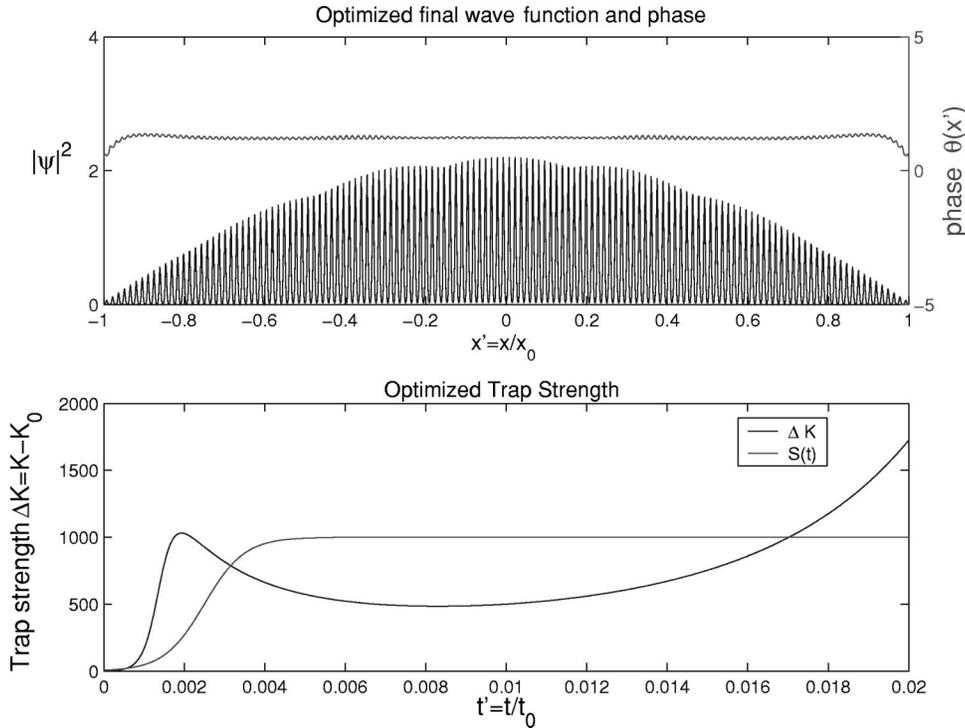
FIG. 4. Optimized trap strength evolution (bottom) and final wave packet at $t=T$ (top). The flat phase is strikingly apparent.

The Hamiltonian of the problem is $H(t,\psi,K,\chi)=K(t)\langle\chi+\frac{1}{2}\sigma(\Delta\psi)|x^2|\psi\rangle$, so that according to the above procedure we improve K at each iteration by $K\rightarrow K+\lambda^{-1}\langle\chi+\frac{1}{2}\sigma(\Delta\psi)|x^2|\psi\rangle$.

C. Optimization results

Following Ref. [16], we transform the NLSE to dimensionless units $t'=t/t_0$ and $x'=x/x_0$, where $x_0=x_{TF}=20.3$ μm and $t_0=mx_0^2/2\hbar=75$ ms. The Thomas-Fermi radius $x_{TF}=\sqrt{2\mu_{TF}/m\omega_{trap}^2}$ gives the size of the condensate, and is defined through a chemical potential $\mu_{TF}\equiv\hbar/t_{NL}$ determined by normalization of the wave function to unity. The wave function too is scaled $\psi\rightarrow\sqrt{x_{TF}}\psi$, and in order to keep the time scales of our 1D model comparable to the 3D reality, we adjust $U\rightarrow CU$ by a factor $C=\sqrt{\pi}/[\Gamma(2+1/2)]=\frac{4}{3}$ (Ref. [16]). We take $t_{NL}=96.2$ μs, optical wavelength $\lambda=589$ nm, and $V_0=10.94E_R$ for the final field intensity. All parameters were taken to resemble the experiments described in Ref. [16]. Performing these transformations, we end up finally with a dimensionless NLSE,

$$i|\dot{\psi}\rangle=\left(-\frac{1}{4}\frac{\partial^2}{\partial x^2}+K(t)x^2+S(t)V\cos^2(kx)+U|\psi|^2\right)|\psi\rangle,$$

(49)

where the trap constant $K=\omega_{trap}^2t_0^2$, the field intensity $V=V_0t_0/\hbar=2\times250^2$, and the nonlinear coefficient $U=\frac{4}{3}\mu_{TF}t_0/\hbar=1039$, such that all space, time, and energy quantities are now expressed in units of x_0, t_0, and $\hbar/t_0$, respectively.

Initially, the wave packet is in an eigenstate of the potential with trap constant $K_0=779$, and is therefore static. The

switching-on function plotted in Fig. 4 turns on the optical potential at a quarter of the optimizing interval ($T/4$), and is constant from there to the final time. With no adaption of the trap constant, a phase develops across the wave function, as shown in Fig. 2. The optimization process decreases the objective monotonically (as plotted in Fig. 3) and yields a strategy of increasing $\Delta K(t)=K(t)-K_0$ to achieve a flat phase at the final time $T=1500$ μs. These striking results are shown in Fig. 4.

IV. CONCLUSIONS

Using a set of general methods developed by Krotov, we have extended the capabilities of the optimal control theory to the NLSE. Although the linearized version of the method is sufficient for the linear Schrödinger equation, the full flexibility of the general method is required for a rigorous treatment of the nonlinear Schrödinger equation. Mention should be made of the interesting recent work of Hornung and de Vivie-Riedle [17], applying optimization techniques to molecule formations in a BEC, although the ϕ function in that work included linear terms only. A parallel approach was pursued by Pötting *et al.* [18] who used genetic algorithms to control the momentum state of a BEC. The significance of this paper is twofold. First, both formal equations and a concrete and efficient algorithm were presented for optimizing the NLSE in cases where the nonlinear terms are significant. Second, the methodology was applied successfully to an interesting physical problem confronting the use of trapped BECs for quantum computing, namely, producing a constant final phase profile across the condensate after an optical lattice is turned on. Further work on understanding analytically the mechanism found by OCT is still underway. We believe the working equations developed here will have many more

---

[1] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa, Science **288**, 824 (2000).

[2] R. J. Gordon and S. A. Rice, Annu. Rev. Phys. Chem. **48**, 601 (1997).

[3] S. A. Rice and M. Zhao, *Optical Control of Molecular Dynamics* (Wiley, New York, 2000).

[4] R. Bartels, S. Backus, E. Zeek, L. Misoguti, G. Vdovin, I.P. Christov, M.M. Murnane, and H.C. Kapteyn, Nature (London) **406**, 164 (2000).

[5] C. Rangan and P.H. Bucksbaum, Phys. Rev. A **3**, 3417 (2001).

[6] V.F. Krotov and I.N. Fel'dman, Engrg. Cybern. **17**, 123 (1983).

[7] A.I. Konnov and V.A. Krotov, Automation Remote Control **60**, 1427 (1999).

[8] Here and throughout the following, vector treatment of the appropriate variables is assumed, although all vector notations are omitted to avoid congestion. Multiplication of vector variables is therefore to be understood as a dot product. An indexed notation of vector components will be used only when unavoidable for clarity, and summation convention will then be implied.

[9] V. F. Krotov, *Global Methods in Optimal Control Theory* (Dekker, New York, 1996).

[10] In the derivation below, the following delicate point should be noted: $R(t,\psi,u)$ is, by definition, a function of three variables, whereas $H(t,\psi,u,p)$ is a function of an additional argument $p=\partial\phi/\partial\psi$. In other words, although p depends on ψ with respect to R, this is not the case with respect to H where ψ and p are to be taken as independent variables. In Eq. (17), we wish to vary R with respect to ψ, which means H must be varied in both ψ and p; $\partial R/\partial\psi \leftrightarrow \{(\partial p/\partial\psi)\partial/\partial p + \partial/\partial\psi\}H$.

[11] We stress that in the following definition, $\chi(t)$ is a function solely of t, and is obtained by inserting the explicit dependence of $\psi_0(t)$ into $\partial\phi(t,\psi)/\partial\psi$. This explains the advance to the third line in Eq. (17) where the total time derivative of $\partial\phi(t,\psi)/\partial\psi|_{\psi_0}$ is translated to a simple time derivative of $\chi(t)$.

[12] J. Somlói, V.A. Kazakov, and D.J. Tannor, Chem. Phys. **172**, 85 (1993).

[13] D. J. Tannor, V. Kazakov, and V. Orlov, in *Time Dependent Quantum Molecular Dynamics, Vol. 299 of NATO Advanced Study Institute, Series B: Physics*, edited by J. Broeckhove and L. Lathouwers (Plenum, New York, 1992).

[14] Another way to map the quantum control problems onto the linear example is by formulating the Schrödinger equation in the density matrix formalism of Liouville space, i.e., $|\dot\rho\rangle\rangle = i\mathcal{L}_h|\rho\rangle\rangle$ with the objective $F(\psi_T)=\langle\langle B\rho\rangle\rangle$, where B is some target state. In this form the equivalence to the linear case becomes self-evident.

[15] The choice of λ^{-1} is not restricted by the algorithm; however, some limit must exist through the physics and numerics of the problem. The λ can also be adjusted at will during the optimization to improve convergence. The choice of the symbol λ intentionally points to the similarity of the role this parameter plays with that of the strength of the penalty imposed above and to the similarity of the resulting forms for choice of the new u. The difference ends up being solely that the proposed method takes $\lambda^{-1}\partial H/\partial u$ to be a correction to the old field instead of an entirely new choice.

[16] M. Trippenbach, Y.B. Band, and P.S. Julienne, Phys. Rev. A **62**, 023608 (2000).

[17] T. Hornung and R. de Vivie-Riedle, (private communication).

[18] S. Pötting, M. Cramer, and P. Meystre, Phys Rev. A. **64**, 063613 (2001).