Apodizer Design to Efficiently Couple Light into a Fiber Bragg Grating

J. Adriazola, R. Goodman

June 8, 2022

Abstract

We provide an optimal control framework for efficiently coupling light in a bare fiber into Bragg gratings with an appreciable Kerr nonlinearity. The light-grating interaction excites gap solitons, a type of localized nonlinear coherent state which propagates with a central frequency in the forbidden band gap, resulting in a dramatically slower group velocity. Due to the nature of the band gap, a substantial amount of light is back-reflected by the grating's strong reflective properties. We optimize, via a projected gradient descent method, the transmission efficiency of previously designed nonuniform grating structures in order to couple more slow light into the grating. We further explore the space of possible grating designs, using genetic algorithms, along with a previously unexplored design parameter: the grating chirp. Through these methods, we find structures which couple a greater fraction of light into the grating with the added bonus of creating slower pulses.

1 Experimental and Technological Context

Dramatic slowing of light has been observed in a wide variety of experimental settings over the past two decades. This phenomenon offers enticing technological applications including efficient optical switches, sensitive interferometry, and optical quantum memory [9]. Of the several experimental platforms which can generate slow light, fiber Bragg gratings (FBGs) offer the considerable advantage of having structural properties which can be tailored specifically to the characteristics of an incoming light source. Already, modern optical communications systems use FBGs as notch filters and as components in optical add-drop multiplexers [3]. Optical fibers transmit information as coherent pulses of light that must be manipulated or redirected as they travel. Therefore, technology that can significantly slow down, or even halt, light is highly desirable.

An FBG is an optical element whose index of refraction varies periodically, see Figure 1. The grating enables the strong dispersion of light over a short distance due to a resonance between the grating's period Λ and electromagnetic wavelengths near the Bragg wavelength $\lambda_{\mathrm{B}} = 2\Lambda$. In so-called chirped gratings, discussed in further detail in Section 2.1 and shown in Figure 1, the period Λ is spatially varying.

The effect of an FBG is to strongly couple forward and backward propagating waves near the resonant wavelength. This creates a photonic bandgap, i.e., an interval of frequencies at

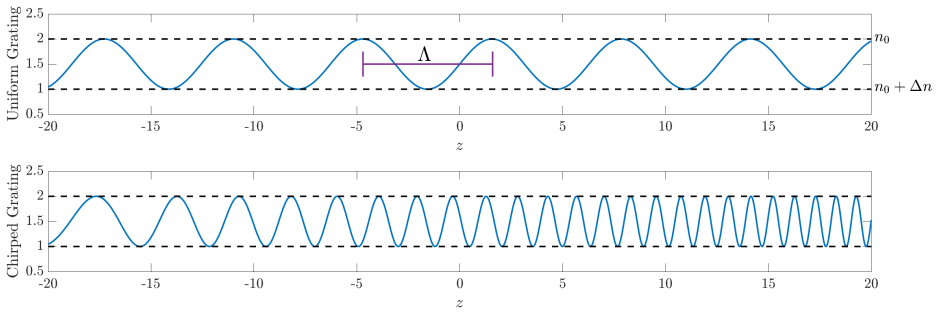


Figure 1: Top: Schematic showing an index of refraction $n(z)$ that periodically varies between a reference index $n_0 = 1$ to $n_0 + \Delta n = 2$. Bottom: Schematic of a chirped grating.

which low-amplitude light is reflected and unable to propagate. This bandgap is centered at the Bragg angular frequency $\omega_B = \frac{\pi c}{\bar{n}\Lambda}$, where c is the speed of light in vacuum.

In materials with an appreciable Kerr nonlinearity, i.e., materials whose index of refraction responds proportionally to the electric field intensity [17], high intensity light shifts the Bragg frequency. For positive nonlinearities, the refractive index increases with intensity, thereby shifting the bandgap downward. Therefore, at the edges of a high intensity region of light, the light is continuously Bragg reflected into the high intensity region which propagates through the grating seemingly unimpeded. Systems, such as these, which exhibit coherent structures arising from a balance between nonlinear effects and dispersion typically support solitary waves.

Indeed, Aceves and Wabnitz constructed a two-parameter class of solitary wave solutions [1], often called Bragg solitons to distinguish them from the classical notion of a soliton [18]. These waves solve evolution equations derived from Maxwell's equations using coupled-mode theory, which we briefly discuss in Section 2, and can travel with a speed anywhere from the speed of light in the medium down to zero. The existence of Bragg solitons demonstrates the possibility of slow light in an FBG. While Bragg solitons can in theory propagate at slow speeds, it is difficult to initialize such waves experimentally: to create a Bragg soliton, one must input light at a frequency inside the bandgap, and such frequencies are strongly reflected. Neglecting nonlinear effects, the FBG essentially acts as a band-stop filter, reflecting wavelengths whose frequency is within the bandgap.

To overcome this, Mok, et al. [16], use a two-pronged strategy to couple light into an FBG. First, they use an apodized grating, i.e., the grating strength is ramped up gradually from zero. Secondly, they input so-called out-gap solitons, wave packets with a mean frequency outside of the bandgap. As a result of the apodization, the light coupled into the grating has its frequency gradually shifted into the bandgap. Although this experiment is the first of its type and generated a pulse with a group velocity 16% that of light in glass, this pulse contained only about 20-30% of the the input energy, while the remaining light was reflected, rendering the setup highly inefficient.

In order to address this inefficiency, Rosenthal and Horowitz [20], designed a two-segment apodization function that allowed the creation of a pulse that retains about 68% of the incident energy and with a speed roughly 3.2% of light speed. Despite this remarkable improvement in efficiency, the authors provide limited mathematical detail about the process

behind discovering such a design. This leaves room for quantitative investigations into the efficiency of generating slow light using FBGs.

In this article, we improve on the method of Rosenthal and Horowitz by mathematically formulating an appropriate optimal control problem whose objective is to design an apodization profile that maximizes the coupling of light into the FBG. By including, an addition, a spatially varying chirp profile in the optimizaiton problem, we achieve a higher coupling efficiency of 82.6% transmission, while further reducing the pulse speed to about 0.5% the speed of light in glass.

This work is organized as follows: in Section 2, we provide the physical model and give precise details of past numerical experiments we aim to improve upon. We attempt to gain intuition behind the experiment by fitting the numerical data to the Aceves and Wabnitz waveform. In Section 3, we formulate the optimal control problem which seeks to maximize the energy transmitted into the fiber while treating the grating structure as the control. We provide the necessary optimality conditions for the control problem, design grating structures using numerical methods discussed in Appendix A, and present numerical results in Section 4, wrapping up in Section 5.

2 The Physical and Numerical Model

2.1 Brief Overview of Coupled-Mode Theory

The evolution of an electric field propagating in an optical fiber can be effectively modeled by the one dimensional nonlinear wave equation

$$\partial_\tau^2 \left(n^2(z, E^2)E\right) = \partial_z^2 E, \tag{1}$$

in dimensionless units where the speed of light $c = 1$, and where z denotes the axial direction of propagation. In the context of Figure 1, we choose the reference index of refraction $n_0 = 1$. Let ε be a small contrast of the index of refraction n so that it can be modeled, following [11], as

$$n = 1 + \varepsilon \left(\nu(\varepsilon z)\cos(2k_{\mathrm{B}} z + 2\Phi(\varepsilon z)) + |E|^2\right). \tag{2}$$

Here the coefficient ν describes the strength of the grating and Φ' describes the chirp, i.e., the local modulation of the grating's wavelength. The final term describes a Kerr nonlinearity with small Kerr coefficient ε.

Using the method of multiple scales with the ansatz,

$$E = \sqrt{\varepsilon}\left(u(\varepsilon z, \varepsilon\tau)e^{i(k_{\mathrm{B}}(z-\tau)+\Phi)} + v(\varepsilon z, \varepsilon\tau)e^{-i(k_{\mathrm{B}}(z+\tau)+\Phi)}\right) + \mathcal{O}\left(\varepsilon^{\frac{3}{2}}\right) \tag{3}$$

and letting $x = \varepsilon z$, $t = \varepsilon\tau$, denote the slow variables, the following system of hyperbolic equations, known as the nonlinear coupled-mode equations (NLCME),

$$\begin{aligned} i\partial_t u + i\partial_x u + \kappa(x)v + \eta(x)u + \left(|u|^2 + 2|v|^2\right)u = 0, \\ i\partial_t v - i\partial_x v + \kappa(x)u + \eta(x)v + \left(2|u|^2 + |v|^2\right)v = 0. \end{aligned} \tag{4}$$

arise as solvability conditions on the forward and backward slowly varying envelopes $u(x,t)$ and $v(x,t)$, respectively [11]. The coefficient $\kappa(x)$ is proportional to the local strength $\nu(x)$

of the grating while $\eta(x)$ is proportional to the local chirp $\Phi'(x)$. We refer to regions where $\kappa(x)$ and $\eta(x)$ vanish as the bare fiber to indicate the absence of the grating.

For low-amplitude light, the NLCME reduce to a set of linear coupled-mode equations

$$\begin{aligned} i\partial_t u + i\partial_x u + \kappa(x)v + \eta(x)u = 0, \\ i\partial_t v - i\partial_x v + \kappa(x)u + \eta(x)v = 0. \end{aligned} \tag{5}$$

When κ and η are constant, this system has a dispersion relation given by

$$\Omega(Q) = \eta \pm \sqrt{Q^2 + \kappa^2}. \tag{6}$$

Introducing a chirp $\eta(x)$ into the grating shifts the center of the bandgap, i.e., the set of frequencies $\Omega \in \left(\eta - \sqrt{Q^2+\kappa^2}, \eta + \sqrt{Q^2+\kappa^2}\right)$ which, for fixed κ and η, do not satisfy the the dispersion relation for any wavenumber Q. We show an example of dispersion relation (6) in Figure 2.

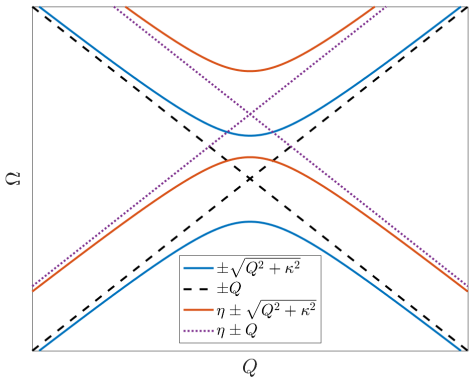


Figure 2: The dispersion relation (6) of the linear coupled-mode equations' (5) both with and without a chirp, demonstrating how the chirp shifts the bandgap.

In the case of a uniform grating, that is, where $\kappa \equiv \kappa_0$ and $\eta \equiv 0$, the NLCME admit a two-parameter family of translationally invariant solitary wave solutions, called Bragg solitons [1],

$$\begin{aligned} u_{\mathrm{B}} &= \sqrt{\frac{\kappa_0(1+c)}{3-c^2}}\left(1-c^2\right)^{1/4} W(X)\exp\left(i\phi(X) - iT\cos\theta\right), \\ v_{\mathrm{B}} &= -\sqrt{\frac{\kappa_0(1-c)}{3-c^2}}\left(1-c^2\right)^{1/4} W^*(X)\exp\left(i\phi(X) - iT\cos\theta\right), \end{aligned} \tag{7}$$

where

$$\begin{aligned} X &= \kappa_0\left(1-c^2\right)^{-1/2}(x-ct), \\ T &= \kappa_0\left(1-c^2\right)^{-1/2}(t-cx), \\ \phi(X) &= \frac{4c}{3-c^2}\arctan\left(\tanh\left(X\sin\theta\right)\tan\frac{\theta}{2}\right), \\ W(X) &= \sin\theta\ \mathrm{sech}\left(X\sin\theta - \frac{i\theta}{2}\right), \end{aligned} \tag{8}$$

with free parameters $0 \leq \theta \leq \pi$ and $-1 < c < 1$. The dependence of the Bragg soliton on the parameters is quite complicated, but we can make a few observations. The parameter c describes the velocity of the pulse, appears in a Lorentz contraction, and, through the factors $(1 \pm c)^{1/2}$, controls the relative amplitude of the forward and backward envelopes. Note that for stationary Bragg solitons, i.e., when $c = 0$, the frequency of the stationary oscillation is given by $\kappa_0 \cos\theta$ so as θ is increased from 0 to π, the frequency of the standing wave moves from the right edge of the band gap to the left edge, while always remaining inside the gap.

The constant coefficient NLCME possess two conserved quantities, an energy

$$E = \int_{-\infty}^{\infty} \left(|u|^2 + |v|^2 \right) dx := \int_{-\infty}^{\infty} \mathcal{E} dx, \tag{9}$$

and a momentum

$$P = i \int_{-\infty}^{\infty} \left(u \partial_x u^\dagger + v \partial_x v^\dagger \right) dx := \int_{-\infty}^{\infty} \mathcal{P} dx, \tag{10}$$

where $\dagger$ denotes complex conjugation. Allowing the coefficients to vary in space breaks the translation invariance, so that conservation of momentum fails to hold, yet energy conservation remains. Because of this fact, we make use of the energy (9) in posing an optimal control problem in Section 3.

2.2 The Numerical Setting

Using the physical modell discussed in Subsection 2.1, we now provide the numerical setting of the Rosenthal and Horowitz experiment. In describing numerics and displaying results, we use SI units and a nondimensionalization of Equations (1) and (4) consistent with work in [20]. Let $\kappa(x)$ be the grating profile that grows from $\kappa = 0$ for $x \leq 0$ to a value κ_0 for $x \geq a > 0$. The pulse is launched with a fixed profile from a point $x = x_{\text{input}} < 0$ in the bare fiber and has a momentum directed toward the apodization interval $0 \leq x \leq a$. The Rosenthal and Horowitz experiment does not include a chirp $\eta(x)$, so we set $\eta \equiv 0$ in this section and postpone discussions of η's role until we define the relevant optimal control problem in Section 3.

The existence of the solitary wave (7) demonstrates that slow light in FBGs is theoretically possible. Mok, et al., use a raised-cosine apodization profil to significantly reduce the speed of an incoming pulse, while propagating only about 20% of the incident energy past the apodization region into the constant-amplitude portion of the grating. Rosenthal and Horowitz significantly improve on this energy transmission to about 66%, by using two-segment apodization profiles of the form

$$\kappa(x) = \begin{cases} \frac{\zeta\kappa_0}{2}\left(1 - \cos\frac{\pi x}{L_1}\right) & 0 < x \leq L_1 \\ \zeta\kappa_0 + \frac{\kappa_0}{L_2}(1-\zeta)(x - L_1) & L_1 < x < a = L_1 + L_2. \end{cases} \tag{11}$$

The Mok, et al., design is simply the case $\zeta = 1$, and we provide a graph of the apodization profile (11), with $\zeta = 0.995$ in Figure 3.

The main numerical experiment in [20] takes the form of a signaling problem, i.e., the solution is initialized by a time-dependent boundary condition at the "input" endpoint. Th

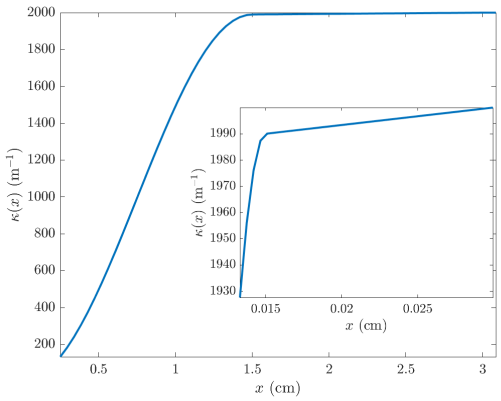


Figure 3: The Rosenthal and Horowitz apodization function (11), with $\zeta = 0.995$, $\kappa_0 = 2\text{mm}^{-1}$, and $L_1 = L_2 = 1.5\text{cm}$.

e signaling data is given by the profile

$$u(x_{\text{input}}, t) = A\text{sech}\left(\frac{t-\phi}{\sigma}\right)e^{-i\Omega t}, \quad v(x,0) = 0, \tag{12}$$

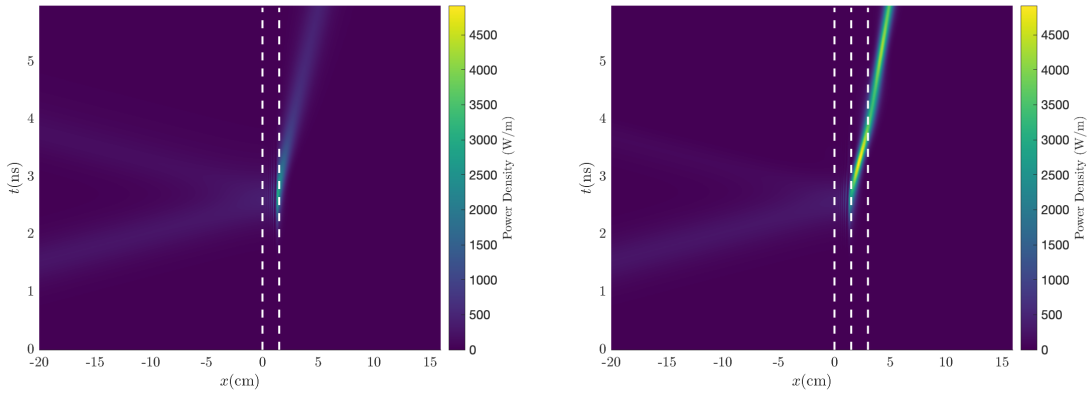
where x_{input} denotes the left-most endpoint of the spatial domain. It propagates through an initial segment of fiber with no grating with an oscillatory frequency outside the band gap of the dispersion relation (6) caused by the grating of amplitude κ_0. The intention of the experiment is for an **out-gap soliton** of the form (12) to interact with the apodization (11), and, through the coupled-mode dynamics (4), transform into a **in-gap soliton** of the form (7).

Figure 4 shows our reproduction of both Mok's and Rosenthals numerical experiments. It demonstrates the vast improvement in energy transmission that results from the choice $\zeta = 0.995$, instead of $\zeta = 1$, in the apodization function (11). The choices of the remaining apodization parameters, in SI units, are $L_1 = 1.5\text{cm}$, $L_2 = 1.5\text{cm}$, $\kappa_0 = 2\text{mm}^{-1}$, while the signalling data parameters are $A = 16.4\text{W}$, $\sigma = 96.4\text{ps}$, $\phi = 4\text{ns}$, and $\Omega = 0.398\text{GHz}$.

We use a second order in time operator splitting method, detailed in Subsection A, to solve Equation (4) with a spatial discretization of 4000 points and temporal discretization of 12000 points. We solve Equation (4) out to 6 ns, set $x_{\text{input}} = -20$ cm, and set the right endpoint to 16 cm. The numerical method and majority of these parameters will remain consistent throughout this chapter, unless otherwise noted.

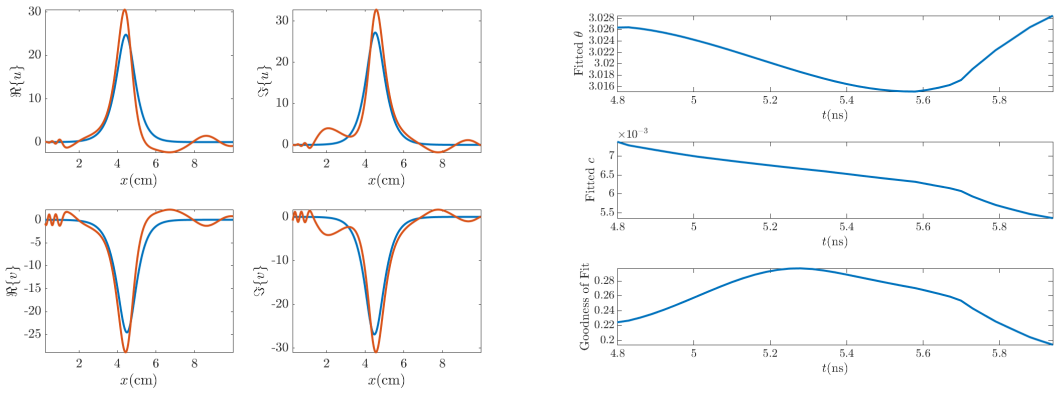
To justify the remarkable performance of the two-segment apodization function, Rosenthal and Horowitz appeal to ideas from soliton perturbation theory. They argue that the second segment of the apodization function *adiabatically* shifts the high-intensity waveform, initially at $x = L_1$, into the band gap, and, as a result, minimizes back-reflection. We believe the nature of this argument to be ad hoc, and that the efficiency of the Rosenthal and Horowitz apodization requires further investigation.

To interpret the result shown in Figure 4(b), we fit spatial solitary waves of the form (7) to the numerical simulation data at specific times and display results in Figure 5. In these fits, we see somewhat convincing evidence the excited waveforms are Bragg solitons, especially in the power density.



(a) (b)

Figure 4: Numerical simulations of Equation (4) with the apodization design (11) consistent with the parameters detailed in the text. Dashed lines provided to help visualize the regions in space over which the two segment apodization varies, cf. (11) and Figure 3. Panel (a) corresponds to the original Mok, et al., design, i.e., $\zeta = 1$, while Panel (b) has $\zeta = 0.995$.



(a) (b)

Figure 5: A least-squares fit of a Bragg soliton of the form (7) to the solution of the Rosenthal experiment at the instant the pulse exits the grating at $t = 4.8$ ns, as shown in Figure 4. **(a)** The real and imaginary components of the coupled-modes u and v. **(b)** The parameters defining the Bragg solitonin Equation (7), showing a near constant value in the phase parameter θ and a slow speed parameter c. The goodness-of-fit over time is also shown.

To demonstrate the potential advantages of using optimal control to further improve transmission efficiency, we first perform a simple optimization over the parameters in the family of apodization profiles given by Equation (11). In particular, we define a parameter $\xi \in (0,1)$,

$$L_1 = 3(1-\xi) \text{ cm}, \quad L_2 = 3\xi \text{ cm}, \tag{13}$$

and optimize the profile over the parameters (xi, ζ), so that the parameters are $(\xi, \zeta)_{\text{RH}} = (0.5, 0.995)$ in Figure 4. A simple minimization finds the optimal parameters $(\xi, \zeta)_* = (0.567, .99244)$, which yield an improved transmission near 68%. We interpret this result as an indication that optimal control theory should be able to further improve the transmission efficiency.

3 Optimal Control Formulation

3.1 Objective

We now precisely formulate an optimal control problem whose objective is to find the grating structures that maximize the transmission of light into the constant grating portion of the optical fiber. To this end, we make use of the following local conservation law

$$\partial_t \mathcal{E} + \partial_x \mathcal{F}_{\mathcal{E}} = 0, \tag{14}$$

where $\mathcal{F}_{\mathcal{E}}$ is the local energy flux. Note that this conservation law is the differential form of energy conservation, where the energy is given by Equation (9).

We treat the grating functions $\kappa(x)$ and $\eta(x)$ as the control functions, and assume that the apodization region is of fixed width $a > 0$, consistent with the Rosenthal and Horowitz apodization function (11). In addition, we use the conventions consistent with those of Figure 3, including the signalling data (12), length of the spatial domain, duration of simulation time, and number of discretization points. The admissible class $\mathcal{C}$ of grating functions we search over is the space of absolutely continuous functions such that

$$\kappa(x) = \begin{cases} 0 & x \leq 0, \\ \kappa_0 & a \leq x, \end{cases} \quad \text{and} \quad \eta(x) = \begin{cases} 0 & x \leq 0, \\ 0 & a \leq x. \end{cases} \tag{15}$$

Now, the optimal control problem we seek to solve is

$$\min_{(\kappa,\eta)\in\mathcal{C}} \mathcal{J} = \min_{(\kappa,\eta)\in\mathcal{C}} \left\{ -\int_0^T \mathcal{F}_{\mathcal{E}}(u, v; a) dt + \frac{\gamma}{2} \int_0^a \left((\partial_x \kappa)^2 + (\partial_x \eta)^2 \right) dx \right\}, \tag{16}$$

subject to the differential equation constraint (4).

Although the first term in the objective $\mathcal{J}$ is a cost that runs over time, the following simple calculation demonstrates that this term can alternatively be written as

$$\int_0^T \mathcal{F}_{\mathcal{E}}(u, v; a) dt = \int_0^T \int_\infty^a \partial_x \mathcal{F}_{\mathcal{E}} dt dx = \int_a^\infty \int_0^T \partial_t \mathcal{E} dt dx = \int_a^\infty \mathcal{E}(x, T) dx, \tag{17}$$

by the fundamental theorem of calculus, Fubini's theorem, and conservation law (14). In this sense, the term which promotes a greater energy flux into the constant grating portion

can be written as a running cost of terminal energy in space. The second term in objective $\mathcal{J}$ is called a Tikhonoff regularization and is taken over space. Such a regularization term is used extensively in studies of ill-conditioned optimal control and inverse problems [12, 24]. Its effect is to penalize rapid variations in the grating which would be infeasible to design, in practice.

3.2 Necessary Optimality Conditions

In order to solve Problem (16), we use a line search discussed in Appendix A.2. Part of the method involves the computation of the gradient which depends on a suitably defined criteria for optimality.

To this end, let $H(x-a)$ denote Heaviside's function, and rewrite the energy flux term in objective (16), assuming the flux vanishes at infinity and is zero initially, as

$$\begin{aligned} -\partial_t \mathcal{F}_{\mathcal{E}}\big|_{x=a} &= \int_{-\infty}^{\infty} H(x-a) \int_0^T \partial_t \mathcal{E}\, dt dx \\ &= \int_{-\infty}^{\infty} H(x-a) \int_0^T \left(u^\dagger \partial_t u + u \partial_t u^\dagger + v^\dagger \partial_t v + v \partial_t v^\dagger \right) dt dx, \end{aligned} \tag{18}$$

by using the fundamental theorem of calculus along with conservation laws (9) and (14). Define the Lagrangian by

$$\begin{aligned} \mathcal{L} =& \operatorname{Re} \left\langle \lambda, i\partial_t u + i\partial_x u + \kappa(x) v + \eta(x) u + \left(|u|^2 + 2|v|^2 \right) u \right\rangle_{L^2([0,T])} \\ &+ \operatorname{Re} \left\langle \mu, i\partial_t v - i\partial_x v + \kappa(x) u + \eta(x) v + \left(2|u|^2 + |v|^2 \right) v \right\rangle_{L^2([0,T])} \\ &+ H(x-a) \int_0^T \left(u^\dagger \partial_t u + u \partial_t u^\dagger + v^\dagger \partial_t v + v \partial_t v^\dagger \right) dt, \end{aligned} \tag{19}$$

where λ and μ are Lagrange multipliers. The objective in optimal control problem (16) can now be written in the form

$$\mathcal{J} = \int_{\mathbb{R}} \operatorname{Re} \left\{ \mathcal{L} \left(u, v, \partial_t u, \partial_t u^\dagger \partial_t v, \partial_t v^\dagger, \partial_x u, \partial_x v, \kappa, \eta, \lambda^\dagger, \mu^\dagger \right) \right\} dx \tag{20}$$

so that the optimization problem is now unconstrained while enforcing that u and v solve NLCME (4). Using typical arguments from the classical calculus of variations [10], we find the desired optimality conditions by taking the appropriate functional derivatives. Setting functional derivatives with respect to the state variables to zero gives

$$\frac{\delta \mathcal{J}}{\delta u} + \frac{\delta \mathcal{J}}{\delta u^\dagger} = 0, \quad \frac{\delta \mathcal{J}}{\delta v} + \frac{\delta \mathcal{J}}{\delta v^\dagger} = 0. \tag{21}$$

These conditions imply the following equations

$$i\partial_t \lambda + i\partial_x \lambda + \left(\eta + 2\mathcal{E} + u^{\dagger 2} \right) \lambda + \left(\kappa + 4v^\dagger \operatorname{Re}\{u\} \right) \mu = 0, \tag{22a}$$

$$i\partial_t \mu - i\partial_x \mu + \left(\eta + 2\mathcal{E} + v^{\dagger 2} \right) \mu + \left(\kappa + 4u^\dagger \operatorname{Re}\{v\} \right) \lambda = 0. \tag{22b}$$

An integration by parts yields boundary terms which must also be set to zero:

$$\left(\partial_{\partial_t u}\mathcal{L} + \partial_{\partial_t u^\dagger}\mathcal{L}\right)\Big|_{t=T} = 0, \quad \left(\partial_{\partial_t v}\mathcal{L} + \partial_{\partial_t v^\dagger}\mathcal{L}\right)\Big|_{t=T} = 0, \tag{23}$$

directly implying that

$$\lambda(x,T) = 2iH(x-a)\,\mathrm{Re}\left\{u(x,T)\right\}, \tag{24a}$$
$$\mu(x,T) = 2iH(x-a)\,\mathrm{Re}\left\{v(x,T)\right\}. \tag{24b}$$

Since variations of the states u and v need not vanish at $t = T$, equations (24) must be satisfied. Indeed, these conditions determine what is in essence an initial condition for the equations (22) which can then be solved backwards in time.

Next, setting functional derivatives with respect to the control variables to zero gives

$$\delta_\kappa \mathcal{J} = \int_0^T \mathrm{Re}\left\{\lambda^\dagger v + \mu^\dagger u\right\} dt - \gamma\partial_x^2\kappa = 0, \tag{25a}$$
$$\delta_\eta \mathcal{J} = \int_0^T \mathrm{Re}\left\{\lambda^\dagger u + \mu^\dagger v\right\} dt - \gamma\partial_x^2\eta = 0. \tag{25b}$$

Equations (25), together with the boundary conditions implied by the admissible class $\mathcal{C}$, gives two-point boundary value problems over the domain $[0, a]$. Lastly, the vanishing of functional derivatives with respect to the costate variables $\lambda^\dagger$ and $\mu^\dagger$ returns the state equations (4), i.e., the NLCME.

We use a second order in time operator splitting method to solve the state equations (4) and to solve costate equations (22) backward from their terminal condition (24). The details of the method, which itself is an improvement on the splitting method used in [19], are given in [2]. Equations (25) are used in the computation of control gradients as required by the computational optimization method in Section A.2.

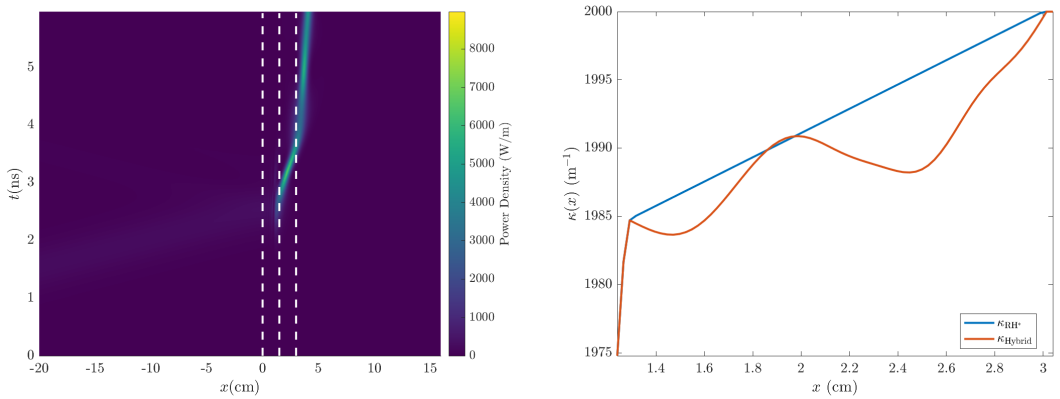
3.3 Numerical Optimization Strategy

In order to solve Problem (16), we use a *hybrid* optimization method; a combination of a global, non-convex method followed by a local, iterative method. The methodology we use in this paper is similar to one used by Sørensen, et al. [21], and allows for the use of a global search routine based on stochastic optimization to overcome non-convexity. Non-convex 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.

The first step in the global search is to use a Chopped Random Basis (CRAB) method [6, 7] which efficiently reduces the numerical optimization problem to a nonlinear programming problem (NLP). This NLP is then solved via a genetic algorithm called Differential Evolution (DE) [22]. Implementation details of this method are given in Appendix A.1. The next step is to refine the result of CRAB/DE using a projected gradient descent due to von Winckel and Borzi [4]. We give details of von Winckel's and Borzi's method in Appendix A.2.

4 Optimization Results

We now present the results of using the previously discussed hybrid numerical optimization strategy on optimal control problem (16). In most of the results presented, we slightly change the optimization problem. We keep the raised-cosine portion of the profile fixed on $[0, L_2]$ and only optimize the second section of the profile on $[L_2, L_2]$. This is based on our initial experiments that found optimization over the first segment made less difference. In all simulations, we find the Tikhonoff parameter γ on the order of 10^{-6} to be satisfactory. All other conventions regarding numerical simulations are consistent with those used so far throughout this work. In our first experiment, we do not allow a chirp (fixing $\eta = 0$). Our hybrid method finds an apodization function which couples 74.1% of the incident light into the grating. This is shown in Figure 6. Because of the stochastic nature of the method, running the simulation again would find a different profile, but we found that if the parameters of the search were held fixed, the resulting efficiency was fairly stable.



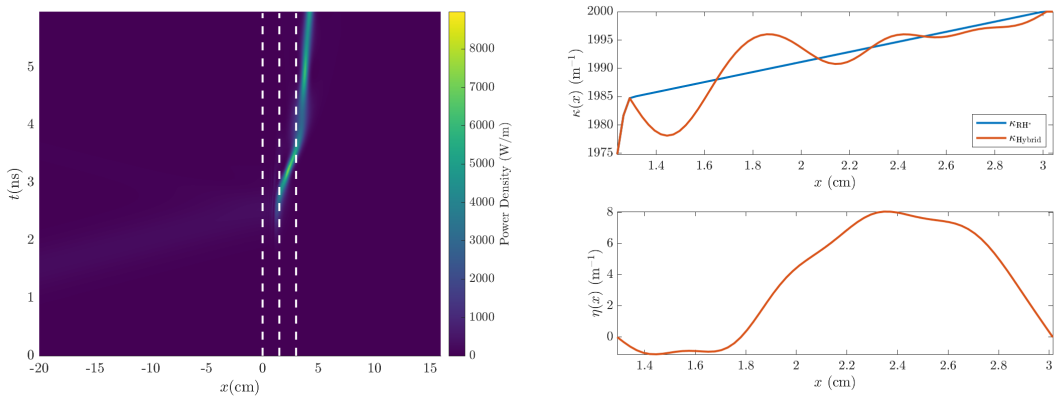
(a) (b)

Figure 6: The result of using the hybrid method to find more efficient apodization functions κ which are nearby the $(\xi, \zeta)_*$ apodization. 74.1% of the incident light is now coupled into the grating.

Next, we include a chirp on the interval (L_1, L_2), and show the computed optimizer in Figure 7: it couples 77.7% of the incident light. From the optimal apodization functions, shown in Figures 6–7, we observe that the nearby efficient apodization functions have a somewhat large negative gradient to the right of the optimization boundary point x_0. This hints at how more of the total apodization region should be allocated toward the shallow, adiabatic portion of the grating.

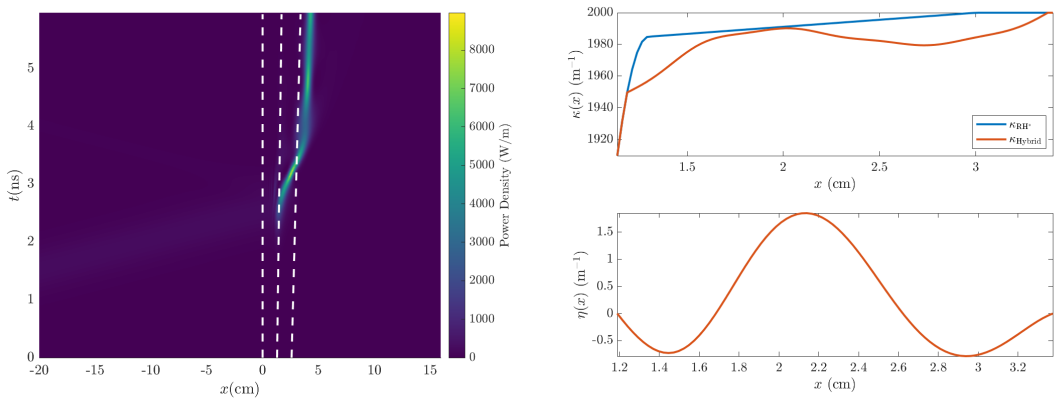
For this reason, we slightly relax the design constraint of a three centimeter apodization region, extend the width of the optimization domain to $[1.19, 3.4]$ cm, and perform the search again. The extension to the left is motivated by the above-mentioned observation on the nearby optimal apodization functions, and the extension to the right is ad hoc. We extend the domain so that the entire grating structure is still reasonably within some technological constraint, i.e., within a margin of 15% the total size of the Rosenthal and Horowitz apodization region, and, moreover, maximize performance. We show, in Figure 8,

grating functions which now successfully couple 82.6% of the incident light. We also find that the resulting in-gap soliton has a group velocity with a magnitude 0.53% that of light speed. We emphasize this unintended, yet fortuitous improvement in the slowdown of the coupled light against the result of the original $(\xi, \zeta)_{\mathrm{RH}}$ apodization visually in Figure 9.



(a) (b)

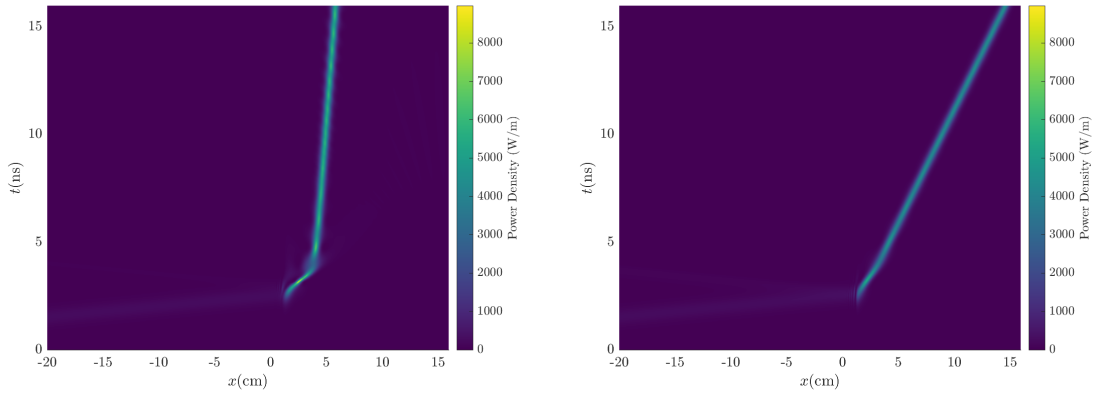
Figure 7: The result of including a chirp in the search for efficient grating functions near the $(\xi, \zeta)_*$ apodization. 77.7% of the incident light is now coupled into the grating.



(a) (b)

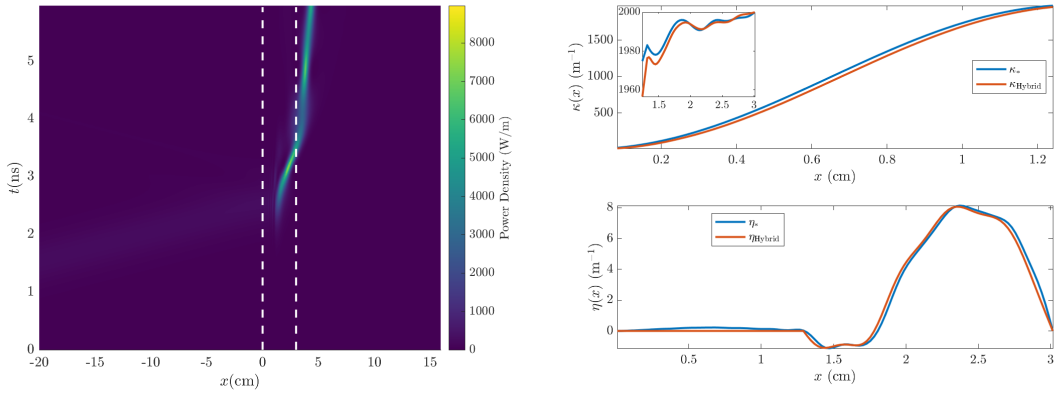
Figure 8: The result of relaxing the optimization domain to be slightly wider, i.e., $x_0 = 1.19$ cm and $a = 3.4$ cm. The FBG couples 82.6% of the light with a group velocity 0.53% the speed of light.

The final result we show, in Figure 10, performs a search on the entire optimization domain $[0, 3]$ cm. After several hours of computation time, the best grating structure found is 81.4% efficient. Moreover, this is found by looking for apodization profiles near the design shown in Figure 7. We believe this result clearly shows the value of Rosenthal and Horowitz's intuition in their design choices, since, without making reference to their design, the best grating structure found by our methods, including a chirp, is about 68% efficient.



(a) (b)

Figure 9: A long time simulation consistent with the results of Figure 8 emphasizing the dramatic reduction in the group velocity of the solitary wave. **(a)** The locally optimal grating structure shown in Figure 7, **(b)** The $(\xi, \zeta)_{\mathrm{RH}}$ apodization from Figure 3.



(a) (b)

Figure 10: A result of using the hybrid method optimizing over the entire domain $x \in [0, 3]$cm. The coupling efficiency of the grating is about 81.4%. The notation κ_* and η_* denote the designs shown in Figure 7.

5 Concluding Remarks

In this work, we formulate a simple and physically motivated optimal control problem aimed at efficiently coupling light into an FBG. By employing standard numerical optimization methods, widely used for example in the quantum control literature [7,13,21], we demonstrate the viability of optimal control theory in the design of FBGs which act as efficient compressors and pulse-delayers. We optimize previously reported designs, and provide guidance on how to explore the space of possible designs. In addition, the methodology used here can be applied, with suitable modifications, to other problems constrained by dispersive equations.

By considering the chirp of the grating as part of the design, we see an improvement in both the transmission of the light and the effectiveness of the grating as a pulse-delayer. Moreover, we find the globally optimal apodization functions for this problem are most likely ones which have features similar to the Rosenthal and Horowitz design. We also observe that in cases where we see significant gains in the transmission of light into the fiber, we find this comes at the cost of decoherence through two possible mechanisms: radiation buildup and solitary wave fissioning.

Evidence of radiation buildup is seen in the contrast between the tails of the power densities shown in Figures 5. Further evidence exists through the temporal power spectra at spatial points well beyond the apodization region in Figure 8 and the associated Bragg soliton fits in Figure 11. We find that the grating structure is only 78.1% effective in terms of *coherent* energy since 94.5% of the total transmitted belongs to the Bragg soliton rendering this result to be less impressive. Some decoherence may be due to solitary wave fissioning, although we do not observe clear evidence of that here. Even if evidence of the onset of fissioning is unconvincing, its presence is certainly possible and well-understood. Fissioning, as a culpable mechanism of decoherence, has been well-studied over the past few decades. Indeed, numerical studies by Mak and Malomed [14] demonstrate the possibilities of Bragg soliton splitting, albeit in a slightly different experimental context. Theoretical insight into this mechanism is based on the inverse scattering formalism and the mechanism has been accounted for in nearly-integrable versions of the Korteweg-deVries equation in work due to Zabusky and Tappert [23] and the Benjamin-Ono equation in work due to Choi [8].

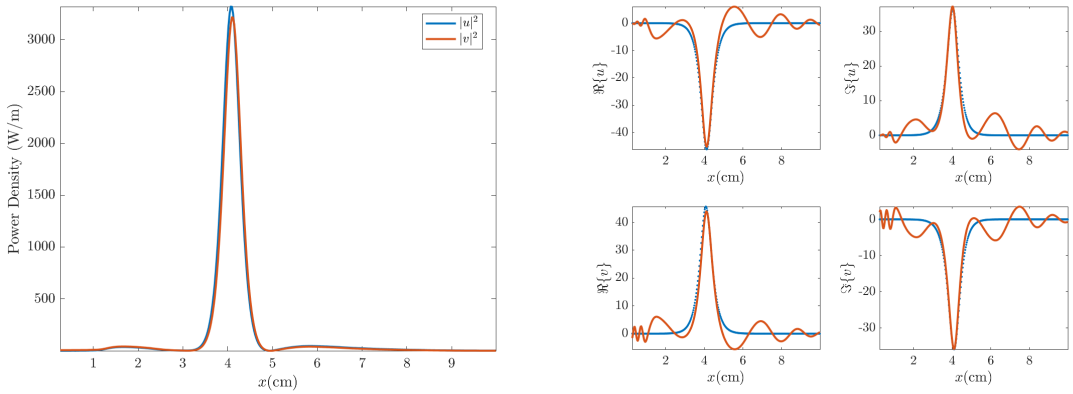
We believe our numerical results, in particular those of Figure 8, serve as an impetus for investigating the coupled-mode dynamics more thoroughly using soliton-specific methods. We leave this investigation as subject for future work.

Acknowledgments

We would like to thank Professors David Shirokoff, Alejandro Aceves, and Richard Moore for helpful discussions.

References

[1] A. B. Aceves and S. Wabnitz, *Self-induced transparency solitons in nonlinear refractive periodic media*, Physics Letters A, 141 (1989), pp. 37–42.



(a) (b)

Figure 11: **(a)**: Solutions of Equation (4) at $t = 4.8$ns corresponding to the optimal grating structures in Figure 8. **(b)**: A Bragg soliton fitting with (dimensionless) parameters $c = 0.0058$, $\theta = 0.1806, \Psi_0 = 3.862$, and $T = 14,162$. Evidence of solitary wave decoherence is present.

[2] J. Adriazola, *Coherent Control of Dispersive Waves*, PhD thesis, New Jersey Institute of Technology, 2021.

[3] G. P. Agrawal and S. Radic, *Phase-shifted fiber Bragg gratings and their application for wavelength demultiplexing*, IEEE Photonics Technology Letters, 6 (1994), pp. 995–997.

[4] A. Borzì, G. Ciaramella, and M. Sprengel, *Formulation and numerical solution of quantum control problems*, SIAM, 2017.

[5] S. Boyd and L. Vandenberghe, *Convex optimization*, vol. 1, Cambridge University Press Cambridge, 2004.

[6] T. Calarco, E. A. Hinds, D. Jaksch, J. Schmiedmayer, J. I. Cirac, and P. Zoller, *Quantum gates with neutral atoms: Controlling collisional interactions in time-dependent traps*, Phys. Rev. A, 61 (2000), p. 022304.

[7] T. Caneva, T. Calarco, and S. Montangero, *Chopped random-basis quantum optimization*, Phys. Rev. A, 84 (2011), p. 022326.

[8] W. Choi, *On the fission of algebraic solitons*, Proceedings: Mathematical, Physical and Engineering Sciences, 453 (1997), pp. 1753–1762.

[9] M. DeSterke, *Propagation through apodized gratings*, Optics Express, 3 (1998), pp. 405–410.

[10] I. Gelfand and S. Fomin, *Calculus of variations*, 1963.

[11] R. H. Goodman, R. E. Slusher, and M. I. Weinstein, *Stopping light on a defect*, Journal of the Optical Society of America B, 19 (2002), pp. 1635–1652.

[12] M. Hintermuller, D. Marahrens, P. A. Markowich, and C. Sparber, *Optimal bilinear control of Gross-Pitaevskii equations*, SIAM Journal on Control and Optimization, 51 (2013), pp. 2509–2543.

[13] U. Hohenester, P. K. Rekdal, A. Borzì, and J. Schmiedmayer, *Optimal quantum control of Bose-Einstein condensates in magnetic microtraps*, Physical Review A, 75 (2007), p. 023602.

[14] W. C. K. Mak, B. A. Malomed, and P. L. Chu, *Slowdown and splitting of gap solitons in apodized Bragg gratings*, Journal of Modern Optics, 51 (2004), pp. 2141–2158.

[15] J. Mennemann, D. Matthes, R. Weishaupl, and T. Langen, *Optimal control of Bose-Einstein condensates in three dimensions*, New Journal of Physics, 2015.

[16] J. Mok, C. M. De Sterke, I. Littler, and B. Eggleton, *Dispersionless slow light using gap solitons*, Nature Physics, (2006).

[17] J. V. Moloney and A. Newell, *Nonlinear Optics*, CRC Press, 2018.

[18] A. C. Newell, *Solitons in mathematics and physics*, SIAM, 1985.

[19] A. Rosenthal and M. Horowitz, *Analysis and design of nonlinear fiber Bragg gratings and their application for optical compression of reflected pulses*, Optics Letters, 31 (2006), pp. 1334–1336.

[20] A. Rosenthal and M. Horowtiz, *Efficient method for launching in-gap solitons in fiber Bragg gratings using a two-segment apodization profile*, Optics Letters, 33 (2008), pp. 678–680.

[21] J. J. Sørensen, M. Aranburu, T. Heinzel, and J. Sherson, *Approaching the quantum speed limit with global-local optimization*, arXiv preprint arXiv:1802.07521, (2018).

[22] R. Storn and K. Price, *Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces*, Journal of Global Optimization, 11 (1997), pp. 341–359.

[23] F. D. Tappert and N. J. Zabusky, *Gradient-induced fission of solitons*, Physical Review Letters, 27 (1971), p. 1774.

[24] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Springer Netherlands, 1995.

[25] L. N. Trefethen, *Spectral methods in MATLAB*, SIAM, 2000.

[26] G. von Winckel and A. Borzi, *Computational techniques for a quantum control problem with H1-cost*, Inverse Problems, (2008).

A Numerical Optimization Methods

A.1 The Global Method

In the first step of the hybrid method, we represent the apodization profile using a Galerkin approximation. This results in a so-called Chopped Random Basis (CRAB) method which reduces the complexity of the optimal control problem so that standard non-convex nonlinear programming (NLP) techniques can be applied. It 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 basis is chosen so that controls remain in the appropriate admissible space $\mathcal{C}$ in the context of the control problem (16).

A common representation in the CRAB method is of the form

$$\beta_r(x) = \mathcal{P}(x; \beta(x_0), \beta(a), x_0, a) + \sum_{j=0}^{N-1} \varepsilon_j \varphi_j(x; x_0, a), \quad x \in [x_0, a]. \tag{26}$$

Here, β denotes either grating function κ or η, $\mathcal{P}$ is a fixed function satisfying the boundary conditions of the admissible class (15), each $\varphi_j(x)$ 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 control ansatz (26) reliably simplifies the optimal control problem.

More specifically, the CRAB ansatz we use throughout this work is

$$\kappa_r(x) = (\kappa_0 - \kappa_{\mathrm{RH}}(x_0)) \sum_{j=1}^{15} \frac{r_\kappa}{j^2} \sin\left(j\pi \frac{x - x_0}{a - x_0}\right) + (\kappa_0 - \kappa_{\mathrm{RH}}(x_0)) \frac{x - x_0}{a - x_0} + \kappa(x_0), \tag{27a}$$

$$\eta_r(x) = \frac{\kappa_0}{100} \sum_{j=1}^{15} \frac{r_\eta}{j^2} \sin\left(j\pi \frac{x - x_0}{a - x_0}\right), \tag{27b}$$

on the optimization domain $x \in [x_0, a]$, where r_κ and r_η are random variables drawn uniformly from $[-1, 1]$, κ_{RH} is a Rosenthal and Horowitz apodization (11) with free parameters (ξ, ζ), and $\kappa_0 = 2\mathrm{mm}^{-1}$, consistent with the constant grating portion of κ_{RH}. From experience with the numerical optimization, we find that chirp functions which are about two orders of magnitude smaller than κ_0 perform well; this is why we include the factor of 100 in the CRAB ansatz for η. A modest number of 15 basis functions has proven adequate here.

To solve the resulting NLP problem, we use differential evolution (DE) [22]. DE is a stochastic optimization method used to search for candidate solutions to non-convex optimization problems. 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, with parameter $p_{\mathrm{weight}} \in (0, 2)$, and a "crossover," with parameter p_{cross} which randomly exchanges "traits", or elements, between agents. We find a population size of 50 agents, a weight $p_{\mathrm{weight}} = 0.8$, and crossover $p_{\mathrm{cross}} = 0.9$ to be effective.

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 which 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. We find 30 iterations to be sufficient in this regard.

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 non-convex optimization problems, they do not make guarantees about 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 in order to generate initial controls, through the CRAB representation (26), for an iterative method which guarantees local optimality up to some threshold. We note that in practice, the best of five or so realizations through the CRAB/DE method is sufficient before moving on to the refinement stage.

A.2 The Local Method

We use a line search strategy due to von Winckel and Borzi [26]. The von Winckel and Borzi (vWB) method is an appropriate generalization of the well-known gradient descent method from $\mathbb{R}^n$ to an appropriate affine function space which automatically preserves the boundary conditions of the admissible class $\mathcal{C}$ mentioned in the context of optimal control problem (16). This method has been frequently applied in the quantum control literature; see for example [13, 15, 21].

For ease of notation, we describe the vWB method for optimizing the apodization function κ, since its extension to η is trivial. Recall the optimal control problem we want to solve is, in unconstrained form, given by Equation (20). The method of gradient descent, in this context, is given by following iteration

$$\kappa_{k+1} = \kappa_k - \alpha_k \nabla_\kappa \mathcal{L}\big|_{\kappa=\kappa_k}, \tag{28}$$

where the linear operator ∇_κ 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 [5].

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 $L^2([x_0, a])$, then it can be identified with the functional derivative of the objective $\mathcal{J}$, which in this case can be shown to be

$$\delta_\kappa \mathcal{J} = \int_0^T \mathrm{Re}\left\{\lambda^\dagger v + \mu^\dagger u\right\} dt - \gamma \partial_x^2 \kappa. \tag{29}$$

Note that this coincides with the Euler-Lagrange equation $\delta_\kappa \mathcal{J} = 0$ given by Equation (25). If this choice is made, however, the increment $\alpha_k \nabla_\kappa \mathcal{L}\big|_{\kappa=\kappa_k}$ would not in general satisfy

the boundary conditions on the control κ_k, and the updated control κ_{k+1} would leave the admissible set $\mathcal{C}$. This problem is avoided by using a different function space X to define the operator ∇_κ.

To this end, consider an arbitrary displacement $\nu \in C_c^\infty([x_0, a])$ and an arbitrary $\varepsilon > 0$. We know Taylor's theorem holds, i.e., the series

$$\mathcal{J}[\kappa + \varepsilon\nu] = \mathcal{J}[\kappa] + \varepsilon \left\langle \nabla_u \mathcal{L}(u), \nu \right\rangle_X + \mathcal{O}(\varepsilon^2) \tag{30}$$

holds term-by-term independently of the choice of the Hilbert space X for sufficiently regular functionals $\mathcal{J}$. The vWB method chooses the homogeneous and traceless Sobolev space $\dot{H}_0^1([x_0, a])$ for X. By equating the directional, or Gateaux, derivatives with respect to $L^2([x_0, a])$ and with respect to $\dot{H}_0^1([x_0, a])$, we see that

$$\begin{aligned} \left\langle \nabla_\kappa \mathcal{L}, \nu \right\rangle_{L^2([x_0,a])} &= \left\langle \delta_\kappa \mathcal{J}, \nu \right\rangle_{L^2([x_0,a])} \\ = \left\langle \nabla_\kappa \mathcal{L}, \nu \right\rangle_{\dot{H}_0^1([x_0,a])} &:= \int_{x_0}^{a} \partial_x \nabla_u \mathcal{L} \partial_x \nu dx = - \left\langle \partial_x^2 \nabla_\kappa \mathcal{L}, \nu \right\rangle_{L^2([x_0,a])}, \end{aligned} \tag{31}$$

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

Since this holds for all displacements $\nu \in C_c^\infty([x_0, a])$, we conclude, by the fundamental lemma of the calculus of variations [10], the strong form of Equation (31)

$$-\partial_x^2 \nabla_\kappa \mathcal{L} = \delta_\kappa \mathcal{J}, \quad \nabla_\kappa \mathcal{L}(x_0) = \nabla_\kappa \mathcal{L}(a) = 0, \tag{32}$$

also holds. This renders an admissible gradient whose homogeneous Dirichlet conditions are induced by choosing increments specifically from the traceless space $\dot{H}_0^1([x_0, a])$. In order to solve the boundary value problem (32) for the control gradient $\nabla_\kappa \mathcal{L}$, we use Chebyshev collocation [25].