

# Inverse Design of Non-Periodic Dual-Pillar Structures for Dielectric Laser Acceleration

## Master's Thesis in Physics

Presented by

**Christian Neureuter**

29<sup>th</sup> of September 2022

Chair for Laserphysics  
Faculty of Science  
Friedrich-Alexander-Universität Erlangen-Nürnberg



Supervisor: Prof. Dr. Peter Hommelhoff



# Danksagung

Ich möchte Prof. Dr. Peter Hommelhoff von der Friedrich-Alexander-Universität und dem Max-Planck-Institut für Physik des Lichts herzlich dafür danken, dass sie sich so entgegenkommend in meiner ungewöhnlichen Situation im Zusammenhang mit COVID-19 gezeigt haben. Ich möchte auch Dr. Roy Shiloh danken, der ein großartiger Betreuer war. Es war ein schwieriger Weg durch diese Dissertation und ich hätte es ohne Roys Unterstützung nicht geschafft. Er hat mich nicht nur mit Fachwissen und Anleitung versorgt, sondern auch moralisch unterstützt. Ich möchte auch meiner Familie für ihre fortwährende Unterstützung und der meines Freundes MJ danken.

# Abstract

Herein we will demonstrate the application of inverse design to dual pillar, non-periodic, dielectric laser accelerator (DLA) structures. Inverse design, as a concept, involves selecting the desired final output and then using an algorithm to numerically optimize a structure which has the same approximate output. The algorithm utilizes gradient descent methods, colloquially seen in machine learning, to allow the computer to traverse the high dimensional parameter space quickly. The adjoint method allows us to describe certain systems with linear operators, then uses linear algebra to produce an easily calculable gradient for the parameter space. A typical gradient descent algorithm uses a two point approximation to create the gradient between stochastically chosen steps. However, with the adjoint method, the gradient is calculated with the forward and adjoint simulation fields and a two point derivative of  $\Delta\epsilon_r$  (relative permittivity). In non-infinitely periodic structures, there are interactions between the cells and hence the system is no longer linearly independent. Due to this, all constraints, bounds, and weights to the variables have to be calculated within the algorithm which constructs the Jacobian. Out-of-the box optimizers, even those which allow you to supply your own Jacobian, are not sufficient because they apply constraints *ex post facto*; thus, a custom optimization algorithm is necessary.

It is possible to optimize for the  $y$ -position of the pillars within their cell using this method. While the longitudinal position of the pillar within its cell (along the electron propagation axis) provides a full  $180^\circ$  phase control in infinitely periodic structures, only an  $\approx \pi/5 = 36^\circ$  is possible in non-periodic structures with fabrication constraints on the system. The concepts of alternating phase focusing (APF) could be easily implemented into this algorithm to provide a full  $360^\circ$  in phase control for non-periodic structures; enabling better phase matching, and more flexibility in the structure design. Non-periodicity of the structures is interesting since the adjoint calculations typically assume periodicity; however, as an acceleration structure, the electron will change velocity, thus changing its relative phase to the laser. This dephasing can be accounted for, and optimized for, by changing the period size of the cells. This thesis is the step between a periodic, non-accelerating structure, and a non-periodic structure with dynamic periodicity set by maintaining a phase relationship with the particle, via particle tracking. This thesis accounts for the acceleration and dephasing, but in discrete intervals, rather than a pseudo-continuum. The custom optimization algorithm created for this thesis was successfully able to match, to ( $> 95\%$ ), complicated design curves for large scale, non-periodic structures, for electrons in the 30-50 keV range (although this should be easily scalable to higher energy particles). It was found that arbitrarily complicated design curves could be optimized up to the rate of change limit of  $\text{Re}\{e_1\} \approx 0.1/10$  periods.

“Nobody ever figures out what life is all about, and it doesn’t matter. Explore the world. Nearly everything is really interesting if you go into it deeply enough.”

- **Richard P. Feynman**

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.0.1	Laser Accelerators . . . . .	2
1.0.2	Inverse Design . . . . .	4
<b>2</b>	<b>Theory</b>	<b>5</b>
2.1	Near Field Acceleration . . . . .	6
2.2	Dual Grating Structures . . . . .	8
2.3	Non-Periodic Structures . . . . .	9
2.4	Adjoint Method . . . . .	11
2.4.1	Adjoint Method for DLA . . . . .	12
2.4.2	Cost Function . . . . .	14
2.4.3	Adjoint Source Radiation . . . . .	16
2.4.4	Adjoint Source and Phase . . . . .	17
2.4.5	Adjoint Inverse Design . . . . .	18
2.4.6	Particle Acceleration and Non-Periodicity . . . . .	19
2.4.7	Fringing . . . . .	21
2.5	Numerical Optimization . . . . .	22
2.5.1	Constrained Optimization . . . . .	23
2.5.2	Non-Convex Optimization . . . . .	25
<b>3</b>	<b>Methods</b>	<b>27</b>
3.0.1	Lumerical . . . . .	27
3.1	Custom Optimization Algorithm . . . . .	27
3.1.1	Periodicity Calculation and Simulation Setup . . . . .	29
3.1.2	Fourier Transform . . . . .	30

---

3.1.3	Adjoint Simulation . . . . .	30
3.1.4	Construction of $\frac{\Delta\varepsilon_r}{\Delta\vec{\phi}}$ . . . . .	32
3.1.5	Custom Optimization . . . . .	36
<b>4</b>	<b>Results</b>	<b>37</b>
4.0.1	Simulation Parameters . . . . .	37
4.1	Custom Optimizer Validation . . . . .	38
4.2	Non-Periodic Structure Optimization . . . . .	42
<b>5</b>	<b>Conclusion</b>	<b>50</b>

# Introduction

Conventional particle accelerators must follow a synchronicity condition between the particle's motion and the oscillation of the accelerating field [1]. This synchronization depends on particle velocity, cell length, types of confining field, and the frequency of the accelerating field. [1]. Different types of accelerators have different synchronicity conditions, but they all follow the same principle of matching the phase of the particle motion to the applied fields. Typical high energy particle acceleration is done via RF waves [2, 3, 1]. Due to the size of waves in the RF-range, the synchronicity condition requires quite large structures. The strength of the accelerating field gradient is on the order of 10s of MeV/m, and thus requires many periods of integration with the applied field in order to achieve high levels of acceleration [4, 1]. Lasers, however, having wavelengths on the order of hundreds of nanometers, have a much smaller synchronicity condition, making them an interesting candidate for particle accelerators. Recent advances in ultra-fast optics have provided femto- and atto-second laser pulses with high peak field intensities. These ultra-short pulses have been demonstrated to provide a maximum acceleration gradient of  $> 1 \text{ GeV/m}$  [5, 6], far higher than a typical RF linear accelerator (linac) [7, 8]. Lasers are also useful because of the amount of control optics allow over the applied beam.

## 1.0.1 Laser Accelerators

The laser was invented in 1962, soon after, Shimoda suggested using optical fields to accelerate electrons [9]. The next year, Lohman patented an idea for a “particle accelerator utilizing coherent light” [10], which included two dielectric gratings facing each other and complementary, counter-propagating, drive lasers. Laser accelerators developed from maser pumped accelerators with metallic gratings (1987) [11], to laser pumped metallic gratings, finally to dielectric laser acceleration (2013) [12, 13].

DLA involves using a dielectric material, in our case silicon, to diffract the incoming laser light and use the near fields to accelerate a charged particle. Although the idea for dual grating, dual driven, DLAs has been around for quite some time, it wasn't until the last decade that they were successfully tested. It took nano-fabrication advancements to facilitate the use of lasers as accelerators. Due to the wavelength of the applied field, structure features are typically sub-micron. This means extremely precise lithographic techniques had to be developed before a useful structure could be made. The benefits of dielectric over metals, though, is that dielectrics provide transparency to the field as compared to metal gratings.

This increases the damage threshold because the incident beam won't couple to the structure, allowing access to higher acceleration gradients compared to RF accelerators [13]. Metallic gratings are usually operated at field strengths in the 25-50 MV/m range, whereas DLA structures have been demonstrated with acceleration gradients greater than 1 GV/m [5, 14]. DLAs have even been shown to withstand field strengths of up to 9 GV/m [6], limited only by the damage threshold and higher order optical effects. This implies that a DLA could achieve similar acceleration to a traditional RF linac, but scaled down by almost two orders of magnitude [15].

Due to the high field intensity of the ultra short laser pulse, the required focus strengths, necessary to control the electron-beam, are too high for conventional magnets [13]. With a dual pillar structure, illuminated by symmetric counter propagating lasers (or a reflected laser [16]), the focusing component is integrated into the accelerator structure [7]. A dual pillar structure provides a cosh acceleration profile as seen in Fig. 2.2b. [17, 16, 7] Uniformity of the accelerating field, within its given cell, is necessary for DLAs, in order to minimize the deflection forces and enable longer accelerators [16]. Dual sided laser illumination provides the greatest field symmetry [16] and, as such, is the setup used in this thesis.

Since the synchronicity condition depends on the velocity of the particle ( $\lambda_p = m \cdot \beta \lambda_l$ ) and this is an accelerator, it would be assumed that the periodicity  $\lambda_p$  would be a function of time. However, most DLA research has been done with periodic cells, and thus long structures would require compensation for the dephasing that would be worsening as the particle gains energy. Introducing a non-periodic array of  $\lambda_p$  was a goal of this thesis.

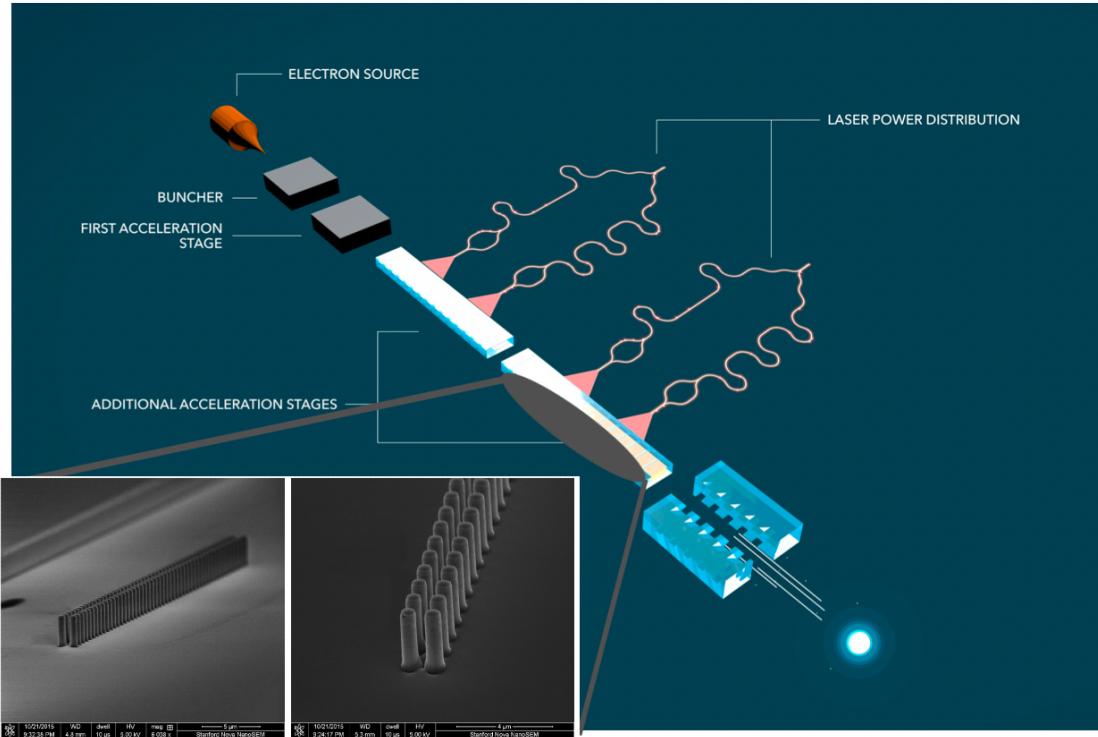


Figure 1.1: Model of an accelerator on a chip design, with an inset of an SEM image of a dual pillar accelerator structure like the ones designed in this thesis (reprinted from [18]).

## 1.0.2 Inverse Design

Photonic devices are typically designed by using intuition for a device design, parameterization of the device and then optimization of those parameters based on the performance of a desired metric [19]. Inverse design is a very different optimization method; although it has been used in other scientific fields, it has only relatively recently gained attention in photonics [19]. The desired performance metric of the device is defined by an objective function, which is then computationally optimized. Thus, by using the inverse design method, device design can be automated and more finely optimized, often yielding smaller, better performing devices [19]. Further, this allows us to traverse a much larger parameter space, leading to the discovery of structures beyond what someone could reasonably come up with [19]. Since the ultimate goal for laser accelerators is to miniaturize them enough to fit the entire photonic structure on a single chip, this makes inverse design an ideal method for optimizing these accelerator structures.

### Coordinates and Variables

The coordinates and variable definitions corresponding to the setup used in this thesis are as follows.

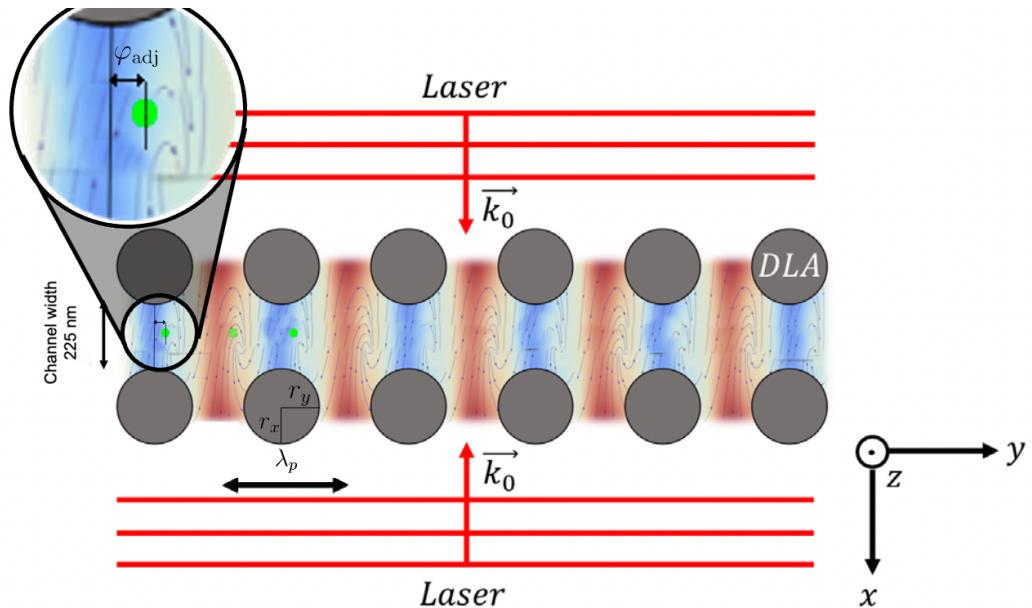


Figure 1.2: Graphical representation of the setup used in this thesis, with corresponding global parameters such as laser wavelength and channel width. Figure edited from [20] to match our definitions.

# Theory

The Lawson Woodward-Theorem states that “no combination of far fields, in an otherwise field free vacuum, can produce first order net acceleration.” [14] This essentially means that a charged particle cannot be accelerated by an electromagnetic wave (far field) in a vacuum. It also assumes a high energy particle, and that non linear (ponderomotive) forces are neglected. [17] Conceptually, this theorem comes from the idea that light is transversely polarized; hence, parallel to a particle, the particle will only feel oscillating transverse fields that cannot accelerate it. [17] If the wave is at an angle with respect to the particle, then there will be longitudinal field components, but they again will average to zero as the particle crosses the beam, leaving zero net acceleration.[21] One way around this is to use near fields to accelerate particles.

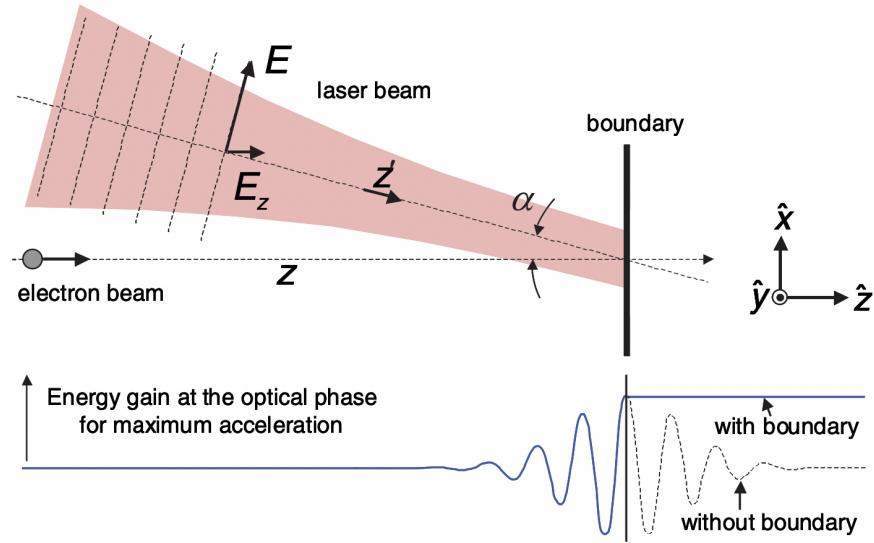


Figure 2.1: Illustration demonstrating the net zero acceleration of the Lawson Woodward Theorem. Reprinted from [21].

## 2.1 Near Field Acceleration

The inverse Smith-Purcell effect (ISP), occurs when an electron beam moves parallel and close to a metallic optical diffraction grating, in a direction perpendicular to the grating rulings [11]. The near fields accelerate the electrons, and from the Larmor Formula, we know that accelerating charged particles radiate [22]. The relationship is a synchronous one between the electrons, incident light, and grating periodicity.[11] The mathematics of this acceleration of a particle by an evanescent field are insightful in our design of DLAs. Near a single grating with periodicity  $\lambda_p$  and grating vector  $k_p = 2\pi/\lambda_p$ , the diffracted incident wave produces spatial harmonics with wave vectors  $k_{\parallel}^n = K + nk_p$ , where  $K$  is the projection of the incident wave in-plane with the particle and  $n = 0, 1, 2, \dots$  are the spatial modes [19]. The EM field of each mode can be described by

$$\mathcal{A}(\mathbf{r}, t) = \mathcal{A}_n e^{i(k_{\perp}^n \cdot z + k_{\parallel}^n \cdot \mathbf{r} - \omega t + \phi_0)} \quad (2.1)$$

Thus, the total field near the grating is a Fourier series of all spatial harmonics. If we define our particle as traveling parallel to the grating (perpendicular wrt to the 'teeth') with trajectory  $\vec{r}(t) = \vec{v}(t) \cdot t$  with a velocity  $v = |\vec{v}| = \beta c$ , then the spatial harmonic will have an in-plane phase velocity of [19]

$$\mathbf{v}_{\text{ph}} = \frac{\omega \mathbf{k}_{\parallel}}{k_{\parallel}^2} \quad (2.2)$$

With an incident wave of angular frequency  $\omega$ , continuous control of the particle requires

$$\mathbf{v}_{\text{ph}} \cdot \frac{\mathbf{v}_{\text{particle}}}{v} = \frac{\omega}{k_{\parallel}} \cos(\varphi) \stackrel{!}{=} \beta c \quad (2.3)$$

This means the projection of the accelerating mode, on the particle's trajectory, has to match the particle's velocity [19]. Here  $\varphi$  being the angle between them. This is where we derive the synchronicity condition from.

$$k_{\parallel} = \frac{\omega}{\beta c \cos(\varphi)} = \frac{k_0}{\beta \cos(\varphi)} = \frac{k_0}{\beta \cos(\varphi)} \quad (2.4)$$

where we used  $k_0 = \omega/c$ . We see from this that the grating wave-vector must be proportional to the in plane projection of the incident laser's wave-vector. Likewise, if we use  $k_0 = 2\pi/\lambda_0$  we can convert to wavelengths, and with some rearranging, we get the synchronicity condition in the form used in this thesis.

$$\lambda_p = m \cdot \beta \cdot \lambda_l \quad (2.5)$$

Here we define  $\lambda_p$  as the periodicity of the structure (cell size),  $\lambda_l = \lambda_0$  is the wavelength of the incident laser, and we have added the diffraction order  $m$ , which will be derived in a later section. In order to experience efficient acceleration, the particle must pass the grating

within a distance on the order of [19]

$$d_x \approx \frac{i}{k_{\perp}} = \frac{\beta\sqrt{1-\beta^2}\lambda_l}{2\pi} = \frac{\lambda_p\sqrt{1-\beta^2}}{2\pi} \quad (2.6)$$

our ‘channel width,’ between pillars, that the electron traverses, is chosen based on this distance.

We are only interested in the synchronous modes in this thesis, since they are the only ones which can provide acceleration. This can be shown by looking at the fields in a vacuum, where they must satisfy the wave equation

$$\left( \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathcal{A}(\mathbf{r}, t) = 0 \quad (2.7)$$

Using the field from Eq. 2.1, we get

$$k_{\perp} = k_l \sqrt{1 - \frac{1}{\beta^2 \cos^2(\varphi)}} \quad (2.8)$$

Since beta is necessarily less than 1, the perpendicular wave vector is imaginary, thus our field from Eq. 2.1 is no longer an oscillating imaginary exponential (a wave) it is now a negative exponential; meaning, the field dies off exponentially, perpendicular to the grating, and thus it is these evanescent fields which account for the acceleration. Assuming our incident laser is polarized parallel to the particle’s trajectory (perpendicular to the gratings  $\cos(\varphi) = 1$ ), the evanescent field can be determined by using Ampere’s Law, Faraday’s Law, Eq. 2.1, and

$$\mathbf{k} = k_l \begin{pmatrix} 1/\beta \\ 0 \\ i\sqrt{1-\beta^2}/\beta \end{pmatrix} \quad (2.9)$$

to calculate the fields. The only in plane fields, then, will be

$$\mathbf{E} = \begin{pmatrix} -cB_z/\beta \\ icB_z\sqrt{1-\beta^2}/\beta \\ E_z \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} E_z/(\beta c) \\ -iE_z\sqrt{1-\beta^2}/\beta \\ B_z \end{pmatrix} \quad (2.10)$$

The force from the fields is defined by the Lorentz force,

$$\begin{aligned} \mathbf{F} &= q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \\ &= q \begin{pmatrix} -cB_z(1-\beta^2)/\beta \\ icB_z\sqrt{1-\beta^2}/\beta \\ 0 \end{pmatrix} \end{aligned} \quad (2.11)$$

Showing that a longitudinal force on the particle will result in transverse deflection that is

90° out of phase. To confine the beam, methods such as alternating phase focusing must be employed [23]. This will be one of the next steps beyond this thesis. In our case, only the TM mode contributes to the acceleration [7].

## 2.2 Dual Grating Structures

Single grating structures are limited in their applicability because they produce an asymmetric, skew, force field as seen in Fig. 2.2a. Dual grating structures provide a homogeneous field by canceling out the deflecting forces of each other, therefore creating a symmetric force around the axis of the accelerator channel.

Ref. [7] states that their 3D finite-difference time-domain (FDTD) simulations of DLA structures, including the mesa, confirm that it is a good approximation to assume the z-axis is semi-infinite and thus we can work in our pseudo-2D environment. The transverse-magnetic field component  $B_z$  from Eq. 2.11, for the fundamental grating spatial mode, can be described by [7, 24]

$$B_z(x, t) = (Ce^{-x/\Gamma} + Ce^{x/\Gamma}) e^{(ik_y y - i\omega_l t + i\theta)} \quad (2.12)$$

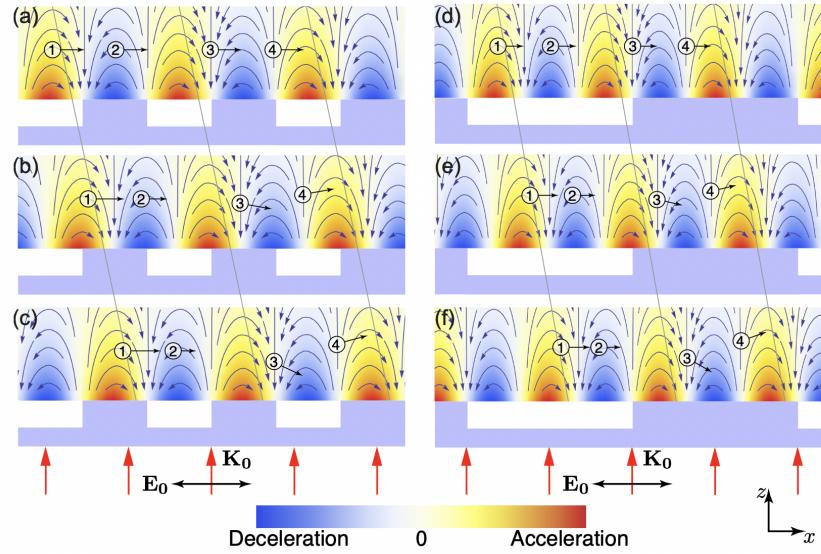
where C is the electric field coefficient of the first spatial mode,  $\theta$  is the optical phase of side 1 (hereon we assume  $\theta = 0$  since it is a matter of tuning your system and doesn't change anything conceptually),  $\Gamma = \lambda_l \gamma / 2\pi = 1/k_x$  is the optical decay length,  $\gamma = (1 - \beta^2)^{-1/2}$  is the Lorentz factor, and  $k_y = 2\pi/\lambda_p$  is the gratings wave vector [7].

Plugging equation 2.12 into equation 2.11, we can write the force in a hyperbolic form as,

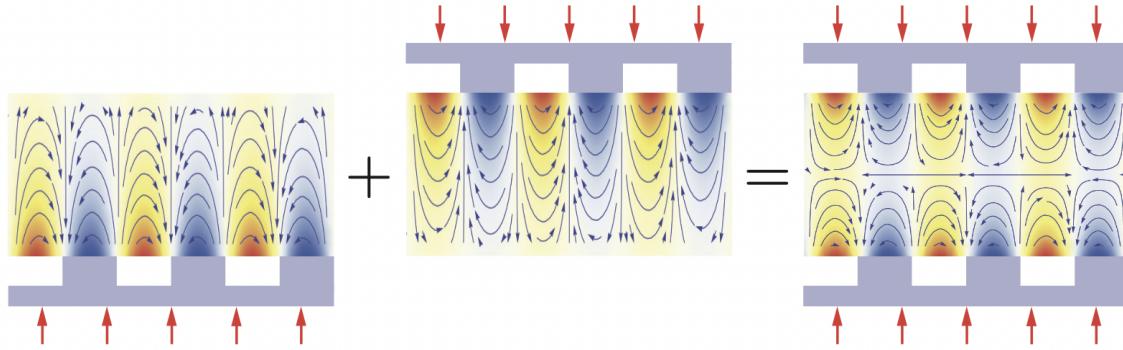
$$F_{\cosh} = \begin{pmatrix} 0 \\ 2C \cosh(k_x x) \\ 0 \end{pmatrix} \quad F_{\sinh} = \begin{pmatrix} 0 \\ -2C \sinh(k_x x) \\ 0 \end{pmatrix} \quad (2.13)$$

In the dual-drive case, the relative phase difference between the optical mode on each side can change the grating mode between a hyperbolic cosine accelerator mode with a uniform longitudinal gradient and a hyperbolic sine mode with a uniform transverse gradient [7].

These modes being separated by phase shifts of  $2n\pi$  for cosh acceleration and  $(1 + 2n)\pi$  for sinh deflection. As can be seen in Fig. 2.3 and 2.2b, the cosh mode provides a pure acceleration mode which is also a speed of light mode, meaning these equations hold for  $\beta \rightarrow 1$  [24].



(a) Depiction of the deflecting forces of a single grating accelerator structure. Reprinted from [24].



(b) By combining two gratings, the deflecting forces will cancel out, leading to a uniform acceleration gradient. Reprinted from [17].

Figure 2.2: Figures demonstrating the advantage of a dual pillar structure.

## 2.3 Non-Periodic Structures

Since these are accelerating structures, the velocity of the particles will change. This change means that our synchronicity Eq. 2.5, is actually a function of time.

$$\lambda_p(t) = \beta(t) \cdot \lambda_l \quad (2.14)$$

Therefore, in a structure with a constant periodicity, the particle will gradually fall more and more out of phase with the applied field. We can calculate the maximum length of an acceleration structure, to just before the particle falls so far out of phase that it begins to

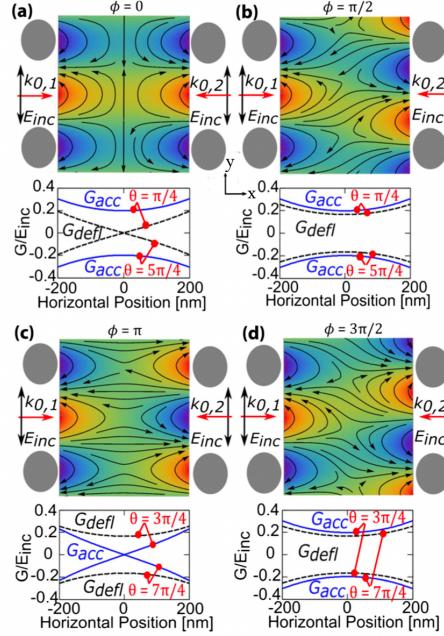


Figure 2.3: Plots showing the cosh and sinh accelerating modes of a dual drive, dual grating DLA structure with the relative phase difference between the drive lasers needed to produce each. Reprinted from [7].

experience a net deceleration. Ref. [17] provides a great derivation and explanation for the calculation of this *dephasing length*, using the force derived in Eq. 2.11. The maximum dephasing length is calculated to be

$$y_{\text{deph}} = \sqrt{\frac{\beta \lambda_l E_{\text{kin}} (\frac{E_{\text{kin}}}{m_0 c^2} + 1) (\frac{E_{\text{kin}}}{m_0 c^2} + 2)}{4G_{\max}(y_0)}} \quad (2.15)$$

where

$$G_{\max}(y_0) = \frac{ec|B_z|}{\beta\gamma} e^{-k_l y_0 / (\beta\gamma)}$$

is the maximum acceleration gradient felt by the particle. Although lasers are able to provide pulses with peak field intensities of 10s of GV/m, we are only able to harness a portion of that, due to the efficiency of turning the energy into the first (evanescent) diffraction mode. A quick *back-of-the-envelope* calculation for an electron with injection velocity  $\beta c = 0.4c$  and an incident laser wavelength  $\lambda_l = 2\mu\text{m}$ , gives us a maximum dephasing length on the order of a micron. This length is barely more than the synchronicity condition, thus we only get about one grating period of acceleration, drastically limiting their applications as accelerators. This is why we must introduce non-periodic acceleration structures which account for the particle acceleration and subsequent dephasing by sequentially adapting the size of each ‘cell’ of the structure to the phase velocity of the particle. This was a primary goal of this thesis.

## 2.4 Adjoint Method

For problems described by linear systems of differential equations, the adjoint method can be used to calculate the gradient. In the adjoint method, the problem is arranged in a form of

$$A(\phi)\mathbf{x} = \mathbf{b} \quad (2.16)$$

where  $A$  is a sparse matrix consisting of the parameters  $\phi$  describing the system and  $b$  is some ‘driving source’ [19]. From this we can define a complex objective function  $J = J(\mathbf{x})$ , which similarly depends on our complex fields  $\mathbf{x}$ . If we rearrange Eq. 2.16 and then differentiate with respect to  $\phi$ , we get

$$\begin{aligned} \frac{d\mathbf{x}}{d\phi} &= \frac{\partial A^{-1}}{\partial \phi} \mathbf{b} \\ &= -A^{-1} \frac{\partial A}{\partial \phi} A^{-1} \mathbf{b} \\ &= -A^{-1} \frac{\partial A}{\partial \phi} \mathbf{x} \end{aligned} \quad (2.17)$$

where we have used  $K^{-1'} = K^{-1}K'K^{-1}$ . Since our fields  $\mathbf{x}$  are complex, we define the complex conjugate as

$$\begin{aligned} \mathbf{x} &= \text{Re}\{\mathbf{x}\} + i \text{Im}\{\mathbf{x}\} \\ \mathbf{x}^* &= \text{Re}\{\mathbf{x}\} - i \text{Im}\{\mathbf{x}\} \end{aligned}$$

If we similarly differentiate the objective function  $J$  with respect to  $\phi$ , and then take the real part—because that is what will contribute to the acceleration gradient—we get <sup>1</sup>

$$\frac{dJ}{d\phi} = -2 \text{Re} \left\{ \frac{\partial J}{\partial \mathbf{x}} A^{-1} \frac{\partial A}{\partial \phi} \mathbf{x} \right\} \quad (2.18)$$

The factor of 2 comes from  $\mathbf{x}$  being complex-valued, hence when differentiating  $J$ , one must account for both  $\mathbf{x}$  and  $\mathbf{x}^*$  [19].

We now define a second *adjoint* system

<sup>1</sup>

$$\begin{aligned} \frac{dJ}{d\phi} &= \frac{\partial J}{\partial \mathbf{x}} \frac{d\mathbf{x}}{d\phi} + \frac{\partial J}{\partial \mathbf{x}^*} \frac{d\mathbf{x}^*}{d\phi} \\ &= \text{Re} \left\{ -\frac{\partial J}{\partial \mathbf{x}} A^{-1} \frac{\partial A}{\partial \phi} \mathbf{x} \right\} + i \text{Im} \left\{ \frac{\partial J}{\partial \mathbf{x}} A^{-1} \frac{\partial A}{\partial \phi} \mathbf{x} \right\} + \text{Re} \left\{ -\frac{\partial J}{\partial \mathbf{x}^*} A^{-1} \frac{\partial A}{\partial \phi} \mathbf{x} \right\} - i \text{Im} \left\{ \frac{\partial J}{\partial \mathbf{x}^*} A^{-1} \frac{\partial A}{\partial \phi} \mathbf{x} \right\} \\ &= -2 \text{Re} \left\{ \frac{\partial J}{\partial \mathbf{x}} A^{-1} \frac{\partial A}{\partial \phi} \mathbf{x} \right\} \end{aligned}$$

$$A^T \mathbf{x}_{\text{adj}} = -\frac{\partial J^T}{\partial \mathbf{x}} \quad (2.19)$$

Since this has to be a linearly independent system,  $A$  will be symmetric and thus  $A^T = A$ . Then the adjoint solution can be rearranged as <sup>2</sup>

$$\begin{aligned} \mathbf{x}_{\text{adj}} &= -A^{-1} \frac{\partial J^T}{\partial \mathbf{x}} \\ \mathbf{x}_{\text{adj}}^T &= -\frac{\partial J}{\partial \mathbf{x}} A^{-1} \end{aligned} \quad (2.20)$$

We can see then, from Eqs. 2.18 and 2.20 that

$$\frac{dJ}{d\phi} = 2 \operatorname{Re} \left\{ \mathbf{x}_{\text{adj}}^T \frac{\partial A}{\partial \phi} \mathbf{x} \right\} \quad (2.21)$$

Typical inverse design algorithms utilize methods such as numerical “finite difference” differentiation, which is essentially a two point approximation of the gradient vector  $\nabla_\phi J$ . In these FD optimizations, it requires the perturbation of each parameter simulated independently. Hence, the number of simulations required scales with the number of parameters. With the adjoint method, though, as can be seen in Eq. 2.16, the only thing that depends on  $\phi$  is  $A$ , and  $\frac{\partial A}{\partial \phi}$  is typically easy to compute. As Ref. [19] explains, since the simulation of the fields is computationally expensive, the adjoint method offers an orders-of-magnitude improvement in computation time over traditional FD optimization methods, which require recalculating the fields for each parameter.

#### 2.4.1 Adjoint Method for DLA

Maxwell’s equations can be written in the frequency domain as

$$\nabla \times \nabla \times E(\vec{r}) - k_l^2 \varepsilon_r(\vec{r}) E(\vec{r}) \equiv AE(\vec{r}) = -i\mu_0 \omega J(\vec{r}) \quad (2.22)$$

where  $J(\vec{r})$  is the electric current distribution,  $k_l$  is the wave vector of the laser,  $\varepsilon_r$  is the relative permittivity, and for non magnetic materials, the susceptibility  $\mu = \mu_0$ .

Since we seek to maximize the acceleration gradient which is the amount of energy gain per unit length, we need to know this gradient over one optical period. In a typical DLA we

<sup>2</sup>

$$\begin{array}{lll} A^T \mathbf{x}_{\text{adj}} = -\frac{\partial J^T}{\partial \mathbf{x}} & \rightarrow & A \mathbf{x}_{\text{adj}} = -\frac{\partial J^T}{\partial \mathbf{x}} \\ \mathbf{x}_{\text{adj}} = -A^{-1} \frac{\partial J^T}{\partial \mathbf{x}} & \rightarrow & \mathbf{x}_{\text{adj}}^T = -\frac{\partial J}{\partial \mathbf{x}} A^{-1} \end{array}$$

can write this as [19]

$$G = \frac{1}{T} \int_0^T E_{\parallel}(\vec{r}(t), t) dt \quad (2.23)$$

where  $E_{\parallel}$  is the projection of the incident field on the particle's trajectory, sampled at  $\vec{r}(t)$ .

As stated in section 2.2, we are working in two dimension; with an incident laser with frequency  $\omega$ , the fields are of the form

$$E(\vec{r}, t) = \operatorname{Re} \left\{ \tilde{E}(\vec{r}) e^{i\omega t} \right\} \quad (2.24)$$

where  $\tilde{E}$  is the complex valued electric field. In this thesis we use monochromatic beams, however, the next step would be to follow the calculations in Ref. [19] and generalize this to broadband light.

Our ‘test’ particle is set to move along the line  $x = 0$ , with velocity  $\vec{v} = \beta c \hat{y}$ , where  $c$  is the speed of light in vacuum and  $0 < \beta \leq 1$ . The particle’s position is then given by,

$$y(t) = y_0 + \beta ct \quad (2.25)$$

For an infinitely long structure (or  $T \rightarrow \infty$ ), our acceleration gradient can be interpreted as being an integral over one spatial period <sup>3</sup>. With our synchronicity condition in Eq. 2.5, we can rewrite Eq. 2.23 as [19]

$$G = \frac{1}{\lambda_p} \operatorname{Re} \left\{ e^{i\varphi_{\text{adj}}} \int_0^{\lambda_p} \tilde{E}_y(0, y) e^{i\frac{2\pi}{\lambda_p} y} dy \right\} \quad (2.26)$$

where  $\varphi_{\text{adj}} = 2\pi y_0 / \lambda_p$  is the *adjoint phase*—also referred to as the ‘injection phase’—which is the relative phase of the particle with the oscillation of the incident field, as it enters the cell. This will be expanded upon in section 2.4.4.

Following Ref. [19], we define an inner product of vectors  $\vec{a}$  and  $\vec{b}$  over a single period of volume  $V'$  as

$$\langle \vec{a}, \vec{b} \rangle = \int_{V'} (\vec{a} \cdot \vec{b}) dV = \int_0^{\lambda_p} \int_{-\infty}^{\infty} (\vec{a} \cdot \vec{b}) dx dy \quad (2.27)$$

Then the acceleration gradient becomes

$$G = \operatorname{Re} \left\{ \langle \vec{E}, \vec{\eta} \rangle \right\} \quad (2.28)$$

where we can see from Eq. 2.26,  $\eta$  is defined as <sup>4</sup>

$$\vec{\eta}(x, y) = \frac{1}{\lambda_p} e^{i\frac{2\pi}{\lambda_p} y} \delta(x) \hat{y} \quad (2.29)$$

---

<sup>3</sup>**Note:** this is an important assumption that is made for our adjoint method calculations. The consequences of this assumption will be referenced in a later section.

<sup>4</sup>We have defined our axis at  $x = 0$ , thus our delta function is  $\delta(x - 0) = \delta(x)$ .

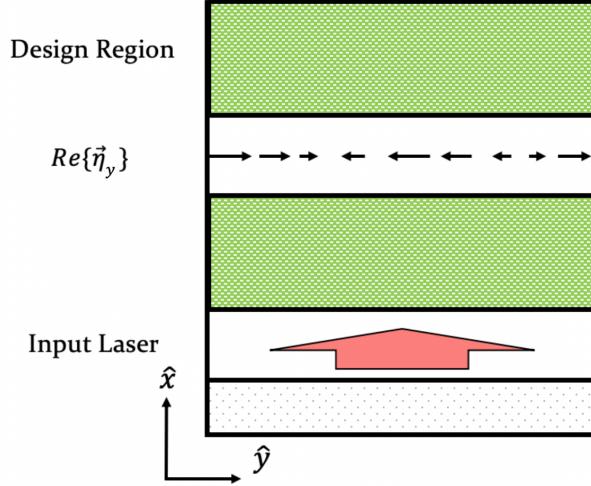


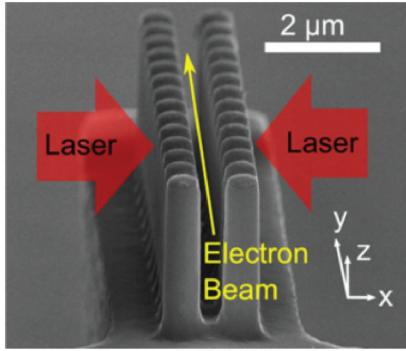
Figure 2.4: Conceptual representation of our adjoint source  $\vec{\eta}$ . Reprinted from [19].

In our DLA structure, we have dual pillar gratings as our ‘structure features’ and the structure is made of silicon. In this case, the structure features can be described mathematically as a spatially varying dielectric, defined by the complex dielectric constant  $\varepsilon_r(\vec{r})$ . The entire cell can be discretized to a grid of points, each of which will have an  $\varepsilon_r(\vec{r})$  value of either 1 (vacuum) or 3.44 (silicon).<sup>5</sup> We then define our parameters,  $\phi$ , which will change the relative permittivity of each grid point. With these definitions, our problem of maximizing the acceleration gradient now becomes a matter of seeing which  $\varepsilon_r$  value, 1 or 3.44, for each grid point, maximizes the acceleration gradient. This seems like it should be a simple problem to optimize by just doing a brute grid search [19]. While this is possible, it is certainly not very feasible. This method becomes too computationally expensive, very quickly. For example, even a small grid of  $10 \times 10 = 100$  cells would require  $2^{100} \approx 10^{30}$  simulations. Since simulation times also scale with the size of the simulation region, clearly this is almost immediately an infeasible method. Even other generalized mathematical and numerical optimization techniques are still outperformed by the adjoint method because they generally require more simulations to traverse the parameter space [19].

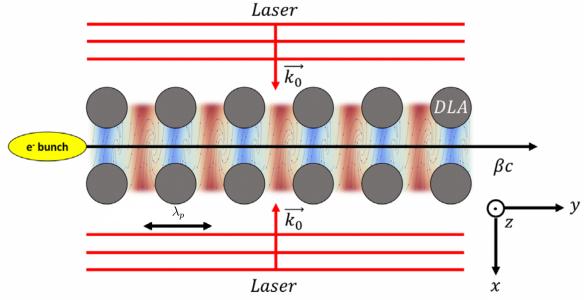
### 2.4.2 Cost Function

In order to calculate the mathematical gradient from the aforementioned adjoint method, we must define a cost function  $J(\mathbf{x})$  which is a function of the fields  $\mathbf{x}$ . Since we seek to maximize the acceleration gradient, it is only natural for us to call our cost function  $G$ . In order to derive  $G$ , we define our coordinate system as noted in Fig. 1.2, where the electron propagates in  $y$  and the  $y$ -polarized laser counter-propagates in the  $+x$  and  $-x$  directions. Fig. 2.5 provides a physical and graphical representation of the setup and coordinate definitions used in this thesis.

<sup>5</sup>Values provided at our laser wavelength of  $1.932\mu\text{m}$  from [25].



(a) SEM image of a dual pillar structure like the ones optimized in this thesis. Updated to our coordinates from [7].



(b) Graphical representation of the setup used in this thesis. Updated to our coordinates from [20].

Figure 2.5: Dual pillar dual illumination DLA structures.

From the synchronicity and Eq. 2.1, we know that the field near the gratings can be described by a Fourier series of all spatial harmonics.

$$E_y(x, y, z) = \sum_{m=-\infty}^{\infty} e_m(x, y) e^{-im \frac{2\pi}{\lambda_p} y} \quad (2.30)$$

$$e_m(x, y) = \frac{1}{\lambda_p} \int_{-\lambda_p/2}^{\lambda_p/2} E_y(x, y, z) e^{im \frac{2\pi}{\lambda_p} dy} \quad (2.31)$$

with  $m$  the number of the harmonic and  $\lambda_p$  the period of the structure in the  $y$ -direction [26].

Fig. 2.6 shows the electric field's spatial Fourier coefficients. We can see that it has a strong first harmonic and a much weaker second harmonic that is out of phase with it. So, for the purpose of maximizing the acceleration gradient, we choose to optimize only for the first harmonic. Selecting only the first harmonic,  $m = 1$ , (hereon omitting it from equations) we can plug Eq. 2.30 and Eq. 2.29 into Eq. 2.28 and get our objective function

$$G = \text{Re}\{\langle e_1 \cdot e^{i\varphi_{\text{adj}}}\rangle\} = |e_1| \cos(\varphi_{\text{adj}} + \phi_{e_1}) \quad (2.32)$$

Since optimization methods generally require a scalar value to optimize, and our goal is to control the acceleration gradient of a multi-period structure, it is only natural to use a

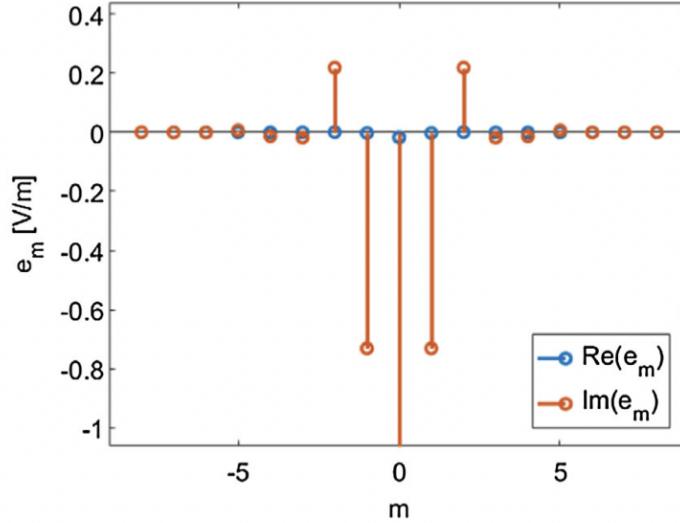


Figure 2.6: Plot of the of spatial harmonics in a DLA structure. Reprinted from [26]

least square function. This gives us our final cost value that will be optimized in this thesis.<sup>6</sup>

$$\mathbf{G} = \sum_n (G_n - d_n)^2 \quad (2.33)$$

where  $G_n$  is the value of the single-pillar objective function of the n-th cell and  $d_n$  is the value of the design curve for this cell [27]. This function is then minimized, so that the curve of the  $G_n$ -values follow the design curve as closely as possible.

### 2.4.3 Adjoint Source Radiation

In order to calculate the gradient of this cost function, we need to fully define our adjoint source. Since we have a charged particle moving, it makes sense to think of the adjoint source as a current. This radiating current density can be described by [19]

$$\vec{J}_{\text{rad}}(x, y; t) = q\beta c \delta(y - y_0 - c\beta t) \delta(x) \hat{y} \quad (2.34)$$

---

<sup>6</sup>Eq. 2.32 then becomes  $G_n = \text{Re}\{\langle e_{1n} \cdot e^{i\varphi_{\text{adj}}} \rangle\}$  and  $\mathbf{G}$  is reserved for our total, scalar cost function, defined in Eq. 2.33.

Fourier transforming Eq. 2.34 <sup>7</sup>

$$\begin{aligned}
 \vec{J}_{\text{rad}}(x, y; \omega_l) &= q\beta c\delta(x)\hat{y} \int_{-\infty}^{\infty} e^{i\omega_l t} \delta(y - y_0 - c\beta t) dt \\
 &= qe^{i\frac{\omega_l(y-y_0)}{c\beta}} \delta(x)\hat{y} \\
 &= qe^{i\frac{2\pi}{\beta\lambda_l}y} e^{-i\varphi_{\text{adj}}} \delta(x)\hat{y}
 \end{aligned} \tag{2.35}$$

where again,  $\varphi_{\text{adj}} = 2\pi y_0/\lambda_p$  is the *adjoint/injection* phase. From Eqs. 2.22 and 2.29, we see that our adjoint source for Maxwell's equations can be defined in terms of our adjoint field ( $\eta$ ) as

$$\vec{J}_{\text{adj}} = \frac{-i}{\mu_0\omega_l} \vec{\eta} \tag{2.36}$$

From this, we can see that

$$\vec{J}_{\text{adj}} = \frac{-ie^{i\varphi_{\text{adj}}}}{2\pi\mu_0\omega_l q} \vec{J}_{\text{rad}} \tag{2.37}$$

enlightening us that the adjoint field  $E_{\text{adj}}$  corresponds (up to a complex constant) to the field radiating from a particle traversing the accelerator. What this means for us is that we must simulate our structure as an accelerator ( $A\vec{E} = -i\omega\mu_0\vec{J}_{\text{acc}}$ ) and a radiator ( $A\vec{E}_{\text{adj}} = -i\omega\mu_0\vec{J}_{\text{adj}}$ ) [19].

#### 2.4.4 Adjoint Source and Phase

In the simulations for this thesis, there were 2 primary ways to implement the adjoint source into the FDTD solver, a collection of successive dipoles, or a plane wave with a  $y$  varying phase [27]. Both simulate the particle's relative phase with the field as it traverses the structure. The plane wave approach was selected for this thesis since it is more easily applicable to the non-periodic structures. From Eqs. 2.29 and 2.36, the phase envelope of the adjoint source can be defined, up to a complex constant, by <sup>8</sup>

$$E_{y_{\text{adj}}}(y) = e^{i2\pi(\frac{y+\lambda_p/2}{\lambda_p})} \tag{2.38}$$

where we have used  $\frac{y+\lambda_p/2}{\lambda_p}$  to ensure that our spatial period is centered within the cell. The phase of this source is determined by the periodicity of the structure  $\lambda_p$ .

To change the relative phase between the particle and field, we look at Eqs. 2.32 and 2.37 and see that we can either change the adjoint phase  $\varphi_{\text{adj}}$  or the phase of the first harmonic  $\phi_{e_1}$ . In a multi-cell structure, we need to remember to specify that this will be the *injection* phase, and thus we will specifically need to change  $\phi_{e_1}$  (the phase of the first harmonic

<sup>7</sup>

Forward Fourier transform  $H(\omega) = \mathcal{F}(h(t)) = \int_{-\infty}^{\infty} f(t)e^{i\omega t} dt$

Inverse Fourier transform  $h(t) = \mathcal{F}^{-1}(H(\omega)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(t)e^{-i\omega t} dy$

<sup>8</sup>Our complex constant here is  $\frac{-i}{2\pi\mu_0\beta c}$ . We do not include this constant in our subsequent derivations, as it does not change the performance or operation of the optimizer.

of the **first cell**). Since we defined the adjoint phase as the phase difference between the incident laser and the injection of the particle into the structure, we can simply adjust the phase of the source by adding a constant shift to the exponential in Eq. 2.38.

$$E_{y_{\text{adj}}} = e^{i2\pi(\frac{y+\lambda_p/2}{\lambda_p})+i\varphi_{\text{adj}}} \quad (2.39)$$

For periodic structures, one can also shift the entire adjoint source; since it depends on  $y$ , changing  $y$  will introduce a relative phase shift.

We may also change the phase of the first harmonic, which will introduce a phase offset. As Ref. [27] shows, the  $y_{\text{pos}}$  of the pillar within its (semi-)periodic ‘cell’ principally shifts the fields along the  $y$ -axis. This will shift the relative phase between the particle’s trajectory and the phase of the fields in a given cell.

### 2.4.5 Adjoint Inverse Design

Now, we will detail how to use the adjoint method and inverse design, in our specific case of this thesis—dual pillar DLA structures. Eq. 2.18 provides us with a formula for the gradient of a cost function in the adjoint method. Using our cost function  $G$  (Eq. 2.32), we know the forward and adjoint fields from Eq. 2.21, in their respective forms from Eqs. 2.31 and 2.29. The only thing then, that we need to calculate to be able to find the gradient of our cost function is  $\frac{\partial A}{\partial \vec{\phi}}$ . Since our structure features are dual pillars made of silicon, as discussed in Section 2.4.1, we use a grid of  $\varepsilon_r$  to describe the system. Our parameters  $\vec{\phi}$  are now the pillar radii ( $r_x$  and  $r_y$ ) and their position within their cell ( $y$ ). Our matrix then becomes  $A(\vec{r}) \rightarrow A(\varepsilon_r(\vec{r}))$ , hence

$$\frac{\partial A(\vec{r})}{\partial \vec{\phi}} = \frac{\partial A(\vec{r})}{\partial \varepsilon_r} \frac{\partial \varepsilon_r(\vec{r})}{\partial \vec{\phi}} \quad (2.40)$$

Comparing this to Eq. 2.22, we see that

$$\frac{\partial A(\vec{r})}{\partial \varepsilon_r} = -k_l^2 \quad (2.41)$$

For  $\frac{\partial \varepsilon_r(\vec{r})}{\partial \vec{\phi}}$ , we can use a simple two point derivative

$$\frac{\partial \varepsilon_r(\vec{r})}{\partial \vec{\phi}} = \frac{\Delta \varepsilon_r(\vec{r})}{\Delta \vec{\phi}} \quad (2.42)$$

finally yielding

$$\frac{\partial A(\vec{r})}{\partial \vec{\phi}} = -k_l^2 \frac{\Delta \varepsilon_r}{\Delta \vec{\phi}} \quad (2.43)$$

which comprise the entries of a symmetric matrix. Now our gradient of the cost function is

$$\frac{dG}{d\vec{\phi}} = -2k_l^2 \operatorname{Re} \left\{ \mathbf{x}_{\text{adj}}^T \frac{\Delta \varepsilon_r}{\Delta \vec{\phi}} \mathbf{x} \right\} \quad (2.44)$$

Plugging in the equations for our fields gives us the gradient calculation used in this thesis for DLA inverse design.<sup>9</sup><sup>10</sup>

$$\frac{\partial G_n}{\partial \vec{\phi}} = -2k_l^2 \operatorname{Re} \left\{ \int_S e_{1n}(0, y) e^{i\varphi_{\text{adj}}} \frac{\Delta \varepsilon_r(x, y)}{\Delta \vec{\phi}} dS \right\} \quad (2.45)$$

where  $S$  is the simulation region [19]. Extending to the multi-pillar case is simply a matter of differentiating Eq. 2.33 with Eq. 2.45 as the single cell gradient.

$$\frac{\partial \mathbf{G}}{\partial \vec{\phi}} = 2 \sum_n (G_n - d_n) \frac{\partial G_n}{\partial \vec{\phi}} \quad (2.46)$$

Here it should be clarified that, although, the notation remained  $\vec{\phi}$  for the single and multi-pillar gradient calculations, there is an important difference in the extension to more cells.

$$\vec{\phi}_{\text{1-pillar}} \neq \vec{\phi}_{\text{multi-pillar}}$$

Since our matrix  $A$  consists of elements operating on the whole field, the multi-pillar parameter vector actually has each cell's parameters as individual entries, e.g.

$$\vec{\phi}_{\text{multi-pillar}} = (\phi_{1a}, \phi_{1b}, \phi_{2a}, \phi_{2b}, \phi_{3a}, \dots)$$

So,  $\partial G_n / \partial \vec{\phi}_{\text{multi-pillar}}$  is actually the gradient of each cell with respect to the entire structures parameters. Although, in a linearly independent system, all elements from outside the cell end up going to zero, it is enlightening to understand the math conceptually. Our system is not linearly independent, the consequences of which will be discussed more in subsequent sections.

#### 2.4.6 Particle Acceleration and Non-Periodicity

As referenced in footnote 3, our calculations and derivations up to this point have assumed a semi-infinite periodic structure. However, as noted in section 2.4.4, our aim of particle acceleration is necessarily incongruous with this assumption due to the aforementioned dephasing.

---

<sup>9</sup>In Eqs. 2.27 and 2.28, we assumed a semi-infinite periodicity in  $x$ . However, its mathematical purpose is to use the delta function to shift the source to the desired axis, here  $x = 0$ . So, we can shift the function and integrate over the simulation area.

<sup>10</sup>

$$\begin{aligned} G &= \frac{1}{\lambda_p x} \int_0^{\lambda_p} \int_{-x/2}^{x/2} e_1(x, y) e^{-i \frac{2\pi}{\lambda_p} y} \cdot e^{i \frac{2\pi}{\lambda_p} y} \delta(x) dx dy \\ \frac{\partial G}{\partial \vec{\phi}} &= \frac{1}{\lambda_p x} \int_0^{\lambda_p} \int_{-x/2}^{x/2} e_1(0, y) e^{i\varphi_{\text{adj}}} \frac{\Delta \varepsilon_r(x, y)}{\Delta \vec{\phi}} dx dy \\ &\quad \frac{1}{\lambda_p x} \int_0^{\lambda_p} \int_{-x/2}^{x/2} dx dy \rightarrow \int_S dS \end{aligned}$$

We can find the energy gain of the particle by integrating the Lorentz force over the distance it is applied ( $\int F dy = Pt \rightarrow (Pt_2 - Pt_1) = W_2 - W_1 = \Delta W$ ), then

$$\begin{aligned}\Delta W(x, y; s) &= q \int_{-\lambda_p/2}^{\lambda_p/2} E_y(x, y; t = (y + s)/v) dy \\ &= q \int_{-\lambda_p/2}^{\lambda_p/2} \operatorname{Re}\{E_y(x, y)e^{i\omega_l(y+s)/v}\} dy\end{aligned}\quad (2.47)$$

The variable  $s$  denotes the relative position of the particle behind an arbitrarily defined reference particle moving at  $y = vt$  [26]. Using the field defined in Eq. 2.30, the energy gain becomes

$$\Delta W(x, y; s) = q \operatorname{Re}\left\{ e^{i\frac{2\pi s}{\beta\lambda_l}} \sum_{m=-\infty}^{\infty} e_m(x, y) \lambda_p \operatorname{sinc}\left(\frac{\lambda_p}{\beta\lambda_l} - m\right) \right\} \quad (2.48)$$

If the argument of the sinc function is anything other than zero, the energy gain will average to zero, hence we once again find our synchronicity condition 2.5 [26]. With that, Eq. 2.48 simplifies to

$$\begin{aligned}\Delta W(x, y; s) &= q \lambda_p \operatorname{Re}\left\{ e_1 \cdot e^{i2\pi\frac{s}{\beta\lambda_l}} \right\} \\ &= q \lambda_p |e_1| \cos\left(2\pi\frac{s}{\beta\lambda_l} + \varphi_1\right)\end{aligned}\quad (2.49)$$

where  $\varphi_1 = \arctan(\operatorname{Im}\{e_1\}/\operatorname{Re}\{e_1\})$  is the phase of the first spatial harmonic of the first cell.

The energy gain is maximized at  $s_{\text{opt}} = -\varphi_1 \beta \lambda_l / 2\pi$  for positively charged particles,  $s_{\text{opt}} = -(\varphi_1 + \pi) \beta \lambda_l / 2\pi$  for negatively charged particles, and zero energy gain is found at  $s_{\text{opt}} = -(\varphi_1 \pm \frac{\pi}{2}) \beta \lambda_l / 2\pi$  [26]. Since we can control the phase of the particle's trajectory, as discussed in section 2.4.4, we can control our arbitrary particle's position  $s$  relative to the field's phase with the adjoint source phase  $\varphi_{\text{adj}}$ . Unless otherwise noted, the particle was set to be in-phase with the field, hence the cos term is 1 and our acceleration calculation used, for each cell in this thesis, was

$$\Delta W_n = q \lambda_{p_n} \operatorname{Re}\{e_{1_n} \cdot e^{i2\pi\lambda_{p_n}/\beta_n(\vec{r})\lambda_l}\} \quad (2.50)$$

were the subscript  $n$  is used to denote the cell number. In order to make  $\lambda_p \rightarrow \lambda_p(\vec{r})$ , we see from Eq. 2.14 (note: we convert from  $\beta(t) \rightarrow \beta(\vec{r})$  with the linear relation  $\vec{r}(t) = \vec{v}(t)t$ ) that we need to find  $\beta(\vec{r})$ . Well, we know  $\beta(r_0)$  because that is simply the injection energy of the particle, but to find the velocity at a later time, given the change in energy, we must find the function of the energy

$$W_{n+1} = W_n + \Delta W_n \quad (2.51)$$

here we work with the discrete cells labeled by  $n$ . In the case of particle tracking, the sum in Eq. 2.48 would simply be an integral over the structure length (technically just a more

refined sum because it is still an FD simulation). From this, we back calculate  $\beta$ <sup>11</sup>

$$\beta_n(\vec{r}) = \sqrt{1 - \left( \frac{1}{W_n/m_e + 1} \right)^2} \quad (2.52)$$

which plugging into Eq. 2.14 gives us our non-periodic structure equation

$$\lambda_{p_n}(\vec{r}) = \beta_n(\vec{r})\lambda_l \quad (2.53)$$

Due to the interdependence of the periodicity, velocity, and field component (here  $\text{Re}\{e_1\}$ ), the periodicity is not a trivial parameter to try to make a dynamic variable for the purposes of adjoint optimization. This peculiarity will be discussed further in section 4.2.

For the purposes of this thesis, the above equations were used in the calculation of the periodicity based on the desired  $e_1$  design curve. This set the periodicity of the structure as a fixed, predetermined, ‘optimal’ parameter. In line with this approach, we also view the electron as accelerating in discrete intervals—each cell—with no relative dephasing.

In a real accelerator structure, of course, the acceleration will be continuous and so there will be some dephasing over the course of each cell. Real-time particle tracking algorithms are the next step in this research, as they would allow the calculation of the dephasing and thus a more accurate calculation of the acceleration and velocity, allowing for better phase tuning and longer accelerators.

## 2.4.7 Fringing

One difficulty with optimization of non-periodic structures is the loss of periodic boundary conditions. As such, our model of a plane wave impinging upon a semi-infinite periodic structure, must be adjusted. In principle we still assume a plane wave-front, however, we no longer have an infinitely periodic structure; meaning, there will be diffraction of the wave at the ends of the structure. We know from Huygen’s principle that every point on a diffracted wave-front is itself a point source [28]. Therefore, at the edges, we will see waves that propagate parallel to the grating causing interference, colloquially known as ‘edge effects.’

---

<sup>11</sup>Here we note that we use the electron mass defined as  $0.511 \text{ MeV}/c^2$ , hence why  $c$  is omitted. We also note that we use the relativistic energy where  $\gamma = \frac{1}{\sqrt{1-\beta^2}}$  is the Lorentz factor.

$$\begin{aligned} \beta &= \sqrt{1 - \left( \frac{1}{\frac{W}{m_e c^2} + 1} \right)^2} \\ W = (\gamma - 1) \left( \frac{0.511E6}{c^2} \right) c^2 \rightarrow \beta &= \sqrt{1 - \left( \frac{1}{\frac{W}{0.511E6} + 1} \right)^2} \end{aligned}$$

In simulation terms, these are often called transient sources [29]. For this thesis, we simply ignore the first and last few periods, and set the injection phase to that of first non-interference cell. The next step would be to also optimize for the ends of the structure. Since this would require a different cost function, the ends of the structure would likely have to be separately optimized then added as a constant/fixed structure, to the optimization simulations of the central majority of the structure.

## 2.5 Numerical Optimization

Gradient descent is a rather straightforward method of mathematical optimization which just involves moving in the direction of the gradient—negative for minimization, positive for maximization—until an extrema is reached. In this sense, the gradient plays the same role as a potential [30]. Our step in this space is

$$\Delta \vec{\phi} = \pm \kappa \frac{\partial G}{\partial \vec{\phi}} \quad (2.54)$$

where  $\kappa$  is the step size (a constant scaling factor). If the step size is too large you may overshoot extrema and if it is too small, you may get stuck in local extrema and optimization takes much longer, so it must be chosen appropriately [30].

To efficiently traverse this parameter space randomly generated ‘noise’ is added to the step in parameter space, making Eq. 2.54 go to

$$\Delta \vec{\phi} = \pm \kappa \frac{\partial G}{\partial \vec{\phi}} + \xi_{\text{noise}} \rightarrow \pm \kappa \langle \frac{\partial G}{\partial \vec{\phi}} \rangle \quad (2.55)$$

The noise allows us to explore the parameter space while, in theory, averaging to zero over the course of the many step optimization (Fig. 2.8) [30].

More advanced gradient descent methods apply *momentum* to the gradient to more efficiently reach the extrema. One of the most popular [30, 31] is known as adaptive moment estimation (adam). Traditional optimization involves using the two methods to randomly traverse the parameter space. For large and/or complicated parameter spaces, this may not be a very efficient or effective method. More advanced use of these kinds of functions involves providing them with the gradient and the cost function, thereby eliminating the need for the stochastic stepping (depending on initial guess). Unfortunately, because our system is not linearly independent, the gradient in our case is dependent on the matrix  $\frac{\Delta \varepsilon_r}{\Delta \vec{\phi}}$  (Eq. 2.45) which is constructed for the whole structure. We cannot use any out-of-the-box optimizers since they apply their optimization algorithms after the gradient has been calculated, which does not account for the coupling. To clarify, our fields have effects well beyond the bounds

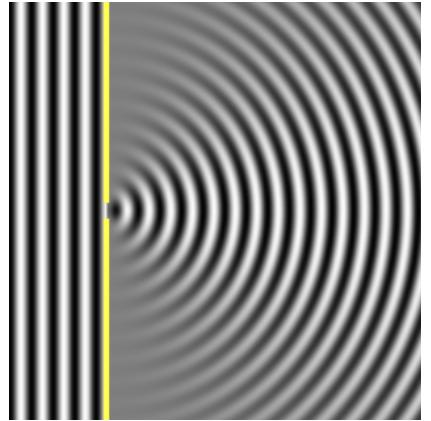


Figure 2.7: Graphical representation of the diffraction effects occurring at the edges of the structure. Reprinted from [28].

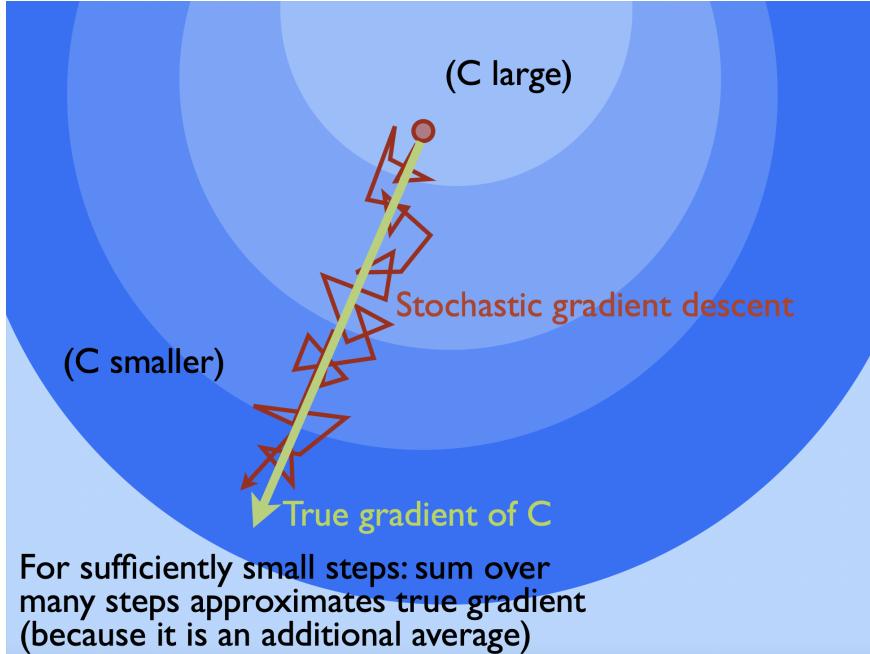


Figure 2.8: Illustration of how stochastic gradient descent works, conceptually. Reprinted from [30]

of their respective cells, introducing a significant amount of interaction between cells. That is, a change to the parameter of one cell will effect several of its neighbors. This interaction means that our system is not linearly independent, hence  $A$ —or now  $\frac{\varepsilon_r}{\Delta \vec{\phi}}$ —is not diagonal. We may still use the adjoint method if the parameter space is symmetric; however, this is not necessarily the case. Our matrix is calculated for the whole structure and as such the gradient cannot be adjusted after its construction because this *ex post facto* manipulation of the gradient does not account for these interactions. For this thesis, a custom optimization algorithm needed to be developed, which will be explained in more detail in the next section.

### 2.5.1 Constrained Optimization

Our problem involves the imposition of bounds and constraints. Although there have been amazing strides made in the nano-fabrication industry, there are still fabrication and material constraints. Our boundaries come from fabrication limits and the geometrical limits of the dual grating structure design.

A general constrained optimization problem can be expressed as

$$\begin{aligned}
 & \text{find} && \min_{x_i} f(x) \\
 & \text{subject to} && g_j(x) \leq 0 \quad j = 1, \dots, n_g \\
 & && h_l(x) = 0 \quad l = 1, \dots, n_h \\
 & && \underline{x}_i \leq x_i \leq \bar{x}_i \quad i = 1, \dots, n_x
 \end{aligned} \tag{2.56}$$

where  $g(x)$  is the vector of inequality constraints,  $h(x)$  is the vector of equality constraints, and  $\underline{x}$  and  $\bar{x}$  are lower and upper design variable bounds (also known as bound constraints) [32].

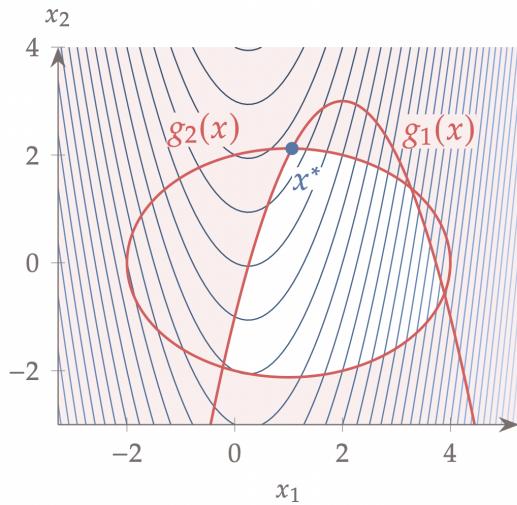


Figure 2.9: Graphical representation of a constrained problem, the contours indicating the objective function, while the two constant curves are the constrain functions. The white area is the only section of the parameters space that contains a solution fulfilling both constraints. Reprinted from [32]

Both objective and constraint functions can be nonlinear, but they should be  $C^2$  continuous to be solved using gradient-based optimization algorithms. The inequality constraints are expressed as “less than” without loss of generality because they can always be converted to “greater than” by putting a negative sign on  $g$  [32].

As was mentioned in the previous section and will be detailed more in section 3.1.4, our adjoint method required a custom optimization algorithm that would apply all bounds, constraints, and weights to the  $\frac{\Delta \varepsilon_r}{\Delta \phi}$  matrix before the calculation of the gradient.

For example, one of the constraints that we are limited by is the fabrication resolution—referring to the smallest feature size that can be successfully constructed. In our case, this was set to be 10nm. While you would ideally like to be able to make all of your steps in discrete steps of exactly 10nm, due to the aforementioned  $\varepsilon_r$  matrix, this is not how the gradient is calculated, and imposing such a constraint is a non-trivial matter.

Ref. [33] provides a possible solution to this problem in the form of level-set parameterizations, which can be used to define a discrete structure. With a level-set parameterization, a binary permittivity distribution is defined via a level-set function. Whenever the function is above 1/2, the permittivity has one value, and whenever the function is below 1/2, the permittivity has another value. In other words, the contour of the device is defined by where the function crosses zero. An example of is shown in Fig. 2.10.

According to Lumerical—the EM solver software used for this thesis—this is how the Yee cells are programmed too [29]. We have a discrete grid of points which represent a space of binary permittivity values, which similarly rounds depending on the value of the cell. One way to implement the level-set parameterization would be to run the matrix through a convolutional filter to ‘downscale’ to the 10nm constraint.

Ref. [33] provides more detail on the construction and use of level-set functions as penalties for the optimization gradient. Care would need to be taken in the implementation of a method like this because it would, again, need to be applied to the construction of the  $\varepsilon_r$  matrix.

### 2.5.2 Non-Convex Optimization

Another complication with this system is the parameter space is non-convex. In a convex parameter space, there is one global minimum and the gradient is consistently and strongly in the direction of that global minimum (e.g. a parabola). In a non-convex parameter space, there are many local extrema and so it becomes increasingly more difficult to find the global minimum. In Fig. 2.11a you can see how a simple optimization algorithm can easily get stuck in local minima, or overshoot other minima. Fig. 2.11b shows a conceptual image of the non-convex, constrained optimization being done in this thesis.

Since we are working in a very high dimensional parameter space  $n \times m$ , where  $n$  is the number of free parameters and  $m$  is the number of periods, and the fields for every cell are coupled, our parameter space is significantly more complicated. An intuitive model of the parameter space, though, is what I will refer to as an ‘egg carton’ parameter space because there are many local minima—see Fig. 2.11c for a visual example. To efficiently and effectively traverse this parameter space, we use what is known as a basin-hopping algorithm.

Basin-hopping is a two part method that uses a global stochastic stepping algorithm to randomly step around the parameter space. Each one of these ‘basins’ that the global algorithm starts in, is locally minimized. Then, the lowest minimum is then taken as the global minimum. The concept was designed to mimic the natural process of energy minimization of clusters of atoms [34]. Essentially, the basin-hopping algorithm picks random initial values (within the bounds and constraints) runs the optimization, keeps the minimum, picks new random initial values, runs the optimizer again, and so on; quickly and more effectively traversing the non-convex parameter space.

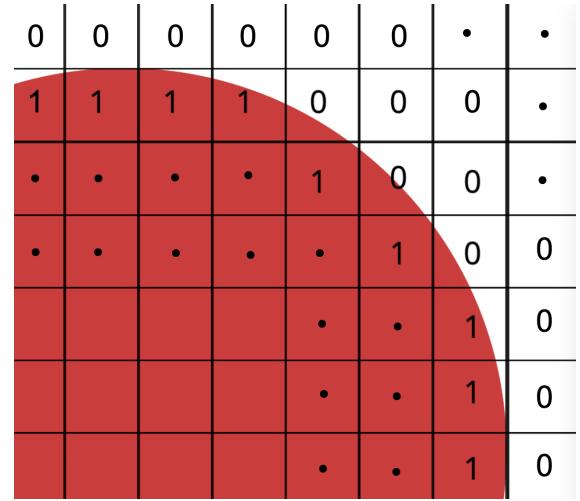
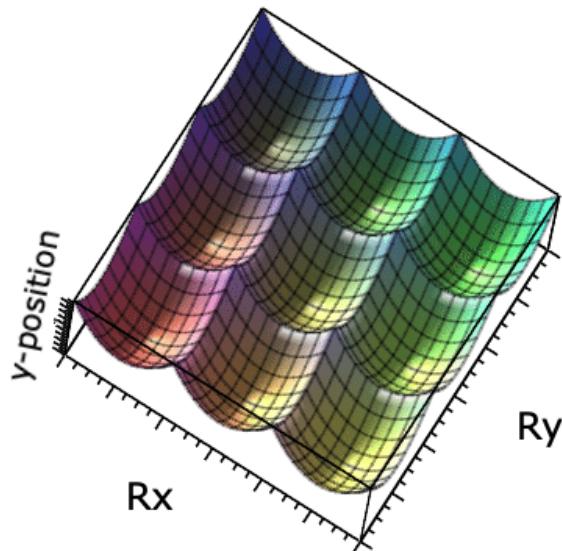
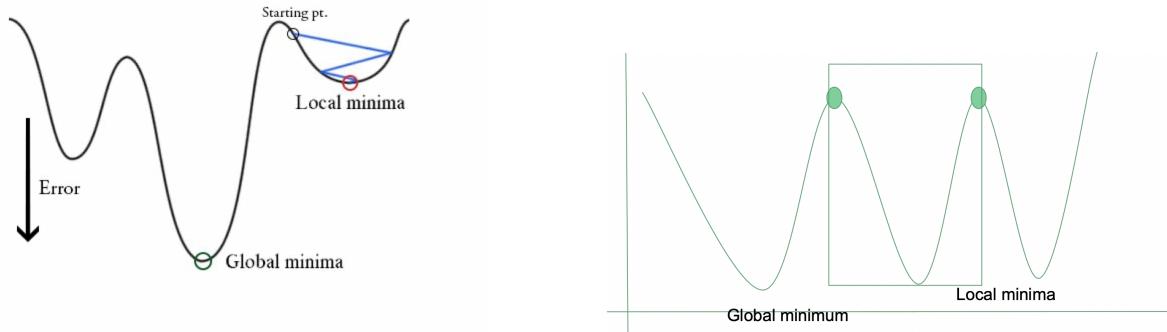


Figure 2.10: Graphical representation of how a simple level-set function determines grid values. If a grid element is  $< 0.5$  it is set to 0, if it is  $\geq 0.5$  it is set to 1.



(c) Note: not an actual plot of the parameter space used. Purely illustrative depiction of the kind of complicated parameter space seen in these optimizations. [36]

Figure 2.11: Constrained optimization of a non-convex parameter space.

# Methods

In this section, we will detail the optimization algorithm used in this thesis. There will be definitions of functions as well as explanations behind the development of the algorithm. As noted before, this custom optimization algorithm became a primary focus of this thesis, due to its necessity within our approximation of the adjoint method.

## 3.0.1 Lumerical

The EM field simulations were done with Ansys/Lumerical's finite-difference time-domain (FDTD) field solver. In this type of solver, the simulation area is divided into a discrete grid and Maxwell's equations are solved at these discrete points, in discrete steps in time [25]. Although Maxwell's equations are solved in time, the software Fourier transforms them to the frequency domain in order to view the system like Eq. 2.22. This software has several API (Application Programming Interface) options, but python was chosen as the scripting language for its flexibility and open source development. Python's scientific library's make it a natural choice for our linear algebra based adjoint calculations and gradient descent based optimization.

## 3.1 Custom Optimization Algorithm

Fig. 3.1 gives a brief but insightful view of the algorithm. We will discuss the algorithm in the computer-logical order in which it is called. While perhaps not the most concise order to go in, it is the most enlightening and closely follows the flowchart. First the user sets global constants like the laser wavelength, particle injection energy, etc. Then the user creates their desired design curve. The design curve has to correlate to a property of the field harmonic ( $e_1$ ), whether that be the magnitude or phase or a related quantity like the energy. Ultimately, if not the magnitude or phase of ( $e_1$ ), the curve must be transformed to one of those quantities because of how we defined our cost function Eq. 2.32. In a non-periodic structure, higher ( $e_1$ ) values are possible at higher  $\beta$  (and thus larger  $\lambda_p$ ). This dependence of the maximum  $|e_1|$  on  $\lambda_p$  is quite complicated and analytically beyond the scope of this thesis. Finding the maximum can only be done numerically. The results of this will be

discussed in the results section, but suffice to say for now that the maximum possible  $\text{Re}\{e_1\}$  values follow a curve, one which can be applied to the design curve to compensate for the difference with a periodic structure.

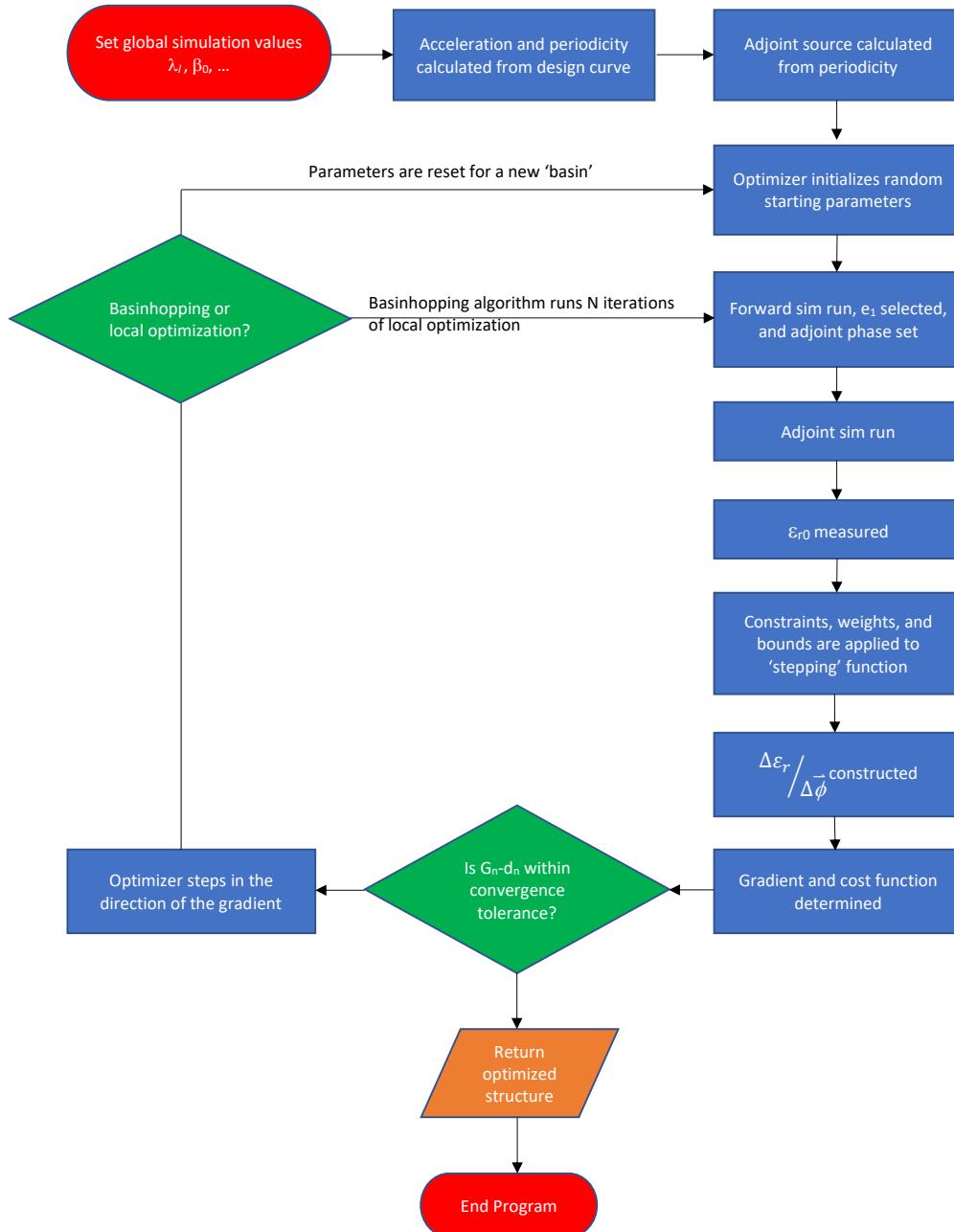


Figure 3.1: Flowchart of the basic functioning of the custom optimization algorithm developed and used for this thesis.

### 3.1.1 Periodicity Calculation and Simulation Setup

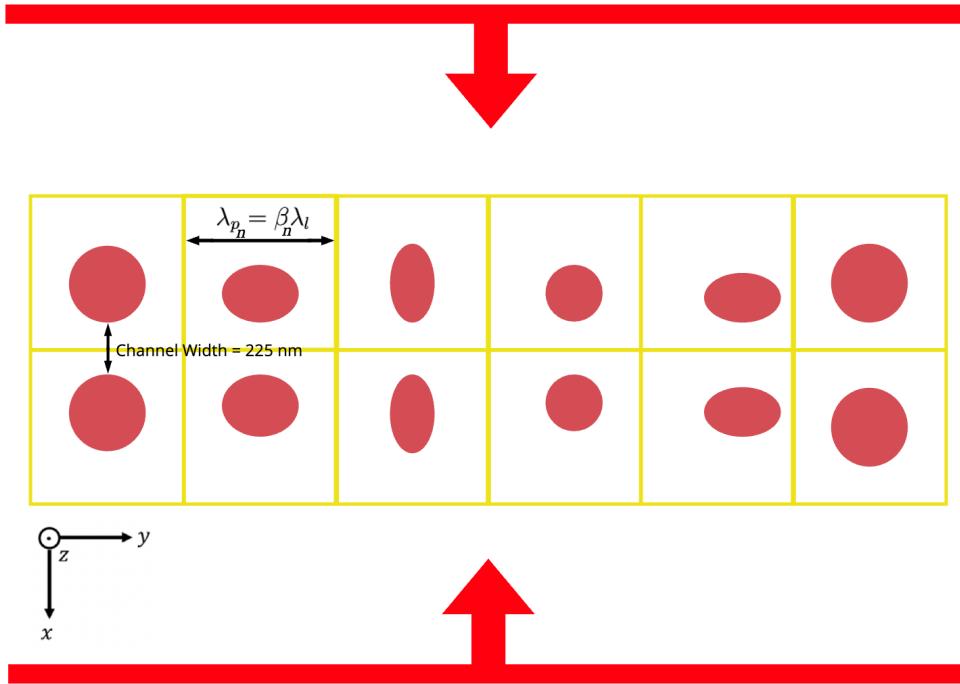


Figure 3.2: Depiction of the simulation setup used in this thesis. Edited version, original from [27].

As noted in section 2.4.6, we use Eqs. 2.52 - 2.55 to pre-determine the optimal periodicity of the structure, based on the design curve  $d_n$  from Eq. 2.33. From this periodicity array, we derive our simulation area parameters.

The basin-hopping algorithm works by re-setting certain optimization parameters, then randomly generating a set of initial parameters  $\vec{\phi}$  (within the bounds) before finally feeding that random guess into the actual ‘optimizer’. By creating the random initial guesses, one is able to stochastically traverse the parameter space. Each initial guess serves as a starting point, from which the optimizer runs only a few, large step iterations, allowing for efficient, effective traversal of the parameter space.

The actual optimizer runs a set number of calls to the objective function which returns  $G_n$  and  $\frac{\partial G_n}{\partial \vec{\phi}}$  (from Eqs. 2.32 and 2.46 respectively), calculates  $\frac{\partial \mathbf{G}}{\partial \vec{\phi}}$  with Eq. 2.33 and  $G$  from Eq. 2.32. The user sets a predefined step size  $\kappa$  (Eq. 2.55) and the optimizer steps around the parameter space.

The main function that is run by the optimizer has the parameter vector  $\vec{\phi}$  as an input and returns  $G_n$  and  $\frac{\partial G_n}{\partial \vec{\phi}}$ . It opens Lumerical, via the lumapi python API library, and sets up an FDTD simulation area every iteration. It runs the forward simulation, adjoint simulation, gathers the epsilon matrix and calculates  $\frac{\partial G_n}{\partial \vec{\phi}}$  from Eq. 2.45.

Inside the main function, the structure is built using the input design parameters  $\vec{\phi}$ , and

field & index monitors are set based on the periodicity array. The monitors, then, are used to define the cells of the structure (be they periodic or non-periodic). Then a forward simulation is run and the field values  $\tilde{E}(\vec{r})$  returned. Since our values are an array made up of the grid points in our FDTD simulation region, divided into arrays according to the periodicity; in non-periodic structures, these arrays will not be of equal length. To account for this, all arrays are linearly interpolated to the size of the array of the largest cell.

### 3.1.2 Fourier Transform

We now have the full field, but since we seek to optimize the first spatial harmonic, we must perform the Fourier transform from Eq. 2.31. We use the python package, scipy's, fast Fourier transform (fft) function. This is a 1D discrete Fourier transform defined by

$$e_m(x, y; m) = \sum_{n=0}^{N-1} E(x, y; t) e^{-i2\pi \frac{m \cdot n}{N}} \quad (3.1)$$

This function returns results in the following arrangement [37]

	Negative Frequency	DC Offset	Positive Frequency
N even	$e_{m=\frac{N}{2}} \dots e_{m=N-1}$	$e_{m=0}$	$e_{m=1} \dots e_{m=\frac{N}{2}-1}$
N odd	$e_{m=\frac{N+1}{2}} \dots e_{m=N-1}$	$e_{m=0}$	$e_{m=1} \dots e_{m=\frac{N-1}{2}}$

We want the first non-DC spatial harmonic, hence we select  $e_{m=1}$  which nicely matches our previous notation. This is a complex-valued function. Fourier series coefficients are bounded by 1, since this would indicate a single frequency source. Since we use a dual drive system in this thesis, the returned  $e_1$  value must be divided in half to retain the normalization of the Fourier series.

### 3.1.3 Adjoint Simulation

In section 2.4.4, we explain that the adjoint phase is defined as the phase between the incident laser and the particle's velocity. Since we typically want the particle injected perfectly in phase with the laser, this value should be set to  $-\phi_{e_1}$  (i.e. minus the phase of the first harmonic of the first cell) to maximize the cos in Eq. 2.32. Due to the interference and edge effects of the ends of a non-periodic (any non-infinitely long) structure, in this thesis the first several periods were essentially 'ignored' and the injection phase (*adjoint phase*) was set to the phase of a later period (generally  $e_{1,5}$ ). This is a reasonable approximation to make because once the optimization of the ends of the structure or particle tracking are implemented, in future work, the ends of the structure can be matched to the interior cells, thereby not changing the validity of the results in the interior of the structure. From Eq. 2.32, we use  $\phi_{e_1} = \arctan\left(\frac{\text{Im}\{e_1\}}{\text{Re}\{e_1\}}\right)$  to define our phase.

Once the adjoint phase is set, the adjoint source needs to be calculated. From Eq. 2.38 we

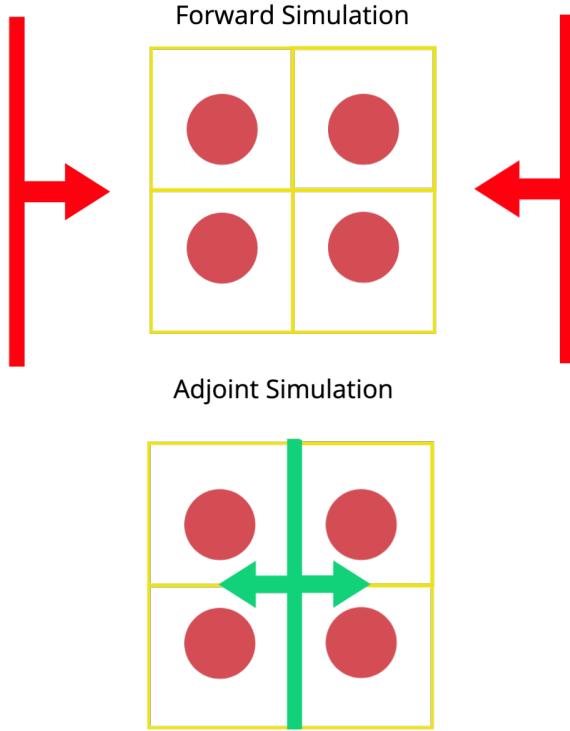


Figure 3.3: Illustration of the forward and adjoint simulation setup. Edited version, original from [27].

see that the adjoint source envelope is spatially varying; however, in an accelerating structure, the periodicity is now also a function of position  $\lambda_p(y)$ . Since the goal of this thesis was to examine the optimization of non-periodic structures, the acceleration was approximated as being at instantaneous intervals of the cell edges. The phase then is calculated for each cell's given periodicity. This produces multiple, discontinuous sinusoids of increasing frequency (due to increased velocity and thus  $\lambda_p$ ). The discontinuities are dealt with by simply concatenating them. Producing, in effect, an exponential with an instantaneously changing frequency at the ends of each cell. Once particle tracking is added to the algorithm, this phase can be made pseudo-continuous by calculating the phase and periodicity at each grid point in the simulation region.

Fig. 3.3 provides an idea of how we set the adjoint source along the axis of the structure and have it radiate outward. The adjoint source is imported into the FDTD solver and the adjoint simulation run. The results are similarly interpolated for size.

### 3.1.4 Construction of $\frac{\Delta \varepsilon_r}{\Delta \vec{\phi}}$

At this point, we have the forward and adjoint fields from Eq. 2.45 , leaving only the  $\frac{\Delta \varepsilon_r}{\Delta \vec{\phi}}$  matrix to be calculated. First, the initial  $\varepsilon_{r_0}$  is retrieved from the index monitors.

To construct the  $\frac{\Delta \varepsilon_r}{\Delta \vec{\phi}}$  matrix, we need to increment each parameter<sup>1</sup>, individually, for the whole structure. To clarify, in our case, our free parameters are the x-radius ( $r_x$ ) and y-radius ( $r_y$ ) of our pillars, and the y-position ( $y$ ) of each pillar within their cell—here we define the  $y$ -position of the pillar by its axial center. So, we ‘step’ in  $r_x$  for the first cell, replace the pillar with the incremented parameter, record its relative permittivity array of ( $\varepsilon_{r_{j=0}}$ ), delete the pillar, reset the parameters, step in  $r_y$  for cell 1, and so on collecting each  $\varepsilon_{r_j}$  for every parameter for every cell in the structure ( $j = 0, 1, 2, \dots, N$ ),  $N$  being the length of  $\vec{\phi}$ . This gives an  $\varepsilon_r$  array with respect to each parameter. We simply subtract each array,  $\varepsilon_{r_j} - \varepsilon_{r_0}$  which is then ‘block-diagonalized’ by separating the arrays based on the periodicity of the structure.

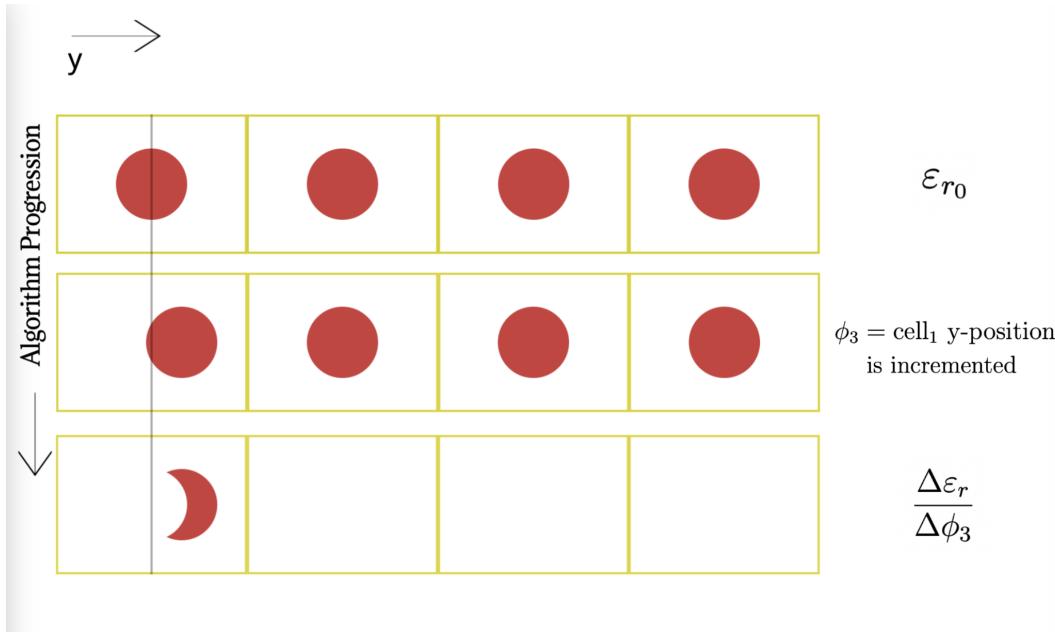


Figure 3.4: Illustration of the construction of the  $\frac{\Delta \varepsilon_r}{\Delta \vec{\phi}}$  matrix.

<sup>1</sup>Parameters must all be incremented in the same direction (+/-) because each  $\varepsilon_r$  matrix is for the whole structure and there are interactions between the cells. Our system and parameter space are also not symmetric, nor centered around 0. For example, an increase in  $y$  of a pillar will cause an increase in the phase; however, a decrease in  $y$ -position will not necessarily cause a negative phase. Hence, the minus signs from  $(G_n - d_n)$  and the decrement of the parameter may improperly cancel out, causing erroneous optimization gradient calculations.

## Bound Constrained Optimization

Although we have everything we need to calculate our gradient in Eq. 2.45, we must now take an aside to address one of the difficult aspects of this kind of optimization. Because of how the gradient is calculated, all constraints, weights, or modifications to the gradient, for optimization purposes, has to be done here, during the construction of  $\frac{\Delta \varepsilon_r}{\Delta \phi}$ .

Another quick back-of-the-envelope calculation tells us that the interaction length between cells is on the order of the wavelength of the laser in all directions.

$$L \approx \pm \lambda_l = \pm \frac{1}{\beta} \lambda_p \quad (3.2)$$

As can be seen from this equation, since  $\beta \leq 1$ , the interaction length will be at least a few cells in both directions. The fields of the structure are coupled and therefore cannot be solved independently—each cell influencing several of its neighbors. **Note:** A more thorough and accurate calculation of the interaction length could be calculated accounting for how far away the EM-waves could ‘catch up’ to the particle and applying the inverse square power reduction etc. However, for the purpose of brevity, it is sufficient to understand that there are strong interactions between the cells. Therefore, our system is not linearly independent, nor truly symmetric, thus the adjoint method is merely an approximation. Typical gradient based optimization applies bounds and constraints after the gradient has been calculated. If the elements of the Jacobian are independent of each other, this is not an issue, but in our case where they are not independent, our step in parameter space will be skewed, thus not moving us in the direction of the minimum because the gradient is calculated for the entire parameter space and cannot be modified after the fact.

This means the only place we can add bounds and constraints without losing information is during the construction of  $\frac{\Delta \varepsilon_r}{\Delta \phi}$ . In practice this was accomplished by

- (1) Setting a default step size  $\xi$

**Note:** Since we are constraining the step, and our minimum feature size is 10 nm, we chose  $\xi = 20$  nm as this default step

- $\xi_i$  then is the step size for each individual parameter for each cell of the structure

- (2) Applying weights with the previous iteration’s  $G_n - d_n$

- The weighting function is a piece-wise sigmoid function that is set to 0 if  $G_n - d_n < \text{tol}$ , where ( $\text{tol}$ ) is some predefined tolerance value, thereby ‘locking in’ the parameters for a given set of cells

- The weights are applied based on the sum of  $G_n - d_n$  for several periods to account for the interactions between cells

$$\mathbf{w}_k \xi_{i:i+3}$$

- (3) Applying constraints based on the current parameters

- These inequality constraints are applied with a scaled sigmoid function of the form  $\mathcal{S}(x) = \frac{1}{1+e^{-bx}}$ , where b is a constant scaling factor to sharply slope the curve around its center
  - For greater than constraints, the argument  $x$  is of the form  $(\frac{r_{x_{\max}}}{r_x} - 1)$
  - For less than constraints, the argument  $x$  is of the form  $(\frac{r_x}{r_{x_{\min}}} - 1)$
- Each parameter,  $r_x$ ,  $r_y$ ,  $y$  has multiple constraints on it
  - e.g.  $r_{x_{\min}}$ ,  $r_{x_{\max}}$ ,  $\Delta r_x$
- Each constraint is calculated, added together, then applied to the step size as

$$\xi_i = \xi_i - \sum_j \mathcal{S}_j(x) \xi$$

where  $j$  denotes the individual constraints on each parameter

- if the sum of all constraints on a parameter is  $> 1$ , the step is set to zero ( $\xi_i = 0$ )
- Here the sigmoid function is used to scale down the step size, e.g. a zero on the sigmoid function corresponds to no constraining of the step, while a 1 conversely sets the step size to 0
- The sigmoid function applies a ‘gradient pressure’ to the step size based on the current parameters, with respect to their constraints
  - See Fig. 2.9 for a conceptual idea of how the constraints are used to bound the system
  - The sigmoid is smoothly differentiable, which is useful for maintaining a smooth gradient pressure
    - \* This was inspired by the method of Lagrange Multipliers for constrained optimization (see Ref. [38] for a more detailed explanation on that method)
- In this thesis, our constraints were on the minimum and maximum radius of the pillars ( $r_{x_{\min}}, r_{x_{\max}}, r_{y_{\min}}, r_{y_{\max}}$ ), the maximum difference between radii of neighboring cells—a fabrication constraint—( $\Delta r_x, \Delta r_y$ ), and the y-position inside of the cell—so the pillar does not cross into a neighboring cell—( $y_{\min}, y_{\max}$ )
  - The design constraints are influenced based on fabrication, computational/resolution, and material (damage threshold) constraints

#### (4) Applying adam optimization algorithm

- The adam algorithm applies ‘momentum’ to the gradient based on previous gradients
  - See Ref. [31] for a more detailed explanation on the math of the adam algorithm

- Since the constraints on the steps may not correlate with the magnitude of the gradient—which is, instead, determined by the parameter space—we cannot use the previous constrained step to determine the momentum
- We also cannot apply adam after the gradient has been calculated, as discussed previously
- The solution in this thesis was to create a new variable called *gradam*, which is the final gradient  $\frac{\partial \mathbf{G}}{\partial \vec{\phi}}$  (Eq. 2.33) that is fed into the adam algorithm to have momentum added to it
- This *gradam* value is then used as a weight applied to the constrained step size (in our *stepping function*) of the next iteration
  - Since our final gradient is calculated after we have  $\frac{\Delta \varepsilon_r}{\vec{\phi}}$ , we must apply the *gradam* value to the subsequent iteration of the optimizer

$$\xi_i^{\text{iter}=l} = \text{gradam}_i^{\text{iter}=l-1} \cdot \xi_i^{\text{iter}=l}$$

(5) Incrementing parameters

- The now weighted, constrained step-size array with momentum, is added to the current parameter array

$$\vec{\phi} + \vec{\xi}_{\text{constrained}}$$

(6) Bounding parameters

- Bounds are applied as a rigid boundary for each parameter by using

$$\phi_i \gtrless \begin{matrix} \text{high} \\ \text{low} \end{matrix} \text{ bound} \rightarrow \phi_i = \begin{matrix} \text{high} \\ \text{low} \end{matrix} \text{ bound}$$

- The bounds on the  $y$ -position have to be dynamic to adapt to the changing  $r_y$  value
  - They are defined as the cell's beginning plus the  $y$ -radius and the cell's end minus the  $y$ -radius ( $\lambda_{p_{n-1}} + r_y$  and  $\lambda_{p_n} - r_y$ )

(7) Constructing  $\frac{\Delta \varepsilon_r}{\Delta \vec{\phi}}$

- The bound, constrained, weighted, momentum parameters are returned from the stepping function ( $\vec{\phi}$ )
- The pillars are replaced for each of these parameters and the  $\varepsilon_{r_{\phi_i}}$  matrix is measured
- $\varepsilon_{r_0}$  is subtracted from each  $\varepsilon_{r_{\phi_i}}$ , yielding  $\frac{\Delta \varepsilon_r}{\Delta \vec{\phi}}$ 
  - If we rearrange  $\frac{\Delta \varepsilon_r}{\Delta \vec{\phi}}$  from a vector  $i$  components long (each element being a  $\Delta \varepsilon_r$  matrix) to an  $n \times j$  matrix (periods  $\times$  parameters, still with each element as a  $\Delta \varepsilon_r$  matrix), we get certain level of ‘block-diagonality’ as discussed before

### 3.1.5 Custom Optimization

We now return to our optimization algorithm. We have the forward and adjoint fields and  $\frac{\Delta \varepsilon_r}{\Delta \phi}$ . Using Eq. 2.45, we can calculate the gradient which will now be within our bounds and constraints.

Once we have the gradient  $\frac{\partial G_n}{\partial \phi}$ , all we have to do is use Eq. 2.46 to complete our gradient calculation, while Eq. 2.33 provides us with our final cost value.

Due to the interference effects of the transient sources, discussed in section 2.4.7, we use a simple rolling mean of 3 elements to smooth the  $G_n$  curve, as well as temporal apodization, which essentially allows the simulation to run for its usual time, but only records the field results in a window in order to also reduce this ringing around the ends of the structure [29].

The main function returns the cost value  $\mathbf{G}$ ,  $G_n$  values, and the gradient to the optimizer, which simply compares if  $\mathbf{G} < \mathbf{G}_{\min}$ . If yes, then  $\mathbf{G}$ ,  $G_n$ , and the parameters are saved. If no, then  $G_{\min}$  remains unchanged. Since our optimizer is not necessarily going to converge—especially during the coarse, basin-hopping optimizations—this ensures we always save the best results from each ‘optimization.’

The basinhopping algorithm saves all of the outputs from the optimizer, the best of which is then used in the fine optimization. The fine optimization is done with a smaller optimizer step size  $\kappa$  and more steps. This finally yields our suspected, global minimum, within our defined bounds and constraints. Hopefully it is now clear why a custom optimization algorithm is necessary for the constrained optimization of adjoint inverse design of dual pillar DLA structures.

# Results

The primary goals of this thesis were to add the  $y$ -position of the pillars, within their cells, as an optimizable parameter, and to show that non-periodic structures could be optimized via inverse design. Non-periodic structures being of particular interest due to the dephasing of the particle as it accelerates (discussed in detail in section 2.4.6). In the following section, we will demonstrate the results of the solutions to these problems. As detailed in the previous section, a large focus of this thesis became the development of a custom, bound, constrained, non-convex optimization algorithm. We will take a look at the performance of the optimizer on various systems with explanations.

## 4.0.1 Simulation Parameters

All simulations were done with an incident laser with  $\lambda_l = 1.932\mu\text{m}$ . For our calculation of  $\Delta W$ , we used a field intensity of  $\text{Re}\{e_1\} \cdot 100 \text{ MV/m}$ . Since the function used for the constraints allows for a small step outside of them, they should be set just inside of the hard bounds. Our constraints on the system were the minimum radius size ( $r_{x_{\min}}$  and  $r_{y_{\min}}$ ), maximum radius size ( $r_{x_{\max}}$  and  $r_{y_{\max}}$ ), change in radii between neighboring cells ( $\Delta r_x$  and  $\Delta r_y$ ), minimum  $y$ -position within its cell ( $y_{\min}$ ), and maximum  $y$ -position ( $y_{\max}$ ). These constraints primarily come from the damage threshold and fabrication limitations. Here we define a ‘normalized’ (0-1)  $e_1$  value as  $\text{Re}\left\{\frac{e_1}{2 \cdot 100 \text{ MV/m}}\right\}$ , where we account for the dual drive laser power to regain our normalization of the Fourier coefficient.

Table 4.1: Global Bounds and Constraints

	Constraints	Bounds
$r_{x_{\min}}$	$\frac{\lambda_{p_n}}{2} \cdot 0.25$	100 nm
$r_{y_{\min}}$	$\frac{\lambda_{p_n}}{2} \cdot 0.25$	100 nm
$y_{\min}$	$\lambda_{p_n} \cdot 0.25$	$\lambda_{p_{n-1}} + r_y$
$r_{x_{\max}}$	$\frac{\lambda_{p_n}}{2} \cdot 0.75$	$\frac{\lambda_{p_n}}{2} \cdot 0.9$
$r_{y_{\max}}$	$\frac{\lambda_{p_n}}{2} \cdot 0.75$	$\frac{\lambda_{p_n}}{2} \cdot 0.9$
$y_{\max}$	$\lambda_{p_n} \cdot 0.75$	$\lambda_{p_n} - r_y$
$\Delta r_x$	20 nm	N/A
$\Delta r_y$	20 nm	N/A

## 4.1 Custom Optimizer Validation

### Parameter Space Map

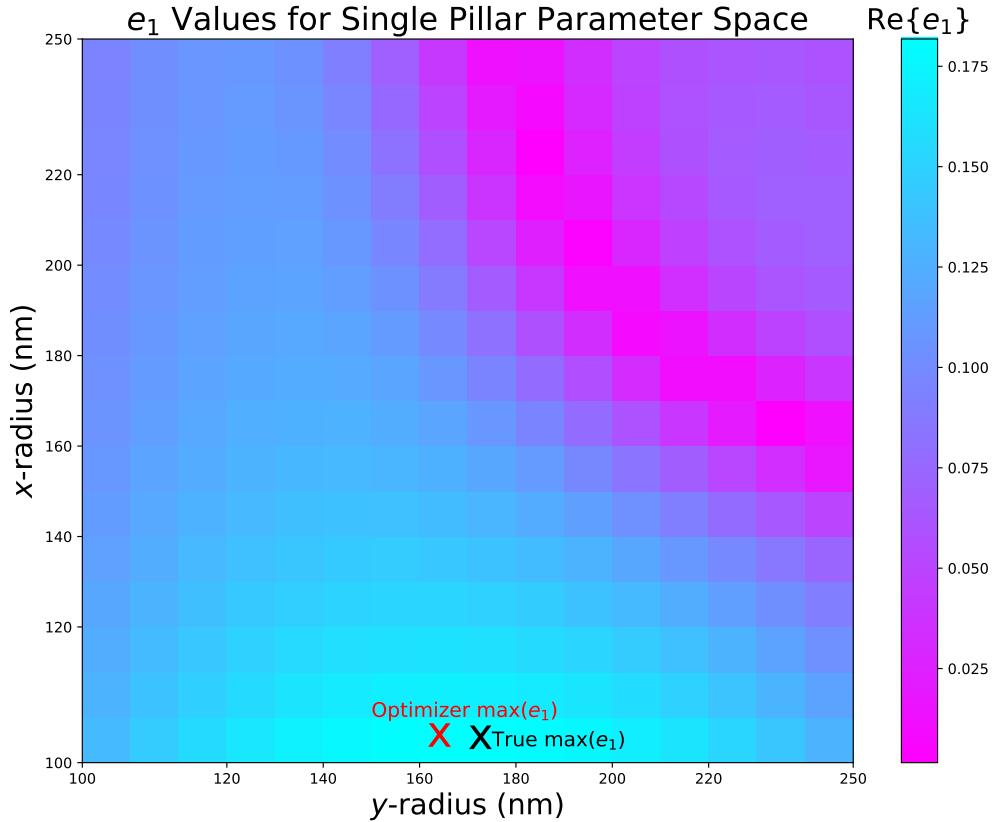


Figure 4.1: Brute force map of the parameter space for a single cell in an infinitely periodic structure with an injection energy of 30.30 keV.

The first step was to test the optimizer. To do this, we did a brute force mapping of the parameter space and then ran the optimizer to see if it converged to the same minimum. For the brute force mapping, we set the adjoint phase to  $-\phi_{e_1}$  in order to maximize  $\text{Re}\{e_1 \cdot e^{i\varphi_{\text{adj}}}\}$ . Fig. 4.1 shows the parameter space for a single cell of an infinitely periodic structure. This mapping was done with the radius of the major and minor axis of an ellipsoid pillar as the parameters ( $r_x$  and  $r_y$ ). The steps in the mapping were set at our fabrication resolution of 10 nm. The maximum possible  $\text{Re}\{e_1\}$  value was found at  $r_x = 100$  nm and  $r_y = 170$  nm, while the optimizer converged at  $r_x = 100$  nm and  $r_y = 160$  nm. Although the optimizer did not find the true global maximum, it can be seen from the figure that the difference between the points is nearly indistinguishable, hence, the optimizer did successfully find the true maximum within its convergence tolerance.

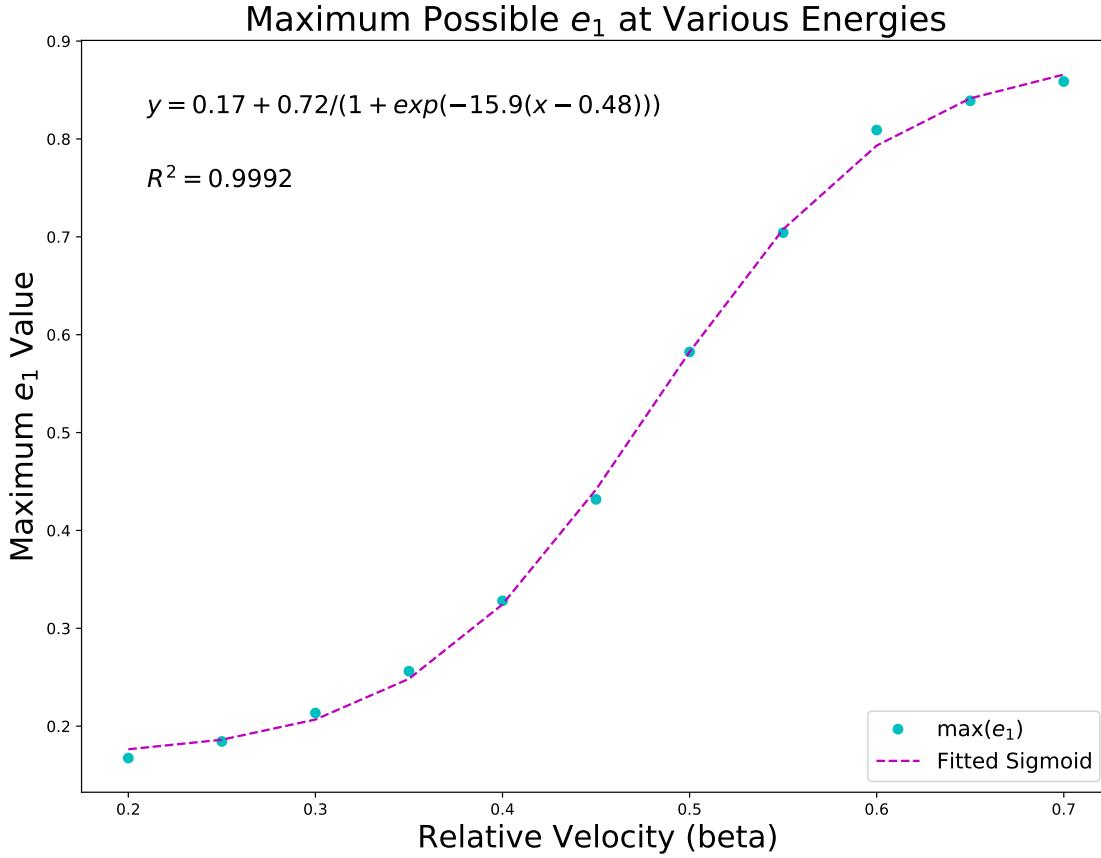


Figure 4.2: Maximum possible  $\text{Re}\{e_1\}$  value at various electron energies.

### $e_1$ vs $\beta$

There is a complicated relationship between the maximum value of the first spatial harmonic and periodicity ( $e_{1,\max} \propto \lambda_p$ ). The exact dynamic of the relationship is beyond the scope of this thesis, however, it can be numerically solved for. This is the plot seen in Fig. 4.2, where we did a brute force search of the 1 pillar maximum for  $\text{Re}\{e_1\}$  at various energies. The maximum  $e_1$  value was plotted against the electron energy and a sigmoid fitted to it. This fitted sigmoid then serves as a sort of scaling function for the design curve  $e_1$  values, to match their true maximum value (if the user so chooses). This allows us to account for the fact that as the particle accelerates down the structure, you can achieve a greater amount of field energy in the first spatial harmonic. Providing us with greater acceleration gradients and ultimately a higher final particle energy.

## Dynamic $y$ -position Parameter

Next, we wanted to test the effect of the  $y$ -position on  $e_1$ . In Fig. 4.3, the magnitude of  $e_1$  never changes by more than 1%. We do see, however, that the phase of  $e_1$  changes as the pillar moves within its cell, as expected. Both Fig. 4.3 and 4.4 show that we are able to get a nearly 180° phase shift from moving the pillar from the bottom of its cell to the top (*bottom* being closer to  $y = 0$ , *top* being further from  $y = 0$  in our coordinate system).

This is a vital parameter to be able to control. Being able to manipulate the phase of  $e_1$  provides greater control and flexibility over the designs of our structures and the complexity of our design curves. This is very useful for non-periodic structures, as the purpose of having the increasing cell size is to maintain the field phase with that of the particle's trajectory. Thus, having control over the phase of  $e_1$  lets us maintain a greater level of synchronicity and avoid unwanted dephasing. It should be noted, though, that the level of control and the range of phase angle available is not necessarily as good as is shown in the aforementioned plot. This is due to the interaction between cells and their effects on each other, since, in a real structure, it is not an infinitely periodic object.

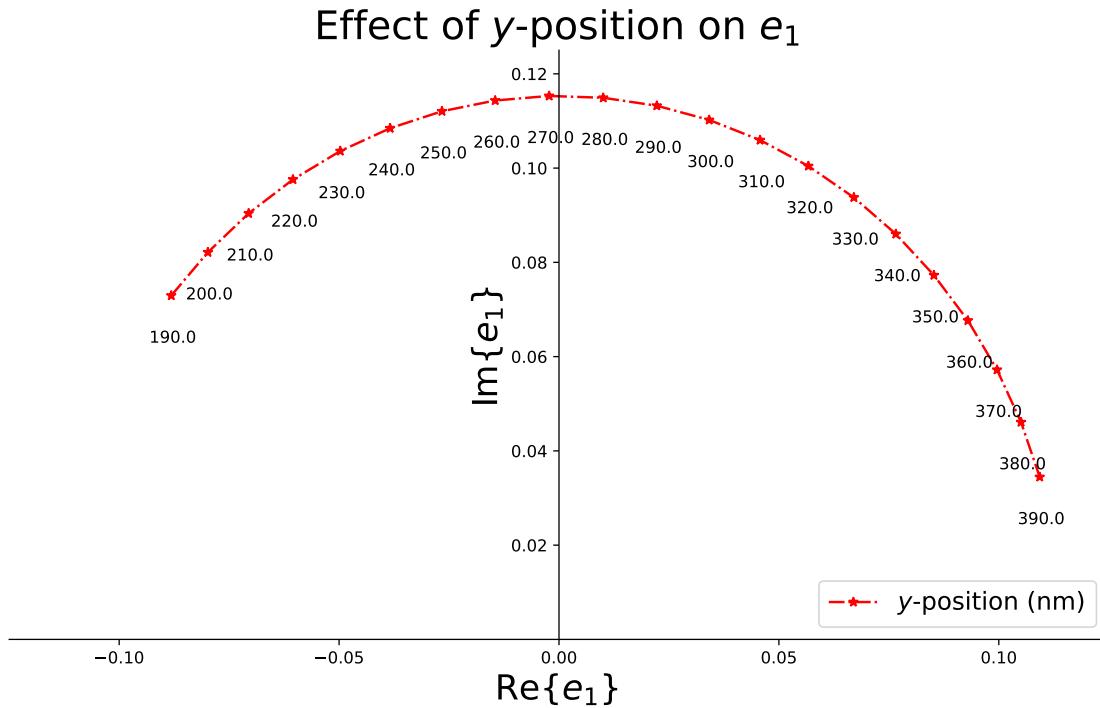


Figure 4.3: Effect of the  $y$ -position on the complex  $e_1$  harmonic of the field.

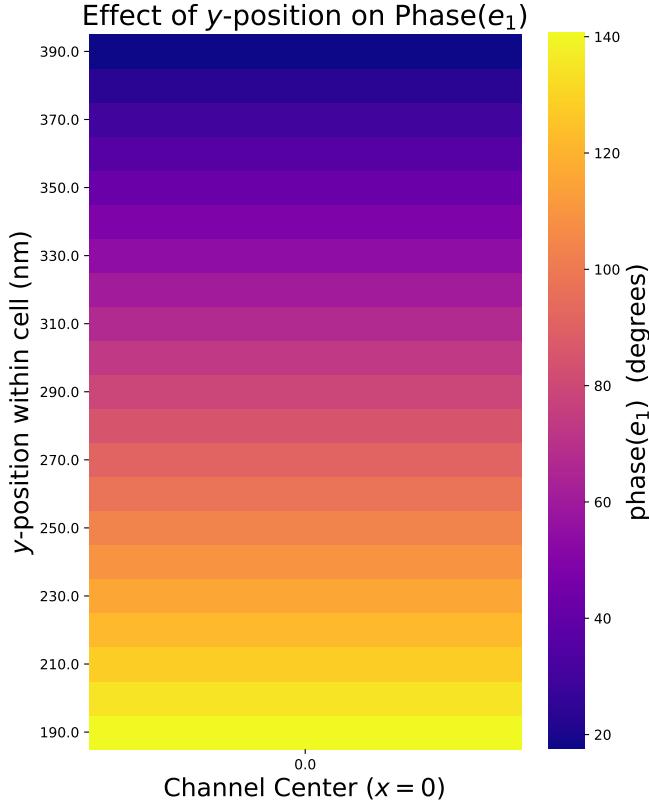


Figure 4.4: Plot of the phase of  $e_1$  as the pillar is moved from the bottom to the top of the cell.

## Multi-Cell Optimization

The optimizer was then tested on a multi-period, periodic structure. From Ref. [27], we know that this adjoint inverse design optimization could already successfully find the extrema with  $r_x$  and  $r_y$  as free parameters. So, using a similar simulation setup was a good test to verify the custom optimization algorithm. Fig. 4.5 shows that the optimizer was still able to successfully converge on a multi-period structure with the newly added  $y$ -position parameter. The shaded region is calculated from the maximum possible  $\text{Re}\{e_1\}$  value at the given *injection* energy. Typically, we might use a normal percentage as our metric by which to judge the optimization. However, since only a portion of the field energy is available for us to utilize, it makes more sense to use the maximum possible  $\text{Re}\{e_1\}$  to set our scale. Here we use an interval, extended 5% beyond the design curve, as our tolerance. This visually shows us where we are within 5% of the design curve, hence, within this green band values match the design curve by 95%+.

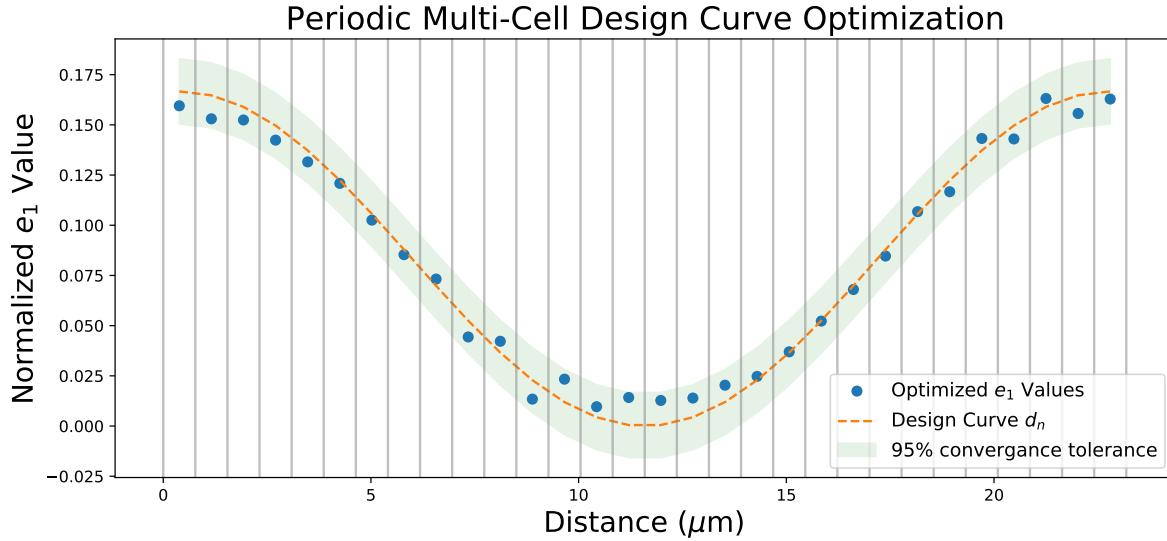


Figure 4.5: Optimization of a periodic, multi-cell DLA structure.

## 4.2 Non-Periodic Structure Optimization

### Small Scale Structure

#### Simulation Parameters

$\beta$	0.4
Injection energy $W$	46.55 keV
Number of periods	50
First and last periods ignored	10
Length of structure	38.73 $\mu\text{m}$

At this point, we were finally ready to use the algorithm on a non-periodic structure. We started with a small scale structure to validate its results. We see in Fig. 4.6, that the optimization was able to very accurately converge to the design curve. Due to computational limitations, the simulations were not run with very many cells. This leads to a relatively small net acceleration and hence small change in periodicity  $\lambda_p$ . Here, we only see a 3 nm increase in periodicity, but this is a good proof-of-concept for the optimization algorithm, which can be easily expanded on a more powerful computer. In a mere 50 periods, we still achieved an energy gain of  $\approx 500\text{eV}$ ,  $W = 46.55 \text{ keV} \rightarrow 1.01 \cdot W = 47.02 \text{ keV}$ . Thus, we could expect to be able to double an electron's energy in only 5000 cells.

The periodicity was attempted to be made into a dynamic parameter but this proved to be beyond the scope of this thesis. The primary difficulty lies in the link between  $e_1$ ,  $\beta$ , and  $\lambda_p$ . To expound,  $\beta_{n+1}$  is determined by the product  $e_{1n} \cdot \lambda_{pn}$ , while  $\lambda_{pn+1}$  is then determined by  $\beta_{n+1}$ . We see that for every change in  $e_{1n}$ , there will necessarily need to be a change in

$\lambda_{p_{n+1}}$ . Similarly, every change in  $\lambda_{p_n}$  will cause a change in  $\lambda_{p_{n+1}}$ . However, if there is no change in  $\lambda_{p_{n+1}}$ , this will necessarily change  $e_{1_{n+1}}$  from the previous iteration. Conceptually this leads to the optimizer being a ‘step behind’ itself—changing  $\lambda_p$  based on the previous iteration which is not accurate to the current structure’s parameters.

Another possible implementation would be to update  $\lambda_{p_{n+1}}$  progressively. Since  $\beta_{n+1}$  depends on  $\lambda_{p_n}$  though, and the adjoint source is supposed to mimic the particle phase velocity, we would have to run a forward and adjoint simulation for the incrementing of each cell. Conceptually, if we run a forward and adjoint simulation, then increment the periodicity of period 1, now  $\beta_{n+1}$  is different and this is propagated down the structure due to the recursive-like relationship between  $\beta$ ,  $\lambda_p$ , and  $e_1$ . Hence, we would need to run forward and adjoint simulations for the incrementing of each successive period. For large structures, it is easy to see that this begins to defeat the purpose of the adjoint method because you are requiring many time-consuming and computationally expensive simulations. For these reasons it was chosen, in this thesis, to define the periodicity based on the ideal case, here the input design curve,  $d_n$ . This method allows the adjoint source to be set as a facsimile of the desired particle trajectory. This resolves the issue mentioned above about having the optimizer being a step behind the adjoint source, since it stays the same. This also maintains the 2-simulation/iteration requirement from the adjoint method.

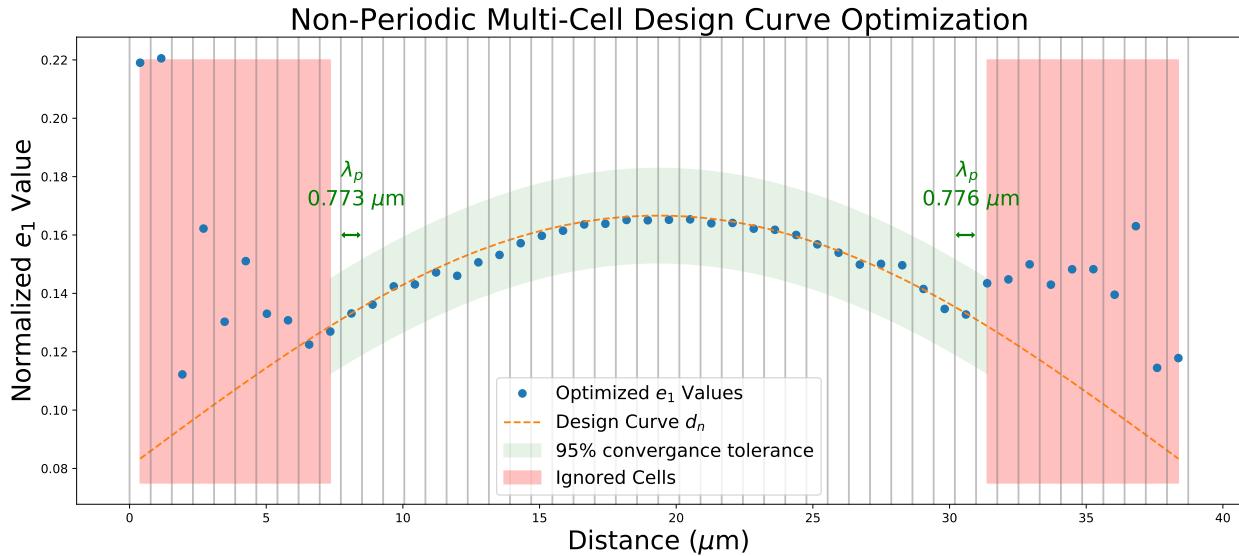


Figure 4.6: Optimization of a non-periodic, multi-cell DLA structure.

### Scaled-Ramped Accelerator

Now that we have seen that we can successfully optimize (in-phase) non-periodic structures, we must consider large structures with very high acceleration. For this we choose a ‘ramped’ design curve, here a half period of a sin function, to smoothly accelerate the particle. Given that we aim to greatly increase the particle’s velocity, we should consider the different maximum  $\text{Re}\{e_1\}$  values possible at the different energies. As discussed above, we can put our

### Simulation Parameters

$\beta$	0.35
Injection energy W	34.50 keV
Number of periods	300
First and last periods ignored	15
Structure length	206.92 $\mu\text{m}$

design curve into the fitted function in Fig. 4.2 for the different maximum  $\text{Re}\{e_1\}$  values. This will scale our design curve accordingly and allow us to optimize for the true maximum acceleration available to us.

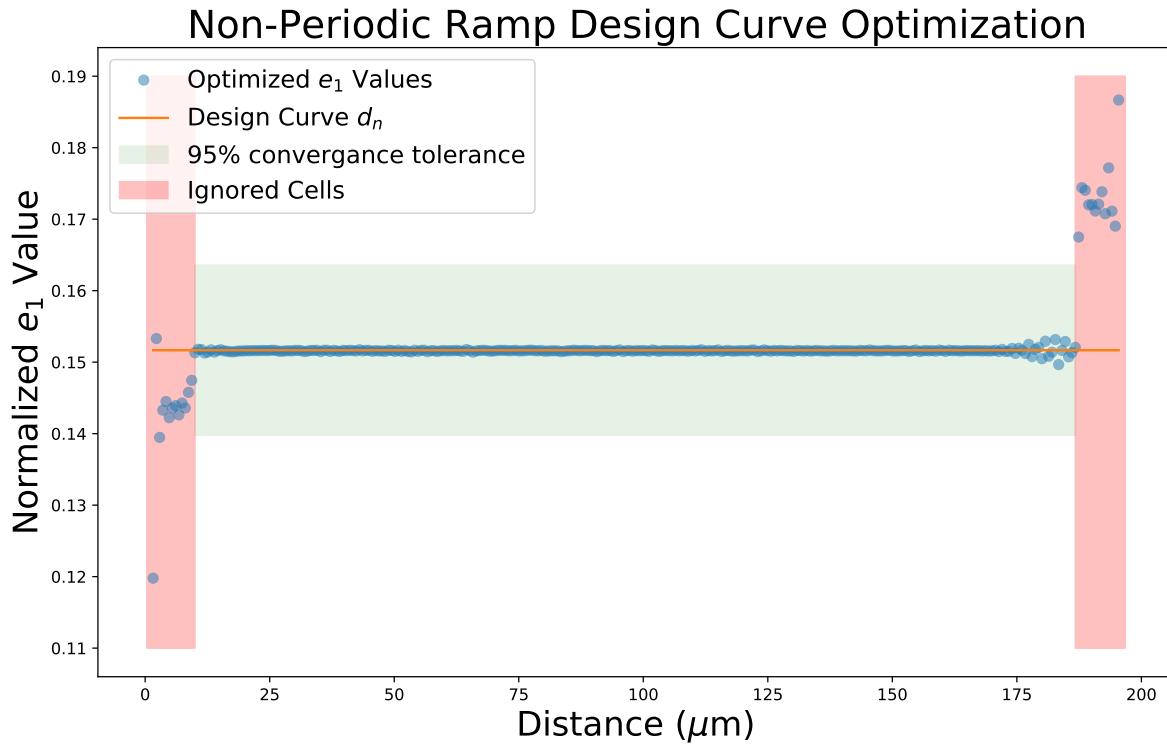


Figure 4.7: Optimization of a non-periodic structure with a scaled, smoothly ramped sine acceleration design curve.

We desired to double the energy of the electron but this would require some 3000 cells (physically quite feasible), but due to computation limitations, we settled on increasing the energy from 34.50 keV to 37.97 keV. Our sin ramp from 34.50 keV to 37.97 keV was then scaled according to the sigmoid function for the energy, discussed above. This then gave us a ‘ramped’ design curve for  $e_1$  values necessary to accelerate an electron from  $0.350c$  to  $0.365c$ . This is very promising, since, with these result, the expansion to larger structures should allow us to very quickly accelerate particles to very high energies.

We see in Fig. 4.7 that the curve looks like a line. Since we were only able to optimize for a relatively small structure, the net energy gain (and thus, increase in cell size  $\lambda_p$ ) is

very small. However, as before, this is still a good proof-of-concept for the optimization algorithm. Although difficult to see, it is a positively sloped *line*. The figure shows that the optimizer was extremely accurate on the interior of the structure. What we see on the ends are cells that were not optimized for, due to the aforementioned fringing and edge effects at the ends of the structure.

### Non-Periodic Sinusoid Optimization

#### Simulation Parameters

$\beta$	0.35
Injection energy W	34.50 keV
Number of periods	300
First and last periods ignored	15
Structure length	204.84 $\mu\text{m}$

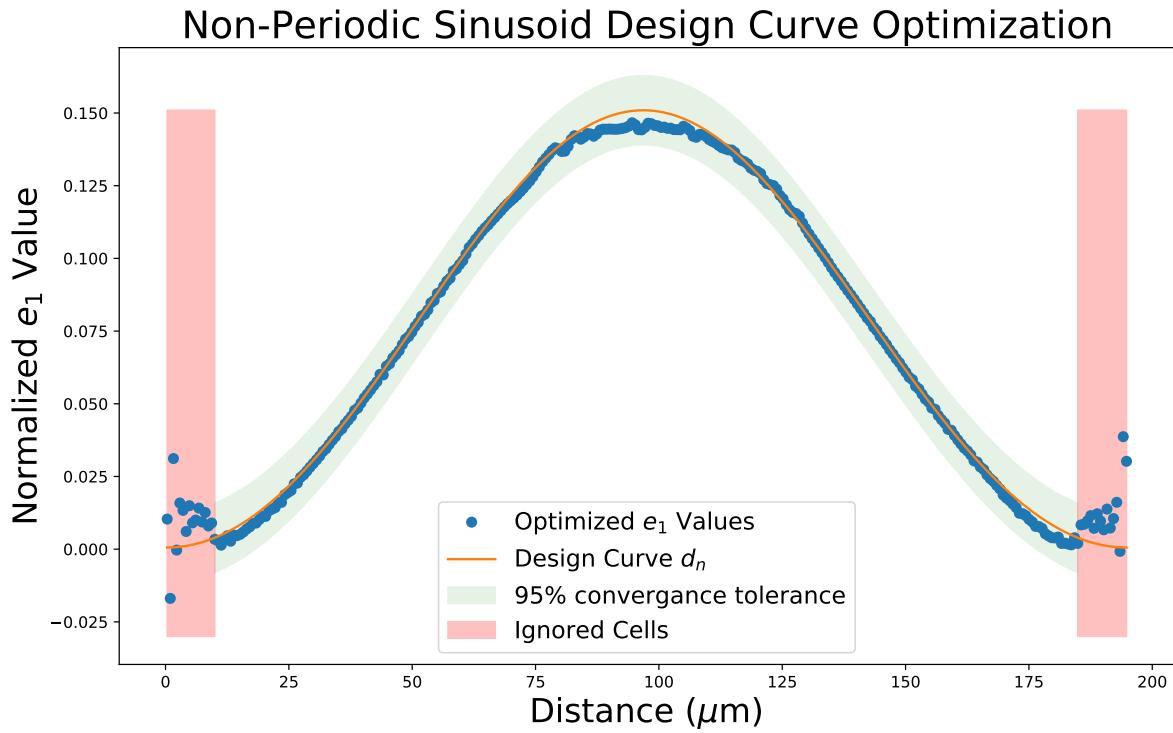


Figure 4.8: Optimization of a non-periodic structure with a sinusoid design curve.

Next we wanted to test the flexibility of the optimizer, so 3 progressively more complicated design curves were created and set for non-periodic structures. Fig. 4.8 shows the optimization result of a non-periodic structure with a sinusoidal design curve, similar to that shown in Fig. 4.5 and 4.6. This was done as a check of the validity of the optimization. Since we knew it could successfully converge on sinusoidal design curves for smalls scale structures (both periodic and non-periodic), we needed to verify that the algorithm could be easily

scaled to larger structures. As we can see in the figure, the algorithm was very successful at matching the design curve within the tolerance (and the constraints). This structure is 300 periods ( $204.84\mu\text{m}$ ) and will accelerate our particle from 34.50 keV to 36.01 keV.

### Superposition and Sawtooth Curves

#### Superposition Parameters

$\beta$	0.4	$\beta$	0.35
Injection energy W	46.55 keV	Injection energy W	34.50 keV
Number of periods	150	Number of periods	150
First and last periods ignored	10	First and last periods ignored	15
Structure length	$116.23\mu\text{m}$	Structure length	$101.84\mu\text{m}$

#### Sawtooth Parameters

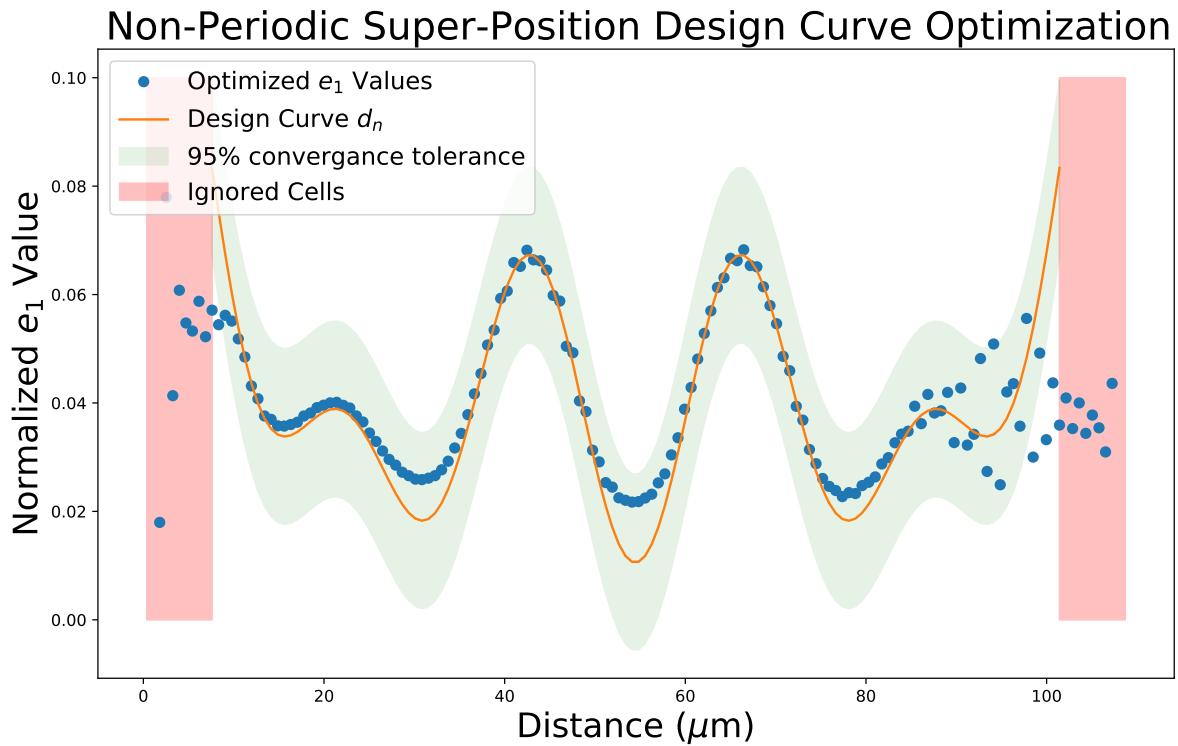


Figure 4.9: Optimization of a non-periodic structure with multiple sinusoids superimposed to generate the design curve.

If the algorithm could find structures that could match various sinusoidal design curves, it begged the question, could it match a combination of sinusoids? That is, could we use a super position of sine waves to make arbitrarily complicated design curves? Yes, and No—arbitrary up to a point. Fig. 4.9 shows that the optimizer was quite good at matching a superposition design curve (minus some outliers further towards the outside of the structure). It should be noted, that this curve was chosen with *a priori* knowledge. From experience

with the simulations, there was a level of intuition about the limit to the complexity that the optimizer could successfully handle. This was tested more thoroughly in Fig. 4.10. From the tests with sawtooth wave design curves, it was found that the upper limit was a change of  $\text{Re}\{e_1\} \approx 0.1/10$  periods. Fig. 4.10 shows that the optimizer is still very accurate up to this limit, however, much beyond this rate of change in  $\text{Re}\{e_1\}$  and the optimizer will tend to ‘run away.’ This is true for any design curve that is not physically possible, therefore more development is needed to keep the optimizer as close to the best, physically possible, solution, without running away, chasing a possibly unobtainable design curve.

One major limiting factor is our fabrication constraint of a change in neighboring pillars parameters  $\Delta_{r_x}$  and  $\Delta_{r_y}$  not being able to exceed 20 nm. With this system, though, we can, in theory, optimize for arbitrarily complicated design curves up to this limit. This is an exciting result as it shows that we can have great control over the particle as it traverses the structure, providing huge flexibility in the functionality of the accelerator. With different constraints on the system, perhaps with advancements in fabrication, there could be even more control and flexibility.

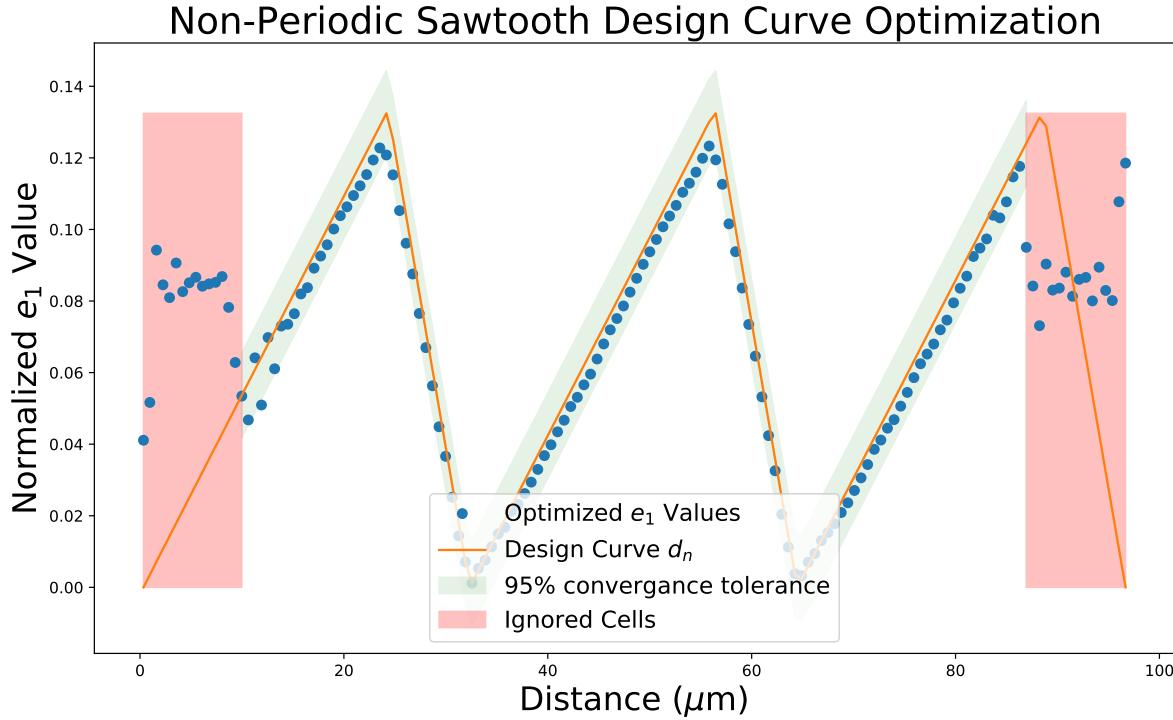


Figure 4.10: Optimization of a non-periodic structure with a sawtooth design curve, demonstrating the maximum rate of change of  $\text{Re}\{e_1\}$  between cells.

### e<sub>1</sub> Phase Optimization

As discussed in section 4.2, we can get a near 180° phase change in  $e_1$ —relative to the particle trajectory. With this and the arbitrarily complicated design curves optimized for in the previous section, we wondered if it were possible to fully change the phase of the field

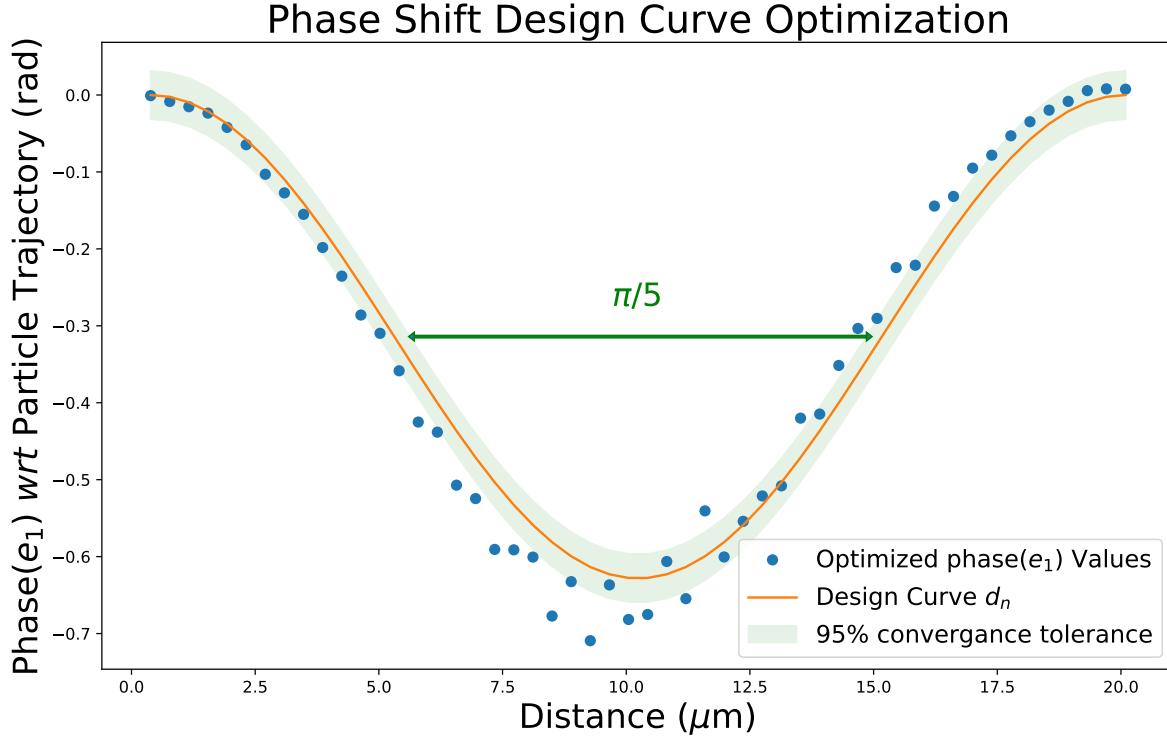


Figure 4.11: Phase optimization of a multi-cell structure. Since we set the adjoint phase to be  $-\phi_{e_1}$ , the relative phase of  $e_1$  will necessarily be zero. Each successive period, then, the phase is changed from  $e_1$ .

with the particle trajectory, from  $-\pi$  to  $\pi$ . As we can see in Fig. 4.11, the total phase shift was only  $\pi/5$ . This seems to be counter-intuitive. How can we achieve an  $e_1$  phase change of nearly  $\pi$  for a single pillar but only  $1/5$  of that for a full structure? Approximately 30 pillars was the minimum to achieve a similar  $\pi/5$  shift (however less successfully converged than the 50 periods used in Fig. 4.11). This did not scale with the number of periods, as we thought it would. Even for larger structures, of hundred of periods, it still could only manage a roughly  $\pi/5$  phase shift. It is unclear as to why this is, although it could be due to the constraints placed on the system, as well as the interaction between cells.

As mentioned above, we do not actually have access to the full  $180^\circ$  phase shift per cell, as we did in section 4.2, in a non-periodic structure. The interactions between cells limits the actual amount by which you can change the phase within our constrained system. Fig. 4.12 provides an intuitive, likely explanation of why we only see a limited phase shift.

In an infinitely periodic structure, when we shift a pillar, they all shift, thus the phase of the entire structure is changed by  $\pi$ . However, in a finite structure (periodic or non-periodic), where we can manipulate cells individually, the phase primarily changes in its cell ( $n$ ), but also affects several of its neighbors in both directions. Conceptually, this phase accumulation has to go somewhere, hence when we use a sinusoid design curve, the phase delay is being compressed in from both sides, leading to the likely explanation for the  $\pi/5$  limit in cyclical phase change in a finite structure.

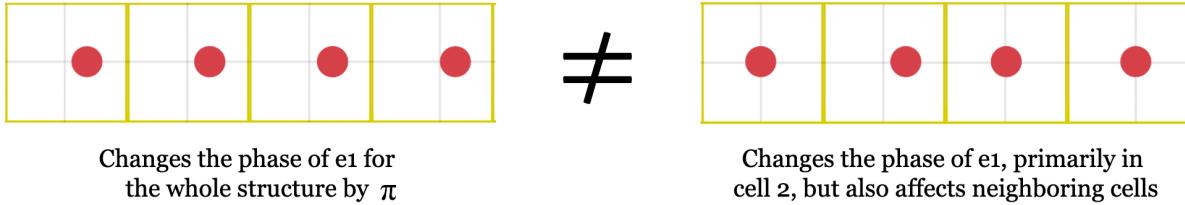
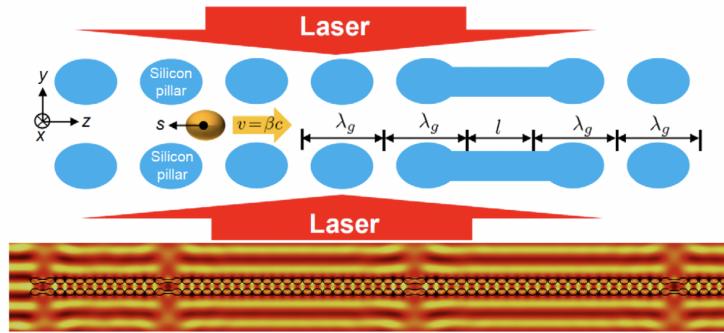
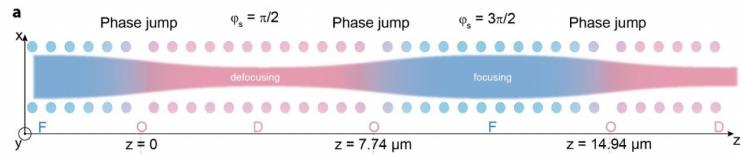


Figure 4.12: Conceptual difference between the results in Fig. 4.4 and 4.11. On the left, an infinitely periodic structure, where, when we shift a pillar, they all shift; hence, changing the phase of the entire structure by  $\pi$ . On the right, a finite structure, where we are able to shift a single pillar, which primarily changes the phase of its cell ( $n = 2$ ), but also affects several of its neighbors in both directions. These interactions between cells, and the conceptual/mathematical differences between infinitely periodic and finite structures, are likely explanations for the limited phase control in finite, non-periodic structures.

Ref. [23] provides a scheme by which extended gaps in the dual pillar structures can be used to create a *phase jump*. This happens because the gaps are non-accelerating and if they are set based on our synchronicity condition (Eq. 2.5), then, by time the particle traverses the gap, it will now be  $\pi$  radians out of phase with the incident field. This method is referred to as alternating phase focusing (APF) since it is primarily used to focus the electron beam. They show that they can achieve a  $\pi$  phase jump with each one of these extended APF structures. One possibility is that the APF scheme and the dynamic  $y$ -position from this thesis could be combined to provide control over a full  $2\pi$  range. This would allow for a lot of flexibility in design curves, optimization for the full phase of  $e_1$ , and for greater control over the phase of the fields relative to the particle trajectory. This will be useful once particle tracking is added and simulations are extended to 3D, since it will allow for greater control over the dephasing of the particle as it traverses very long structures.



(a) An illustration of the APF scheme. Reprinted from [23].



(b) The APF scheme allows for a  $\pi$  phase jump. Reprinted from [39].

Figure 4.13: Figures demonstrating alternating phase focusing (APF).

# Conclusion

In this thesis we aimed to use the adjoint method to inverse design a dual pillar DLA structure. More specifically, the primary goals were to add the  $y$ -position of the pillar, within the cell, as a dynamic, optimizable parameter, and to optimize non-periodic structures to account for the dephasing as the particle accelerates. In trying to add these new features and parameters to the inverse design optimization, it was found that in order to correctly apply constraints to the system, a custom optimization algorithm was necessary to work with the adjoint method. This is due to the interactions between cells and thus asymmetry of our  $A$  operator from Eq. 2.16.

## Adjoint Inverse Design for DLA

Inverse design is a process by which (in our case) photonic devices can be designed using the desired output of the structure as a curve to *fit*, via numerical optimization. Typical numerical optimization methods are some form of gradient descent. These involve taking a stochastic step, calculating the two point gradient, then moving in the direction of that gradient (with user defined step size). This approach, however, becomes computationally expensive for large parameter spaces because it is not a particularly efficient way of navigating high-dimensional, non-convex parameter spaces.

The adjoint method provides a way of calculating the gradient of a linear system with only one operator that depends on the parameters. This allows us to quickly and effectively calculate the gradient of the system with only a forward and adjoint source needed to simulate.

It turns out, that our system cannot truly be defined by the adjoint equation (Eq. 2.16). The adjoint method requires the system to be described by a parameter matrix,  $A$ , which is symmetric. This symmetry gives  $A = A^T$  and let's us define our adjoint system as in Eq. 2.20. Linearly independent systems would be diagonal (meaning no interactions between parameters) and thus symmetric. However, as we can see in the parameter space map (Fig. 4.1), the space is not symmetric for even a single cell. Thus,  $\frac{\Delta\epsilon_r}{\Delta\phi}$  (and by extension  $A$ ) is not symmetric. There is some level of symmetry, though, in the figure, so the adjoint method is

a reasonable approximation. An example of a single cell Jacobian construction:

$$(\partial_{r_{x_1}}, \partial_{r_{y_1}}) \begin{pmatrix} \varepsilon_r & 0.6\varepsilon_r \\ 0.3\varepsilon_r & \varepsilon_r \end{pmatrix} = \begin{pmatrix} \partial_{r_{x_1}}\varepsilon_r + 0.3\partial_{r_{y_1}}\varepsilon_r \\ 0.6\partial_{r_{x_1}}\varepsilon_r + \partial_{r_{y_1}}\varepsilon_r \end{pmatrix} \quad (5.1)$$

where each  $\varepsilon_r$  is the epsilon matrix of the entire structure. Here the 0.3 and 0.6 are added (somewhat proactively), since in our construction of  $\frac{\Delta\varepsilon_r}{\Delta\phi}$ , if we step the same amount,  $\Delta\varepsilon_r$  will be the same for all the parameters; their difference coming from the field at their location. To clarify, the elements of  $\frac{\Delta\varepsilon_r}{\Delta\phi}$  will be arrays of the epsilon matrix for the whole structure, each modified by  $\phi_i$ . Then, each would be multiplied by the fields as in Eq. 2.32, meaning their relative contribution to the off-diagonal terms would be dependent on the field strength at their position. In a multi-cell structure, be it periodic or non-periodic, the interactions die off further from the given cell. This means that for larger structures, the Jacobian might look something like:

$$(\partial_{r_{x_1}}, \partial_{r_{y_1}}, \partial_{r_{x_2}}, \partial_{r_{y_2}}, \partial_{r_{x_3}}, \partial_{r_{y_3}}) \begin{pmatrix} \begin{pmatrix} \varepsilon_r & 0.6\varepsilon_r \\ 0.3\varepsilon_r & \varepsilon_r \end{pmatrix} & 0.2\varepsilon_r & 0.05\varepsilon_r & 0.01\varepsilon_r & 0 \\ 0.2\varepsilon_r & \begin{pmatrix} \varepsilon_r & \cdot \\ \cdot & \varepsilon_r \end{pmatrix} & 0.1\varepsilon_r & 0.01\varepsilon_r & 0 \\ 0.5\varepsilon_r & 0.3\varepsilon_r & \cdot & \cdot & \cdot \\ 0.2\varepsilon_r & 0.2\varepsilon_r & \cdot & \cdot & \cdot \\ 0.1\varepsilon_r & 0.01\varepsilon_r & \cdot & \cdot & \begin{pmatrix} \varepsilon_r & \cdot \\ \cdot & \cdot \end{pmatrix} \\ 0 & 0 & \cdot & \cdot & \end{pmatrix} \quad (5.2)$$

meaning we retain some level of symmetry in the epsilon matrix, in a block diagonal form. Hence, the adjoint method is still a good approximation of the system, especially when looking at the interior of the structure where there will be greater symmetry. For a non-periodic structure, especially those with high accelerations, we must look into these neighboring interactions because they will scale with the acceleration. When we are ready to optimize the ends of the structures, we will also have to look at the neighboring interactions because they'll be stronger and less symmetric. As we can see this approximation is getting worse.

So, for 1 cell, a periodic structure, or a non-periodic structure, technically, the adjoint method is an approximation and not a true description of the system. However, calculating  $\frac{\Delta\varepsilon_r}{\Delta\phi}$  for the entire structure, for every parameter (and thus cell) and using this in the field integral in Eq. 2.32; then, Eqs. 5.1 and 5.2 show us how to construct the true gradient for every cell with their given interactions. It should be noted that the off-diagonal components will need to be scaled given the interaction length of Eq. 3.2. This will approximate the true gradient quite well. Problems arise when we apply strong constraints to the system, or *ex post facto* weights. Since the delta epsilon matrix is now for the entire structure, for each parameter, the gradient is for the entire structure. Thus, if we change the gradient after it has been calculated, (e.g. in the optimizer) for any parameter, we are losing information and not heading in the direction of the true gradient. This explains the need for a custom optimization algorithm. Trying complicated design curves increases the likelihood of constraints being imposed, and a weighting function was used to ‘lock-in’ cells, as the optimizer ran down the structure, which is why this thesis uncovered these issues, where other had not. This became a large part of this thesis and is detailed in section 3 (Methods).

While there are several out-of-the-box optimization algorithms, they apply their bounds and constraints to the gradient after it has been calculated. That is, they calculate the gradient, apply the step size as in Eq. 2.55, *then* apply constraints and bounds. In our case, since our system is not linearly independent, by manipulating the gradient after it has been calculated, information is lost and the gradient is no longer accurate. One of the key findings of this thesis, then, is that due to this “bending” of the adjoint method, you must impose all bounds, constraints, weights, etc. in the construction of the  $\frac{\Delta \varepsilon_r}{\Delta \phi}$  matrix for the gradient calculation (Eq. 2.45). Advanced gradient descent methods, such as adam, can be implemented here too—with some algorithmic manipulation, of course (explained in Methods). The optimization algorithm developed for this thesis also uses a basing-hopping algorithm to more effectively traverse the ‘egg-carton’ parameter space without getting stuck in local minima.

## Outcomes

The interest in adding the  $y$ -position as an optimizable parameter was because, to manipulate the phase of  $e_1$ , is vital to be able to maintain the field phase with that of the particle’s trajectory. The dependency of acceleration on the synchronicity condition  $\lambda_p = \beta \lambda_l$  (Eq. 2.5) comes from the phase synchronicity between the incident field and the particle’s trajectory (Eq. 2.3). While section 4.2 shows that the  $y$ -position of the pillar, within its cell, can achieve a nearly  $180^\circ$  phase shift, Fig. 4.12 gives a likely explanation as to why we only see a  $\pi/5$  ( $36^\circ$ ) phase shift in the structure. Given that non-equivalence, it is clear that we may need another way to achieve full  $360^\circ$  phase control of the structure. A solution to this is proposed in the next section.

Non-periodicity of the structures is of interest because the adjoint calculations typically assume periodicity; however, as an accelerating structure, the particle will change velocity, thus changing its relative phase to the laser. By changing the size of the cells—introducing the non-periodicity—we can account for this dephasing, thus allowing for larger and more complicated acceleration structures. We successfully showed that non-periodic structures can be optimized via the adjoint method of inverse design. With the relationship between the maximum normalized  $\text{Re}\{e_1\}$  value and periodicity ( $\lambda_p$ ), we can scale our design curves with the sigmoid shown in Fig. 4.2. We can achieve a high level of control over the structures with adjoint inverse design. It was found that, given our constraints on the system, we can optimize arbitrarily complicated design curves up to the rate of change limit of  $e_1 \approx 0.1/10$  periods. This is seen in Fig. 4.10 and 4.9. The custom optimization algorithm created for this thesis was successfully able to match to ( $> 95\%$ ) complicated design curves, and large scale non-periodic structures.

This thesis successfully demonstrates that the adjoint method can be used to inverse design a dual pillar dielectric laser accelerator structure with full phase control and non-periodic synchronicity condition ( $\lambda_p(y) = \beta(y)\lambda_l$ ). Despite peculiarities with the adjoint method due, primarily, to the interactions between cells, the optimization algorithm, developed herein, successfully applied bounds, constraints, weights, and a modified adam algorithm. We showed that, given our constraints on the system (Table 4.1) it is possible to opti-

mize arbitrarily complicated design curves up to the rate of change limit of  $\text{Re}\{e_1\} \approx 0.1/10$  periods.

## Future Considerations

This thesis is the step between a periodic, non-accelerating structure, and a non-periodic structure with dynamic periodicity set by maintaining a phase relationship with the particle, via particle tracking. The next step would be to implement particle tracking software with the optimization algorithm developed for this thesis. Then, the optimizer would be able to account for the particle's true phase, even as it traverses the individual cells.

Another easy addition that could be made to the algorithm would be pillar angle as a dynamic, optimizable parameter. This could provide further control over the magnitude and phase of  $e_1$ . By pillar angle, we mean the angle that the major axis of the pillar makes with the  $x$ -axis.

In this thesis we used monochromatic lasers in our simulation, and they were modeled as a plane wave. Since real DLA structures require the high peak field intensities of ultra-short laser pulses—from the uncertainty principal—we know this means that the beam necessarily is not monochromatic. The next step, then, would involve generalizing our method for broadband light, following the derivation in Ref. [19]. This will involve optimizing harmonics of the different wavelengths in the pulse carrier envelope.

As mentioned previously, the cell size (periodicity) was attempted to be made into a dynamic parameter. The issues with it are detailed in section 4.2. Although, it could be possible to devise a clever way to implement the periodicity as a dynamic, optimizable parameter. As was detailed in section 2.5.1, in a real structure, we are bound by fabrication constraints. While it is relatively straightforward to impose our boundary constraints, it is non-trivial to make the optimizer use step sizes that correspond directly to our fabrication resolution (here 10 nm). Ref. [33] provides a way of using, so-called, *level-set functions* to impose these kinds of step function constraints. For example, in this thesis, the optimizer would converge and have values such as  $173.6 \text{ nm} \times 198.1 \text{ nm}$ . Where, due to our fabrication resolution constraints, the closest we could get to that would be  $170 \text{ nm} \times 200 \text{ nm}$ . However, given the complexity of the parameter space, this may not always be the best solution (e.g. perhaps the minimum is  $173.6 \text{ nm} \times 198.1 \text{ nm}$ , but the second best, fabricable structure might be  $180 \text{ nm} \times 190 \text{ nm}$ ). Hence the need for some development on these kinds of constraints within this optimizer.

Finally, as noted, the adjoint method is not a true definition of our system since our matrix  $A$  is not symmetric. We have provided a possible way to calculate  $\frac{\Delta e_r}{\Delta \phi}$  for the entire structure, for every parameter (and thus cell) by using this in the field integral in Eq. 2.32. If we then use Eq. 5.2 to construct the true gradient for every cell with their given interactions, all that is left, to scale the off-diagonal components based the interaction length of Eq. 3.2 and the inverse square law of field power. This could be a possible solution, but more work needs to be done looking into the off diagonal elements in matrix  $A$  of the adjoint method, as it concerns our system of dual pillar DLA structures.



# Bibliography

- [1] Helmut Wiedemann. *Particle Accelerator Physics*. Advanced Texts in Physics. Springer, Berlin, Heidelberg, 2003.
- [2] Sami G Tantawi, Valery Dolgashev, Aaron Jansen, Michael Fazio, Mark Kemp, Jeff Neilson, and Zenghai Li. High Gradient RF Acceleration at SLAC. page 16.
- [3] SLAC | laboratory, Menlo Park, California, United States | Britannica.
- [4] J W Wang and G A Loew. MEASUREMENTS OF ULTIMATE ACCELERATING GRADIENTS IN THE SLAC DISK-LOADED STRUCTURE. page 17.
- [5] M. Kozák, N. Schönenberger, and P. Hommelhoff. Ponderomotive Generation and Detection of Attosecond Free-Electron Pulse Trains. *Phys. Rev. Lett.*, 120(10):103203, March 2018.
- [6] D. Cesar, S. Custodio, J. Maxson, P. Musumeci, X. Shen, E. Threlkeld, R. J. England, A. Hanuka, I. V. Makasyuk, E. A. Peralta, K. P. Wootton, and Z. Wu. High-field nonlinear optical response and phase control in a dielectric laser accelerator. *Commun Phys*, 1(1):1–7, August 2018. Number: 1 Publisher: Nature Publishing Group.
- [7] Kenneth J. Leedle, Dylan S. Black, Yu Miao, Karel E. Urbanek, Andrew Ceballos, Huiyang Deng, James S. Harris, Olav Solgaard, and Robert L. Byer. Phase-dependent laser acceleration of electrons with symmetrically driven silicon dual pillar gratings. *Opt. Lett.*, OL, 43(9):2181–2184, May 2018. Publisher: Optica Publishing Group.
- [8] Joshua McNeur, Martin Kozák, Norbert Schönenberger, Kenneth J. Leedle, Huiyang Deng, Andrew Ceballos, Heinrich Hoogland, Axel Ruehl, Ingmar Hartl, Ronald Holzwarth, Olav Solgaard, James S. Harris, Robert L. Byer, and Peter Hommelhoff. Elements of a dielectric laser accelerator. *Optica, OPTICA*, 5(6):687–690, June 2018. Publisher: Optica Publishing Group.
- [9] Koichi Shimoda. Proposal for an Electron Accelerator Using an Optical Maser. *Appl. Opt., AO*, 1(1):33–35, January 1962. Publisher: Optica Publishing Group.
- [10] A. W. Lohmann. Particle accelerator utilizing coherent light.

- [11] K. Mizuno, J. Pae, T. Nozokido, and K. Furuya. Experimental evidence of the inverse Smith–Purcell effect. *Nature*, 328(6125):45–47, July 1987. Number: 6125 Publisher: Nature Publishing Group.
- [12] E. A. Peralta, K. Soong, R. J. England, E. R. Colby, Z. Wu, B. Montazeri, C. McGuinness, J. McNeur, K. J. Leedle, D. Walz, E. B. Sozer, B. Cowan, B. Schwartz, G. Travish, and R. L. Byer. Demonstration of electron acceleration in a laser-driven dielectric microstructure. *Nature*, 503(7474):91–94, November 2013. Number: 7474 Publisher: Nature Publishing Group.
- [13] N. Schonenberger and P. Hommelhoff. Dielectric Laser Acceleration, August 2020. arXiv:2008.03958 [physics].
- [14] C. Joshi. The Los Alamos Laser Acceleration of Particles Workshop and beginning of the advanced accelerator concepts field. *AIP Conference Proceedings*, 1507(1):61–66, December 2012. Publisher: American Institute of Physics.
- [15] R. Shiloh, T. Chlouba, P. Yousefi, and P. Hommelhoff. Particle acceleration using top-illuminated nanophotonic dielectric structures. *Opt. Express, OE*, 29(10):14403–14411, May 2021. Publisher: Optica Publishing Group.
- [16] Peyman Yousefi, Norbert Schönenberger, Joshua Mcneur, Martin Kozák, Uwe Niedermayer, and Peter Hommelhoff. Dielectric laser electron acceleration in a dual pillar grating with a distributed Bragg reflector. *Opt. Lett., OL*, 44(6):1520–1523, March 2019. Publisher: Optica Publishing Group.
- [17] John Breuer. *Dielectric laser acceleration of non-relativistic electrons at a photonic structure*. PhD thesis, Ludwig–Maximilians–Universitat Munchen, July 2013.
- [18] Uwe Niedermayer, A. Adelmann, S. Bettoni, M. Calvi, M. Dehler, E. Ferrari, F. Frei, D. Hauenstein, B. Hermann, N. Hiller, R. Ischebeck, C. Lombosi, E. Prat, S. Reiche, L. Rivkin, R. Aßmann, U. Dorda, I. Hartl, W. Kuropka, F. Lemery, B. Marchetti, F. Mayet, H. Xuan, J. Zhu, D. S. Black, P. N. Broaddus, R. L. Byer, A. Ceballos, H. Deng, S. Fan, J. Harris, T. Hirano, T. W. Hughes, Y. Jiang, T. Langenstein, K. Leedle, Y. Miao, A. Ody, A. Pigott, N. Sapra, O. Solgaard, L. Su, S. Tan, J. Vuckovic, K. Yang, Z. Zhao, O. Boine-Frankenheim, T. Egenolf, E. Sk r, D. Cesar, P. Musumeci, B. Naranjo, J. Rosenzweig, X. Shen, B. Cowan, R. J. England, Z. Huang, H. Cankaya, M. Fakhari, A. Fallahi, F. X. K rtner, T. Feurer, P. Hommelhoff, J. Illmer, A. Li, A. Mittelbach, J. McNeur, N. Sch nenberger, R. Shiloh, A. Tafel, P. Yousefi, M. Kozak, M. Qi, Y. J. Lee, Y.-C. Huang, and E. Simakov. Challenges in simulating beam dynamics of dielectric laser acceleration. *Int. J. Mod. Phys. A*, 34(36):1942031, December 2019. Publisher: World Scientific Publishing Co.
- [19] Tyler Hughes. *ADJOINT-BASED OPTIMIZATION AND INVERSE DESIGN OF PHOTONIC DEVICES*. PhD thesis, Stanford University, August 2019.
- [20] Dylan S. Black, Zhixin Zhao, Kenneth J. Leedle, Yu Miao, Robert L. Byer, Shanhui Fan, and Olav Solgaard. Operating modes of dual-grating dielectric laser accelerators.

- Phys. Rev. Accel. Beams*, 23(11):114001, November 2020. Publisher: American Physical Society.
- [21] T. Plettner, R. L. Byer, E. Colby, B. Cowan, C. M. S. Sears, J. E. Spencer, and R. H. Siemann. Visible-Laser Acceleration of Relativistic Electrons in a Semi-Infinite Vacuum. *Phys. Rev. Lett.*, 95(13):134801, September 2005. Publisher: American Physical Society.
- [22] Larmor formula.
- [23] Uwe Niedermayer, Thilo Egenolf, and Oliver Boine-Frankenheim. Threedimensional Alternating-Phase Focusing for Dielectric-Laser Electron Accelerators. *Phys. Rev. Lett.*, 125(16):164801, October 2020. arXiv:2004.05458 [physics].
- [24] John Breuer and Peter Hommelhoff. Laser-based acceleration of non-relativistic electrons at a dielectric structure. *Phys. Rev. Lett.*, 111(13):134803, September 2013. arXiv:1308.0464 [physics].
- [25] Adam Reid. Photonic Inverse Design using the Adjoint Method, March 2019.
- [26] Uwe Niedermayer, Thilo Egenolf, and Oliver Boine-Frankenheim. Beam dynamics analysis of dielectric laser acceleration using a fast 6D tracking scheme. *Phys. Rev. Accel. Beams*, 20(11):111302, November 2017. Publisher: American Physical Society.
- [27] Manuel Konrad. *Inverse design of dual pillar structures for dielectric laser acceleration*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, January 2022.
- [28] Huygens–Fresnel principle, July 2022. Page Version ID: 1097518977.
- [29] Understanding time apodization in frequency domain monitors.
- [30] F. Marquardt. Machine Learning and Quantum Devices, 2021. SciPost Phys. Lect. Notes 29.
- [31] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization, January 2017. arXiv:1412.6980 [cs].
- [32] Joaquim R. R. A. Martins and Andrew Ning. *Engineering Design Optimization*. Cambridge University Press, 2021.
- [33] Dries Vercruyse, Neil V. Sapra, Logan Su, Rahul Trivedi, and Jelena Vučković. Analytical level set fabrication constraints for inverse design. *Sci Rep*, 9(1):8999, June 2019. Number: 1 Publisher: Nature Publishing Group.
- [34] Brian Olson, Irina Hashmi, Kevin Molloy, and Amarda Shehu. Basin Hopping as a General and Versatile Optimization Framework for the Characterization of Biological Macromolecules. *Advances in Artificial Intelligence*, 2012:1–19, December 2012.
- [35] Issam Laradji. Non-convex optimization.

- [36] Ole Trinhammer. *Intrinsic quantum mechanics V. Neutrino mass predictions from beta decay.* April 2017.
- [37] J. Makhoul. A fast cosine transform in one and two dimensions. *IEEE Trans. Acoust., Speech, Signal Process.*, 28(1):27–34, February 1980.
- [38] Adrian Tam. Lagrange Multiplier Approach with Inequality Constraints, August 2021.
- [39] R. Shiloh, J. Illmer, T. Chlouba, P. Yousefi, N. Schönenberger, U. Niedermayer, A. Mittelbach, and P. Hommelhoff. Electron phase-space control in photonic chip-based particle acceleration. *Nature*, 597(7877):498–502, September 2021. Number: 7877 Publisher: Nature Publishing Group.



# **Erklärung**

Ich versichere, dass ich die Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Christian Neureuter

Erlangen, den 29. September 2022