

A Reduction-Based Strategy for Optimal Control of Bose-Einstein Condensates

J. Adriazola and R. H. Goodman
New Jersey Institute of Technology
(Dated: November 22, 2021)

Applications of Bose-Einstein Condensates (BEC) often require that the condensate be prepared in a specific complex state. Optimal control is a reliable framework to prepare such a state while avoiding undesirable excitations, and, when applied to the time-dependent Gross-Pitaevskii Equation (GPE) model of BEC in multiple space dimensions, results in a large computational problem. We propose a control method based on first reducing the problem, using a Galerkin expansion, from a PDE to a low-dimensional Hamiltonian ODE system. We then apply a two-stage hybrid control strategy. At the first stage, we approximate the control using a second Galerkin-like method known as CRAB to derive a finite-dimensional nonlinear programming problem, which we solve with a differential evolution (DE) algorithm. This search method then yields a candidate local minimum which we further refine using a variant of gradient descent. This hybrid strategy allows us to greatly reduce excitations both in the reduced model and the full GPE system.

I. INTRODUCTION AND EXPERIMENTAL CONTEXT

Quantum optimal control is concerned with the control of N -body quantum systems [1, 2]. One important example is the reshaping of a dilute atomic Bose-Einstein condensate (BEC). Since they were observed in laboratory experiments in 1995 [3–5], BECs, an ultra-cold quantum fluid whose mean dynamics resemble that of a single atom [6], have proven to be an experimentally reliable and versatile platform for high-precision quantum metrology [7–9], high-precision storage, manipulation, and probing of interacting quantum fields [10, 11]. Future quantum computation and simulation technologies will likely require fast manipulation of BECs [12, 13].

Experimentalists over the last two decades have achieved remarkably high, yet sub optimal degrees of control of BECs by using empirical rules of thumb and intuition gained from significantly reduced models admitting closed-formed solutions [14]. Meanwhile, optimal control theory provides a computational framework for systematically finding highly efficient control policies [15–18]. The success of optimal control theory is demonstrated numerically in three spatial dimensions by Mennemann, et al. [19]. Our work interpolates between these two approaches by applying a general optimization strategy to simpler, ordinary differential equations (ODE) which are, in some sense, still faithful to the partial differential equations (PDE) that model BEC.

Mennemann, et al., apply optimal control to reshape the support of a BEC, reorienting the magnetic field concentrated along one axial direction to another, which in turn reorients the density distribution of the condensate. A two dimensional schematic is shown in Figure 1 with normalized Gaussian wavefunctions of the form

$$\psi_{\text{Gauss}} = \sqrt{\frac{ab}{\pi^2}} e^{-ax^2 - by^2}, \quad (1)$$

where $a, b > 0$ and $(x, y) \in \mathbb{R}^2$. Another manipulation Mennemann, et al., consider is to change the topology of the wavefunction's support. An example of this is

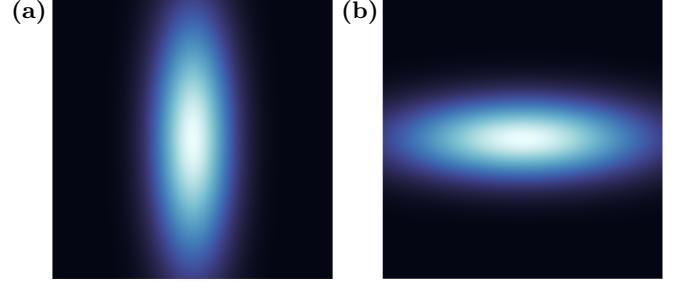


FIG. 1: Reshaping a BEC. (a) The density distributions $|\psi_{\text{Gauss}}|^2$ axially aligned along the vertical with $(a, b) = (10, 1)$ in Equation (1). (b) The density distribution axially aligned along the horizontal with $(a, b) = (1, 10)$.

shown in Figure 2 where a Gaussian wavefunction (1), with $a = b = 1$, is mapped to the normalized toroidal wavefunction

$$\psi_{\text{Toroid}} = \frac{2}{\sqrt{3\pi}} r^2 e^{-r^2}, \text{ where } r^2 = x^2 + y^2.$$

Manipulating the condensate excites oscillations that prevent the transformed distribution from matching the desired distribution after the control process has terminated. In this paper we use optimal control to perform these manipulations while minimizing such oscillations and mismatch.

The, now standard, optimal control problem, first proposed by Hohenester, et al. [20], is to maximize the fidelity between an evolving field ψ at a final time $T > 0$ and an experimentally-desired state ψ_d , subject to a control function u . This problem, expressed in dimensionless form, is

$$\inf_{u \in \mathcal{U}} J = \frac{1}{2} \inf_{u \in \mathcal{U}} [J^{\text{infidelity}}(u) + J^{\text{regular}}(u)], \quad (2)$$

where

$$J^{\text{infidelity}}(u) = 1 - |\langle \psi_d(x), \psi(x, T) \rangle|_{L^2(\mathbb{R}^3)}^2 \quad \text{and} \\ J^{\text{regular}}(u) = \gamma \int_0^T |u|^2 dt, \quad (3)$$

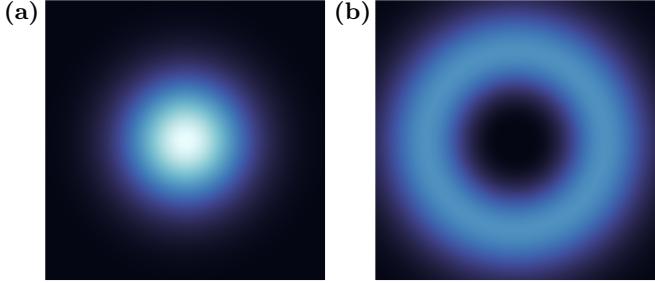


FIG. 2: Topologically changing the condensate’s support.
 (a) The initial distribution $|\psi_{\text{Gaussian}}|^2$ with $a = b = 1$.
 (b) The desired distribution $|\psi_{\text{Toroid}}|^2$.

subject to

$$i\partial_t\psi + \frac{1}{2}\nabla^2\psi - V(x, u(t))\psi - |\psi|^2\psi = 0, \quad (4a)$$

$$\psi(x, 0) = \psi_0(x) \in H^1(\mathbb{R}^3), \quad (4b)$$

$$\|\psi(x, \cdot)\|_{L^2(\mathbb{R}^3)} = 1, \quad (4c)$$

where $\gamma > 0$, $t \in [0, T]$, the wavefunction $\psi(x, t)$ belongs to $L^2([0, T]; H^1(\mathbb{R}^3))$, ψ_0 is some initial state, ∇^2 is the Laplacian operator, $V(x, u)$ models the geometry of confinement as a potential energy parameterized by the control $u(t)$, \mathcal{U} is an admissible class of control functions $u \in H^1([0, T])$ with fixed initial and terminal conditions, $L^2(\Omega)$ is the space of square Lebesgue-integrable functions over the measurable set Ω , and $H^1(\Omega)$ is the Sobolev space of $L^2(\Omega)$ functions whose first weak derivatives are also in $L^2(\Omega)$.

The dynamical constraint (4a) is known as the Gross-Pitaevskii equation (GPE), and the confining potential $V(x, u(t))$ arises due to an applied optical or magnetic field. How the constraint arises as a model of the mean-field dynamics of BEC is discussed in detail in [6]. The terms $J^{\text{infidelity}}$ and J^{regular} are known as the infidelity and regularization terms.

In the language of optimal control theory [21, 22], the infidelity is a type of *terminal cost* which penalizes control policies that miss the desired wave function ψ_d . The regularization is a type of *running cost* which penalizes the usage of physically undesirable controls with fast variations and ensures that Hohenester’s optimal control problem remains well-posed. This is shown for a more general control problem defined by the Hohenester objective (2) and mean field constraint (4a), along with a running cost which also penalizes the amount of work done by the control, in work due to Hintermuller, et al. [23].

Mennemann, et al., numerically study experimentally motivated transformations of $\psi(x, t)$, such as those in Figures 1 and 2, by solving the associated optimal control problem, after setting $\gamma = 10^{-6}$, with a projected gradient method called Gradient Pulse Engineering (GRAPE). Their work is the source of inspiration for this paper.

We begin with two primary questions: can we gain further physical intuition of the condensate dynamics as it is controlled, and can we use this physical insight to

implement optimization strategies in some easier, i.e., finite-dimensional, computational setting? To this end, we introduce a Galerkin ansatz which incorporates the time-dependence of the confining potential, and use this to study two model problems in one space dimension: squeezing a BEC in a steepening quadratic potential and splitting a BEC with a time-dependent barrier. These model problems abstract the salient features of the reorientation problem illustrated in Figure 1 and the splitting problem illustrated in Figure 2.

A. Structure of the Paper

In Section II we describe the Galerkin reduction of the squeezing and splitting problems, following, for example [24, 25]. We assume the reshaping potential is product-separable in space and time, the problem is even-symmetric, and initial conditions are small-amplitude superpositions of eigenfunctions of the associated linear Schrödinger operator so that the dynamics are weakly nonlinear. Using these assumptions, we reduce the dynamics of the controlled condensate a non-autonomous one degree of freedom Hamiltonian system using Galerkin reductions and canonical transformations. We validate this reduction by comparing numerical solutions of the GPE with a specified time-dependent potential with solutions of the reduced system.

We then pose an optimal control problem, in Section III, constrained by the Hamiltonian dynamics whose objective to minimize that same Hamiltonian function and thereby minimize oscillations which persist after the control process is terminated. We then provide the necessary optimality conditions for this class of Hamiltonian control problems.

In Section IV, we outline the numerical methods used to solve the nonconvex optimization problem of Section III. We proceed in two steps, using a so-called hybrid method. Because the objective function is non-convex, it may have many local minima, and we first must search for the best among many candidates. We use a method due to Calarco, et al. [26, 27], called the Chopped Random Basis (CRAB) method to reduce the search space to finite dimensions by considering controls within a space of Galerkin approximations. This space is searched using a global, nonconvex method due to Storn and Price called Differential Evolution (DE) [28]. The second step is to refine the best candidate using a local descent method. The GRAPE method allows us to perform this descent among controls satisfying desired boundary conditions.

To validate the proposed approach, in Section V we solve the Gross-Pitaevskii equation (4a) again, using the controls resulting from the methods of Section IV. We find that this approach both suppresses undesirable persistent oscillations and minimizes the infidelity in the Hohenester objective functional (2).

II. MODEL PROBLEMS FROM A GALERKIN REDUCTION

In this section, we outline the derivation of model Hamiltonian problems via a Galerkin truncation. We apply this truncation to a GPE in one spatial dimension which we assume depends on a stationary potential $V_s(x)$ and a reshaping potential $V_r(x)$, i.e.,

$$i\partial_t \psi = -\frac{1}{2}\partial_x^2 \psi + V_s(x)\psi + u(t)V_r(x)\psi + |\psi|^2\psi. \quad (5)$$

We use a Galerkin expansion of the form

$$\psi(x, t) = \sum_{n=0}^{\infty} c_n(t) \varphi_n(x; u(t)), \quad (6)$$

where each of the basis functions $\varphi_n(x; u(t))$ is an instantaneously normalized eigenfunction of the equation

$$-\frac{1}{2}\partial_x^2 \varphi_n + (V_s(x) + uV_r(x)) \varphi_n = E_n \varphi_n,$$

i.e., the linear Schrödinger equation with u -dependent potential.

By choosing initial conditions which are in the form of (6) with $|c_n|$ relatively small, nonlinear effects remain relatively weak throughout the control process. This allows us to truncate the expansion (6) at a low order. This large reduction of dimension due to the Galerkin greatly simplifies the dynamics and is justified through numerical studies in Section II C. We show the coefficients $c_n(t)$ evolve under a Hamiltonian system whose dynamics motivates the control strategy discussed in Section III.

A. The Squeezing Problem

We first address the problem of squeezing and elongation discussed in Section I and shown in Figure 1. As a model problem, we consider the squeezing of a stationary wave packet centered about the origin and trapped in a reshaping quadratic potential, i.e. $V_r(x) = \frac{1}{2}x^2$, $V_s(x) \equiv 0$ in Equation (5), with the endpoints of the control fixed as $u(0) = u_0 > 0$ and $u(T) = u_T > u_0$.

In this case, each of the $\varphi_n(x; u(t))$ in Expansion (6) satisfies

$$-\frac{1}{2}\partial_x^2 \varphi_n + \frac{1}{2}ux^2 \varphi_n = E_n \varphi_n.$$

The eigenfunctions $\varphi_n(x; u)$ are the well-known Hermite functions and can be generated by the Rodrigues formula

$$\varphi_n(x; u) = (-1)^n \frac{\pi^{-1/4}}{\sqrt{2^n u^{n/4} n!}} u^{1/8} e^{\frac{u^{1/4} x^2}{2}} \partial_x^n e^{-u^{1/4} x^2}. \quad (7)$$

The first three are

$$\begin{aligned} \varphi_0(x; u) &= \xi e^{-\frac{1}{2}\sqrt{u}x^2}, \\ \varphi_1(x; u) &= \sqrt{2}\xi u x e^{-\frac{1}{2}\sqrt{u}x^2}, \\ \varphi_2(x; u) &= \sqrt{2}\xi u (2\sqrt{u}x^2 - 1) e^{-\frac{1}{2}\sqrt{u}x^2}, \end{aligned}$$

where $\xi = \pi^{-1/4} u^{1/8}$. We truncate expansion (6) after the third term, and discard the single odd term involving $\varphi_1(x; u)$ because we assume the initial conditions obey an even symmetry which is invariant under GPE. For convenience, we relabel these first two even eigenstates and their time dependent coefficients as the $n = 0, 1$ states.

To derive the equations governing the time-dependent coefficients present in expansion (6), we substitute the expansion into the GPE (5) and project onto each mode using the standard $L^2(\mathbb{R})$ inner product. Letting \dagger denote complex conjugation and overhead dots denote time derivatives, the resulting ODE system is

$$\begin{aligned} i\dot{c}_0 &= \frac{\sqrt{u}c_0}{2} + \xi^2 \left(\frac{|c_0|^2 c_0}{\sqrt{2}} + \frac{|c_1|^2 c_1}{32} + \frac{3|c_1|^2 c_0}{4\sqrt{2}} \right. \\ &\quad \left. - \frac{|c_0|^2 c_1}{2} + \frac{3c_1^2 c_0^\dagger}{8\sqrt{2}} - \frac{c_0^2 c_1^\dagger}{4} \right) - \frac{iuc_1}{4\sqrt{2}u}, \end{aligned} \quad (8a)$$

$$\begin{aligned} i\dot{c}_1 &= \frac{5\sqrt{u}c_1}{2} + \xi^2 \left(\frac{41|c_1|^2 c_1}{64\sqrt{2}} - \frac{|c_0|^2 c_0}{4} + \frac{3|c_0|^2 c_1}{4\sqrt{2}} + \right. \\ &\quad \left. + \frac{|c_1|^2 c_0}{16} + \frac{c_1^2 c_0^\dagger}{32} + \frac{3c_0^2 c_1^\dagger}{8\sqrt{2}} \right) + \frac{iuc_0}{4\sqrt{2}u}. \end{aligned} \quad (8b)$$

This can be written in the Hamiltonian form

$$i\dot{c}_n = +\partial_{c_n^\dagger} \mathcal{H}, \quad i\dot{c}_n^\dagger = -\partial_{c_n} \mathcal{H}, \quad n = 0, 1, \quad (9)$$

where the Hamiltonian $\mathcal{H}(c_0, c_0^\dagger, c_1, c_1^\dagger; u)$ is given by

$$\begin{aligned} \mathcal{H} &= \xi^2 \left(\frac{|c_0|^4}{2\sqrt{2}} + \frac{41|c_1|^4}{128\sqrt{2}} + \frac{3|c_0|^2 |c_1|^2}{4\sqrt{2}} + \frac{3 \operatorname{Re} \{c_0^2 c_1^\dagger\}^2}{8\sqrt{2}} \right. \\ &\quad \left. - 2 \operatorname{Re} \{c_0 c_1^\dagger\} \left(|c_0|^2 - \frac{|c_1|^2}{8} \right) \right) \\ &\quad + \frac{\sqrt{u}}{2} \left(|c_0|^2 + 5|c_1|^2 \right) - \frac{\dot{u}}{2\sqrt{2}u} \operatorname{Im} \{c_0 c_1^\dagger\}. \end{aligned} \quad (10)$$

Note that, the dynamics conserve the “discrete” mass

$$M_d(t) = |c_0(t)|^2 + |c_1(t)|^2. \quad (11)$$

Next, we reduce the the squeezing Hamiltonian (10) to one and a half degrees of freedom using canonical transformations. This allows the use of phase plane techniques which provide further insight into the problem.

We first convert to action-angle coordinates through the canonical transformation

$$c_0 = \sqrt{\rho_0} e^{-i\theta_0}, \quad c_1 = \sqrt{\rho_1} e^{-i\theta_1}.$$

Hamiltonian (10) then becomes

$$\begin{aligned}\mathcal{H} = & \frac{\sqrt{u}}{2} (\rho_0 + 5\rho_1) - \frac{\dot{u}}{4u} \sqrt{2\rho_0\rho_1} \sin\phi + \\ & \frac{3\xi^2}{8\sqrt{2}} \rho_0\rho_1 \cos(2\phi) + \frac{\sqrt{2}\xi^2}{256} (64\rho_0^2 + 96\rho_1\rho_0 + \\ & 41\rho_1^2 + (8\rho_1^{3/2} \sqrt{2\rho_0} - 56\rho_0^{3/2} \sqrt{2\rho_1}) \cos(\phi)),\end{aligned}$$

where $\phi = \theta_0 - \theta_1$. In these coordinates, the discrete mass (11) is given by $M_d = \rho_0 + \rho_1$. We make the choice to set the discrete mass to one, and introduce the change of variables $\rho_0 = 1 - J$ and $\rho_1 = J$ so that

$$\begin{aligned}\mathcal{H} = & \frac{\xi^2}{128\sqrt{2}} (9J^2 - 32J + 64) - \frac{\dot{u}}{2\sqrt{2}u} \sqrt{(1-J)J} \sin(\phi) \\ & + \frac{\sqrt{u}}{2} (1 + 4J) + \frac{\xi^2}{16} \left(\sqrt{(1-J)J} (8 - 7J) \cos(\phi) \right. \\ & \left. + 3\sqrt{2}(1-J)J \cos(2\phi) \right).\end{aligned}$$

In these coordinates, $J = 0$ indicates that all of the mass is in the ground state while $J = 1$ indicates that all of the mass is in the excited state.

A further canonical transformations facilitate visualization of the phase portrait. Defining $q + ip = \sqrt{2}Je^{i\phi}$ yields

$$\begin{aligned}\mathcal{H}(q, p, u) = & \sqrt{u} \left(q^2 + p^2 + \frac{1}{2} \right) \\ & + \frac{\xi^2}{64} \sqrt{2 - p^2 - q^2} \left(9q^3 - 16q + 9p^2q - \frac{8\sqrt{2}\pi\dot{u}}{u}p \right) \\ & + \frac{\xi^2}{512\sqrt{2}} (57p^4 - 160p^2 + 18p^2q^2 - 39q^4 + 32q^2 + 256).\end{aligned}\quad (12)$$

B. The Splitting Problem

We refer to the problem of topologically changing the support of the condensate, mentioned in Section I and shown in Figure 2, as the “splitting” problem. In the case of one spatial dimension and “splitting” potential $V_r(x) = \delta(x)$, together with quadratic stationary potential $V_s(x) = \frac{1}{2}x^2$, the linear Schrödinger equation is exactly solvable for each value of u .

In order to construct the Galerkin ansatz, we provide brief details on solving the eigenvalue problem

$$-\frac{1}{2}\partial_x^2\varphi_n + \frac{1}{2}x^2\varphi_n + u\delta(x)\varphi_n = E_n\varphi_n. \quad (13)$$

A more thorough computation and discussion is given by Viana-Gomes and Peres [29]. First, note that integrating (13) in a neighbourhood about the origin leads to a jump condition on the derivative:

$$\lim_{\varepsilon \rightarrow 0} \partial_x\varphi(x)|_{-\varepsilon}^{+\varepsilon} = 2u\varphi|_{x=0}. \quad (14)$$

Since all odd $C^1(\mathbb{R})$ functions satisfy the jump condition, the odd-parity states are given by the Rodrigues formula (7) with $u = 1$. Only the even-parity states are modified by the delta function at the origin.

By letting $\varphi = e^{-x^2/2}w(x)$, $z = x^2$, and $E = \nu + \frac{1}{2}$, $\nu \in \mathbb{R}$, Equation (13) and condition (14) become

$$z\partial_z^2w + \left(\frac{1}{2} - z\right)\partial_zw + \frac{\nu^2}{2}w = 0, \quad z > 0, \quad (15a)$$

$$\partial_zw|_{z=0} = uw|_{z=0}. \quad (15b)$$

Equation (15a) is called Kummer’s equation and admits solutions of the form

$$w(z) = A_\nu U\left(-\frac{\nu}{2}, \frac{1}{2}, z\right),$$

where

$$\begin{aligned}U(a, b, z) = & \frac{\Gamma(1-b)}{\Gamma(a+1-b)} M(a, b, z) \\ & + \frac{\Gamma(b-1)}{\Gamma(a)} z^{(1-b)} M(a+1-b, 2-b, z),\end{aligned}\quad (16)$$

is Tricomi’s confluent hypergeometric function whose definition involves the gamma function, $\Gamma(z)$, and Kummer’s function

$$M(a, b, z) = \sum_{n=0}^{\infty} \frac{a^{(n)}z^n}{b^{(n)}n!},$$

with $(.)^{(n)}$ denoting the rising factorial defined by

$$a^{(n)} := \prod_{k=0}^{n-1} (a+k).$$

The coefficient A_ν is a normalization constant.

Applying boundary condition (15b) to $w(z)$ leads to the nonlinear equation

$$\nu - u \frac{\Gamma(1 - \frac{\nu}{2})}{\Gamma(\frac{1}{2} - \frac{\nu}{2})} = 0 \quad (17)$$

for ν . For $u=0$, Equation (17) implies $\nu=0$, and we recover the even Hermite basis given by the Rodrigues formula (7). In general, a numerical solution of Equation (17), demonstrated in [29], shows there is a countable sequence of solutions $\{\nu_n\}$ each satisfying $\nu_{n+1} = \nu_n + 2$. Thus it suffices to solve Equation (17) on the interval $[0,1]$, the interval containing the ground state value of ν , since this determines all other solutions. Therefore, we m restriction ν to $[0, 1]$, so that the first two even eigenfunctions can be written as

$$\begin{aligned}\varphi_j(x; \nu) = & N_j(\nu) e^{-\frac{x^2}{2}} U\left(-\frac{\nu + 2j}{2}, \frac{1}{2}, x^2\right) \\ & := N_j(\nu) e^{-\frac{x^2}{2}} U_j(x^2, \nu), \quad j = 0, 1,\end{aligned}\quad (18)$$

where $N_j(\nu)$ are ν -dependent normalization constants given by

$$N_j^{-2}(\nu) = \int_{\mathbb{R}} e^{-x^2} U_j^2(x^2, \nu) dx, \quad j = 0, 1.$$

These eigenfunctions $\varphi_j(x; \nu)$ will serve as basis functions in the Galerkin expansion.

Note that as $\nu \rightarrow 1$, $u \rightarrow \infty$, since $\Gamma(z)$ has a pole at the origin. In this case, the first two even eigenfunctions reduce to the simple form of “split” wavefunctions

$$\begin{aligned} \varphi_0(x; 1) &= 2^{\frac{1}{2}} \pi^{-\frac{1}{4}} |x| e^{-\frac{x^2}{2}}, \\ \varphi_1(x; 1) &= 2\pi^{-\frac{1}{4}} 3^{-\frac{1}{2}} \left(|x|^3 - \frac{3}{2}|x| \right) e^{-\frac{x^2}{2}}, \end{aligned}$$

shown in Figure 3.

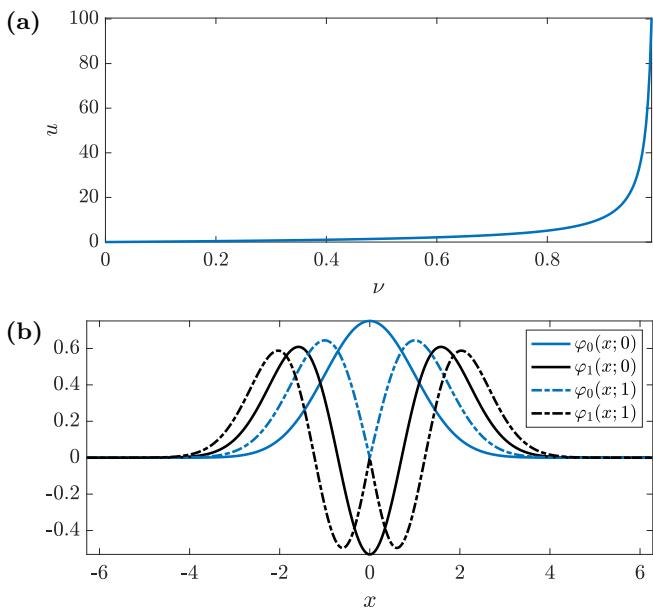


FIG. 3: (a) The function $u(\nu)$ as determined by Equation (17). Large values of u are needed to achieve ν close to 1. (b) The first two even states in (18) for the values of $\nu = 0, 1$.

Proceeding as Section II A, we have, after projecting onto each mode using the $L^2(\mathbb{R})$ inner product,

$$\begin{aligned} i\dot{c}_0 &= \alpha_0 c_0 + \alpha_1 c_1 + u(\beta_0 c_0 + \beta_1 c_1) + \gamma_0 |c_0|^2 c_0 \\ &\quad + \gamma_2 |c_1|^2 c_1 + 2\gamma_3 |c_0|^2 c_1 + \gamma_3 c_0^2 c_1^\dagger \\ &\quad + 2\gamma_4 c_0^\dagger c_1^2 + \gamma_4 |c_1|^2 c_0 - i\Delta c_0, \end{aligned} \quad (19a)$$

$$\begin{aligned} i\dot{c}_1 &= \alpha_1 c_0 + \alpha_2 c_1 + u(\beta_1 c_0 + \beta_2 c_1) + \gamma_3 |c_0|^2 c_0 \\ &\quad + \gamma_1 |c_1|^2 c_1 + 2\gamma_4 c_0^2 c_1^\dagger + \gamma_2 c_0^\dagger c_1^2 \\ &\quad + \gamma_4 |c_0|^2 c_1 + 2\gamma_2 |c_1|^2 c_0 + i\Delta c_1, \end{aligned} \quad (19b)$$

where

$$\begin{aligned} 2\alpha_0 &= \langle \varphi_0, x^2 \varphi_0 - \partial_x^2 \varphi_0 \rangle, \quad 2\alpha_1 = \langle \varphi_0, x^2 \varphi_1 - \partial_x^2 \varphi_1 \rangle, \\ 2\alpha_2 &= \langle \varphi_1, x^2 \varphi_1 - \partial_x^2 \varphi_1 \rangle, \end{aligned} \quad (20)$$

$$\begin{aligned} \beta_0 &= \langle \varphi_0, \delta(x) \varphi_0 \rangle, \quad \beta_1 = \langle \varphi_0, \delta(x) \varphi_1 \rangle, \\ \beta_2 &= \langle \varphi_1, \delta(x) \varphi_1 \rangle, \\ \gamma_0 &= \|\varphi_0^4\|, \quad \gamma_1 = \|\varphi_1^4\|, \quad \gamma_2 = \langle \varphi_0, \varphi_1^3 \rangle, \\ \gamma_3 &= \langle \varphi_1, \varphi_0^3 \rangle, \quad \gamma_4 = \langle \varphi_0^2, \varphi_1^2 \rangle, \\ \Delta &= \langle \varphi_0, \partial_t \varphi_1 \rangle = -\langle \partial_t \varphi_0, \varphi_1 \rangle. \end{aligned} \quad (21)$$

The Hamiltonian associated with ODE (19) for this problem is given by

$$\begin{aligned} \mathcal{H} &= (\alpha_0 + u\beta_0) |c_0|^2 + (\alpha_2 + u\beta_2) |c_1|^2 \\ &\quad + 2(\alpha_1 + \beta_1) \operatorname{Re} \{c_0 c_1^\dagger\} + \frac{1}{2}\gamma_0 |c_0|^4 + \frac{1}{2}\gamma_1 |c_1|^4 \\ &\quad + 2(\gamma_3 |c_0|^2 + \gamma_2 |c_1|^2) \operatorname{Re} \{c_0 c_1^\dagger\} \\ &\quad + \gamma_4 (|c_0|^2 |c_1|^2 + 2 \operatorname{Re} \{c_0^2 c_1^\dagger\}) + 2\Delta \operatorname{Im} \{c_0^\dagger c_1\}. \end{aligned} \quad (22)$$

Applying the same canonical transformations as in Section II A, we arrive at

$$\begin{aligned} \mathcal{H} &= \alpha_0 + \frac{\gamma_0}{2} + \frac{1}{2}(p^2 + q^2)(\alpha_2 - \alpha_0 - \gamma_0 + (\beta_2 - \beta_0)u) \\ &\quad + \sqrt{2 - p^2 - q^2} \\ &\quad \left(q \left(\alpha_1 + \gamma_3 + \frac{1}{2}(\gamma_2 - \gamma_3)(p^2 + q^2) + \beta_1 u \right) + \Delta p \right) \\ &\quad + \frac{\gamma_4}{2}(3q^2 - p^2) + \frac{1}{8}(\gamma_0 + \gamma_1)(p^4 + q^4) \\ &\quad + \frac{\gamma_4}{4}(p^4 - 3q^4) + \frac{1}{4}(\gamma_0 + \gamma_1 - 2\gamma_4)p^2 q^2 + \beta_0 u. \end{aligned} \quad (23)$$

C. Numerical Experiments

In this section, we simulate both the GPE (5) and the Galerkin-truncated systems of ODE describing both the squeezing problem (8) and the splitting problem (19), and compare the results. For both problems we fix the time domain $t \in [0, T]$, $T > 0$, and solve both the GPE and the reduced models using a control of the form

$$u(t) = (u_T - u_0) \frac{t}{T} + u_0, \quad (24)$$

with $u_0 = 1$, $u_T = 100$, and $T = 2.5$ for the squeezing experiment, and $u_0 = 0$, $u_T = 30$, and $T = 10$ for the splitting experiment.

In the ODE models we choose the initial conditions $(c_0(0), c_1(0))$ to minimize the associated Hamiltonian subject to the choice $M_d = 1$, where \dot{u} is set to zero in the definition of \mathcal{H} in Hamiltonians (10) and (22). Thus the initial conditions are taken to be the fixed point of

the system before the control is applied. The GPE is initialized as a superposition of the first two even states,

$$\psi_0(x) = c_0(0)\varphi_0(x, u_0) + c_1(0)\varphi_1(x, u_0),$$

so that it represents the same initial state.

We solve the GPE (5) using a second-order-in-time split-step Fourier method using the midpoint method to integrate the time-dependence on the potential $u(t)V_r(x)$ and solve the ODE systems using MATLAB's `ode45`, i.e., an adaptive-step fourth order Runge-Kutta method. To compare the numerical solution of the GPE system with the numerical solutions to the Galerkin truncated systems, we define the projected solution and the instantaneous Galerkin coefficients by projecting the numerical solution of GPE onto the instantaneous eigenfunctions,

$$\begin{aligned} \psi_{\text{proj}}(x, t) &= \sum_{n=0}^1 \langle \psi_{\text{GPE}}(x, t), \varphi_n(x; u(t)) \rangle \varphi_n(x; u(t)) \\ &:= \sum_{n=0}^1 c_n^{\text{proj}}(t) \varphi_n(x; u(t)). \end{aligned}$$

We may also construct the approximate solution to GPE ψ_{galerkin} by evaluating the Galerkin ansatz using the numerically calculated values of $c_0(t)$ and $c_1(t)$.

Figures 4 and 5 show the results of these numerical experiments. We present false color plots of $|\psi|^2$, $|\psi_{\text{proj}}|^2$, and $|\psi_{\text{galerkin}}|^2$. These show excellent qualitative agreement, especially the last two, showing that the main source of disagreement comes from the truncation. They also show strong agreement between c_j and c_j^{proj} . In particular, we find good visual agreement in the Rabi frequency, i.e., the peak frequency of mass transfer between the first two even modes. This agreement is exhibited by the similar periodic behavior between $c_n(t)$, determined by System (9), and the projected coefficients $c_n^{\text{proj}}(t)$. Finally, they show a discrepancy of at most 3% between the simulated and projected discrete masses in either experiment. As seen in Figures 4 and 5, this discrepancy can be mainly attributed to the tails of the distribution $|\psi_{\text{GPE}}|^2$.

In visualizing the phase portraits associated with Hamiltonians (12) and (22), we use the same numerical setting as that of Figures 4 and 5. The phase portraits, shown in Figure 6, reveal how significant Rabi oscillations present in Figures 4 and 5 are characterized by the distance between the final state $(q(T), p(T))$ and the stable fixed point (q^*, p^*) of Hamiltonian (12). We denote the initial and final Hamiltonians $\mathcal{H}(q(0), p(0), u_0)$, $\mathcal{H}(q(T), p(T), u_T)$ as \mathcal{H}_0 , \mathcal{H}_T , respectively, in Figure 6.

In conclusion, this section provides numerical justification for the large reduction of dynamic complexity, provided by the truncation of the ansatz (6). In addition, the reduced dynamics reveal that a successful control strategy should drive the state of the condensate to the global minimum of its finite dimensional Hamiltonian \mathcal{H}_T . Furthermore, sub-optimality is almost entirely characterized by the amplitude of simple harmonic motion shown in Figure 6.

III. OPTIMAL CONTROL FRAMEWORK

We now state an optimal control problem for systems constrained by Hamiltonian dynamics, motivated by Sections II A and II B. To this end, we use the admissible class of controls $\mathcal{U} = \{u \in H^1([0, T]) : u(0) = u_0, u(T) = u_T\}$, where $u_0, u_T \in \mathbb{R}$ are boundary values for the control u . The optimal control problem we study is

$$\min_{u \in \mathcal{U}} J = \min_{u \in \mathcal{U}} \left\{ \mathcal{H}(q, p, u) \Big|_{t=T} + \frac{\gamma}{2} \int_0^T \dot{u}^2 dt \right\}, \quad (25)$$

subject to Hamilton's equations

$$\dot{q} = \partial_p \mathcal{H}, \quad q(0) = q_0, \quad \dot{p} = -\partial_q \mathcal{H}, \quad p(0) = p_0. \quad (26)$$

Recall, from Section I, the first term in objective J , called the terminal cost, is used to penalize deviations from some desired state, and that the second term, called the running cost, is a Tikhonov regularization on the control u .

Rewriting the terminal cost in (25) as a running cost simplifies the process of computing gradients with respect to the state and control variables. We convert terminal costs into running costs through the fundamental theorem of calculus:

$$\begin{aligned} \mathcal{H}|_{t=0}^{t=T} &= \int_0^T \frac{d\mathcal{H}}{dt} dt \\ &= \int_0^T \left(\frac{\partial \mathcal{H}}{\partial q} \dot{q} + \frac{\partial \mathcal{H}}{\partial p} \dot{p} + \frac{\partial \mathcal{H}}{\partial u} \dot{u} + \frac{\partial \mathcal{H}}{\partial t} \right) dt \\ &= \int_0^T \left(\frac{\partial \mathcal{H}}{\partial q} \frac{\partial \mathcal{H}}{\partial p} - \frac{\partial \mathcal{H}}{\partial p} \frac{\partial \mathcal{H}}{\partial q} + \frac{\partial \mathcal{H}}{\partial u} \dot{u} + \frac{\partial \mathcal{H}}{\partial t} \right) dt \\ &= \int_0^T \left(\frac{\partial \mathcal{H}}{\partial u} \dot{u} \right) dt. \end{aligned}$$

Using Lagrange multipliers, we express the Hamiltonian optimal control problem in unconstrained form as

$$\begin{aligned} \min_{u \in \mathcal{U}} J &= \min_{u \in \mathcal{U}} \left\{ \int_0^T \left[\frac{\partial \mathcal{H}}{\partial u} \dot{u} + \frac{\gamma}{2} \dot{u}^2 + \lambda^\top \left(\dot{q} - \frac{\partial \mathcal{H}}{\partial p} \right) + \mu^\top \left(\dot{p} + \frac{\partial \mathcal{H}}{\partial q} \right) \right] dt \right\} \\ &= \min_{u \in \mathcal{U}} \left\{ \int_0^T \mathcal{L}(q, \dot{q}, p, \dot{p}, u, \dot{u}, \lambda, \mu) dt \right\}, \end{aligned} \quad (27)$$

where $^\top$ denotes the matrix transpose, and where the cost $\mathcal{H}|_{t=0}^{t=T}$ has been dropped since initial values for the state and control variables are specified, therefore fixed when taking derivatives. The necessary conditions for a locally extremal solution to Lagrange problem (27) are given by the Euler-Lagrange equations:

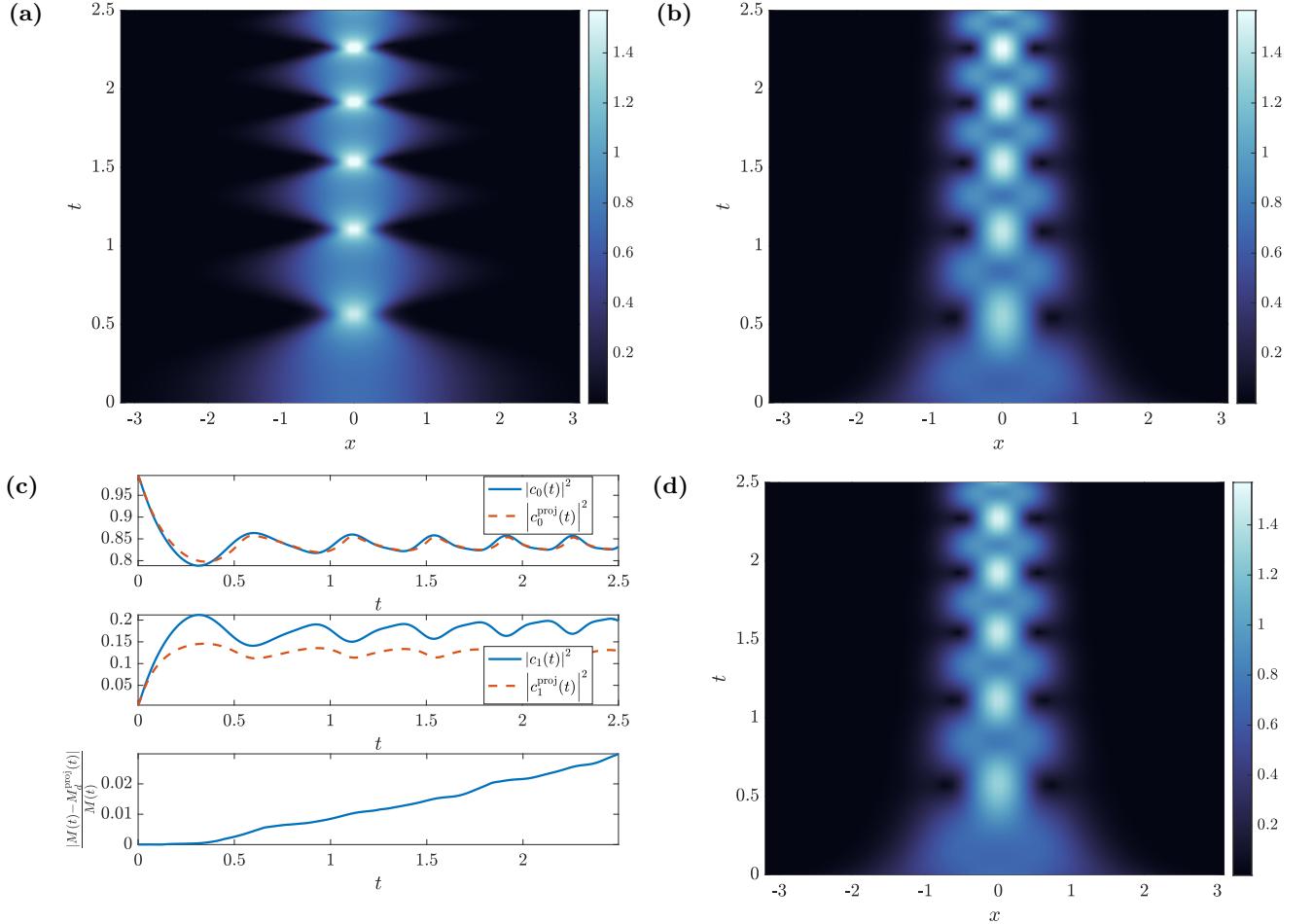


FIG. 4: (a) Full numerical solution to Equation (5) using the quadratic potential with time dependence (24), with $|\psi|^2$ plotted versus x and t . (b) The Galerkin truncation of this solution to the first two even eigenfunctions. (d) An approximate PDE solution constructed from an equivalent solution to ODE system (8). (c) First two rows: $|c_0|^2$ and $|c_1|^2$ computed by solving the ODE system (blue) and solving the PDE system and then projecting (red, dashed). Bottom row: Relative error between the full PDE solution and the projection.

where

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial^2 \mathcal{H}}{\partial q^2} & -\frac{\partial^2 \mathcal{H}}{\partial q \partial p} & \frac{\partial^2 \mathcal{H}}{\partial q \partial u} \\ 0 & 0 & \frac{\partial^2 \mathcal{H}}{\partial p \partial q} & -\frac{\partial^2 \mathcal{H}}{\partial p^2} & \frac{\partial^2 \mathcal{H}}{\partial p \partial u} \\ 0 & 0 & \frac{\partial^2 \mathcal{H}}{\partial u \partial q} & -\frac{\partial^2 \mathcal{H}}{\partial u \partial p} & \frac{\partial^2 \mathcal{H}}{\partial u^2} \end{pmatrix}$$

$$= \left(\begin{array}{c|cc} \mathcal{J} & \mathbf{0}_{2 \times 3} \\ \hline \mathbf{0}_{3 \times 2} & D \left(\frac{\partial \mathcal{H}}{\partial q}, -\frac{\partial \mathcal{H}}{\partial p}, \frac{\partial \mathcal{H}}{\partial u} \right) \end{array} \right),$$

$$\begin{aligned} \frac{d}{dt} \begin{pmatrix} q \\ p \\ \lambda \\ \mu \\ \gamma \dot{u} \end{pmatrix} &= A \begin{pmatrix} \frac{\partial \mathcal{H}}{\partial q} \\ \frac{\partial \mathcal{H}}{\partial p} \\ \lambda \\ \mu \\ \dot{u} \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \frac{d}{dt} \frac{\partial \mathcal{H}}{\partial u} \end{pmatrix}, \\ \begin{pmatrix} q(0) \\ p(0) \\ u(0) \end{pmatrix} &= \begin{pmatrix} q_0 \\ p_0 \\ u_0 \end{pmatrix}, \quad \begin{pmatrix} \lambda(T) \\ \mu(T) \\ u(T) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ u_T \end{pmatrix}, \quad (28) \end{aligned}$$

with \mathcal{J} denoting the corresponding skew-symmetric matrix and D denoting the Jacobian matrix.

From the perspective of optimal control theory, the equations for λ and μ are called the costate equations and are solved backward in time from their respective terminal conditions. The equation for the control $u(t)$, along with the prescribed boundary conditions, is a two-point boundary value problem. Since solving Equation (28) in closed form is not possible, we resort to numerical meth-

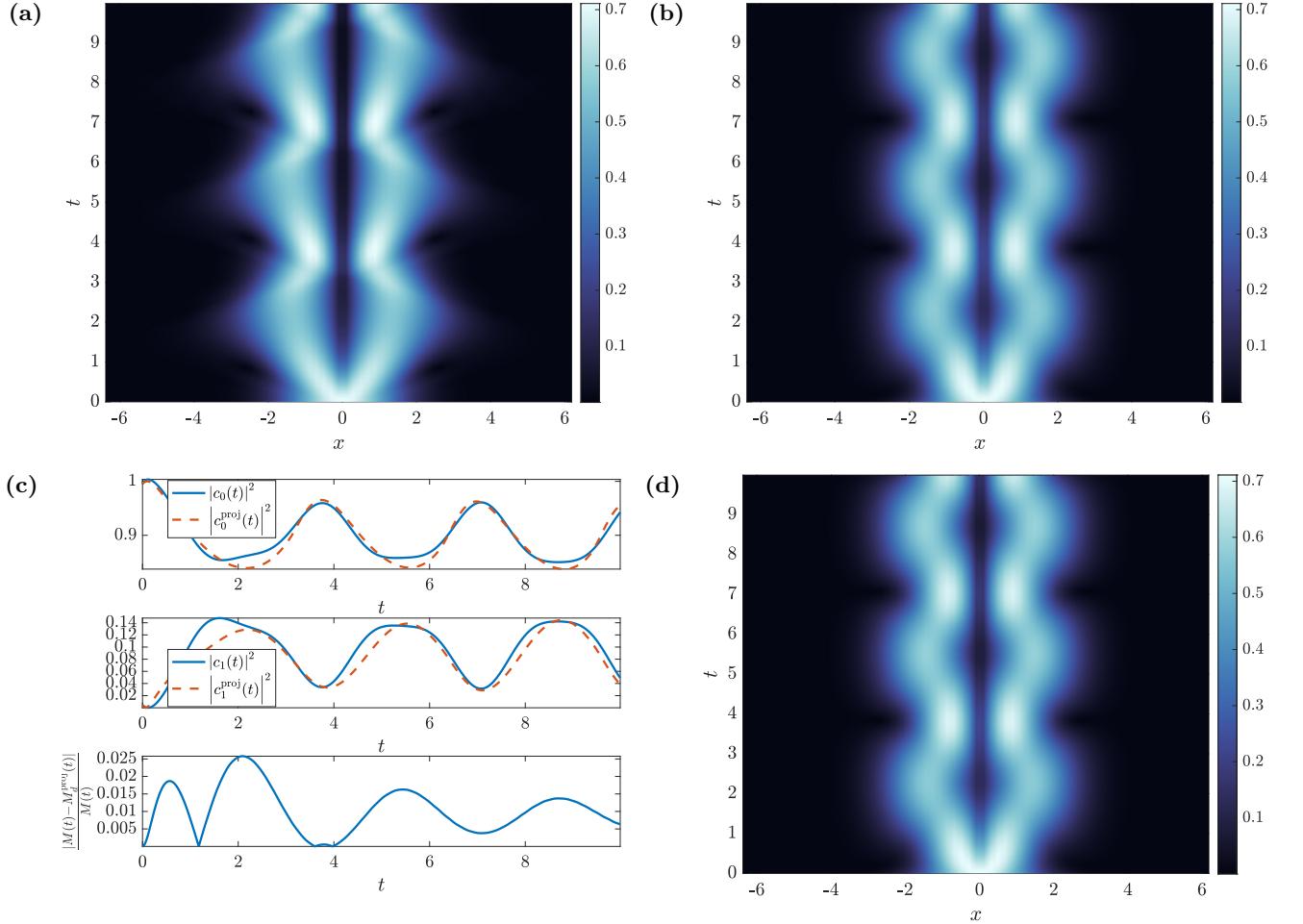


FIG. 5: The splitting experiment. All conventions here are consistent with Figure 4, except in this case, the dynamics are furnished by the Hamiltonian (22).

ods discussed in Section IV in order to solve the optimal control problem (25) by approximating the optimality condition (28).

IV. NUMERICAL OPTIMIZATION

For the squeezing problem, we use a *hybrid method*: a global, non-convex method followed by a local, iterative method. Hybrid methods, when used appropriately, can overcome non-convexity, yet still remain computationally efficient.

In both problems, applying the local iterative method requires differentiating the projection coefficients with respect to the control u . In the case of the splitting problem, differentiating the coefficients (20) with respect to the control u requires an unmanageable implicit differentiation through Tricomi's function (16) and Equation (17). Therefore, we omit the second optimization step in the splitting problem.

The hybrid method here is similar to work by Sørensen, et al. [30], and allows for the use of a global search rou-

tine based on stochastic optimization to overcome non-convexity. Global methods are known to converge slowly near a local minimum [31]. Feeding the result of global methods into a local methods accelerates this slow convergence.

A. The Global Method

The first step in the hybrid method reduces the complexity of the optimal control problem so that standard non-convex Nonlinear Programming (NLP) techniques can be applied. This step, the CRAB method [26, 27], constructs the control from the span of an appropriately chosen finite set of basis functions so that the optimization is performed over a small set of unknown coefficients. We choose the basis to ensure that the controls remain in the admissible space \mathcal{U} of the control problem (25), using

$$u_r(t) = \mathcal{P}(t; u_0, u_T, T) + \sum_{j=0}^{N-1} \varepsilon_j \varphi_j(t; T), \quad t \in [0, T], \quad (29)$$

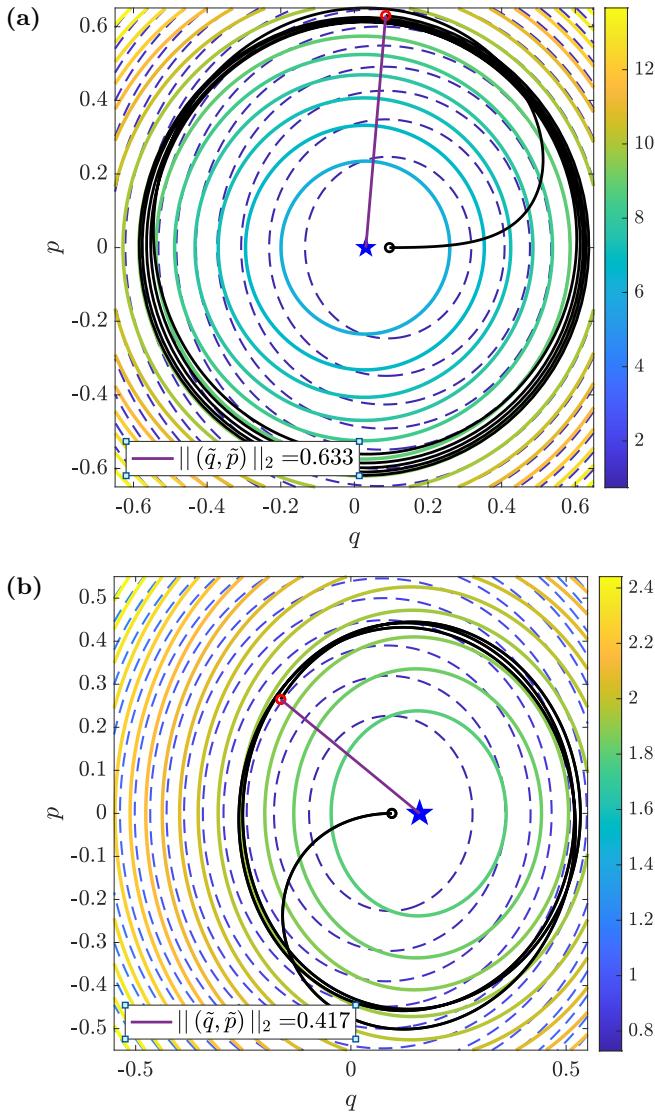


FIG. 6: Phase portraits for the dynamics due to (a) the squeezing Hamiltonian (12), and (b) the splitting Hamiltonian (23). Dashed contours: the initial Hamiltonian, H_0 , initial conditions set at its stable fixed point. Solid contours: the final Hamiltonian, H_T , with a blue star at its minimum. Black lines: numerical trajectories. Red circle: the final state, $(q(T), p(T))$.

where \mathcal{P} is a fixed function, $\{\varphi_j(t)\}_{j \in \mathbb{N}}$ satisfying the boundary conditions defining the admissible class \mathcal{U} , is a set of functions that satisfy homogeneous boundary conditions, and the coefficients ε_j are parameters to be optimized over.

The CRAB method can be viewed as a Galerkin method, so we must choose the number of basis functions N simultaneously large enough to define an accurate approximation, yet small enough so that the overall procedure remains computationally inexpensive. We have found that a set of 15 basis functions works well.

To solve the resulting NLP problem, we use Differential Evolution (DE) [28]. DE is a stochastic optimization

method used to search for candidate solutions to non-convex optimization problems. The idea behind DE is inspired by evolutionary genetics and is thus part of a class of so-called genetic algorithms.

DE searches the space of candidate solutions by initializing a population set of vectors, known as agents, within some region of the search space. These agents are then mutated (see Algorithm 1) into a new population set, or generation. The mutation operates via two mechanisms: a weighted combination and a random “crossover”.

At each generation, Algorithm 1 generates a candidate z to replace each agent y . In the mutation step, it chooses at random three agents a , b , and c to create a new trial agent \tilde{z} through the linear combination

$$\tilde{z} = a + F \cdot (b - c),$$

where $F \in [0, 2]$; see Figure 7. In the crossover step, the candidate vector z is constructed by randomly choosing some components chooses an additional vector d from the current population. The new vector z is constructed by randomly choosing some components \tilde{z} and others from an additional randomly-chosen agent d . If $J(z) < J(y)$, then z replaces y in the next generation.

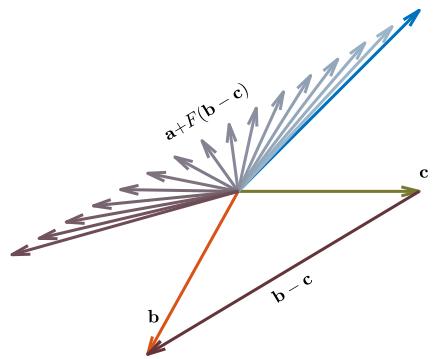


FIG. 7: Schematic of vectors used to construct the mutation function in Algorithm (1). The unlabeled vectors are the linear combinations $a + F(b - c)$, for $0.2 \leq F \leq 1.6$, used in the crossover defined in Algorithm (1) and used by Algorithm (2).

DE ensures that the objective functional J of the optimization problem decreases monotonically with (the optimal member of) each generation. As each iteration “evolves” into the next, inferior agents “inherit” optimal traits from superior agents via mutation, or else are discarded. After a sufficient number of iterations, the best vector in the final generation is chosen as the candidate solution global optimizer.

DE, and genetic algorithms more generally, belong to a class of optimization methods called metaheuristics. Although metaheuristics are useful for non-convex optimization problems, these methods do not guarantee the optimality of candidate solutions. Since the algorithm is stopped after a finite number of iterations, different random realizations return different candidate optimizers. As such, we use DE to search for candidate solutions and

use these candidates as initial conditions for a descent method which guarantees local optimality.

We show, in Figure 8, an example application of DE to minimizing MATLAB's peaks function

$$f_{\text{peaks}}(x, y) = 3(1-x)^2 e^{-(x^2-(y+1)^2)} - \frac{1}{3} e^{-(x+1)^2-y^2} - (2x - 10x^3 - 10y^5) e^{-x^2-y^2} \quad (30)$$

using DE. We see that an initial, random population of vectors converges to the globally optimal regions of the function f_{peaks} . At an intermediate generation, the population vectors compete between two local minima, yet the population vectors eventually converge collectively.

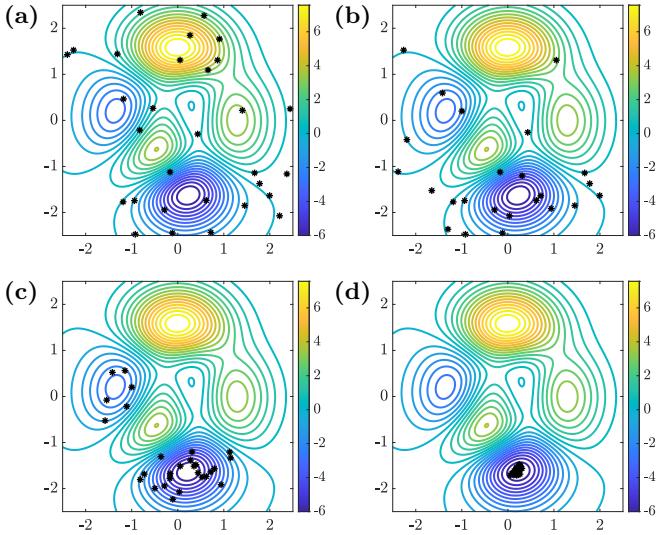


FIG. 8: Iterates of the DE algorithm applied to the peaks function (30): (a) The initial population. (b-d) The population after 1, 10, and 20 iterations. Parameters: $F = 0.6$ and $R_C = 0.9$; number of agents: $N_{\text{pop}} = 20$.

We provide a pseudocode of the general method in Algorithm 2. A more detailed discussion about DE and further implementation and benchmarking details can be found in the book by Storn, et al. [32].

We further demonstrate how DE overcomes non-convexity using a test problem which is much simpler to visualize than the higher dimensional optimal control problem (25). The Ackley function

$$f_{\text{Ackley}}(x, y) = -20e^{-0.2\sqrt{0.5(x^2+y^2)}} - e^{0.5(\cos 2\pi x + \cos 2\pi y)} + e + 20, \quad (31)$$

shown by Figure 9, is non-convex, w many local minima and a global minimum at the origin. Figure 9 shows the convergence of DE, as outlined in Algorithm 2, to the global minimum in less than 40 iterations using parameters $N_{\text{pop}} = 20$, $R_C = 0.9$ and $F = 0.8$

Using DE with the CRAB method requires drawing the coefficients ε_j from the uniform distribution on an appropriately constructed N -dimensional hyperrectangle. We

Algorithm 1: Differential Evolution Mutation

Result: A vector z mutated from agents in a given generation as required by the DE Algorithm (2).

Input: 4 distinct members a, b, c, d from the current generation of agents each with N components, the crossover ratio $R_C \in (0, 1)$, and weight $F \in (0, 2)$.

```

for  $j=1:N$  do
    Compute a random variable rand;
    if rand  $< R_C$  then
        |  $z[j] \leftarrow a[j] + F * (b[j] - c[j])$ 
    else
        |  $z[j] \leftarrow d[j]$ 
    end
end

```

Algorithm 2: Differential Evolution

Result: A vector likely to be globally optimal with respect to an objective J .

Input: A maximum number of iterations N_{max} , crossover ratio $R_C \in (0, 1)$ and weight $F \in (0, 2)$

```

while counter  $< N_{\text{max}}$  do
    Generate a population pop of  $N_{\text{pop}}$  vectors.
    for  $i = 1 : N_{\text{pop}}$  do
        CurrentMember  $\leftarrow$  Pop $i$ ;
        Choose three distinct vectors  $a_i, b_i, c_i$  different from the vector Pop $i$ ;
        Mutate  $a_i, b_i, c_i$ , and the CurrentMember into the mutated vector  $z$  using the mutation parameters  $R_C, F$  and Algorithm 1;
        if  $J(z) < J(\text{CurrentMember})$  then
            | TemporaryPop $i$   $= z$ ;
        end
    end
    Pop  $\leftarrow$  TemporaryPop;
    counter  $\leftarrow$  counter + 1;
end

```

choose the half-length of the j^{th} side of the hyperrectangle to decay quadratically a

$$l_j = \frac{u_T - u_0}{j^2}, .$$

We choose these coefficients to decay quadratically because the Fourier series of an absolutely continuous functions exhibits the same type of decay [33]. In this way, the search space of amplitudes ε_j is not severely restricted, yet the controls generated by the CRAB method remain technologically feasible throughout each generation.

B. The Local Method

For the local search, we use a line search strategy. We introduce here the basic ideas of a line search by discussing the simpler setting of the optimization of a

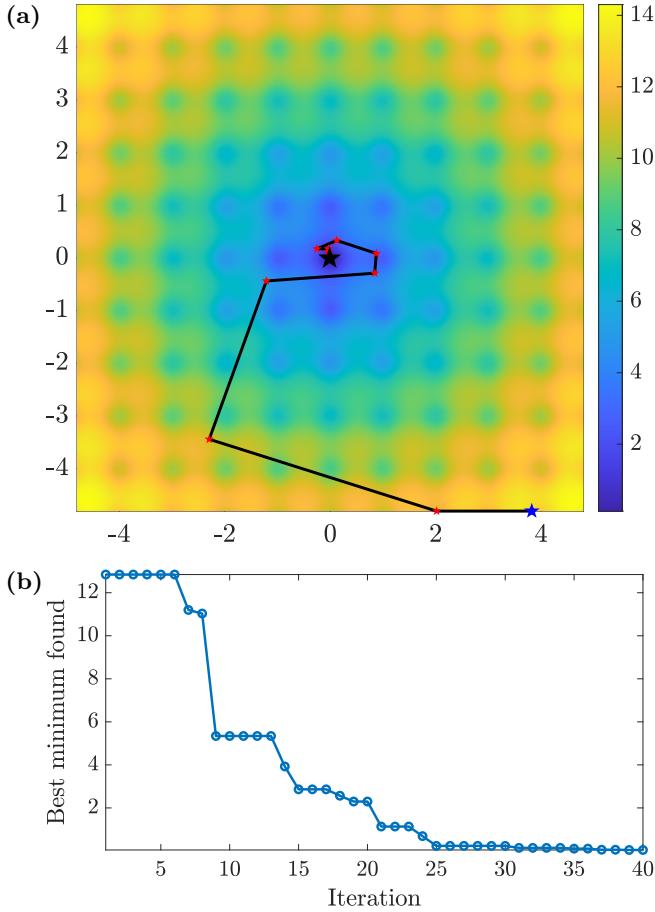


FIG. 9: Minimization of f_{Ackley} (31) using the evolutionary Algorithms 1 and 2. (a) False-color with the optimal member from each iteration of Algorithm 2 denoted by stars. (b) The value at the optimal member of each iteration.

smooth function on \mathbb{R}^n , i.e., $\min_{x \in \mathbb{R}^n} J(x)$. Line searches are iterative methods with two steps per iteration: first, identify a descent direction p_k , and then compute a step size α_k which determines how far x_k should move along p_k at the k^{th} iteration. Put simply, line searches determine p_k and α_k such that

$$J(x_{k+1}) := J(x_k + \alpha_k p_k) < J(x_k). \quad (32)$$

After a Taylor expansion of Inequality (32), we see that

$$\langle p_k, \nabla J(x_k) \rangle_{\mathbb{R}^n} + \mathcal{O}(\alpha_k) < 0.$$

For this to hold uniformly in α_k , we should choose the descent direction p_k such that $\langle p_k, \nabla J(x_k) \rangle_{\mathbb{R}^n} < 0$. The most natural choice is $p_k = -\nabla J(x_k)$, in which case the line search is called a **gradient descent**. Choosing $p_k = -H(x_k)^{-1}\nabla J(x_k)$, where H is the Hessian of J and is assumed to be positive-definite, yields a damped Newton-Raphson method.

The task of determining α_k remains. An *exact* line search chooses α_k to exactly minimize the subproblem

$$\min_{\alpha \in \mathbb{R}} J(x_k + \alpha p_k).$$

This is expensive, and it is usually better to allocate resources d toward computing better search directions p_k and to approximate the stepsize α_k rather than to determine it exactly. A reasonable approach to choosing α_k is to start with some large value and to then continually reduce it until some criteria is met. Observe that

$$J(x_k + \alpha_k p_k) = J(x_k) + \langle \alpha_k, p_k \nabla J(x_k) \rangle_{\mathbb{R}^n} + \mathcal{O}(\alpha_k^2).$$

This suggests it is reasonable to decrease α_k until

$$J(x_k + \alpha_k p_k) \leq J(x_k) + \langle \alpha_k p_k \nabla J(x_k) \rangle_{\mathbb{R}^n}. \quad (33)$$

This *inexact* line search is called backtracking, and Inequality (33) is called the Armijo-Goldstein condition.

We use the method of gradient descent since Newton's method requires a costly computation of J 's second derivatives. Of course, there are many other options to choose from, see e.g. [31], but for our purposes, the basic method of gradient descent with Armijo-Goldstein backtracking suffices. The last thing we require for the hybrid method is to generalize gradient descent from \mathbb{R}^n to an appropriate affine function space. Von Winckel and Borzi introduced the **Gradient Descent Pulse Engineering** (GRAPE) algorithm [34], which automatically preserves the boundary conditions of the admissible class \mathcal{U} for the optimal control problem (25). It has been used frequently in the quantum control literature; see, e.g., [2, 14, 19, 20]. The update in the GRAPE method is

$$u_{k+1} = u_k - \alpha_k \nabla_u \mathcal{L}|_{u=u_k}, \quad (34)$$

where the stepsize α is chosen using backtracking, and the Armijo-Goldstein condition for this problem reads

$$J[u_k - \alpha \nabla_u \mathcal{L}(u_k)] < J[u_k] - \frac{\alpha}{2} \|\nabla_u \mathcal{L}(u_k)\|_{L^2([0,T])}^2. \quad (35)$$

Until condition (35) is satisfied, the value of the stepsize α is decreased by some factor $\phi < 1$. Since the gradient descent (34) depends on the function space in which $\nabla_u \mathcal{L}(u)$ is to be understood, we review some basic facts about calculus on infinite-dimensional (affine) spaces.

The Gateaux derivative of a functional J , evaluated at a point $u \in \mathcal{U}$ in the direction of a displacement vector $v \in C_c^\infty([0, T])$ is defined by

$$d_u J[u; v] := \lim_{\varepsilon \rightarrow 0} \frac{J[u + \varepsilon v] - J[u]}{\varepsilon},$$

and if this exists for all admissible displacement vectors v , the functional J is said to be Gateaux differentiable. Given the uniform bound $\sup_{u \in \mathcal{U}} |\mathcal{L}(u)| \leq M$ for some

finite M , a direct calculation shows

$$\begin{aligned} d_u J[u; v] &= \lim_{\varepsilon \rightarrow 0} \frac{J[u + \varepsilon v] - J[u]}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left(\int_0^T \mathcal{L}(u + \varepsilon v) dt - \int_0^T \mathcal{L}(u) dt \right) \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^T \int_0^1 ds \mathcal{L}(u + s\varepsilon v) ds dt \\ &= \lim_{\varepsilon \rightarrow 0} \int_0^T \int_0^1 \mathcal{L}'(u + s\varepsilon v) v ds dt \\ &= \int_0^T \nabla_u \mathcal{L}(u) v dt := \langle \delta_u J, v \rangle_{L^2([0,T])}, \end{aligned}$$

using the bound on \mathcal{L} in order to invoke the Lebesgue dominated convergence theorem in the last equality.

The gradient of \mathcal{L} with respect to the $L^2([0,T])$ inner product can be identified with the functional derivative $\delta_u J$ calculated and expressed through the last entry in Equation (28), i.e., $\nabla_u \mathcal{L} = \delta_u J$ in the space $L^2([0,T])$. However, were one to perform a gradient descent on an initially admissible control u_k , the increment $\alpha_k \nabla_u \mathcal{L}|_{u=u_k}$ would fail to satisfy the boundary conditions and the updated function would leave the admissible set \mathcal{U} . We can avoid this problem by drawing the update from a more carefully chosen function space.

Since Taylor's theorem must hold for all sufficiently regular functionals on Hilbert spaces, the Taylor series

$$\begin{aligned} J[u + \varepsilon v] &= J[u] + \varepsilon d_u J[u, v] + \mathcal{O}(\varepsilon^2) \\ &= J[u] + \varepsilon \langle \nabla_u \mathcal{L}(u), v \rangle_X + \mathcal{O}(\varepsilon^2) \end{aligned}$$

holds term-by-term for all spaces X . The vWB method relies on choosing X to be the traceless and homogeneous Sobolev space $\dot{H}_0^1([0,T])$, i.e., the vector space of measurable functions, that vanish on the boundary of $[0,T]$ such that the norm $\|\star\|_{\dot{H}^1([0,T])}$ induced by the inner product

$$\langle \star, \star \rangle_{\dot{H}^1([0,T])} := \int_0^T (\partial_t \star)^\dagger (\partial_t \star) dt$$

is finite. This implies, by equating the Gateaux differential with respect to $L^2([0,T])$ and with respect to $\dot{H}_0^1([0,T])$,

$$\begin{aligned} \langle \delta_u J, v \rangle_{L^2([0,T])} &= \langle \nabla_u \mathcal{L}(u), v \rangle_{\dot{H}_0^1([0,T])} \\ &= \int_0^T \partial_t \nabla_u \mathcal{L} \partial_t v dt \\ &= - \int_0^T \partial_t^2 \nabla_u \mathcal{L} v dt \\ &= - \langle \partial_t^2 \nabla_u \mathcal{L}, v \rangle_{L^2([0,T])}, \end{aligned} \quad (36)$$

where integration by parts is used once along with the boundary conditions of v . Since this holds for all displacements $v \in C_c^\infty([0,T])$, we conclude that in order to

perform a gradient descent at the current control u , we must first solve the strong form of (36)

$$\partial_t^2 \nabla_u \mathcal{L} = -\delta_u J, \quad \nabla_u \mathcal{L}(0) = \nabla_u \mathcal{L}(T) = 0 \quad (37)$$

in order to determine the admissible gradient of the objective with respect to the control.

Note that boundary value problem (37) yields a control gradient with homogeneous Dirichlet boundary conditions. This implies that the use of an iterative method which uses this control gradient in an update automatically preserves the boundary conditions of the control, as desired. In order to solve the two-point boundary values problems for the control gradients, we use spectral methods such as Chebyshev collocation [33].

We provide a straightforward extension to the GRAPE method, appropriate for problems where, in addition, Neumann boundary data is specified for the admissible class \mathcal{U} . We encounter a problem of this type in subsection II A. The idea is to use the inner product on $\dot{H}_0^2([0,T])$, so that we are instead tasked with solving an inhomogeneous biharmonic equation with homogeneous boundary data:

$$\begin{aligned} \partial_t^4 \nabla_u \mathcal{L} &= \delta_u J, \\ \nabla_u \mathcal{L}(0) = \nabla_u \mathcal{L}(T) &= \partial_t \nabla_u \mathcal{L}|_{t=0} = \partial_t \nabla_u \mathcal{L}|_{t=T} = 0. \end{aligned}$$

Once again, the gradient $\nabla_u \mathcal{L}(u)$ preserves the appropriate boundary data when using a line search. In fact, it is also clear that the boundary value problem

$$\begin{aligned} \partial_t^{2p} \nabla_u \mathcal{L} &= (-1)^p \delta_u J, \\ \partial_t^j \nabla_u \mathcal{L}|_{t=0} = \partial_t^j \nabla_u \mathcal{L}|_{t=T} &= 0, \quad j = 0, 1, \dots, p-1, \end{aligned} \quad (38)$$

generalizes the GRAPE method to the space $\dot{H}_0^p([0,T])$, for $p \in \mathbb{Z}^+$. This method is summarized by Algorithm (3).

V. NUMERICAL RESULTS

We use the optimization methods of Section III in order to numerically solve the optimal control problem (25). We find that using the Tikhonov parameter $\gamma = 10^{-4}$ is sufficient to render the problem well-conditioned. The global CRAB/DE method, being a search method, provides no guarantee of reaching the global minimizer. Indeed, we have found several results that are competitive in minimizing the objective.

THE SQUEEZING PROBLEM

To demonstrate the necessity of the hybrid method, we first demonstrate what happens when we omit the first step, the global search. We apply the descent method, Algorithm (3) with $p = 1$, directly to the squeezing problem from Section II A in order to compute a local minimum downhill from the linear ramp used in Figure 4.

Algorithm 3: Gradient Descent Method in $\dot{H}_0^p([0, T])$.

Result: Admissible control u which is locally optimal with respect to the objective functional

Input: Initial admissible control u , the objective functional J , tolerance tol , maximum number of iterations Nmax , and reduction parameter $r \in (0, 1)$

```

while error > tol and counter < Nmax do
    Evolve the state variable  $(q, p)$  from  $t = 0$  to  $t = T$ , using Equations (26);
    Evolve the costate variables  $(\lambda, \mu)$  from  $t = T$  to  $t = 0$ , using Equation (28);
    Compute  $\nabla_u \mathcal{L}$  via Equation (38) with source term  $\frac{d}{dt} \frac{\partial \mathcal{H}}{\partial u}$  given by Equation (28);
    while Inequality (35) is false and  $\alpha > \text{tol}$  do
        |  $\alpha \leftarrow r\alpha$ ;
    end
    if  $\alpha < \text{tol}$  then
        | break;
    else
        |  $u \leftarrow u - \alpha \nabla_u \mathcal{L}$ ;
        | error  $\leftarrow J[u] - J[u + \alpha \nabla_u \mathcal{L}]$ ;
        | counter  $\leftarrow$  counter + 1;
    end
end

```

Figure 10, shows the locally optimal control, the locally optimal state dynamics, the corresponding numerical solution of the GPE (5), the corresponding phase portrait for the reduced dynamics, and convergence of GRAPE. The procedure is moderately successful, approximately halving the objective function, but it fails to eliminate a significant oscillation.

We now show the results of the full hybrid method. For the CRAB ansatz, we use 15 sine modes and an admissible linear ramp, setting

$$u_r(t) = u_0 + (u_T - u_0) \frac{t}{T} + \sum_{j=1}^{15} \varepsilon_j \sin\left(\frac{j\pi t}{T}\right) \quad (39)$$

We apply DE to determine effective coefficients ε_j , with parameters $F = 0.8$, $R_C = 0.9$, $N_P = 40$, and $N_{\text{max}} = 30$ in Algorithms (1) and (2).

Note that the value of \mathcal{H}_T depends on the quantity $\dot{u}|_{t=T}$. Since we are interested in the case that \mathcal{H} is constant for $t > T$, the minimum value of the Hamiltonian we are truly interested in is independent of any terms which depend on the derivative of the control. For this reason, we choose to minimize the Hamiltonian with \dot{u} set to 0 at $t = T$. We follow this with a GRAPE descent algorithm in $\dot{H}_0^2([0, T])$, i.e. Algorithm (3) with $p = 2$ in order to preserve both Dirichlet and Neumann data. This allows us to perform a line search for controls that minimize the modified Hamiltonian $\mathcal{H}_T|_{\dot{u}=0}$, rather than the full terminal Hamiltonian \mathcal{H}_T .

The hybrid method performs significantly better, as seen in Figure 11. The value of the terminal Hamiltonian $\mathcal{H}_T|_{\dot{u}=0}$ is one order of magnitude smaller than the

terminal Hamiltonian in Figure 10. To further characterize optimality, we compute the infidelity term from the Hohenester objective (3) between the computed solution of the GPE and the at time $t = T$, and ψ_d the minimizer of the GPE energy. Figure 11 shows that the hybrid method has reduced $J^{\text{infidelity}}$ by an order of magnitude compared the linearly controlled condensate.

We notice the coefficients c_0 and c_1 resulting from the hybrid method, shown in Figure 11, lose a fair amount of regularity at certain moments during the control process. For this reason, we show another, slightly less optimal, result in Figure 12 found by the same methodology, but where the dynamics are smoother. Note that the dynamics lose smoothness at the precise instant that u becomes very small. In a technological setting, these irregularities can be more systematically avoided by appending an inequality constraint to the admissible space \mathcal{U} , or by further using a Tikhonov regularization on the dynamics in the objective (25).

THE SPLITTING PROBLEM

Figure 13 shows similar results for the splitting problem of Section II B. All conventions used here are the same as those of Figure 5. As discussed in Section IV, we perform only the global optimization and not the descent method. Despite not applying a descent method, the global method significantly reduces the oscillations compared with the linear ramp control.

The squeezing problem takes an average of about thirty second on a 2.6 GHz 6-Core Intel i7 Macbook Pro, while the splitting problem takes 3-5 minutes. The splitting problem takes more time since for each evaluation of the objective in Equation (25) it must compute the costly inner products of Equations (20).

VI. CONCLUSION

We have demonstrated that reducing the GPE dynamics to a single non-autonomous degree of freedom Hamiltonian system and assuming a restricted form of the control is an effective and inexpensive approach to optimal control of two problems in the reshaping of a Bose-Einstein condensate. Moreover, we provide a complete characterization of the physics of controlled condensates using standard dynamical systems techniques. The techniques described here can be applied to other control problems constrained by Hamiltonian PDE, and perhaps to problems where posing an optimal control problem is challenging, if not impossible, without a visualization in a low dimensional setting.

Further refinements of this work can be pursued. This includes generalizing the form of the potentials shown in the GPE (5) so that, for example, the optimization is performed both over space and time. Also, a truncation of

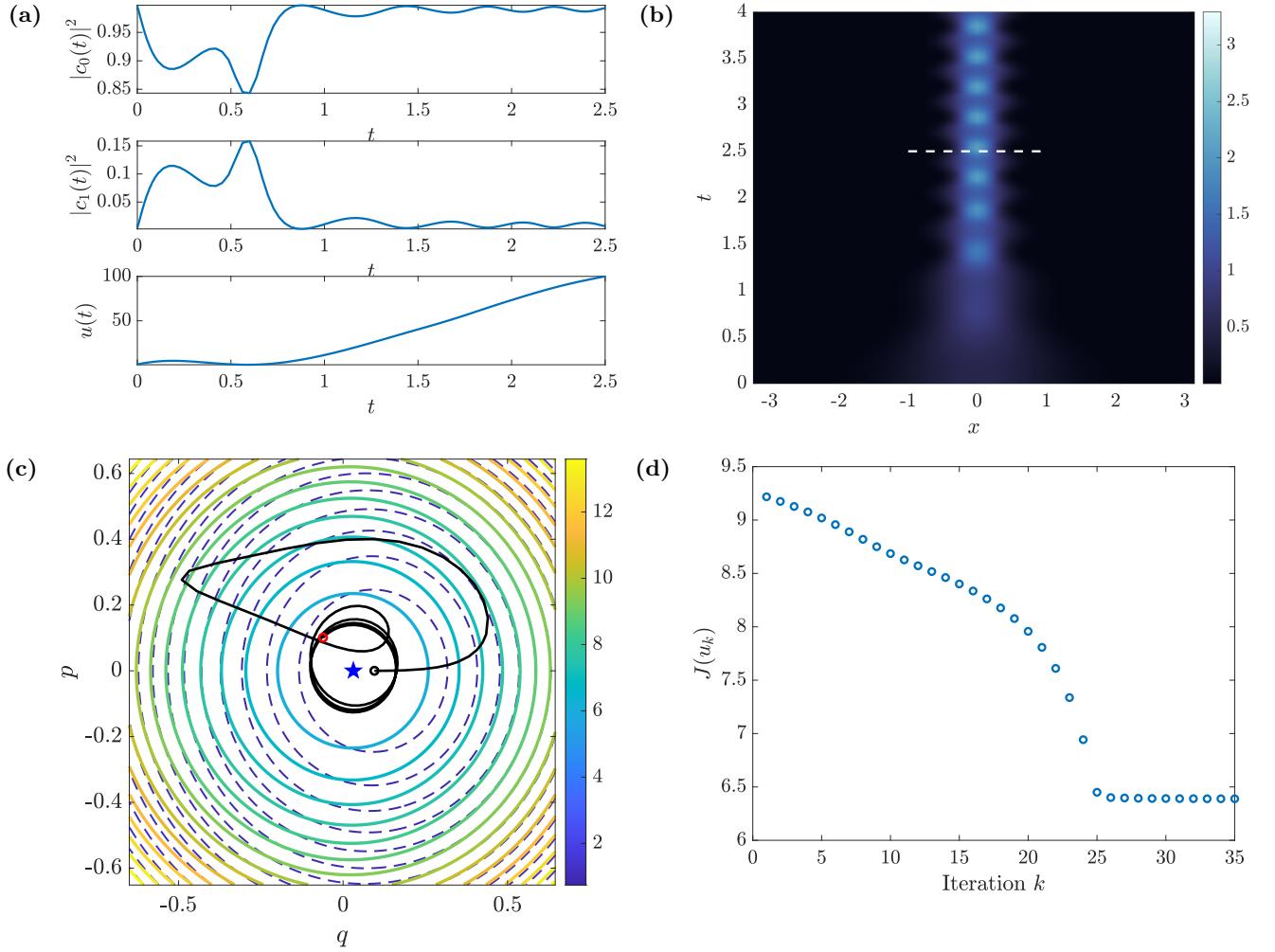


FIG. 10: The result of using the GRAPE algorithm of Section IV in the space $\dot{H}_0^1([0, T])$ with the linear ramp from Figure 4 as an initial control. Conventions used here are similar to conventions used in Figures 4 and 6. (a) The Galerkin coefficients which satisfy Equations (6) with optimal control u . (b) The numerical solution of the GPE with the optimal control u from Panel (a) up until the dashed white line. The persisting dynamics are computed with constant control $u(T)$. (c) The resulting phase portrait with the inset showing the persistent oscillation. (d) The convergence of GRAPE.

the Galerkin expansion (6) at a higher order can be pursued to refine the suppression of excitations that might have been missed by the order of the reduction used in this paper. While we have applied the Galerkin truncation to the GPE in one space dimension, the approach should still be applicable to problems in two or three dimensions. The speedup enabled by this reduction allows the use of DE, which, in turn, permits the exploration of

a large class of controls.

ACKNOWLEDGMENTS

We would like to thank Professors David Shirokoff, Christina Frederick, Alejandro Aceves, Richard Moore, and Panos Kevrekidis for their many helpful comments and discussions.

-
- [1] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen, *et al.*, *Eur. Phys. J. D* **69**, 1 (2015).
- [2] A. Borzi, G. Ciaramella, and M. Sprengel, *Formulation and Numerical Solution of Quantum Control Problems* (SIAM, Philadelphia, 2017).
- [3] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, *Science* **269**, 198 (1995).
- [4] C. C. Bradley, C. A. Sackett, J. J. Tollett, and R. G. Hulet, *Phys. Rev. Lett.* **75**, 1687 (1995).

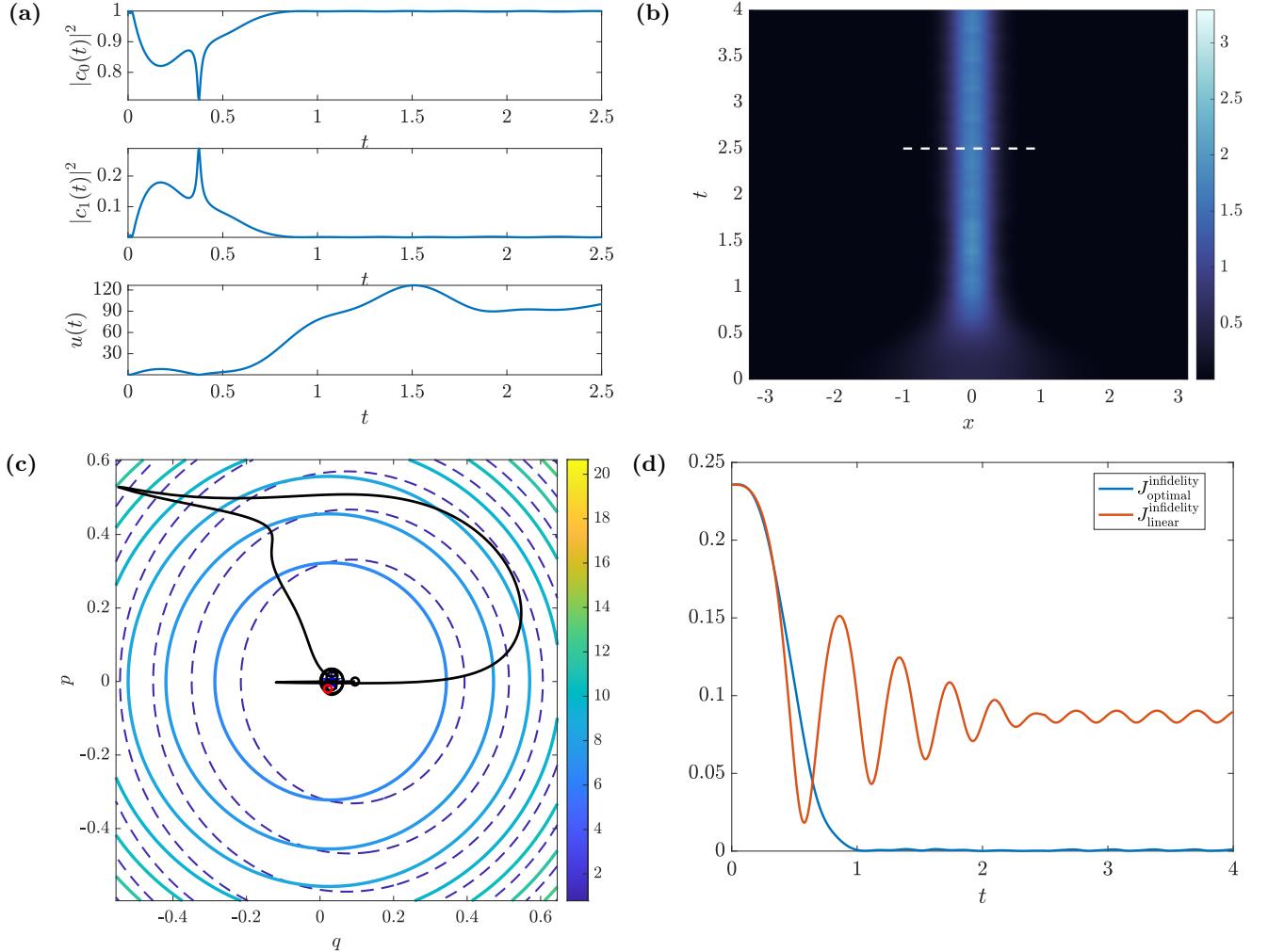


FIG. 11: The result of using the hybrid optimization technique outlined in Section IV on the squeezing problem of Subsection II A. The conventions used here are identical to those used in Figure 10. Panel (d) shows the infidelity (3) of the optimal control and infidelity of the linear control from Figure 4.

- [5] K. B. Davis, M. O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, *Phys. Rev. Lett.* **75**, 3969 (1995).
- [6] L. P. Pitaevskii and S. Stringari, *Bose-Einstein Condensation* (Clarendon Press, Oxford, 2003).
- [7] C. Gross, T. Zibold, E. Nicklas, J. Esteve, and M. K. Oberthaler, *Nature* **464**, 1165 (2010).
- [8] B. Lücke, M. Scherer, J. Kruse, L. Pezzé, F. Deuretzbacher, P. Hyllus, O. Topic, J. Peise, W. Ertmer, J. Arlt, *et al.*, *Science* **334**, 773 (2011).
- [9] M. F. Riedel, P. Böhi, Y. Li, T. W. Hänsch, A. Sinatra, and P. Treutlein, *Nature* **464**, 1170 (2010).
- [10] I. Bloch, J. Dalibard, and W. Zwerger, *Rev. Mod. Phys.* **80**, 885 (2008).
- [11] C. F. Ockeloen, R. Schmied, M. F. Riedel, and P. Treutlein, *Phys. Rev. Lett.* **111**, 143001 (2013).
- [12] T. Calarco, E. A. Hinds, D. Jaksch, J. Schmiedmayer, J. I. Cirac, and P. Zoller, *Phys. Rev. A* **61**, 022304 (2000).
- [13] D. Kielpinski, C. Monroe, and D. J. Wineland, *Nature* **417**, 709 (2002).
- [14] S. van Frank, M. Bonneau, J. Schmiedmayer, S. Hild, C. Gross, M. Cheneau, I. Bloch, T. Pichler, A. Negretti, T. Calarco, and S. Montangero, *Sci. Rep.* **6**, 1 (2016).
- [15] A. P. Peirce, M. A. Dahleh, and H. Rabitz, *Phys. Rev. A* **37**, 4950 (1988).
- [16] C. P. Koch, J. P. Palao, R. Kosloff, and F. Masnou-Seeuws, *Phys. Rev. A* **70**, 013402 (2004).
- [17] D. E. Kirk, *Optimal control theory: An introduction* (Courier Corporation, 2004).
- [18] C. Brif, R. Chakrabarti, and H. Rabitz, *New J. Phys.* **12**, 075008 (2010).
- [19] J. Mennemann, D. Matthes, R. Weishaupl, and T. Langen, *New J. Phys.* **17**, 113027 (2015).
- [20] U. Hohenester, P. K. Rekdal, A. Borzi, and J. Schmiedmayer, *Phys. Rev. A* **75**, 023602 (2007).
- [21] A. E. Bryson and Y.-C. Ho, *Applied optimal control: Optimization, estimation, and control* (Hemisphere Publishing Corporation, 1975).
- [22] E. J. McShane, SIAM (Society for Industrial and Applied Mathematics) *Journal on Control and Optimization* **27**, 916 (1989).

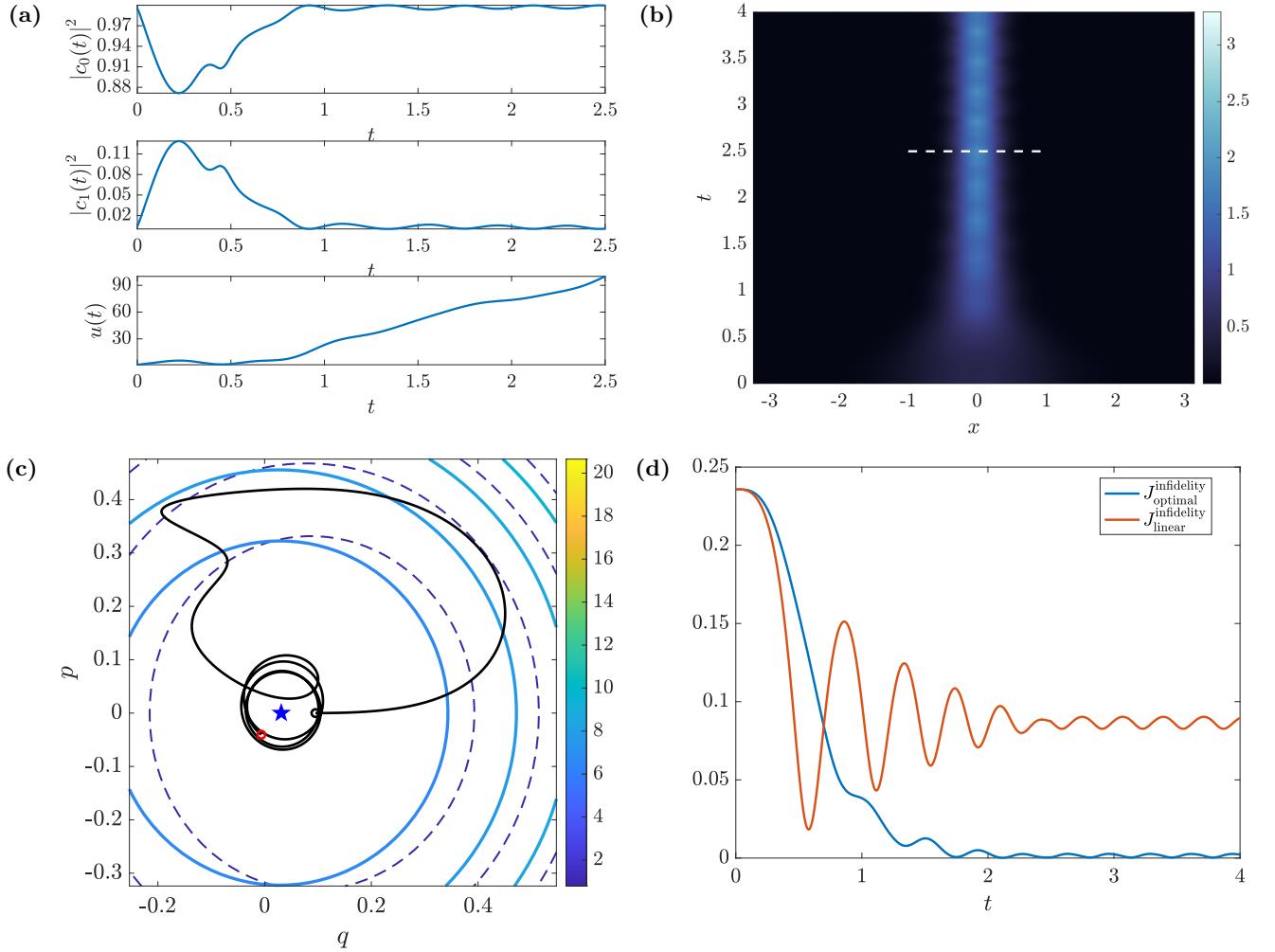


FIG. 12: Another result of using the hybrid optimization technique outlined in Section IV on the squeezing problem of Subsection II A. The conventions used here are identical to those used in Figure 10. Panel (d) shows the infidelity (3) of the optimal control and infidelity of the linear control from Figure 4.

- [23] M. Hintermuller, D. Marahrens, P. A. Markowich, and C. Sparber, SIAM J. Control Optim. **51**, 2509 (2013).
- [24] R. H. Goodman, J. Phys. A: Math. Theor. **44**, 425101 (2011).
- [25] R. H. Goodman, J. L. Marzuola, and M. I. Weinstein, Disc. Cont. Dyn. Sys. A **35**, 225 (2015).
- [26] P. Doria, T. Calarco, and S. Montangero, Phys. Rev. Lett. **106**, 190501 (2011).
- [27] T. Caneva, T. Calarco, and S. Montangero, Phys. Rev. A **84**, 022326 (2011).
- [28] R. Storn and K. Price, J. Global Optim. **11**, 341 (1997).
- [29] J. Viana-Gomes and N. M. R. Peres, Eur. J. Phys **32**, 1377 (2011).
- [30] J. J. W. H. Sørensen, M. O. Aranburu, T. Heinzel, and J. F. Sherson, Phys. Rev. A **98**, 022119 (2018).
- [31] S. Boyd and L. Vandenberghe, *Convex optimization*, Vol. 1 (Cambridge University Press, Cambridge, 2004).
- [32] K. Price, R. Storn, and J. Lampinen, *Differential Evolution: A practical approach to global optimization* (Springer-Verlag Berlin and Heidelberg GmbH & Co. KG, Berlin, 2018).
- [33] L. N. Trefethen, *Spectral methods in MATLAB* (SIAM, New York, NY, 2000).
- [34] G. von Winckel and A. Borzi, Inverse Probl. **24**, 034007 (2008).

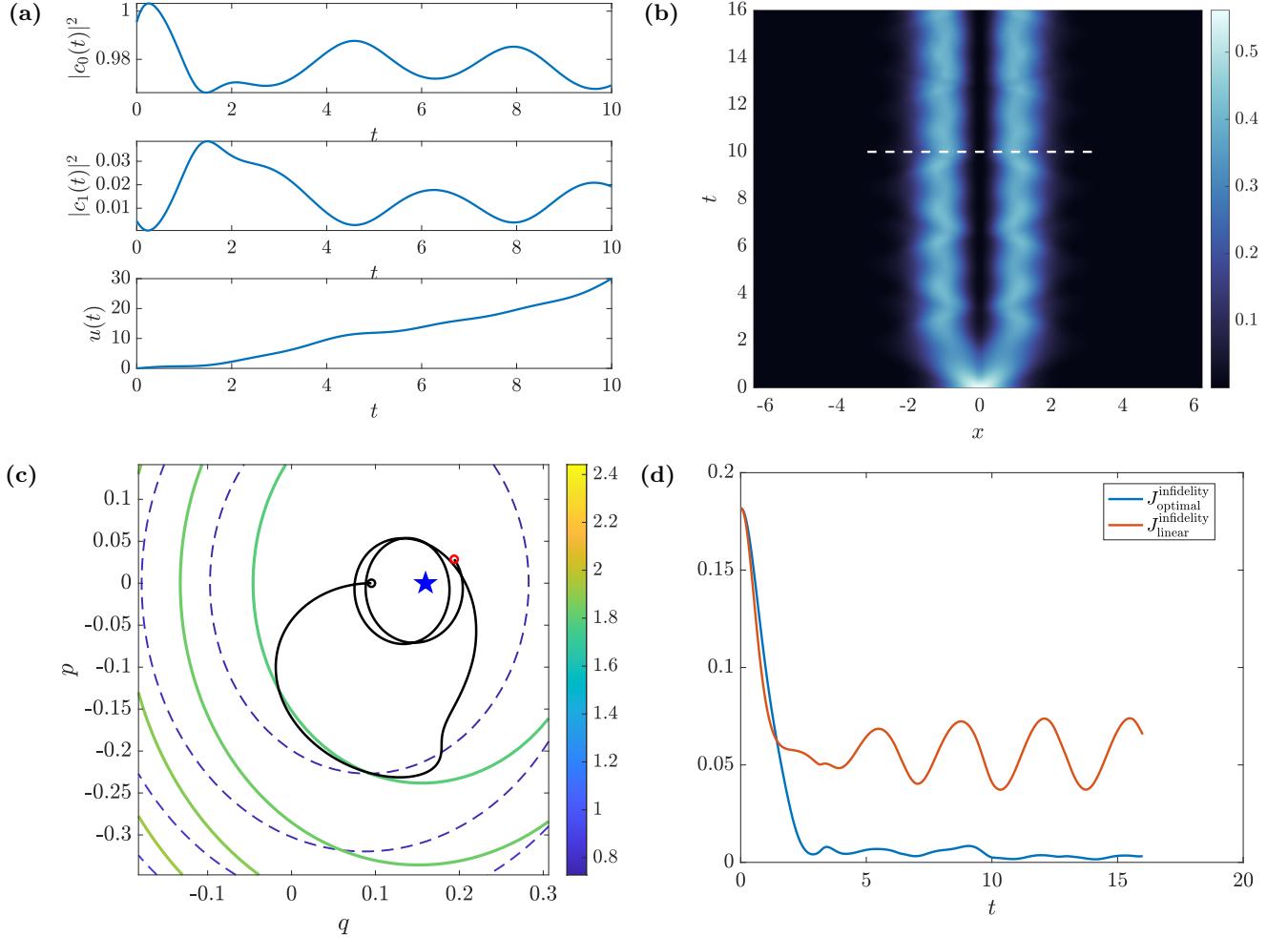


FIG. 13: The result of using the CRAB method on the splitting problem of Section II A. The same conventions of Figures 5 and 11 are used here.