Optimal control approach to gradient-index design for beam reshaping

J. Adriazola* [iD] **AND R. H. Goodman**

Department of Mathematical Sciences and Statistics, New Jersey Institute of Technology, University Heights, Newark, New Jersey, 07102, USA
Corresponding author: ja374@njit.edu

We address the problem of reshaping light in the Schrödinger optics regime from the perspective of the optimal control theory. In technological applications, Schrödinger optics is often used to model a slowly varying amplitude of a para-axially propagating electric field where the square of the waveguide's index of refraction is treated as the potential. The objective of the optimal control problem is to find the controlling potential which, together with the constraining Schrödinger dynamics, optimally reshapes the intensity distribution of Schrödinger eigenfunctions from one end of the waveguide to the other. This work considers reshaping problems found in work by Kunkel and Leger, and addresses computational needs by adopting tools from the quantum control literature. The success of the optimal control approach is demonstrated numerically. © 2022 Optica Publishing Group

1. INTRODUCTION

Humans have been reshaping light for thousands of years, and it remains an active research area to this day, from the ancient Assyrians' introduction of primitive lenses circa 750 BCE [1] to designs based on the sophisticated techniques of optimal transport [2]. Requiring a laser beam to have a specified irradiance distribution has diverse and broad applications that include laser/material processing, laser/material interaction studies, fiber injection systems, optical data image processing, and lithography [3]. Geometric optics is the simplest physical setting in which to study beam reshaping, and one that is often chosen. However, in the presence of diffractive effects, the wave nature of light must be accounted for, as is often the case in nanoscale optical technologies.

In designing gradient index (GRIN) waveguides and free-form optical systems, mathematical representations that afford computational tractability are critical. These representations must be employed due to the large number of design parameters that arise from a discretization of the continuous optical system. One design approach, called wavefront matching (WFM) [4,5] avoids this large number of parameters by using a multilens approach. Another approach, such as in [6,7] uses polynomial representations of the GRIN waveguide. Yet another approach, first used in this context by Kunkel and Leger [8,9], is called the phase retrieval method [10]. The common thread in these approaches is that some computational reduction is made; yet, they are all viable means to numerically construct GRIN optical waveguides that reshape laser beams in the presence of diffraction into light with a desired intensity distribution. An example of the type of reshaping problem that arises in applications is shown in Fig. 1.

A disadvantage of these previous methodologies is that generalizing them to either higher spatial dimensions or generalizing the dynamical constraints, such as including nonlinear effects, may be difficult. Indeed, despite achieving great success, Kunkel and Leger show several necessary adjustments must be made to adapt their previous methodology in two spatial dimensions [8] to the case of three [9]. On the other hand, optimal control theory, an extension of the calculus of variations [11,12], provides a more general alternative method.

The chief advantage of using optimal control theory is in its abstract framework, which easily handles entire classes of optimization problems at once, independent of its dimension or class of constraints. We use this control framework to solve optimization problems constrained by a one spatial dimensional linear equation to more easily showcase the method. We leave the straightforward extensions of the program developed here in handling either nonlinear constraints or a higher spatial dimensional setting as a subject for future study.

In this work, we pose an optimal control problem with an objective functional, first used in the context of high-fidelity quantum fluid manipulations by Hohenester *et al.* [13], constrained by the following standard model for paraxial light beam propagation. Consider an electromagnetic field propagating transversely through a linear waveguide, i.e., a waveguide through which the electrical field responds linearly to the polarization of the propagation media. Asssume the propagating field is time-harmonic, has negligible magnetic field components, and satisfies the hypothesis of the paraxial approximation; namely, that the direction of propagation does not deviate significantly from the axial direction defined by the waveguide. Then,
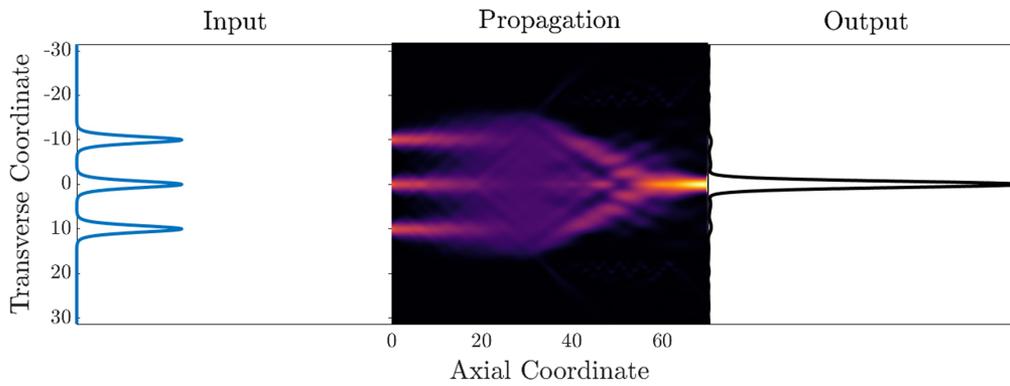
Fig. 1. Example of light reshaping in which three pulses of light are combined into one using the optimal control theory. More detail about the construction of the GRIN component, which combines these pulses, is provided throughout the paper.

one can show that Schrödinger's equation, in dimensionless form,

$$i\psi_z = -\frac{1}{2}\Delta\psi + V(x, z)\psi, \tag{1}$$

arises as a slowly varying amplitude approximation to the variable-coefficient Helmholtz equation [14].

Here, z is the axis of propagation, x is the transverse direction, Δ is the Laplacian in the transverse direction, $V(x, z)$ is proportional to the square of a spatially varying refractive index, and the wavefunction $\psi(x, z)$ is interpreted as a spatially varying complex electric amplitude. We assume the propagation media is lossless; hence, the potential V is a real function of the waveguide coordinates. The paraxial approximation is often studied because the numerical solution of Schrödinger's equation is significantly cheaper computationally, and easier to understand analytically, than the full numerical solution of either Helmholtz's or Maxwell's equations.

In posing the design problem, we must also make reductions, such as the works previously cited. We use the simplifying assumption that the potential can be written in the form $V(x, z) = V(x, u(z), v(z))$, where $V(x, u, v)$ is a two-parameter family of potentials. Thus, the design of a reshaping potential $V(x, z)$ is reduced to a search for 1D optimal controls. For ease of notation, we describe the method for the one-parameter family $V(x, u(z))$ and describe the necessary adjustments in numerical applications.

The light reshaping problem in this paper therefore is: Find the optimal control $u(z)$ that best transforms the intensity distribution of an initial Schrödinger state $\varphi_0(x)$ into the intensity distribution of the desired state $\varphi_d(x)$ satisfying

$$-\frac{1}{2}\Delta\varphi_0(x) + V(x, u(0))\varphi_0(x) = \lambda_0\varphi_0(x), \tag{2a}$$

$$-\frac{1}{2}\Delta\varphi_d(x) + V(x, u(l))\varphi_d(x) = \lambda_l\varphi_d(x), \tag{2b}$$

i.e., the initial and desired states are eigenfunctions, of the time-independent Schrödinger operator $P = -\frac{1}{2}\Delta + V(x, u(z))$, at $z = 0$ and at the end of a specified propagation length l, respectively. Thus, we formulate the problem of designing an optimal coupler between two waveguides with different transverse profiles and their eigenpairs (φ_0, λ_0) and (φ_d, λ_l).

A. Structure of the Paper

In Section 2, we precisely state the eigenfunction reshaping problem considered throughout this work. The problem is similar to quantum optimal control problems previously considered in the literature, e.g., [13,15,16]. We discuss, in detail, our assumptions about the control problem, and provide the optimality conditions given by the Euler–Lagrange equations [11].

In Section 3, we provide an overview of the numerical methods used to solve the control problem posed in Section 2. The procedure is a combination of a global, nonconvex method followed by a local, iterative method. In the context of numerical optimal control, this approach is called a **hybrid method** [17]. Hybrid optimization methods, when used appropriately, can overcome nonconvexity, yet still remain computationally efficient.

Doria *et al.* [18] and Caneva *et al.* [19] were the first to use the type of nonconvex method we use in this work. This method reduces the dimensionality of the control problem so that standard global search routines based on stochastic optimization can be used. Since stochastic methods come at the cost of slow convergence near local minima [20], local methods are then used to accelerate convergence toward the nearest minimum. The local method we use, by von Winckel and Borzì [21], is a gradient descent called GRAPE, which ensures that controls remain in the admissible search spaces used throughout this work.

In Section 4, we address many of the practical and computational aspects arising from the specific beam reshaping problems of interest. We use reductions that greatly simplify the computational complexity of the problem, and greatly aid in efficiently searching the space of reshaping potentials. The success of these reductions, together with the methods detailed in Section 3, is demonstrated numerically on the three reshaping problems, one of which is shown in Fig. 1.

B. Notation and Conventions

We make use of various function spaces when stating the optimal control problem. For example, general Banach spaces are denoted by $\mathcal{B}$. The Lebesgue space denoted by $L^p(\Omega)$, where Ω is a measurable set, is the equivalence class of measurable functions that agree almost everywhere, such that the norm

$$\| f\|_{L^p(\Omega)} := \left(\int_\Omega |f|^p \mathrm{d}\mu \right)^{\frac{1}{p}} \qquad (3)$$

is finite. Similarly, the Sobolev space $H^k(\Omega)$ is the space of k times weakly differentiable functions f, with respect to $x \in \Omega$, whose norm

$$\| f\|_{H^k(\Omega)} := \left(\sum_{j=0}^{k} \left\| \partial_x^j f \right\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \qquad (4)$$

is finite. The space of essentially bounded functions $L^\infty(\Omega)$ is the space where

$$\| f\|_{L^\infty(\Omega)} := \mathrm{ess} \sup_{x \in \Omega} |f(x)| < \infty. \qquad (5)$$

Homogeneous Sobolev spaces, denoted by $\dot{H}^k(\Omega)$, are the spaces of functions such that $\|\partial_x^k f\|_{L^2(\Omega)}$ is finite. A traceless Sobolev space, denoted $H_0^k(\Omega)$, is the space of functions in $H^k(\Omega)$ that vanish on the boundary $\partial\Omega$. The space of k times continuously differentiable functions is denoted $C^k(\Omega)$, and the space of essentially bounded $C^k(\Omega)$ functions is denoted by

$$C_b^k(\Omega) := C^k(\Omega) \cap L^\infty(\Omega). \qquad (6)$$

The notation $\mathcal{B}_1(\Omega_1; \mathcal{B}_2(\Omega_2))$ is understood as the space of functions f such that $\| f(\Omega_1, \cdot)\|_{\mathcal{B}_2(\Omega_2)} \in \mathcal{B}_1(\Omega_1)$. Spaces where each element is compactly supported on Ω are denoted by $\mathcal{B}_c(\Omega)$. Lastly, the notation $\dagger$ denotes Hermitian conjugation.

2. OPTIMAL CONTROL FRAMEWORK

The salient elements of the problem structure we consider are from the work of Hohenester *et al.* [13], which uses the objective functional

$$J = \frac{1}{2} \left(\| \varphi_d(\cdot) \|_{L^2(\mathbb{R}^n)}^4 - \left| \langle \varphi_d(\cdot), \psi(\cdot, l) \rangle_{L^2(\mathbb{R}^n)} \right|^2 \right)$$
$$+ \frac{\gamma}{2} \int_0^l |\partial_z u|^2 \mathrm{d}z, \qquad (7)$$

where $\gamma > 0$ and $z \in (0, l)$ is the axial coordinate, with $l > 0$.

The objective functional J involves the *infidelity*

$$J_{\mathrm{infidelity}} = \frac{1}{2} \left(\| \varphi_d(\cdot) \|_{L^2(\mathbb{R}^n)}^4 - \left| \langle \varphi_d(\cdot), \psi(\cdot, l) \rangle_{L^2(\mathbb{R}^n)} \right|^2 \right), \qquad (8)$$

which penalizes misalignments of the computed function $\psi(x, l)$ with respect to the desired state $\varphi_d(x)$. In the language of optimal control theory [12,22], the infidelity is called a terminal cost. This objective functional disregards the physically unimportant global phase difference between the desired and computed states, a significant advantage over the typical least-squares approach.

The second contribution to the objective, the running cost over $[0, l]$, is a *regularization* of the control function $u(z)$. This penalizes the use of control functions with large $\dot{H}([0, l])$ norms, and is well-known in the literature as a type of Tikhonov regularization [23]. The introduction of this regularization conditions the optimal control problem. Indeed,

Hintermuller *et al.* prove the control framework of Hohenester *et al.* is well-posed with the introduction of a Tikhonov regularization; i.e., there exists a control $u \in H^1([0, l])$ that minimizes the objective J [24].

The optimal control problem we consider in this paper is

$$\inf_{u \in \mathcal{U}} J, \qquad (9)$$

which is subject to Schrödinger's Eq. (1) with the initial and desired states φ_0 and φ_d satisfying Eq. (2). The search for optimal controls is performed over the admissible class $\mathcal{U} = \{u \in H^1([0, l]) : u(0) = u_0, u(l) = u_l\}$. We assume the eigenfunctions φ_d and φ_0 are both in the space $H^1(\mathbb{R}^n)$. We also assume that the eigenfunctions φ_0 and φ_d have unit intensity, i.e., $\| \varphi_0\|_{L^2(\mathbb{R}^n)} = \| \varphi_d\|_{L^2(\mathbb{R}^n)} = 1$, so that the infimum of the infidelity in Eq. (8) is 0. Lastly, we assume the potential $V(x, u(z))$ is in the space $C_b^0([0, l]; H^1(\mathbb{R}^n))$ for every $u \in \mathcal{U}$.

Note that with the above assumptions in place, the regularity of the wavefunction ψ solving Eq. (1) is known [25]: $\psi \in C^1([0, l]; H^1(\mathbb{R}^n))$. Moreover, the control problem with objective functional in Eq. (7) is well-posed for sufficiently large $\gamma > 0$ [24].

By using the method of Lagrange multipliers, we write the optimal control problem in Eq. (9) in unconstrained form as

$$\min_{u \in \mathcal{U}} J = \min_{u \in \mathcal{U}} \int_0^l \mathcal{L}(\psi, \partial_z \psi, \partial_x^2 \psi, \psi^\dagger, p^\dagger, u, \partial_z u) \mathrm{d}z, \quad (10)$$

where the Lagrange density is given by

$$\mathcal{L} = \mathrm{Re}\left\{ \left\langle p, i\partial_z \psi + \frac{1}{2}\Delta \psi - V(x, z)\psi \right\rangle - \langle \varphi_d, \psi \rangle \langle \partial_z \psi, \varphi_d \rangle \right\}$$
$$+ \frac{\gamma}{2} |\partial_z u|^2, \qquad (11)$$

with inner products understood in the sense of $L^2(\mathbb{R}^n)$, and where p is a Lagrange multiplier. It is straightforward to show, using standard arguments from the calculus of variations [11,26], that the optimality conditions of Eq. (10) are given by

$$2i\partial_z \psi = -\frac{1}{2}\Delta \psi + V(x, z)\psi, \quad \psi(x, 0) = \varphi_0(x), \quad \text{(12a)}$$

$$i\partial_z p = -\frac{1}{2}\Delta p + V(x, z)p, \quad ip(x, l) = \langle \varphi_d, \psi(x, l) \rangle_{L^2(\mathbb{R}^n)} \varphi_d, \qquad \text{(12b)}$$

$$\gamma \partial_z^2 u = -\mathrm{Re}\langle p, \partial_u V\psi \rangle_{L^2(\mathbb{R}^n)}, \quad u(0) = u_0, \ u(l) = u_l. \qquad \text{(12c)}$$

Equation (12b) is the adjoint equation of Eq. (12a) and governs the axial evolution of the Lagrange, or costate, multiplier p backward from its terminal condition at $z = l$. The similarity of Eqs. (12a) and (12b) is due to the self-adjoint nature of the Schrödinger operator $P = -\frac{1}{2}\Delta + V(x, z)$. Equation (12c) governs the optimal control u and, together with the boundary conditions defined through the admissible class $\mathcal{U}$, is a boundary value problem on $[0, l]$.

Equations (12a) and (12b) are both solved via a second-order Fourier split-step method, where the z-dependence of the

potential is handled by the midpoint method. We also note that Eq. (12c) will not be solved numerically, but will instead be reinterpreted in the context of the optimization method discussed in Section 3.B.

Consider the so-called reduced objective functional

$$\mathcal{J}:\mathcal{U} \to \mathbb{R}, \quad u \mapsto \mathcal{J}[u] := J\left[\psi(u), u\right]. \quad \textbf{(13)}$$

Let u^* denote an optimal control, and define $\psi^* := \psi(u^*)$, $p^* := p(u)$. Since the optimal control problem in Eq. (9) is well-posed, then for every $u \in \mathcal{U}$,

$$\mathcal{J}[u] \geq \mathcal{J}[u^*] = \min_{u \in \mathcal{U}} \mathcal{J}; \quad \textbf{(14)}$$

i.e., the minimum is attained by the optimal control u^*. In addition, at the minimum, the optimal triple (ψ^*, p^*, u^*) satisfies Eq. (12). For this reason, pursuing numerical approximations of Eq. (12) and the optimality condition in Eq. (14) when searching for the optimal control u^* is meaningful.

3. NUMERICAL OPTIMIZATION METHODS

To solve Eq. (9), we use a *hybrid* optimization method; a combination of a global, nonconvex method followed by a local, iterative method. The methodology we use in this paper is similar to one used by Sørensen *et al*. [17], and allows for the use of a global search routine based on stochastic optimization to overcome nonconvexity. Nonconvex objective functions may, of course, possess many local minima, and a global method seeks to efficiently search for a near-optimal one. By then feeding results from the global method into the local one, convergence near the local minimum is accelerated. We previously used this methodology in [27], and more specific details about the numerical optimization is provided there.

A. Global Method: Chopped Random Basis/Differential Evolution

The first step in the hybrid method is to use a Galerkin method, which reduces the complexity of the optimal control problem so that standard nonconvex nonlinear programming (NLP) techniques can be applied. This step relies on choosing controls from the span of an appropriately chosen finite set of basis functions so that the optimization is performed over a relatively small set of unknown coefficients. The choice of basis is such that controls remain in the appropriate admissible space $\mathcal{U}$ in the context of the control problem in Eq. (9).

We choose to use the representation

$$u_r(z) = \mathcal{P}(z; u_0, u_l, l) + \sum_{j=0}^{N-1} \varepsilon_j \varphi_j(z; l), \quad z \in [0, l], \quad \textbf{(15)}$$

where $\mathcal{P}$ is a fixed function satisfying the boundary conditions of Eq. (12c), $\varphi_j(z)$ is a basis function with vanishing boundary conditions, and the coefficients ε_j are parameters to be optimized over. It is clear that if the polynomial $\mathcal{P}$ and the basis functions φ_j are chosen well enough, then the control ansatz in Eq. (15) reliably simplifies the optimal control problem. An effective Galerkin approximation must be constructed the set of basis functions N simultaneously large enough to define

an accurate approximation, yet small enough so that the overall procedure remains computationally inexpensive. In this work, we use 15 basis functions. This reduces the optimization problem to a small-scale NLP problem that can be solved using standard techniques.

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 nonconvex optimization problems. The idea behind DE is a so-called genetic algorithm that draws inspiration from evolutionary genetics. DE searches the space of candidate solutions by initializing a population set of vectors, known as agents, within some chosen region of the search space. These vectors are then randomly mutated into a new population set, or generation.

The mutation operates via two mechanisms: a weighted combination and a "crossover" that randomly exchanges "traits," or elements, between agents. The method requires three parameters; the weight F, the crossover parameter R_C, and the size of the population N_{pop}. Throughout this work, we find that the parameters $F = 0.8$, $R_C = 0.9$, $N_{\text{pop}} = 40$ work well.

DE ensures that the objective functional $\mathcal{J}$ decreases monotonically with each generation. As each iteration "evolves" into the next, inferior vectors "inherit" optimal traits from superior vectors via mutations. DE only allows mutations that are more optimal with respect to $\mathcal{J}$ to pass to the next generation. After a sufficient number of iterations, the best vector in the final generation is chosen as the candidate solution most likely to be globally optimal with respect to an objective functional.

Genetic algorithms, which require very few assumptions about the objective functional, are part of a wider class of optimization methods called metaheuristics. Although metaheuristics are useful for nonconvex optimization problems, they do not guarantee the global optimality of candidate solutions. Since the algorithm is stopped after a finite number of iterations, different random realizations return different candidate optimizers. For this reason, we use DE to search for candidate solutions and use these candidates to generate initial controls, through the representation in Eq. (15), for a method that guarantees local optimality up to some threshold.

B. Local Method: Projected Gradient Descent

We use a line search strategy from von Winckel and Borzì called GRAPE [29]. The GRAPE method is an appropriate generalization of the well-known gradient descent method from $\mathbb{R}^n$ to an appropriate affine function space that automatically preserves the boundary conditions of the admissible class $\mathcal{U}$ mentioned in the context of an optimal control problem (9). This method has been frequently applied in the quantum control literature; see, for example, [13,15,17].

To describe the GRAPE method, recall the reduced, unconstrained form of the optimal control problem given by Eqs. (10) and (13). The method of gradient descent, in this context, is given by following iteration

$$u_{k+1} = u_k - \alpha_k \nabla_u \mathcal{L}|_{u=u_k}, \quad \textbf{(16)}$$

where the linear operator ∇_u is the gradient, or Fréchet derivative, of the Lagrangian $\mathcal{L}$ with respect to the control u.

The stepsize α_k is chosen adaptively via the Armijo–Goldstein condition [20].

Recall that the definition of a Fréchet derivative depends on the choice of function space in which it is to be understood. If the Fréchet derivative is understood in the sense of $L^2([0, l])$, then it can be identified with the functional derivative of the objective $\mathcal{J}$, which in this case can be shown to be

$$\delta_u \mathcal{J} = -\gamma \partial_z^2 u - \mathrm{Re}\langle p, \partial_u V \psi \rangle_{L^2(\mathbb{R}^n)}. \qquad \textbf{(17)}$$

This coincides with the Euler–Lagrange equation $\delta_u \mathcal{J} = 0$ given by Eq. (12c). If this choice is made, however, the increment $\alpha_k \nabla_u \mathcal{L}|_{u=u_k}$ would not in general satisfy the boundary conditions on the control u_k, and the updated control u_{k+1} would leave the admissible set $\mathcal{U}$. This problem is avoided by using a different function space X defining the operator ∇_u.

To this end, consider an arbitrary displacement $v \in C_c^\infty([0, l])$ and an arbitrary $\varepsilon > 0$. We know Taylor's theorem holds; i.e., the series

$$\mathcal{J}[u + \varepsilon v] = \mathcal{J}[u] + \varepsilon \langle \nabla_u \mathcal{L}(u), v \rangle_X + \mathcal{O}(\varepsilon^2) \qquad \textbf{(18)}$$

holds term-by-term independently of the choice of the Hilbert space X for sufficiently regular functionals $\mathcal{J}$. The GRAPE method chooses the function space $\dot{H}_0^1([0, l])$ for X. By equating the directional, or Gateaux, derivatives with respect to $L^2([0, l])$ and with respect to $\dot{H}_0^1([0, l])$, we see that

$$\langle \nabla_u \mathcal{L}, v \rangle_{L^2([0,l])} = \langle \delta_u \mathcal{J}, v \rangle_{L^2([0,l])} = \langle \nabla_u \mathcal{L}, v \rangle_{\dot{H}_0^1([0,l])}$$

$$:= \int_0^l \partial_z \nabla_u \mathcal{L} \partial_z v \, dz = -\langle \partial_z^2 \nabla_u \mathcal{L}, v \rangle_{L^2([0,l])}, \qquad \textbf{(19)}$$

where an integration by parts is used once along with the boundary conditions on v.

Since this holds for all displacements $v \in C_c^\infty([0, l])$, we conclude, by the fundamental lemma of the calculus of variations [11], the strong form of Eq. (19),

$$-\partial_z^2 \nabla_u \mathcal{L} = \delta_u \mathcal{J}, \quad \nabla_u \mathcal{L}(0) = \nabla_u \mathcal{L}(l) = 0, \qquad \textbf{(20)}$$

also holds. This renders an admissible gradient whose homogeneous Dirichlet conditions are induced by choosing increments specifically from the traceless space $\dot{H}_0^1([0, l])$. To solve the boundary value problem in Eq. (20) for the control gradient $\nabla_u \mathcal{L}$, we use Chebyshev collocation [30].

4. BEAM RESHAPING PROBLEMS

A. Beam Reshaping with the Pöschl–Teller Potential

To demonstrate the beam reshaping problem in Eq. (9) and its numerical solution in a simple setting, we consider the initial and terminal eigenfunctions for which $V(x, 0)$ and $V(x, l)$ can be computed in closed form. It is well-known that the so-called Pöschl–Teller potential,

$$V(x) = -\frac{s(s+1)}{2} \mathrm{sech}^2(x),$$

$s \in \mathbb{N}$, gives Legendre functions as eigenfunctions for the time-independent Schrödinger equation [31]. As a test, we consider the problem of reshaping the ground state eigenfunction for

$s = 1$ to the ground state corresponding to $s = 4$. We find that parametrizing the potential $V(x)$ with a "depth" control $u(z)$ and a "width" control $v(z)$ to be sufficient in our search for an optimal, axially varying Pöschl–Teller potential. More precisely, we assume the following form of the potential

$$V(x, u(z), v(z)) = -\frac{u(z)}{2} \mathrm{sech}^2(v(z)x), \qquad \textbf{(21)}$$

where the initial and terminal eigenfunctions are given by

$$\varphi_0(x) = -\frac{1}{\sqrt{2}} \mathrm{sech}(x), \quad \varphi_d(x) = -\frac{3}{2\sqrt{3}} \mathrm{sech}^2(x), \qquad \textbf{(22)}$$

and the appropriate control boundary conditions are

$$u(0) = 2, \quad u(l) = 20, \qquad \textbf{(23a)}$$

$$v(0) = 1, \quad v(l) = 1. \qquad \textbf{(23b)}$$

This assumption on $V(x, z)$ slightly changes the optimality condition in Eq. (12c) such that

$$2\gamma \partial_z^2 u = -\mathrm{Re}\langle p, V_0 \psi \rangle_{L^2(\mathbb{R}^n)}, \quad u(0) = u_0, \ u(l) = u_l, \qquad \textbf{(24a)}$$

$$\gamma \partial_z^2 v = -\mathrm{Re}\langle p, V_l \psi \rangle_{L^2(\mathbb{R}^n)}, \quad v(0) = v_0, \ v(l) = v_l \qquad \textbf{(24b)}$$

are now the appropriate Euler–Lagrange equations for the controls u and v, while the state and costate Eqs. (12a) and (12b) remain unchanged.

For this test problem, we choose the following domain and discretization parameters: $l = 3$, $l \times 2^7$ discretization points in z, and $x \in [-4\pi, 4\pi]$ with 2^{11} discretization points. To test the GRAPE method, the indirect optimization strategy of Subsection 3.B, we use linear controls u, v satisfying the boundary conditions in Eqs. (23a) and (23b) as an initial guess, as well as a small Tikhonov regularization parameter of $\gamma = 10^{-6}$. We show, in Fig. 2, that the beam reshaping has only marginally improved. Although we see that the gradient descent has converged, we conclude that the use of a global optimization method is generally necessary.

We now turn to the CRAB/DE method of Subsection 3.A. We use 70 iterations of DE and a sine series together with the admissible linear polynomial required by Eq. (15), i.e.,

$$w_r(z) = \sum_{j=1}^{15} \frac{r_w}{j^2} \sin\left(\frac{j\pi z}{l}\right) + (w_l - w_0)\frac{z}{l} + w_0, \qquad \textbf{(25)}$$

where w stands for either u or v. The amplitudes r_w are random variables drawn uniformly from $[-1, 1]$. We choose the coefficients $A_j = \frac{r_w}{j^2}$ to decay quadratically because the Fourier series of an absolutely continuous functions exhibits the same type of decay [30]. In this way, along with the relative smallness of the Tikhonov parameter, the search space for the optimal controls u and v is not severely restricted, yet candidate controls remain technologically feasible throughout each generation of DE.

We see that the CRAB method essentially solves the Pöschl–Teller beam reshaping problem on its own without assistance of a further gradient descent. Although not shown, an application of a gradient descent method does not improve the result to
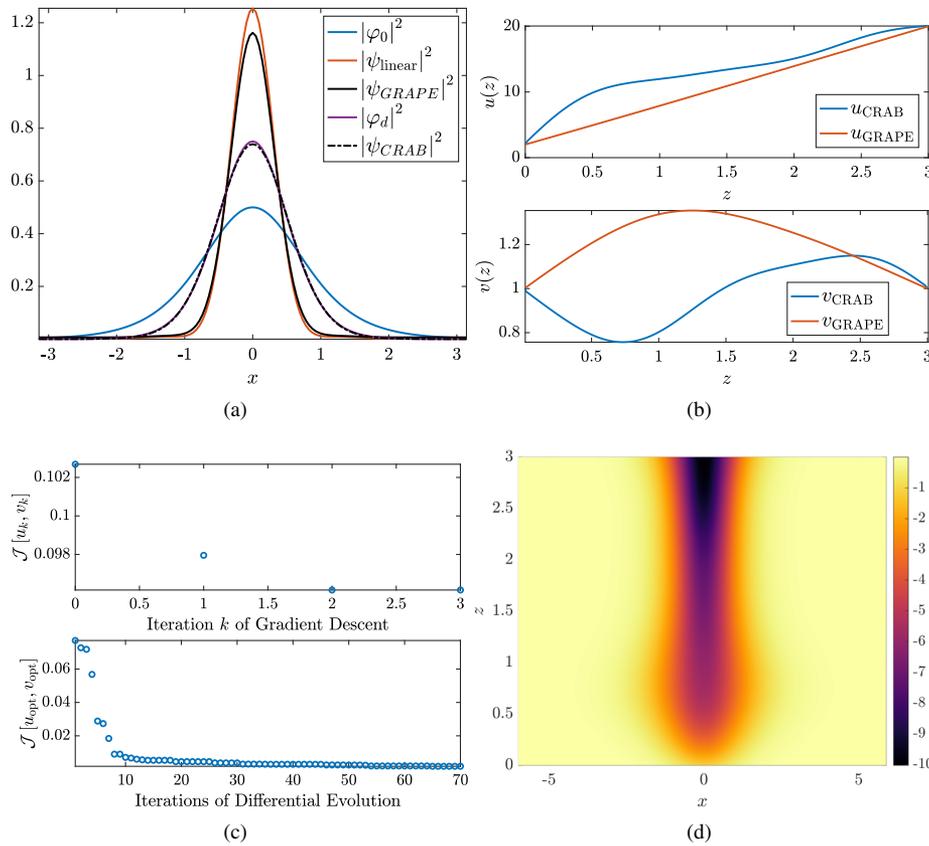
Fig. 2. Result of using the optimization methods in Sections 3.A and 3.B to reshape Pöschl–Teller eigenfunctions defined by Eq. (22). (a) Intensity profiles for the initial and desired eigenfunctions, a profile corresponding to linear controls, and profiles corresponding to the controls shown in (b), which are computed about linear controls via GRAPE and CRAB. (c) Local convergence of the descent method and the infidelity of optimal members from each iteration of differential evolution. (d) Potential $V(x, u(z), v(z))$, cf. Eq. (21), resulting from the CRAB method.

machine precision. For this reason, we conclude the CRAB result, also shown in Fig. 2, is a local minimum and extremely competitive among the many global minimizers that may exist for this optimal control problem. Although the CRAB method is extremely successful on its own for this problem, later reshaping problems will benefit from a refinement through the GRAPE method.

We note that the CRAB method comes at the cost of being much more computationally expensive than GRAPE. For this problem, DE requires $40 \times 70 = 2800$ solves of the Schrödinger Eq. (12a) while GRAPE requires, on average, two orders of magnitudes fewer solves of Eqs. (12a) and (12b). Still, on a standard 2.6 GHz 6-Core MacBook Pro, the CRAB/DE methodology for the reshaping problems in this paper only takes a few minutes.

We have one final comment about the stability of the optimal potential found via the CRAB method: The optimal potential is fairly sensitive to small changes in its defining amplitudes. Indeed, by randomly changing each coefficient within 1% of their relative values, this leads to a 5% average increase in the infidelity over 1000 simulations, while a 5% change in the amplitudes leads to an average increase of 20% in the infidelity. Since the two other reshaping problems considered later on in this paper are defined on larger domains, we suspect that the resulting optimal potentials will be even more sensitive to random perturbations. We do not confirm this, and instead

conclude from this simple stability study that if the physical construction of the GRIN optical elements shown through out this paper are carried out they should be done so with extreme precision.

B. Top Hat Problem

We now show how to solve two beam reshaping problems similar to those considered by Kunkel and Leger [8], with a transverse dimension $n = 1$, but by using the optimal control problem in Eq. (9). In the first problem, we transform the Pöschl–Teller eigenfunction, when $s = 1$, into the "top hat" mode

$$\varphi_{\text{tophat}} = A e^{-a x^m}, \tag{26}$$

where A is a normalization coefficient. For sake of computational demonstration, we choose $a = 10^{-3}$ and $m = 8$. The terminal potential $V_l(x)$, which has φ_{tophat} as its ground state mode, is computed via the least squares problem

$$\min_{V_l(x) \in H_b^1(\mathbb{R})} J = \min_{V_l(x) \in H_b^1(\mathbb{R})} \frac{1}{2} \left\| \varphi_{\text{tophat}}(x) - \varphi_d(x; V_l(x)) \right\|_{L^2(\mathbb{R})}^2 \tag{27}$$

subject to

$$-\frac{1}{2} \partial_x^2 \varphi_d(x) + V_l(x) \varphi(x) = \lambda_l \varphi_d(x). \tag{28}$$
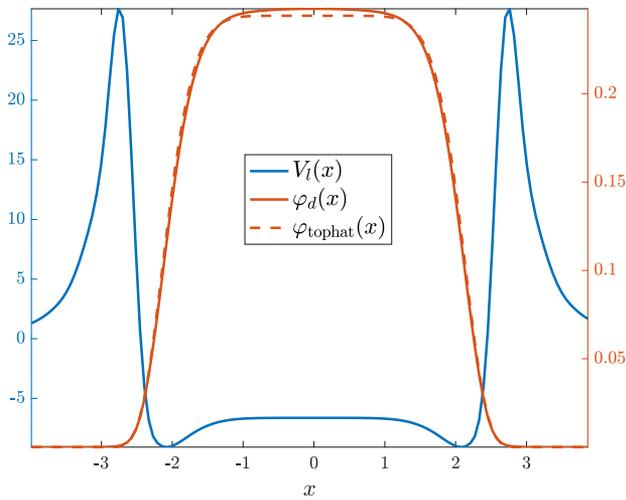
Fig. 3. Top hat potential $V_l(x)$ that solves the inverse scattering problem in Eq. (27) with a top hat eigenfunction (26). The computed eigenfunction $\varphi_d(x)$ is in solid red.

We show the resulting top hat potential $V_l(x)$ and eigenfunction $\varphi_d(x)$ from this procedure in Fig. 3. The computed eigenfunction $\varphi_d(x)$ is then used as a proxy for the true desired eigenfunction φ_{tophat} for the objective in Eq. (7) of the optimal control problem.

With $V_l(x)$ computed, we address the corresponding beam-reshaping problem. We reduce the search space of possible potentials by assuming they take the following form:

$$V(x, u(z), v(z)) = u(z)\,V_0(x) + v(z)\,V_l(x), \qquad \textbf{(29)}$$

where $u(l) = v(0) = 0$ and $u(0) = v(l) = 1$. We show the results of the optimal control problem using the hybrid method shown in Section 3 in Fig. 4. We fix $x \in [-5\pi, 5\pi]$, $z \in [0, 7]$, and use all other conventions consistent throughout previous sections.

C. Beam Addition Problem

Kunkel and Leger [8] also consider the problem of merging several pulses into one; cf. Fig. 1. To this end, we use an initial configuration of three seperated Pöschl–Teller potentials, each with $s = 1$, i.e.,

$$V_0(x) = -\left(\text{sech}^2(x - a) + \text{sech}^2(x + a) + \text{sech}^2(x)\right), \qquad \textbf{(30)}$$

$$\varphi_0(x) = -\frac{1}{\sqrt{6}}\left(\text{sech}(x - a) + \text{sech}(x + a) + \text{sech}(x)\right), \qquad \textbf{(31)}$$
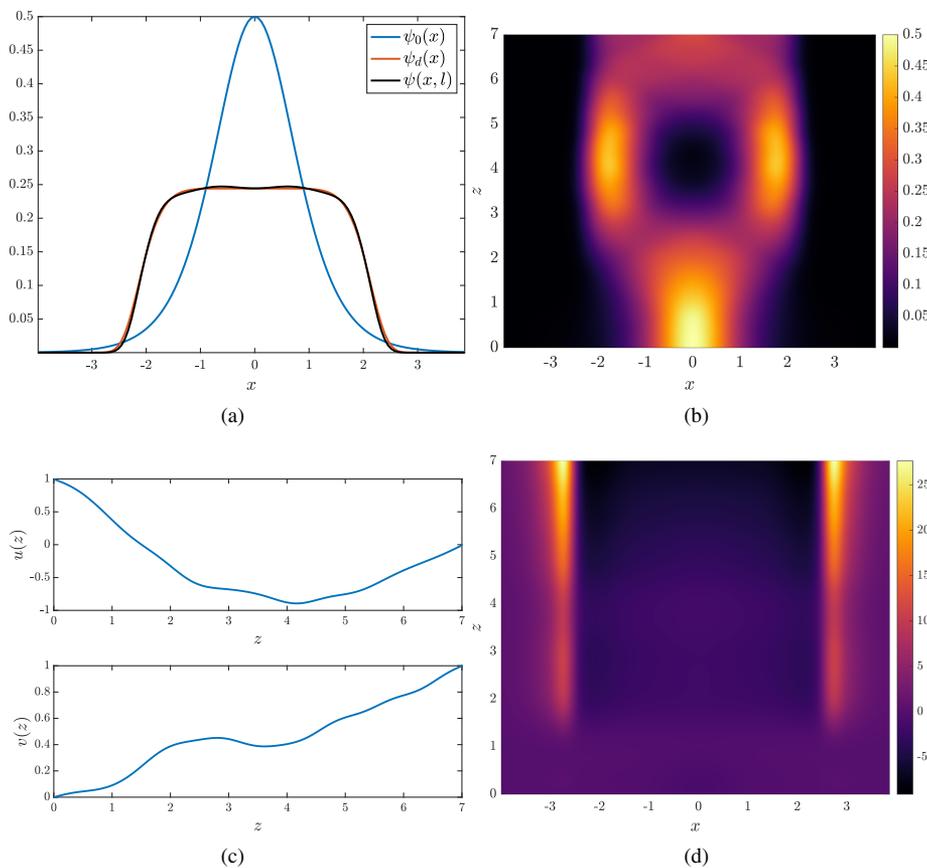
Fig. 4. Numerical solution of the top hat problem. (a) Intensity profiles for the initial, desired, and final computed wavefunctions. (b) Axial evolution of the wavefunction intensity. (c) Computed controls $u(z)$ and $v(z)$ resulting from the hybrid method. (d) Optimal potential resulting from (c) and the assumed form in Eq. (29).
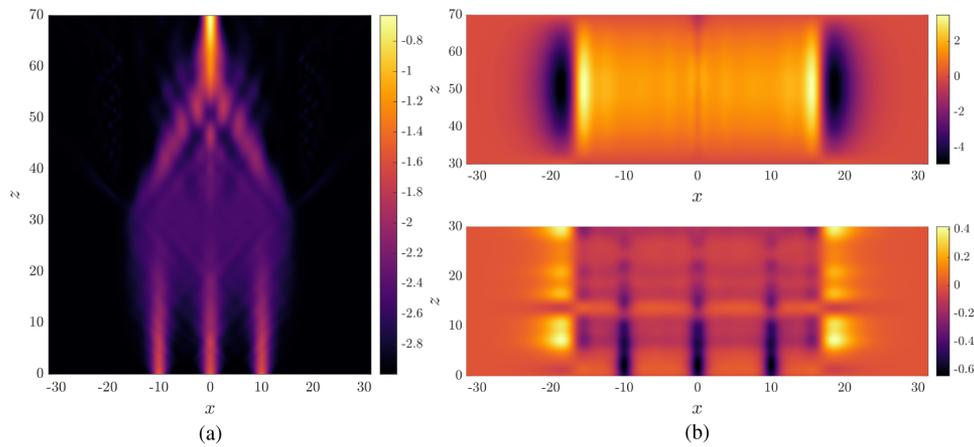
Fig. 5. (a) Schrödinger intensity distribution $|\psi|^2$ on a logarithmic scale. (b) Two stages of the computed optimal potential $V(x, z)$, shown separately because their ranges differ widely.

where the spacing parameter $a > 0$. Although $\varphi_0(x)$ is not exactly an eigenfunction of $V_0(x)$, it approximates an eigenfunction with improving accuracy as a is increased; we use $a = 10$.

We emulate Kunkel and Leger's strategy of partitioning the optimal control problem into two stages. In the context of this problem, we first perform an optimization on the interval [0,30], where we use $V_0(x)$ as an initial potential and use the top hat potential of Fig. 3 as the terminal potential. We then perform an optimization on the interval [30,70], where the terminal data, i.e., the terminal potential and resulting terminal wavefunction, is used as initial data and the now terminal potential is given by a single Pöschl–Teller potential with $s = 3$. Both stages of the optimization are performed using the hybrid method on potentials of the form in Eq. (29) with appropriate boundary conditions, to ensure the continuity of potentials across $z = 30$, and with the parameters γ and r_w the same as they were in Section 4.B.

We further refine our results by relaxing the restriction of the search space from the assumed form in Eq. (29) via a gradient descent on a wider space; that is, we perform a full 2D gradient descent on the potential $V(x, z)$ resulting from the two-stage optimization. To compute the gradient in this case requires a solution of the Dirichlet problem

$$\nabla^2_{x,z} \nabla_V \mathcal{J} = -\delta_V \mathcal{J}, \tag{32a}$$

$$\nabla_V \mathcal{J}|_{\partial \Omega} = 0, \tag{32b}$$

where the inhomogeneity is given by

$$-\delta_V \mathcal{J} = \gamma \nabla^2_{x,z} V + \mathrm{Re}\langle p, \psi \rangle_{L^2(\mathbb{R})}, \tag{33}$$

and where $\nabla^2_{x,z}$ is the Laplacian operator over x and z, and $\partial \Omega$ is the boundary of the computational domain $[-15\pi, 15\pi] \times [0, 70]$. This is the GRAPE method discussed in Subsection 3.B, in the space $\dot{H}^1_0(\Omega)$.

Note that the source term in Eq. (33) in the Poisson Eq. (32a) arises from Eq. (12c) and involves the computation of the Laplacian $\nabla^2_{x,z}$, which itself arises from the proper modification of the Tikhonov regularization in the objective in Eq. (7); i.e.,
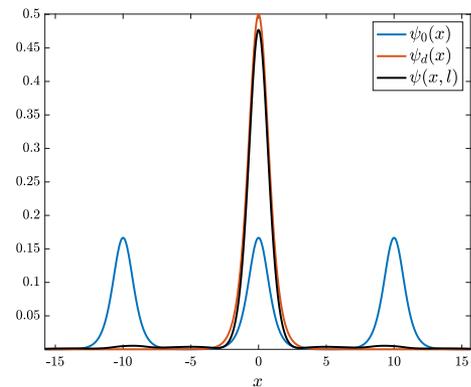
Fig. 6. Initial, desired, and final computed intensity profiles corresponding to Fig. 5.

the cost now also runs over the spatial dimension and penalizes large spatial derivatives of the reshaping potential V. We find this penalization is, on average, two orders of magnitude larger than penalizations that only run over the axial direction z, so we decrease the Tikhonov parameter to $\gamma = 10^{-8}$.

We show, in Figs. 5 and 6, the final result of the GRAPE method in $\dot{H}^1_0(\Omega)$, after inputting the optimal controls computed through the two-stage hybrid optimization strategy. Although the GRAPE refinement is only about 5% more optimal, with respect to the infidelity in Eq. (9) over the controls computed by CRAB, numerical convergence to the nearest local minimum on this wider space of potentials is guaranteed.

5. CONCLUSION AND FUTURE WORK

We have successfully applied the optimal control theory to the design of GRIN fibers that reshape a beam of light into a desired shape. To thoroughly, yet efficiently, search the space of possible designs, we use a combination of a Galerkin reduction of the control, a projected gradient descent method, product separability of the reshaping potentials of the form $V(x, u(z), v(z))$, a partitioning of the control into stages, and finally gradient descents on a wider space to reshaping potentials of the full form $V(x, z)$.

This methodology provides a systematic approach to the design process, but, of course, leaves further room for exploration. Moreover, the examples in this paper are proofs of concept in that we have only applied the methods to waveguides with a single transverse dimension, whereas the phase retrieval method has now been applied to waveguides with two transverse directions [9]. Future work may include extending the methods of this paper to higher dimensions. Fortunately, this extension is straightforward by virtue of the optimal control framework.

REFERENCES

1. The British Museum, "The Nimrud lens," https://www.britishmuseum.org/collection/object/W_-90959 (750BC–710BC).
2. Z. Feng, B. D. Froese, R. Liang, D. Cheng, and Y. Wang, "Simplified freeform optics design for complicated laser beam shaping," Appl. Opt. **56**, 9308–9314 (2017).
3. F. M. Dickey, L. S. Weichman, and R. N. Shagam, "Laser beam shaping techniques," Tech. Rep. (Sandia National Labs, 2000).
4. S. D. Campbell, J. Nagar, and D. H. Werner, "Multi-element, multi-frequency lens transformations enabled by optical wavefront matching," Opt. Express **25**, 17258–17270 (2017).
5. J. Nagar, D. Brocker, S. Campbell, J. Easum, and D. Werner, "Modularization of gradient-index optical design using wavefront matching enabled optimization," Opt. Express **24**, 9359–9368 (2016).
6. T. Yang, N. Takaki, J. Bentley, G. Schmidt, and D. T. Moore, "Efficient representation of freeform gradient-index profiles for non-rotationally symmetric optical design," Opt. Express **28**, 14788–14806 (2020).
7. J. P. Rolland, M. A. Davies, T. J. Suleski, C. Evans, A. Bauer, J. C. Lambropoulos, and K. Falaggis, "Freeform optics for imaging," Optica **8**, 161–176 (2021).
8. W. M. Kunkel and J. R. Leger, "Gradient-index design for mode conversion of diffracting beams," Opt. Express **24**, 13480–13488 (2016).
9. W. M. Kunkel and J. R. Leger, "Numerical design of three-dimensional gradient refractive index structures for beam shaping," Opt. Express **28**, 32061–32076 (2020).
10. J. R. Fienup, "Phase retrieval algorithms: a comparison," Appl. Opt. **21**, 2758–2769 (1982).
11. I. Gelfand and S. Fomin, *Calculus of Variations* (Prentice-Hall, 1963).
12. E. J. McShane, "The calculus of variations from the beginning through optimal control theory," SIAM J. Control Optim. **27**, 916–939 (1989).
13. U. Hohenester, P. K. Rekdal, A. Borzi, and J. Schmiedmayer, "Optimal quantum control of Bose-Einstein condensates in magnetic microtraps," Phys. Rev. A **75**, 023602 (2007).
14. J. W. Goodman, *Introduction to Fourier Optics*, 4th ed. (W.H. Freeman & Co./Macmillan Learning, 2017).
15. J. Mennemann, D. Matthes, R. Weishaupl, and T. Langen, "Optimal control of Bose-Einstein condensates in three dimensions," New J. Phys. **17**, 113027 (2015).
16. S. van Frank, M. Bonneau, J. Schmeidmayer, S. Hild, C. Gross, M. Cheneau, I. Bloch, T. Pichler, A. Negretti, T. Calarco, and S. Montangero, "Optimal control of complex atomic quantum systems," Sci. Rep. **6**, 1–12 (2016).
17. J. J. W. H. Sørensen, M. O. Aranburu, T. Heinzel, and J. F. Sherson, "Quantum optimal control in a chopped basis: Applications in control of Bose-Einstein condensates," Phys. Rev. A **98**, 022119 (2018).
18. P. Doria, T. Calarco, and S. Montangero, "Optimal control technique for many-body quantum dynamics," Phys. Rev. Lett. **106**, 190501 (2011).
19. T. Caneva, T. Calarco, and S. Montangero, "Chopped random-basis quantum optimization," Phys. Rev. A **84**, 022326 (2011).
20. S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University, 2004).
21. A. Borzi, G. Ciaramella, and M. Sprengel, *Formulation and Numerical Solution of Quantum Control Problems* (Society for Industrial and Applied Mathematics (SIAM), 2017).
22. A. E. Bryson and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control* (Hemisphere, 1975).
23. A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems* (Springer, 1995).
24. M. Hintermuller, D. Marahrens, P. A. Markowich, and C. Sparber, "Optimal bilinear control of Gross-Pitaevskii equations," SIAM J. Control Optim. **51**, 2509–2543 (2013).
25. A. Maspero and D. Robert, "On time dependent Schrödinger equations: Global well-posedness and growth of Sobolev norms," J. Funct. Anal. **273**, 721–781 (2017).
26. T. Witelski and M. Bowen, *Methods of Mathematical Modeling*, Spring Undergraduate Mathematics Series (2016).
27. J. Adriazola and R. H. Goodman, "Reduction-based strategy for optimal control of Bose-Einstein condensates," Phys. Rev. E **105**, 025311 (2022).
28. R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," J. Global Optim. **11**, 341–359 (1997).
29. G. von Winckel and A. Borzi, "Computational techniques for a quantum control problem with H^1-cost," Inverse Problems **24**, 034007 (2008).
30. L. N. Trefethen, *Spectral Methods in MATLAB* (Society for Industrial and Applied Mathematics (SIAM), 2000).
31. G. Pöschl and E. Teller, "Bemerkungen zur Quantenmechanik des anharmonischen Oszillators," Z. Phys. **83**, 143–151 (1933).