Topology Optimization of Nanophotonic Devices

Lirong Yang

Kongens Lyngby 2011
FOTONIK-PHD-2011

Technical University of Denmark
Department of Photonics Engineering
Building 343, DK-2800 Kongens Lyngby, Denmark
Phone +45 4525 6352, Fax +45 4525 6581
info@fotonik.dtu.dk
www.fotonik.dtu.dk

Summary

This thesis explores the various aspects of utilizing topology optimization in designing nanophotonic devices. Either frequency-domain or time-domain methods is used in combination with the optimization algorithms, depending on various aims of the designing problems.

The frequency-domain methods are appropriate for problems where the power is to be maximized or minimized at a few frequencies, without regards on the detailed profile of the optical pulse or the need of large amount of frequency samplings. The design of slow light couplers connecting ridge waveguides and the photonic crystal waveguides is showcased here. It is demonstrated both numerically and experimentally that the optimized couplers could improve the coupling efficiency prominently.

With more focus on the time-domain optimization method, the thesis discusses extensively the design of pulse-shaping filters, which greatly exploits the benefits of time-domain methods. Finite-difference time-domain method is used here as the modeling basis for the inverse problem. Filters based on both one-dimensional gratings and two-dimensional planar structures are designed and different issues regarding local minima, black and white design, minimum lengthscale and flexible pulse delay are addressed to demonstrate time-domain based topology optimization's potential in designing complicated photonic structures with specifications on the time characteristics of pulses.

Resumé

Denne afhandling udforsker forskellige aspekter af anvendelsen af topologioptimering ved design af nanofotoniske komponenter. Der anvendes enten en frekvensdomæne- eller en tidsdomænemetode i kombination med optimeringsalgoritmerne, afhængigt af de forskellige mål for designopgaverne.

Frekvensdomænemetoderne er hensigtsmæssige til problemer hvor effekten skal maksimeres eller minimeres ved nogle få frekvenser, uden at tage hensyn til den detaljerede profil af den optiske puls eller behovet for mange frekvenskomposanter. Designet af koblere til "langsomt lys", der forbinder normale kantbølgeledere og fotoniske krystalbølgeledere er behandlet her. Det er demonstreret både numerisk og eksperimentelt at de optimerede koblere kan forbedre koblingens effektivitet markant.

Med fokus på tidsdomæneoptimeringsmetoden diskuterer afhandlingen omfattende designet af pulsformgivningsfiltre, som i høj grad udnytter fordelene ved tidsdomænemetoderne. "Finite-difference time-domain" metoden er anvendt som modelleringens fundament for det inverse problem. Filtre baseret på både én-dimensionelle gitre og to-dimensionelle plane strukturer er blevet designet, og forskellige problemer vedrørende lokale minima, sort/hvidt design, mindste længdeskala og variable pulsforsinkelser er blevet adresseret. Hensigten er at demonstrere potentialet for tidsdomænebaseret topologioptimering ved design af komplicerede fotoniske strukturer til frembringelse af lyspulser med specifikke af tidskarakteristikker.

Publications and conference contributions

Peer reviewed international scientific journal publications:

[1] **L. Yang**, A. Lavrinenko, L. Frandsen, P. Borel, A. Tetu, and J. Fage-Pedersen. Topology optimisation of slow light coupling to photonic crystal waveguides. Electronics Letters, 43(17):923 - 924, 2007.

[2] **L. Yang**, A. V. Lavrinenko, J. M. Hvam, O. Sigmund. Design of one-dimensional optical pulse-shaping filters by time-domain topology optimization. Applied Physics Letters, 95(26):261101 - 261101-3, 2009.

Peer reviewed international scientific conference contributions:

[3] A. Têtu, **L. Yang**, A.V. Lavrinenko, L. H. Frandsen, and P. I. Borel. Enhancement of coupling to the slow light regime in photonic crystal waveguides using topology optimization. Proceedings of CLEO/QELS, CA, May 21-26, 2006.

[4] **L. Yang**, A. V. Lavrinenko, P. I. Borel, L. H. Frandsen, J. F. Pedersen, A. Têtu, J. S. Jensen, O. Sigmund. Improved slow light coupling efficiency to photonic crystal waveguides. Oral presentation. Optical Society of America topical meeting on Nanophotonics (OSA NANO'07), Hangzhou, June 18-21, 2007.

[5] **L. Yang**, A. V. Lavrinenko, O. Sigmund and J. Hvam. 1D grating structures designed by the time domain topology optimization. Poster presentation. Proceedings of the 17^{th} International Workshop on Optical Waveguide Theory and Numerical Modelling (OWTNM 2008), Eindhoven, the Netherlands, pp 41, June 13-14, 2008.

[6] M. Pu, **L. Yang**, L. H. Fradsen, J. S. Jensen, O. Sigmund, H. Ou, K. Yvind, and J. M. Hvam. Topology-optimized slow-light couplers for ring-shaped photonic crystal waveguide. National Fiber Optic Engineers Conference, OSA Technical Digest (CD) (Optical Society of America, 2010), paper JWA30, 2010.

[7] **L. Yang**, A. V. Lavrinenko, O. Sigmund and J. Hvam. Time-domain topology optimization of pulse-shaping filters. Oral presentation. Proceedings of the 18^{th} International Workshop on Optical Waveguide Theory and Numerical Modelling (OWTNM 2010), Cambridge, UK, pp 49, April 9-10, 2010.

Acknowledgements

This thesis could not have existed without the help of many people, with my three supervisors being the utmost crucial factors in making it all possible. I owe greatly to Jørn Hvam for his timely encouragement and moral support, till the very end; Ole Sigmund for his always enlightening academic guidance and inspirational discussions throughout the past many years; and Andrei Lavrinenko for his constant help with all scientific details, large and small. I am also very grateful towards Ole Sigmund and Jakob Jensen at MEK, DTU for letting me use their brilliant in-house program *Topopt* to design the slow light couplers. My discussions with Jonas Dahl at the very beginning of this project was very helpful and encouraging. My experimental collaborators, Lars H. Frandsen, Amélie Têtu and Minhao Pu have all been waving their amazing magic wands in the lab, bringing the designs to life. Last but not the least, I am grateful to my loving family for their unyielding support through all this time.

Contents

Summary **i**

Resumé **iii**

Publications and conference contributions **v**

Acknowledgements **vii**

1 Introduction **1**
1.1 Motivation 1
1.2 Thesis structure 5

2 Topology optimization **7**
2.1 Basics of topology optimization 8
2.2 Comparisons to genetic algorithms 10
2.3 Conclusions 12

3 Maxwell's equations and their numerical solutions **13**
3.1 Maxwell's equations 15
3.2 Finite element method 17
  3.2.1 Helmholtz equation 17
  3.2.2 Discretization 19
3.3 Finite-difference time-domain method 20
  3.3.1 Maxwell's equations reduction to 2D and 1D 20
  3.3.2 The Yee grid and the leap frog scheme 20
  3.3.3 FDTD update equations 22
  3.3.4 Stability criteria 24
  3.3.5 Absorbing boundary conditions 25
3.4 Conclusions 26

4 Frequency-domain topology optimization 27
- 4.1 Rationale . . . 28
- 4.2 Design and fabrication of slow light couplers . . . 29
  - 4.2.1 PhCW with round holes . . . 30
  - 4.2.2 PhCW with ring-shaped holes . . . 33
- 4.3 Conclusions . . . 35

5 1D time-domain topology optimization 39
- 5.1 Problem formulation . . . 39
- 5.2 Sensitivity analysis . . . 40
- 5.3 Proof of concept . . . 41
- 5.4 Optimization of 1D pulse-shaping filters . . . 43
  - 5.4.1 Motivation . . . 43
  - 5.4.2 Objective function . . . 44
    - 5.4.2.1 Envelope objective function . . . 44
    - 5.4.2.2 Sensitivity analysis for the envelope objective function . . . 45
    - 5.4.2.3 Explicit penalization . . . 46
    - 5.4.2.4 Modified objective function . . . 48
  - 5.4.3 Results . . . 48
- 5.5 Conclusions . . . 50

6 Minimum lengthscale control and black/white designs 51
- 6.1 Test problem formulation . . . 53
- 6.2 SIMP . . . 55
- 6.3 Density filters . . . 56
- 6.4 Sensitivity filters . . . 59
- 6.5 Explicit penalization . . . 60
- 6.6 Modified Heaviside filters . . . 62
- 6.7 Conclusions . . . 64

7 2D Time-domain Topology Optimizations of Pulse-shaping Filters 67
- 7.1 The inverse problem . . . 68
- 7.2 Square-pulse filters . . . 68
  - 7.2.1 Original problem formulation . . . 69
  - 7.2.2 Delay variable . . . 69
  - 7.2.3 Transmission efficiencies for the filters . . . 72
  - 7.2.4 Minimum length-scale control and black/white design . 73
  - 7.2.5 Results . . . 73
- 7.3 Saw-tooth filters . . . 75
- 7.4 Pulse-splitting filters . . . 76
- 7.5 Thresholded performance . . . 76
- 7.6 Conclusions . . . 78

8 Conclusions and future work 79

A Sensitivity analysis for topology optimizations based on finite-difference time-domain method 83
- A.1 Sensitivities for 1D problems 83
  - A.1.1 Problem formulation 83
  - A.1.2 Definition of sensitivities 85
  - A.1.3 The finite difference method for calculating sensitivities 85
  - A.1.4 1D sensitivity analysis by using the adjoint-variable method 86
    - A.1.4.1 Derivation of the implicit sensitivity term . . . 86
    - A.1.4.2 Derivative residual 89
    - A.1.4.3 The adjoint problem 91
    - A.1.4.4 The adjoint current 93
    - A.1.4.5 Implementation of sensitivity analysis using the adjoint-variable method 94

Chapter 1

Introduction

1.1 Motivation

The inventions of semiconductor lasers and optical fibers in the 1960s and 1970s mark the inception of the photonics research. As the under-sea optical fibers convey gigabytes of data with light signals across the globe every second, photonic devices gradually took over the stage of telecommunication which was previously dominated by their electronic counterparts. Although telecommunication became the prime arena for photonics, other non-communication applications, including fiber sensors, non-linear optics and bio-optics, also benefitted from this new field of research. Guided-wave devices started to gain the attention of academia and industry for their low loss and high bandwidth characteristics. In a planar waveguide, light is confined by total internal reflection (TIR) in a small modal region inside the high-index semiconductor materials instead of being guided by discrete lenses and mirrors as in bulk optics. Planar waveguides transform photonic devices into compact chip sets with more stability and less power consumption and are widely deployed instead of traditional optical components in emission, transmission, amplification, detection and modulation of light.

Researchers looked into different semiconductor materials in order to find a good platform for realizing various photonic functionalities. III-V semiconducting compounds and other crystals like lithium niobate ($LiNbO_3$) were the prime candidates in the early years, either due to their direct band gaps for light emission and detection, or for the Pockels effect crucial in modulation and switching. On the other hand, silicon has long been established as the dominant material in the electronics industry. It is a cheap crystal with robust quality, and its complementary metal-oxide-semiconductor (CMOS) foundry technology has been well-established and is capable of high volume manufacturing. Naturally, it would be a cost-effective and an elegant solution if silicon photonic components would be readily available to integrate with the existing material platform of electronics. In the 1980s, the potential of silicon photonics surfaced with the material's newly recognized transmission transparency in the telecommunication wavelength ($1.3\mu m \sim 1.55\mu m$). Moreover, waveguides built on silicon-on-insulator (SOI) platform are able to guide light at a very low propagation loss due to the large index contrast between the waveguide core and its silica claddings. Gradually, silicon becomes a prominent candidate for photonic devices. Various efforts were devoted to design and fabrication of silicon-based photonic devices that are compatible with the standard CMOS technology for electronics. Even though bulk crystal silicon does not have a direct band gap for easier light emissions or Pockels effect for enabling switching functions, alternative properties of silicon are being explored to devise silicon-based photonic components including switches, modulators and detectors. On another end, progress in heterogeneous integration between active materials and SOI [1][2] also makes it possible to group function blocks of different materials on the same photonic chip. As of today, silicon has become the dominant photonic material for optoelectronic integrated chips (OEIC) and photonic integrated chips (PIC), and the progress in silicon photonics exhibits the potential to finally combine best of two worlds: electronics and photonics.

With the ever increasing internet traffic in this multimedia era, larger bandwidth is required on the existing fiber network. Researchers are striving to achieve both higher transmission speed per wavelength channel as well as a bigger number of wavelength channels transmitted per fiber. As these high-speed systems are being developed, the electronic components are pushed towards their speed limit. Devices like switches which can operate extremely fast became the new research direction, where there will be less need to convert light to electricity and vice versa. As the trend for further integration of electronics and photonics progresses, the need for additional reduction of the sizes of photonic components strengthens. Even though the cross-section for silicon waveguides has reduced significantly due to improvement of surface roughness in the fabrication process, traditional waveguides structure still faces a road block. Total internal reflection, which allows light to propagate along the waveguide, requires large incident angles as light zigzags inside the chamber. This fundamentally puts a lower limit on the curvature of waveguide bends, and hence obstructs further miniaturization of photonic devices.

Photonic crystals (PhC) came into sight in 1987 [3][4], and gradually gained significant attention after 2000. Dielectric materials are arranged periodically in specific lattice patterns, much like atoms in the crystalline structures in solids. By simulating the crystal structures and expressing them on a more macroscopic level, these artificial crystals then acquire a bandgap for light in a certain frequency range, similar to the electronic bandgap in semiconductors.

This photonic bandgap forbids photons to propagate through the bulk crystal. Hence by making a line defect in a bulk PhC, light will be trapped inside the line defect, forming an effective waveguide. Since light is strongly confined in the defect, waveguides can now have much sharper turns without loosing too much light due to scattering. This invention potentially paved the way for designing compact photonic circuit layout in a small chip area, thus effectively counters the size issues of optoelectronic integrated circuits.

Interests have been garnered around further improving these PhC components in regards to lower loss, higher bandwidth and other desirable properties. The large amount of scatters existing in these devices naturally provides a rich possibility of re-arranging the rods and holes to fine tune the device performances. Various attempts have been made at adjusting the lattice structure locally to improve the device performance, based on some basic physical arguments. For example, coupled-mode theory is applied in designing efficient PhC-based Y-junctions [5] and high-transmittance waveguide bends [6][7][8]. Small holes, either uniform or adiabatically-arranged, are introduced along the defect or in the vicinity of the bends to assist gradual modal conversion. Known frequency shifts between the crystals' different lattice periodicities or various propagation modes give rise to a mean of manipulating the band diagrams of the structure by dislocating parts of the lattice [9] or by inserting a roll of small holes midst of a waveguide to prohibit a multimode from forming [10]. Resonance cavities are produced around the bends or splitter junctions to properly couple the light from the input waveguide to the output waveguide(s) around a waveguide bend [11] or a power splitter [12]. Apart from these physical arguments, intuitive geometrical assumptions are also used to create functional features in the structure. For a more efficient PhC waveguide bend, critical holes/rods that are originally located on the lattice points, are rearranged around the bends [13] or join together [14] to form a 'smoother' corner for light to pass through. For most of the above-mentioned applications, the details of the geometrical maneuver, i.e. the size of the new holes, the extent of the lattice dislocations or the exact cavity geometry, are selected empirically and largely determined in a trial-and-error process. Physical arguments used to envision the functional geometries, while useful at times, do not guarantee optimal performance. The transmission, bandwidth and the reflection are highly sensitive to small variations in the geometry, which calls for a more rigourous design methodology. The procedure also lacks generality, which prohibits its further extension to more complex function blocks. More systematic measures are also available. Instead of choosing which holes to move around based on crude arguments, sensitivity analysis can be used where small variations are exerted to the positions, sizes or material composition of specific lattice site and the device performances are evaluated accordingly [15]. Such a method quantitatively determines the most influential geometrical features to which the device performance is most susceptible. On another end, stochastic optimizations (simulated annealing, evolutionary algorithms, etc.) are utilized to find the optimal sizes or locations of the holes/rods around the bends or splitter joints in order to improve the transmission in a PhC waveguide bend [16] and a PhC-based power divider

[17]. Although effective in finding a better layout with improved performances, the number of design variables allowed in stochastic optimization methods is usually very limited (see section 2.2 for arguments). For one-dimensional grating design problems, these methods are adequate if relatively few layers of the gratings are needed [18]. For two-dimensional design problems where the geometries are more complex than their one-dimensional counterparts, the optimization processes are often reduced to simplified shape optimizations. The design variables are sizes or material distributions of the lattice sites instead of the complete design domain where neither the boundaries of the features nor their connectiveness is known a priori. The full topology is mapped by projecting these few design variables to the whole domain. Intuitively, we may assume that there exists a better solution with a topology containing more irregular shapes than round holes. In 2004, Sigmund and Jensen proposed using topology optimization (TO) to optimize the PhCW bends [19]. By using a systematic algorithm, a more optimal solution which contains topologies not confined by predetermined shapes was found. Soon, TO was utilized to design more PhCW-based devices [20][21][22][23][24]. More application areas including designing photonic crystal cell geometries with optimal planar bandgap structures [25] as well as high Q-factor PhC microcavities [26] also emerged. TO has been proven an efficient tool to optimize a design region as part of the whole PhC component in order to improve the device performance without compromising the bandgap properties of the original device.

Previously, TO of photonic devices was mainly based on frequency domain method. In this thesis, we explore the possibilities of designing nanophotonic devices using the combination of TO and the finite-difference time-domain method (FDTD). FDTD-based TO was first exemplified by Nomura in the design of broadband dielectric resonator antennas [27]. To further examine the scope and feasibility of this optimization method, we aim at designing two-dimensional (2D) planar pulse-shaping filters and focus on the temporal conversions between the input and output pulses.

Various pulse-shaping filters were used in telecommunications, nonlinear optics and biomedical imaging. For example, in high-speed optical communication systems, well-defined temporal square wave pulses as switching signals are essential in counteracting timing jitter problems. The most-employed technique for ultrafast pulse shaping is Fourier synthesis [28]. It is based on spatial filtering of optical frequency components and is implemented by a relatively intricate system comprised of discrete optical components like diffraction gratings, lenses and phase/amplitude spatial masks. Another more intuitive method is by combining interferometers and delay lines [29]. By coherently and successively delaying the Gaussian-like input pulse and then superimposing the delayed pulses, arbitrary pulse shapes can be achieved depending on the amount of delay. Apart from discrete systems, fiber gratings-based filters have also become prime candidates since they are more stable and more coupling-friendly with the planar waveguide systems. Several methods exists for designing fiber gratings. Electromagnetic inverse scattering is used by matching the spacial refractive index modulation profile to that of the spectral impulse response of the desired transfer function between the input and the output pulse. The modulation profile is then expressed in a superstructured fiber Bragg grating that acts as a spatial filter for shaping pulses [30]. The layer peeling method was developed by geophysicists to examine the physical properties and structures of the layered media where waves propagate. It was later incorporated by Feced [31] and Skaar [32] to determine the layered structure of fiber gratings that have a specific spectral response. However, both methods are confined to one-dimensional (1D) layered systems and cannot be applied to designing filters based on two-dimensional planar structures.

FDTD-based TO is used here to design the layout of pulse-shaping filters based on 2D planar SOI waveguides. Such devices have the potential to be directly integrated with other waveguide systems on OEICs and PICs.

1.2 Thesis structure

The thesis is structured as follows.

We start out by briefly describing the basic concepts and advantages of topology optimization in chapter 2. Chapter 3 deals with the modeling perspectives of our implementation, and gives a practical account of finite element method (FEM) and finite-difference time-domain method. Frequency-domain TO is showcased in chapter 4 to optimize slow-light couplers between ridge waveguides and PhCW, with round holes as well as ring-shaped holes in the lattice. Chapter 5 to Chapter 7 present the methodology and results of topology optimization based on FDTD. 1D grating design problems are presented and discussed in chapter 5. In order to achieve practical designs that can eventually be fabricated on the 2D SOI platform, Chapter 6 details the technical tools we use to ensure minimum length-scale control and black/white design. The results for 2D pulse-filtering designs are presented in Chapter 7. In chapter 8, the results for the thesis are summarized and the conclusions drawn. In the appendix, detailed derivation sensitivity expressions using the adjoint-variable method is presented.

CHAPTER 2

Topology optimization

Just as the microstructures in a material determines its physical properties, specific geometrical layouts of a macroscopic structure also decide its behaviors and performances to a great extent. Researchers have long been using mathematical programming to find optimal compositions of materials to improve various characteristics of structures. Techniques for structural optimization have stridden from linear programming to nonlinear programming, from optimizing shapes and sizes to optimizing layouts where sizes, shapes and connectivity of the features are all unknowns, and from small-scaled and simple mechanical models to large and multiphysics problems.

Two main strategies exist in optimizing topological features of a structure. One of them is to generate a set of individual solutions based on certain heuristic algorithms, and to evaluate them in order to select the best ranked solution. The generation of these solutions is usually based on a stochastic process. Among this class of probabilistic optimizers [33][34][35][36][37], genetic algorithms (GA) which is inspired by evolutionary processes is a strong contenders. The other type of optimization resorts to a continuous process, where intermediate solutions are produced according to the gradient information from the previous iteration. These intermediate solutions do not necessarily present physical structures on their own, but with proper penalization and controls, they gradually converge to a final physical solution, which is considered an optimum. Of the latter, topology optimization (TO) is a popular method that has been proven its efficacy in many problems. It was first introduced by Bendsøe and Kikuchi [38] in 1988 on material distribution problems using composite materials. By distributing material freely in the design domain, TO has been utilized in optimizing various physical quantities (compliance, displacement, stress and etc.) in mechanical structures. Since its inception, the technique has undergone great development and has been expanded to multiple physics problems including Stokes flow problems, heat conduction problems, wave propagation problems and etc. [39]. Its versality lies in the utilization of the adjoint method [40] to retrieve sensitivities in an efficient manner.

As the modern fabrication technology advances, production of artificial microscopic features down to the size of several nanometers becomes feasible. It not only provides the human beings with new dimensions of controlling objects and energy in a minuscule way not fathomable before, but also naturally broadens the realm in which TO is applicable. For example, TO has been used to optimize electrothermomechanical in a microelectromechanical system (MEMS) [41]. Photonic crystal waveguide termination has been design to have a much larger directional emission [42]. TO has also revealed some interesting link between the optimal cell structure design for PhCs and geometrical tessellation methods [25]. Apart from the common formulation where frequency-domain methods are used, TO has also extended to time-domain method. Nomura has used FDTD-based TO in antenna design [27], and Dahl did a pilot study of a transient topology-optimization approach in one-dimensional photonic devices [43]. For a more comprehensive review over TO applications in designing photonic devices, please refer to [44].

In this chapter, we familiarize the readers with basic concepts of topology optimization. A brief comparison between TO and one of the other major optimization algorithms, genetic algorithms, is also presented.

2.1 Basics of topology optimization

In this section, we first list the terminologies of some basic concepts in TO to assist the readers with understanding of this thesis. An flow chart for the general TO process is drawn afterwards.

Design domain : The design domain is the geometric volume, area or distance where the optimization algorithm distributes material within and is a part of the total calculation domain. The domain is discretized into elements or grid points which are not only the basic building blocks for the numerical modeling process, but also manifest the updated physical properties in each intermediate topology.

Densities ($\boldsymbol{\rho}$) : This is a vector of N variables that are being directly updated by the optimization algorithm, N being the total number of design variables in the design problem. In most cases, each design variable (ρ_i) corresponds to the physical properties of a single element/grid point (x_i) in the design domain through a certain material interpolations. The most straight forward material interpolation, for a two-phases dielectric material design problem for example, renders the relationship between the densities (in this case the electrical permittivities) and the local permittivities as such: $x_i = \varepsilon_2^r + (\varepsilon_2^r - \varepsilon_1^r)\rho$. Here, ε_2^r and ε_1^r are the higher and lower relative permittivities of the two design materials, and i is the order of the element/grid point in the design domain. The local material property takes the form of the high refractive index material if ρ equates to 1, low refractive index material if ρ is 0, and linearly scales in between the two materials when ρ is otherwise.

Objective function ($F(\boldsymbol{\rho})$) : The objective function is a function depending explicitly and/or implicitly on the design variables. It evaluates the global fitness of the current solution. For mechanical problems, it can be the compliance of the structure, the displacement at a certain structural point, or in a more complicated case, the crashworthiness of a car. In a wave propagation problem, it can be the energy flow through a certain port, or the band gap size of a bulk photonic crystals. The aim of the optimization can be to minimize or maximize the objective function value, which should gradually converge through the optimization process.

System equations : The system equations are what the numerical modeling of the structure must adhere to. For wave propagation problems, they can be e.g. Helmholtz equations or Maxwell's equations.

Constraints : The minimization or maximization of the objective function value is usually not without constraints for mechanical problems. Such constraints are usually constituted of volume, stress, or displacement. In wave propagation problems, volumetric constraints are less pertinent since there is marginal difference in how much dielectric material is present in the final design as long as the design domain is fixed. However, constraints might be added as a numerical maneuver, e.g. to improve convergence.

Sensitivities ($\frac{\partial F}{\partial \boldsymbol{\rho}}$) : Explicit derivatives of the objective function ($F(\boldsymbol{\rho})$) and other constraints with respect to $\boldsymbol{\rho}$. It is a quantitative measure of how individual design variables impact the design goals. According to the theory of adjoint-variable analysis [45], at most two system analyses are needed to compute all sensitivity information in a structure to a certain response.

Figure 2.1 is a flow chart for a typical TO process. Here we use the method of moving asymptotes (MMA) as the mathematical programming tool to update the design variables [46]. MMA approximate the smooth, non-linear optimization problems with a sequence of simpler convex subproblems. These subproblems are constructed based on sensitivity information at the current iteration as well as the previous few iterations. MMA has been used with TO techniques in many applications and has demonstrated its efficiency and stability in solving optimization problems with many design variables and very few constraints [39].

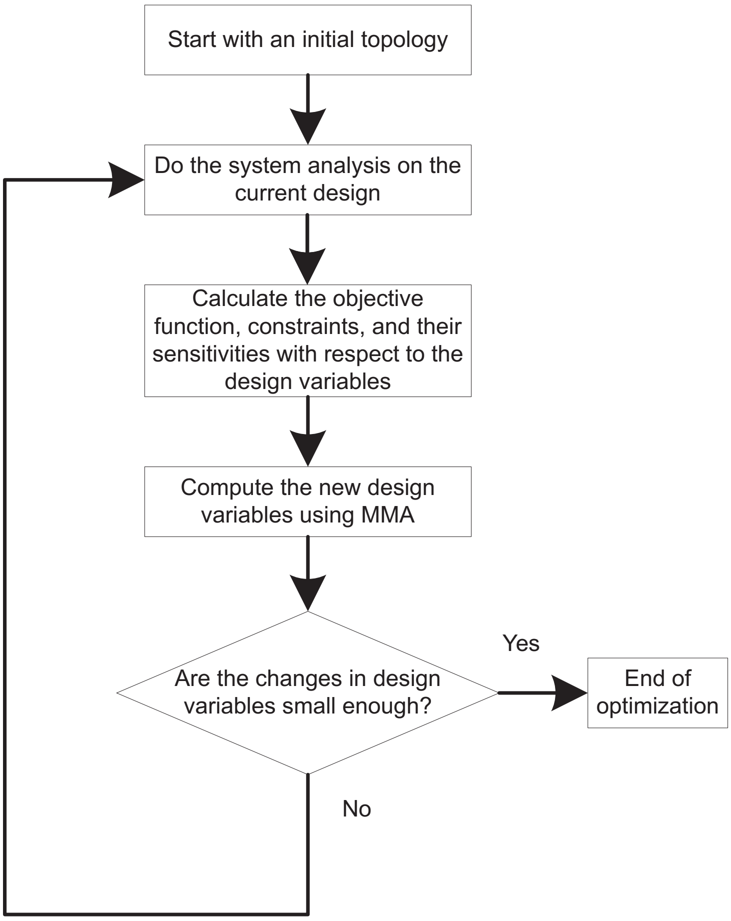

Start with an initial topology
Do the system analysis on the current design
Calculate the objective function, constraints, and their sensitivities with respect to the design variables
Compute the new design variables using MMA
Are the changes in design variables small enough?
Yes
End of optimization
No

Figure 2.1: The flow chart of a typical topology optimization process.

2.2 Comparisons to genetic algorithms

Genetic algorithm (GA) is one of the major contenders in solving inverse problems [34][33]. It is an evolutionary optimization method based on Darwinian survival-of-fittest principle. The design variables are assembled into a vector as one candidate solution, termed an individual. The population consists of a number of individuals, which are usually generated randomly in the beginning of the optimization. In each generation, the fitness of each individual is evaluated by a fitness function. Individuals who perform well on this evaluation will be selected to breed a new generation. There are several genetic operators which transform the current selected individuals in order to render the next generation. The most used operators are: 1. *crossover (mating)*, where two or more of the individuals in the selected population are combined according to certain rule to form a new individual, much like the mating process in nature.

This process makes sure that the good genomes are kept throughout the generations, so the average 'fitness' among the population is guaranteed to improve; 2. *mutation*, where parts of the individual are swapped to different values. Mutation maintains the diversity of the population, which helps the optimization look beyond the nearby local minima and hopefully reach towards the global minimum.

Several pioneering experiments have been done to apply GAs to the design of photonic devices. Goh et. al. proposed using genetic optimization for one-dimensional (1D) and two-dimensional (2D) photonic bandgap structures [47]. In 1D, the widths of 20 or so dielectric stacks are being optimized, while in 2D, the radii of the 9 holes in a subcell are optimized in order to design a large-bandwidth bulk material.

GAs have the following advantages:
1). The bitstring representation of the solutions (chromosomes) fits well with binary optimization problems. No special care needs to be taken to ensure the final design consisting of only two distinctive materials (0/1 design).

2). The problem formulation is flexible. As soon as a fitness function can be defined, it can be used as a merit function to evaluate a design and ultimately guide the optimization to an optimal design. For example, if the electrical (E) fields can be calculated for a design, optimizations can be done directly to alter the distribution of E, or its Fourier transformation in the frequency space. Since gradient information is not needed here, more complex objective function can be applied without regarding whether its derivatives to the design variables exist.

However, GAs also have several major disadvantages:

1. The greatest setback of GAs is their intimidating computational expenses. The number of system analyses needed in the optimization process of GA is the product of the population size (the number of candidate solutions in each generation) and the total number of generations. A sufficiently large S to explore enough solution space is needed, and a certain number of generations are also necessary for the optimization to converge to a reasonable design. For large topological problems where the number of design variables is usually in the order of hundreds of thousands, both population size and generations needed grow exponentially with the problem size, making the computational load astronomical. This drawback largely limits the range of problems GAs can solve in the field of nano-photonics. For example, while a unit cell structure with varying sizes of holes can be optimized by GAs by projecting a few design variables to a full array of periodic cells, a full 2D/3D inverse scattering problem where the periodicity is to be broken is far too computationally heavy to be solved by GAs.

2. Tuning of the parameters. GA is quite sensitive to several parameters,

e.g. the population size, the rate of mutation, the crossover probabilities and etc.. These parameters are generally problem specific and thus can be tedious to adjust properly. For TOs, when used with a robust SLP algorithm, the optimization usually converges well as long as the objective function is properly designed and scaled.

Even though TO and GA are vastly different optimization methods, one problem is common to them: both are easy to fall prey to local minima if the problem is non-convex. These iterative methods are generally "short sighted", hence devising a good objective function/fitness function that is well regulated is crucial in obtaining a good design.

2.3 Conclusions

We gave a brief overview of the basic concepts of topology optimization method and its merits. By using intermediate values, a piece-wise constant variable is formulated as a continuous variable, which makes the optimization a continuous process. The adjoint-variable method is used to derive the sensitivity information from just two system analyses, which makes efficient topology optimization a possibility. The method of moving asymptotes, a mathematical programming tool, has been proved efficient to work with typical topology optimization problems with many design variables but few constraints. A short analysis was presented to compare TO with GA, a popular stochastic optimization method. GA provides a flexibility when formulating a design problem, since it doesn't require the underlying problem to be differentiable, and no sensitivity expressions need to be derived. It also naturally fits the scope of multi-phase optimizations, since each one of its candidate solutions is already a physical topology and thus no need for further procedures to make sure that intermediate materials are eliminated. However, the number of GA's fitness evaluations has an exponential dependence on the number of design variables, which largely limits the scope of applications where GA is feasible.

Chapter 3

Maxwell's equations and their numerical solutions

Ever since James Clark Maxwell's seminal paper in 1861 [48], Maxwell's equations have been deemed as the governing equations of interactions between electric- and magnetic-fields (EM) around their sources. Solutions to Maxwell's equations guide scientists in understanding and exploring natural phenomena as well as spearheading many of the most exhilarating inventions in human history, among them telephone, radar and modern telecommunication. However, until 1960s, the solutions to Maxwell's equations were mainly analytical ones. The availability of these solutions as well as the feasibility of solving them depend greatly on the complexities and the sizes of the structures of interest. Numerical solutions, while being the clear candidate for its potential in solving problems with more complicated boundary conditions and parameters, are essentially impossible to implement due to limited computational means. This renders it difficult to study the EM wave problems full vectorially and limits the further understanding of optically large and complex structures. Fortunately, with the advent of powerful digital computers and advanced programming languages, researchers are able to implement various numerical solutions to study EM wave problems with intricate geometries.

Two main classes of EM solvers exist, categorized by the forms in which Maxwell's equations are formulated. One of them solves the integral form of Maxwell's equations, and includes methods like method of moments (MoM) [49], fast multipole method (FMM) [50], and plane-wave time-domain method (PWTD) [51]. These methods only require discretizations on the surface of the structure instead of the whole volume, thus decreasing the complexity of the solution. However, many of these solvers (e.g. MoM and FMM) depend on the calculation of Green's functions on each subdomains, which limits their generality in more complicated scattering problems. The other class of solvers are based on the partial differential equation (PDE) form of Maxwell's equations, wave equations or Helmholtz equations. These include finite-difference time-domain method (FDTD) [52][53][54], finite-element method (FEM) [55][56], finite-element time-domain method (FETD) [57] [58] and finite-volume time-domain method (FVTD) [59][60]. These methods discretize the space volumetrically.

The above mentioned PDE solvers can be further classified into time-domain methods (FDTD, FETD and FVTD), and frequency-domain methods (FEM). The frequency-domain methods solve one frequency at a time, and is faster if only a few frequencies are requested for solutions. Meanwhile, the time-domain methods solves a wide band of frequencies in one go, and is naturally more efficient when broadband calculations are needed.

Introduced in 1966 by Yee [52] and heralded by Taflove [54], FDTD method has proved its efficacy in simulating wave propagations and scatterings in the optics domain. One of the main challenges for FDTD when it comes to modeling complicated structures is that it uses the uniform Cartesian grid for discretization. For layouts where curved material boundaries are abundant, e.g. photonic crystals, a staircasing scheme is usually taken to approximate these boundaries in a saw-tooth manner. This kind of approximation destroys the second-order accuracy of the algorithm [61]. On the other hand, FEM, FETD and FVTD all work with unstructured grids. These grids are especially desirable if different resolutions are needed across the calculation domain. In an unstructured grid, finer subdomains can be allocated around the irregular discontinuities to improve the accuracy of the approximation, while larger elements are used elsewhere to maintain the computational efficiency. In many finite-element implementations, the meshes can be generated in such an adaptive manner automatically. However, compared to FDTD, FEM, FETD and FVTD methods are not as efficient when it comes to computation resources. Take FETD for example, it requires the solution of a sparse linear system at each time step, which produces a bottleneck for the solver when the size and complexity of the problem scales up. Various efforts have been attempted to speed up the matrix solution, e.g. by using mass lumping [62][63]. Compared to matrix-based methods, FDTD also has the benefits of being highly parallelizable. Moreover, as a time-domain method, nonlinearity and time-varying scatters are much more easily implemented in FDTD than in a frequency-domain method, e.g. FEM. Hybrid methods which combine FDTD with other methods exist, where FDTD on uniform grids is used in large homogeneous volumes and FEM/FETD/FVTD on unstructured grids is applied near complex material boundaries [64][65][66]. Subgridding can also be applied, where parts of the domain are discretized by finer Cartesian grids, thus reserving the structured nature of FDTD [67][68]. This strategy, however, introduces spurious reflections and is not very stable. Adaptive meshes/subgriddings, however, are difficult to apply efficiently in topology optimization processes, where the layout of the structure alters in each optimization iteration. The constant changes render the original adaptive mesh or subgridding invalid due to the change of locations of the discontinuities. Re-meshing in between the iterations is possible, but it is computational expensive and thus counteracting the improved calculation efficiencies brought about by adaptive meshing. More importantly, by refining mesh around material boundaries, more detailed features are encouraged to appear in these areas, making convergence of the optimization difficult.

In this chapter, we introduced the basis of FEM and FDTD methods which are the underlying modeling techniques for the topology optimizations presented in this thesis. Both methods are derived from the differential form of Maxwell's equations, but FEM solves the de-coupled Helmholtz equation which is time-independent, while FDTD solves both electric- and magnetic-fields in the time domain. Since the main focus of this thesis is on topology optimization based on time-domain methods, FEM is only going to be very briefly addressed, while more detailed aspects of FDTD are presented and discussed here.

3.1 Maxwell's equations

Maxwell's equations are described as follows:

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} - \mathbf{M} \qquad (Faraday's\ law) \tag{3.1a}$$

$$\frac{\partial \mathbf{D}}{\partial t} = \nabla \times \mathbf{H} - \mathbf{J} \qquad (Ampere's\ law) \tag{3.1b}$$

$$\nabla \cdot \mathbf{D} = \rho \qquad (Gauss's\ law\ for\ the\ electric\ field) \tag{3.1c}$$

$$\nabla \cdot \mathbf{B} = 0 \qquad (Gauss's\ law\ for\ the\ magnetic\ field) \tag{3.1d}$$

where,

$\mathbf{E}$ is the electric field (in $[V/m]$)
$\mathbf{H}$ is the magnetic field (in $[A/m]$)
$\mathbf{D}$ is the electric flux density (in $[C/m^2]$)
$\mathbf{B}$ is the magnetic flux density (in $[Wb/m^2]$)
μ is the magnetic permeability (in $[H/m]$)
ε is the electric permittivity (in $[F/m]$).
$\mathbf{J}$ is the electric current density (in $[A/m^2]$)
$\mathbf{M}$ is the equivalent magnetic current density (in $[V/m^2]$)
ρ is the electric charge density (in $[C/m^3]$).

Note that all of the above fields, current density and flux variables have dependence on time (t). However, the time dependence is eliminated from the notation for convenience in the following text.

For linear, isotropic and nondispersive material, the fluxes and fields assume the following relationships:

$$\begin{aligned}\mathbf{D} &= \varepsilon\mathbf{E} \\ \mathbf{B} &= \mu\mathbf{H}\end{aligned} \tag{3.2}$$

By allowing materials with isotropic, nondispersive electric and magnetic losses that attenuate **E** and **H** fields via conversion to heat energy, we have:

$$\begin{aligned}\mathbf{J} &= \mathbf{J}_{source} + \sigma\mathbf{E} \\ \mathbf{M} &= \mathbf{M}_{source} + \sigma^*\mathbf{H}\end{aligned} \tag{3.3}$$

where σ is the electric conductivity (in $[S/m]$), and σ^* is the equivalent magnetic loss (in $[\Omega/m]$).

By applying the relations in equations 3.2 and 3.3 onto equations 3.1a and 3.1b, the Maxwell curl equations boil down to:

$$\frac{\partial\mathbf{H}}{\partial t} = -\frac{1}{\mu}(\nabla\times\mathbf{E}) - \frac{1}{\mu}(\mathbf{M}_{source} + \sigma^*\mathbf{H}) \tag{3.4a}$$

$$\frac{\partial\mathbf{E}}{\partial t} = \frac{1}{\epsilon}(\nabla\times\mathbf{H}) - \frac{1}{\epsilon}(\mathbf{J}_{source} + \sigma^*\mathbf{E}) \tag{3.4b}$$

By expanding the vector components of the right hand side of the above two equations, we have the following set of scalar equations:

$$\frac{\partial H_x}{\partial t} = -\frac{1}{\mu}[\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} - (M_{source_x} + \sigma^* H_x)] \tag{3.5a}$$

$$\frac{\partial H_y}{\partial t} = -\frac{1}{\mu}[\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} - (M_{source_y} + \sigma^* H_y)] \tag{3.5b}$$

$$\frac{\partial H_z}{\partial t} = -\frac{1}{\mu}[\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} - (M_{source_y} + \sigma^* H_z)] \tag{3.5c}$$

$$\frac{\partial E_x}{\partial t} = \frac{1}{\epsilon}[\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - (J_{source_x} + \sigma E_x)] \tag{3.5d}$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\epsilon}[\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} - (J_{source_y} + \sigma E_y)] \tag{3.5e}$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon}[\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - (J_{source_z} + \sigma E_z)] \tag{3.5f}$$

3.2 Finite element method

3.2.1 Helmholtz equation

The finite element method solves the decoupled Helmholtz equation in the frequency domain.

Consider a lossless, source free, linear, isotropic and non-dispersive medium, Eqn. 3.4b becomes:

$$\frac{1}{\varepsilon}(\nabla \times \mathbf{H}) = \frac{\partial \mathbf{E}}{\partial t} \tag{3.6}$$

By taking the curls of both sides of the above equation, we have:

$$\nabla \times (\frac{1}{\varepsilon}\nabla \times \mathbf{H}) = \nabla \times \frac{\partial \mathbf{E}}{\partial t} = \frac{\partial}{\partial t}(\nabla \times \mathbf{E}) \tag{3.7}$$

Substitute the curl of **E** in the above equation with Eqn. 3.1a and we have:

$$\begin{aligned} \nabla \times (\frac{1}{\varepsilon}\nabla \times \mathbf{H}) &= \frac{\partial}{\partial t}(-\mu\frac{\partial \mathbf{H}}{\partial t}) \\ &= -\mu\frac{\partial^2 \mathbf{H}}{\partial t^2} \end{aligned} \tag{3.8}$$

Similarly for the electric field, we have:

$$\nabla \times (-\frac{1}{\mu}\nabla \times \mathbf{E}) = \varepsilon\frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{3.9}$$

For dielectric materials which are of the main interest of this thesis, the permeability μ stays constant throughout the domain, and can thus be taken out from the first curl on the LHS of the above equation. Thus, Eqn. 3.9 can be rewritten as:

$$\nabla \times \nabla \times \mathbf{E} = -\mu\varepsilon\frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{3.10}$$

Now the electrical and magnetic fields are decoupled, unlike in the original Maxwell's equations. Let us assume that the fields have a harmonic dependence on time. Take **H** for example:

$$\mathbf{H} = \mathbf{H}_0 e^{-j\omega t}, \frac{\partial \mathbf{H}}{\partial t} = -j\omega\mathbf{H}_0 e^{-j\omega t} = -j\omega\mathbf{H}, \frac{\partial^2 \mathbf{H}}{\partial t^2} = (-j\omega)(-j\omega\mathbf{H}) = -\omega^2\mathbf{H} \tag{3.11}$$

where ω is the angular frequency (in $[rad/s]$).

By inserting Eqn. 3.11 into Eqn. 3.8 and canceling out the time-dependence terms from both sides, the equation for $\mathbf{H}$ field can eventually be written as:

$$\nabla \times (\frac{1}{\varepsilon}\nabla \times \mathbf{H}) = \mu\omega^2\mathbf{H} \tag{3.12}$$

Here $\mathbf{H}$ is shorthanded for $\mathbf{H}(\mathbf{x}, \mathbf{y}, \mathbf{z})$ where the dependence on time t is removed.

Consider the following vector calculus identity:

$$\nabla \times (\psi\mathbf{A}) = \psi\nabla \times \mathbf{A} + \nabla\psi \times \mathbf{A} \tag{3.13}$$

where ψ is a scalar field and $\mathbf{A}$ is a vector. By replacing ψ by $\frac{1}{\varepsilon}$ and $\mathbf{A}$ by $\nabla \times \mathbf{H}$, the above relation renders Eqn. 3.12 as follows:

$$\frac{1}{\varepsilon}\nabla \times (\nabla \times \mathbf{H}) + \nabla\frac{\mathbf{1}}{\varepsilon} \times (\nabla \times \mathbf{H}) = \mu\omega^{\mathbf{2}}\mathbf{H} \tag{3.14}$$

The triple vector product identity gives:

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C} \tag{3.15}$$

where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are all vectors. By using this identity on both the first and the second terms on the left hand side of Eqn. 3.14, it can be rewritten as below:

$$\frac{1}{\varepsilon}\left[\nabla(\nabla \cdot \mathbf{H}) - \nabla^{\mathbf{2}}\mathbf{H}\right] + \nabla(\nabla\frac{1}{\varepsilon} \cdot \mathbf{H}) - \nabla\frac{\mathbf{1}}{\varepsilon} \cdot \nabla\mathbf{H} = \mu\omega^{\mathbf{2}}\mathbf{H} \tag{3.16}$$

Now consider the 2D TE_z case where the structure is invariant in the z-direction and extends to infinity along the z-axis. The magnetic field is reduced to one non-zero component H_z and the permittivity (ϵ) has only dependence on the x and y axes. Since H_x and H_y are both 0 while H_z is invariant along the z-axis, the divergence of the $\mathbf{H}$ field ($\frac{\partial H_x}{\partial x} + \frac{\partial H_y}{\partial y} + \frac{\partial H_z}{\partial z}$ is 0, rendering the term $\nabla(\nabla \cdot \mathbf{H})$ null. Moreover, the gradient $\nabla\frac{1}{\varepsilon}$ has only components in the x and y directions and is thus orthogonal to the $\mathbf{H}$ field which has only z component. Hence the term $\nabla(\nabla\frac{1}{\varepsilon} \cdot \mathbf{H})$ is also 0. By removing the zero terms and replace the vectorial magnetic field $\mathbf{H}$ by its component H_z, Eqn. 3.16 can be rewritten for the 2D TM_z case as follows:

$$\frac{1}{\varepsilon}\nabla^2 H_z + \nabla\frac{1}{\varepsilon} \cdot \nabla H_z = -\mu\omega^2 H_z \tag{3.17}$$

It is easy to recognize that the left hand side of the above equation fits the right hand side of the vector calculus identity stated below by replacing ψ by $\frac{1}{\varepsilon}$ and $\mathbf{A}$ by ∇H_z:

$$\nabla \cdot (\psi\mathbf{A}) = \psi\nabla \cdot \mathbf{A} + \nabla\psi \cdot \mathbf{A} \tag{3.18}$$

By using the above relation, the 2D TE_z Helmholtz equation can finally be written in the compact form of:

$$\nabla \cdot (\frac{1}{\varepsilon} \nabla H_z) + \mu \omega^2 H_z = 0 \tag{3.19}$$

Using similar approaches, the 2D TM_z Helmholtz equation can also be derived. It results in the following form:

$$\nabla^2 E_z + \mu \varepsilon w^2 E_z = 0 \tag{3.20}$$

3.2.2 Discretization

In Jensen and Sigmund's FEM modeling of the 2D photonic crystal problem for their topology optimization technique [69], the computation domain is discretized into rectangular subcells. Each subcell contains a field unknown u_e, which are collected into the global field unknown vector $\mathbf{u}$. Edge elements are used in electromagnetic problems like ours in order to eliminate spurious modes [70]. The transverse field across a subcell (u_e) can be expressed as the superposition of the related edge elements weighted by basis functions. By using a weak form (integral form) of the governing equation and a standard Galerkin method for discretization, the problem results in a set of linear equations:

$$(-w^2 \mathbf{M} + iw\mathbf{C} + \mathbf{K})\mathbf{u} = \mathbf{f} \tag{3.21}$$

where $\mathbf{f}$ is the load term modeling the incident wave. Matrix $\mathbf{K}$ is the global stiffness matrix and matrix $\mathbf{M}$ is the global mass matrix. Both matrix are corresponding terms to the original Helmholtz equations of Eqn. 3.19 and Eqn. 3.20, and are assembled from the element matrices. The detailed formations of these matrices are presented in [69]. Matrix $\mathbf{C}$ incorporates the absorbing boundary conditions (ABCs) and material damping that are not present in the original governing equation. Perfectly matched layers (PML) are used as ABCs to decrease the reflections from the computational domain truncations. Artificial material damping is also used in Sigmund and Jensen's work in order to avoid resonance-based local maxima and grey elements.

The general Galerkin method as well as FEM techniques can be referred to in [71][55][56].

3.3 Finite-difference time-domain method

3.3.1 Maxwell's equations reduction to 2D and 1D

Assuming the structure extends to infinity in the z direction with uniform transverse cross section, the z-derivatives in Maxwell's equations Eqn. 3.5a can be removed, resulting in two sets of 2D equations each of which contains only three field components instead of six.

For transverse-magnetic mode with respect to z-axis (TM_z mode), the equations involve only H_x, H_y and E_z:

$$\frac{\partial H_x}{\partial t} = -\frac{1}{\mu}[\frac{\partial E_z}{\partial y} + M_{source_x} + \sigma^* H_x] \tag{3.22a}$$

$$\frac{\partial H_y}{\partial t} = -\frac{1}{\mu}[\frac{\partial E_z}{\partial x} - (M_{source_y} + \sigma^* H_y] \tag{3.22b}$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon}[\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - (J_{source_z} + \sigma E_z)] \tag{3.22c}$$

For transverse-electric mode with respect to z-axis (TE_z mode), the equations involve only E_x, E_y and H_z:

$$\frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon}[\frac{\partial H_z}{\partial y} - J_{source_x}] \tag{3.23a}$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon}[-\frac{\partial H_z}{\partial x} - J_{source_y}] \tag{3.23b}$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu}[\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} - M_{source_z}] \tag{3.23c}$$

For the 1D problem where the geometry has neither variations in y nor in z directions, derivatives with respect to either y or z are removed. Maxwell's equations becomes:

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu}[\frac{\partial E_z}{\partial x} - (M_{source_y} + \sigma^* H_y)] \tag{3.24a}$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon}[\frac{\partial H_y}{\partial x} - (J_{source_z} + \sigma E_z)] \tag{3.24b}$$

3.3.2 The Yee grid and the leap frog scheme

in 1966, Kane Yee introduced the Yee grid lattice [52] where the electric and magnetic fields are positioned half a grid size apart from the neighboring fields

(Fig. 3.1). The staggered manner of the field positions makes it natural to use the central-difference scheme to approximate the partial derivatives of the fields, and the results of this combination is a divergence free mesh in the absence of free electric and magnetic charge (see Chapter 3 in [54]).

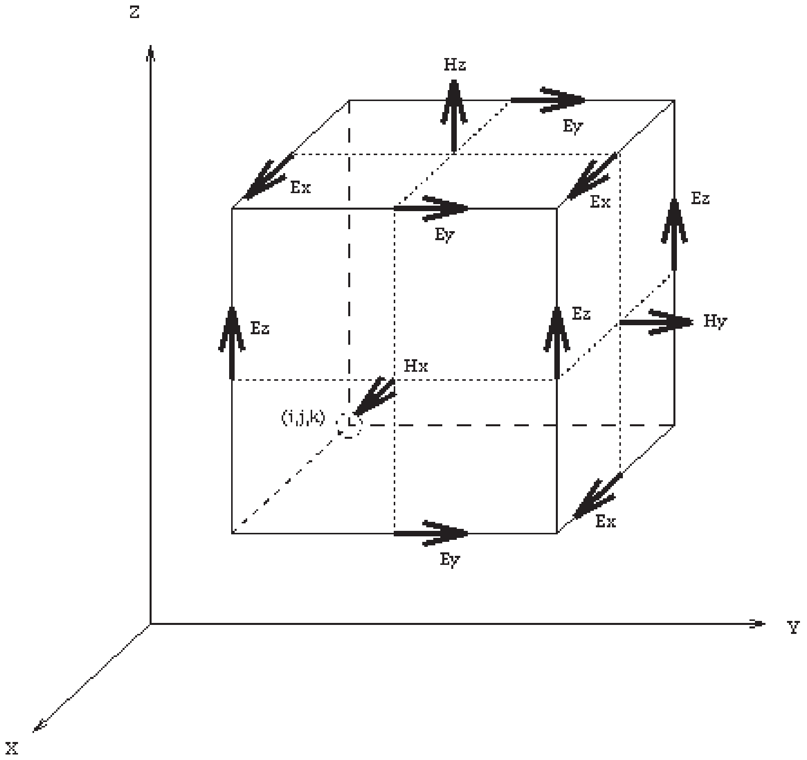


Figure 3.1: Electric and magnetic field positions on the 3D staggered Yee grid lattice.

The fields are then updated in the time domain using a leapfrog scheme (Fig. 3.2), where all the E fields are calculated by using the previously stored H fields data from half a time step ago, and vice versa for calculating the H fields. The FDTD update scheme for x-directed 1D case is illustrated in Fig. 3.2 where the locations of $\mathbf{H}$ and $\mathbf{E}$ in both time and space are staggered apart. The field component at time step $n\Delta t$ and grid point $i\Delta x$ are denoted as u_i^n, where u is either the electric- or magnetic- field and Δt, Δx are the time step size and grid spacing, respectively.

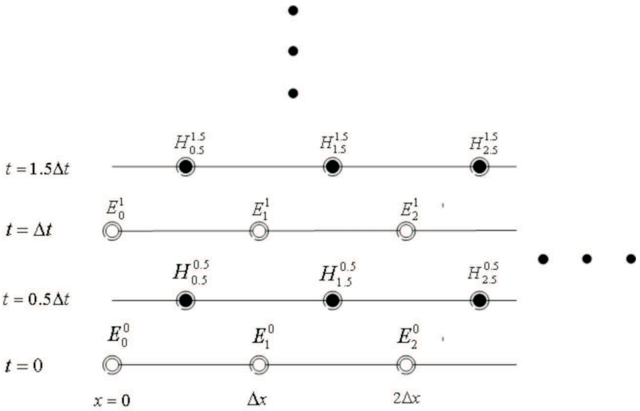

$H^{1.5}_{0.5}$
$H^{1.5}_{1.5}$
$H^{1.5}_{2.5}$
$t = 1.5\Delta t$
E^1_0
E^1_1
E^1_2
$t = \Delta t$
$H^{0.5}_{0.5}$
$H^{0.5}_{1.5}$
$H^{0.5}_{2.5}$
$t = 0.5\Delta t$
E^0_0
E^0_1
E^0_2
$t = 0$
$x = 0$
Δx
$2\Delta x$

Figure 3.2: 1D leap frog update scheme on Yee grid.

3.3.3 FDTD update equations

For 2D cases, assume a square lattice where the grid spacings in both directions are the same: $\Delta x = \Delta y = \Delta$. The locations of the three field components for TM_z mode are illustrated in Fig. 3.3. The E_z field is denoted as $E_z \big|^{n\Delta t}_{i\Delta,j\Delta}$; the H_x field is denoted as $H_x \Big|^{(n+\frac{1}{2})\Delta t}_{i\Delta,(j+\frac{1}{2})\Delta}$; and the H_y field is denoted as $H_y \Big|^{(n+\frac{1}{2})\Delta t}_{(i+\frac{1}{2})\Delta,j\Delta}$. For simplicity reasons, the increment symbols Δ and Δt are removed from the notation so the fields are shorthanded as: $E_z \big|^{n}_{i,j}$, $H_x \Big|^{n+\frac{1}{2}}_{i,j+\frac{1}{2}}$ and $H_y \Big|^{n+\frac{1}{2}}_{i+\frac{1}{2},j}$.

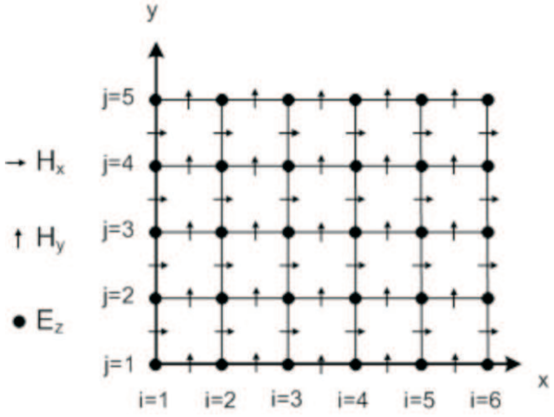

y
→ H_x
↑ H_y
● E_z
j=5
j=4
j=3
j=2
j=1
i=1
i=2
i=3
i=4
i=5
i=6
x

Figure 3.3: 2D Yee grid for TM_z mode.

Using the central difference approximation, the partial differential of a field at coordinates $x = i\Delta$, $y = j\Delta$ and time step n becomes:

$$\frac{\partial u}{\partial t}\Big|_{i,j}^{n} = \frac{u_{i,j}^{n+1/2} - u_{i,j}^{n-1/2}}{\Delta t} + O[(\Delta t)^2] \tag{3.25}$$

By inserting the central difference expression into Eqn. 3.22a, we have:

$$\frac{H_x\Big|_{i,j+1/2}^{n+1/2} - H_x\Big|_{i,j+1/2}^{n-1/2}}{\Delta t} = -\frac{\Delta}{\mu_{i,j+1/2}}\Bigg(\frac{E_z\big|_{i,j+1}^{n} - E_z\big|_{i,j}^{n}}{\Delta} + M_{source_x}\Big|_{i,j+1/2}^{n+1/2} + {\sigma^*}_{i,j+1/2}H_x\Big|_{i,j+1/2}^{n}\Bigg) \tag{3.26}$$

Since H_x field is only saved at half integer time steps ($0.5\Delta t$, $1.5\Delta t$, etc.), $H_x\Big|_{i,j+1/2}^{n}$ is not readily available. By using a semi-implicit approximation (see Chapter 3 in [54]), the value for the integer time step H_x can be deemed as:

$$H_x\Big|_{i,j+1/2}^{n} = \frac{H_x\Big|_{i,j+1/2}^{n+1/2} + H_x\Big|_{i,j+1/2}^{n-1/2}}{2} \tag{3.27}$$

Substitute Eqn. 3.27 into Eqn. 3.26 and rearrange the equation, the value for $H_x\Big|_{i,j+1/2}^{n+1/2}$ can be derived as:

$$\begin{aligned} H_x\Big|_{i,j+1/2}^{n+1/2} = &\left(\frac{2\mu_{i,j+1/2} - {\sigma^*}_{i,j+1/2}\Delta t}{2\mu_{i,j+1/2} + {\sigma^*}_{i,j+1/2}\Delta t}\right) H_x\Big|_{i,j+1/2}^{n-1/2} \\ &- \frac{2\Delta t}{2\mu_{i,j+1/2} + {\sigma^*}_{i,j+1/2}\Delta t}\Bigg(\frac{E_z\big|_{i,j+1}^{n} - E_z\big|_{i,j}^{n}}{\Delta} \\ &+ M_{source_x}\Big|_{i,j+1/2}^{n+1/2}\Bigg) \end{aligned} \tag{3.28a}$$

Similarly, equations 3.22b and 3.22c can be treated the same way and rewritten as:

$$\begin{aligned} H_y\Big|_{i+1/2,j}^{n+1/2} = &\left(\frac{2\mu_{i+1/2,j} - {\sigma^*}_{i+1/2,j}\Delta t}{2\mu_{i+1/2,j} + {\sigma^*}_{i+1/2,j}\Delta t}\right) H_y\Big|_{i+1/2,j}^{n-1/2} \\ &+ \frac{2\Delta t}{2\mu_{i+1/2,j} + {\sigma^*}_{i+1/2,j}\Delta t}\Bigg(\frac{E_z\big|_{i+1,j}^{n} - E_z\big|_{i,j}^{n}}{\Delta} \\ &- M_{source_y}\Big|_{i+1/2,j}^{n+\frac{1}{2}}\Bigg) \end{aligned} \tag{3.28b}$$

$$\begin{aligned}E_z\Big|_{i,j}^{n+1} &= \left(\frac{2\varepsilon_{i+1/2,j}-\sigma_{i,j}\Delta t}{2\varepsilon_{i+1/2,j}+\sigma_{i,j}\Delta t}\right)E_z\Big|_{i,j}^{n}\\ &+\frac{2\Delta t}{2\varepsilon_{i,j}+\sigma_{i,j}\Delta t}\Bigg(\frac{H_y\Big|_{i+1/2,j}^{n+1/2}-H_y\Big|_{i-1/2,j}^{n+1/2}}{\Delta}\\ &-\frac{H_x\Big|_{i,j+1/2}^{n+1/2}-H_x\Big|_{i,j-1/2}^{n+1/2}}{\Delta}\\ &-J_{source_z}\Big|_{i,j}^{n+1}\Bigg) \end{aligned} \tag{3.28c}$$

For 1D cases (equations 3.24a and 3.24b), the following update equations can be written by using the grid and time locations shown in Fig. 3.2:

$$\begin{aligned}H_y\Big|_{i+1/2}^{n+1/2} &= \left(\frac{2\mu_{i+1/2}-{\sigma^*}_{i+1/2}\Delta t}{2\mu_{i+1/2}+{\sigma^*}_{i+1/2}\Delta t}\right)H_y\Big|_{i+\frac{1}{2}}^{n-\frac{1}{2}}\\ &+\frac{2\Delta t}{2\mu_{i+\frac{1}{2}}+{\sigma^*}_{i+1/2}\Delta t}\Bigg(\frac{E_z\big|_{i+1}^{n}-E_z\big|_{i}^{n}}{\Delta}\\ &-M_{source_y}\Big|_{i+\frac{1}{2}}^{n+\frac{1}{2}}\Bigg)\end{aligned} \tag{3.29a}$$

$$\begin{aligned}E_z\Big|_{i}^{n+1} &= \Bigg(\frac{2\varepsilon_i-\sigma_i\Delta t}{2\varepsilon_i+\sigma_i\Delta t})E_z\big|_{i}^{n}\\ &+\frac{2\Delta t}{2\varepsilon_i+\sigma_i\Delta t}\Bigg(\frac{H_y\big|_{i+1}^{n}-H_y\big|_{i}^{n}}{\Delta}\\ &-J_{source_y}\Big|_{i}^{n+1}\Bigg)\end{aligned} \tag{3.29b}$$

3.3.4 Stability criteria

Explicit updating scheme is used in FDTD, rendering the method conditionally stable. The maximum time-step allowed in FDTD is inversely proportional to the minimum grid step size among all directions.

A Courant stability bound is established as follows:

$$\xi = c\vartriangle t\sqrt{\frac{1}{(\vartriangle x)^2}+\frac{1}{(\vartriangle y)^2}+\frac{1}{(\vartriangle z)^2}} \leq 1 \tag{3.30}$$

where ξ is defined as the Courant number or stability factor, and $\vartriangle x$, $\vartriangle y$ and $\vartriangle z$ are the grid spacings in the three dimensions respectively. A Courant number higher than 1 would cause the field to grow exponentially (proved in Chapter 4, [54]). Hence an upperbound of the time step size is easily determined once the grid is set.

For a highly intricate optical layout where the minimum grid step size is bounded by the lengthscale of the minimum geometrical feature, the time-step size becomes small, resulting in a large number of total time steps needed. Weak or non-conditionally stable methods exist, e.g. the alternating-direction-implicit (ADI) method [72]. Implicit updating is used in ADI where the time step size is no longer bounded by the Courant stability criteria. However, a high Courant number, while alleviating the computational cost of FDTD, introduceds large dispersion and truncation errors. Moreover, ADI requires to solve tridiagonal matrices during each time step, which makes it less efficient compared to the matrix-free operation of the explicit updating scheme.

3.3.5 Absorbing boundary conditions

While it is necessary to truncate the computation domain, the outer lattice boundary must simulate the extension to infinity in order to study unbounded regions. Hence, creating artificial absorbing boundary conditions (ABCs) where incident waves are absorbed instead of reflected back into the calculation domain becomes crucial in computational electromagnetics. Effective ABCs should be able to absorb incident waves within a large bandwidth, with little reflection, disregarding the incident angles.

In 1994, Berenger introduced the perfectly matched layers [73] where plane waves of arbitrary incidence, polarization and frequency are matched at the boundary. In this 2D formulation, the magnetic field component H_z in a TE_z plan wave impinging on the boundaries is split into two orthogonal waves, H_{zx} and H_{zy}. These two field components, together with E_x and E_y, continue to propagate into the PML slab after exiting the physical domain. By configuring the electric conductivities (σ_x and σ_y) and the magnetic conductivities (σ^*), the impedances of both sides of the boundary can be matched for each of these field component, making the boundary reflectionless. 3D PML was later developed by Katz [74].

Although PMLs have theoretically zero reflections for Maxwell's equations, spurious reflections occur due to the discretizations in the actual implementation of FDTD. Consider an x-directed wave impinging normally upon a PML slab, a polynomial grading can be introduced to gradually increase the PML losses along the x axis:

$$\sigma_x(x) = (x/d)^m \sigma_{x,max} \tag{3.31}$$

Here x is the distance to the boundary and d is the thickness of total the PML layers. A polynomial constant $3 \leq m \leq 4$ is usually used. Such a distribution of σ_x results in a low absorption rate at the beginning of PML, but the loses quickly grow deeper inside the layers.

Numerical experiments show that PML exhibits an excellent absorbing performance in 2D, with a global error 7 orders of magnitude smaller than earlier ABCs like Mur ABC (see Chapter 7 in [54]).

Uniaxial PML (UPML) is also developed where an anistropic absorbing medium is configured in the PML slab to absorb the fields propagating along both directions [75][76]. UMPL is shown to be akin to the original split-field PML in effectiveness, and since no field splitting is needed is needed, hence improving the computational efficiency of PML layers.

3.4 Conclusions

We introduced the underlying modeling methods used (FEM and FDTD) for the topology optimization cases presented in this thesis. The general perspectives of the two methods as well as the details regarding the Yee grid, stability criteria and boundary conditions for FDTD implementations are addressed. Although the structured and uniform grid necessary in FDTD presents a challenge in modeling extremely complex geometries, the method still offers great benefits in efficient computing and the capacity of massive parallelization, compared to matrix-based methods. Moreover, as a time-domain method, FDTD provides the natural ability to incorporate nonlinearity modeling without much difficulty. Though nonlinearities are not covered in this thesis, the potential of our method to extend to such regimes would certainly be interesting and inspiring future works.

Chapter 4

Frequency-domain topology optimization

In this chapter, we review the rationale as well as some design examples of the frequency-domain TO. Based on the time-harmonic two-dimensional finite element (FE) modelling of the photonic devices, the optimization redistributes the two-phase materials in the design domain. The methodology was first published in 2004 by Jensen and Sigmund for the design of a transmission-efficient 90° bend in a two-dimensional photonic crystal waveguide (PhCW) [19]. The fabrication and characterization of a TO-designed Z-bend PhCW was carried out by Borel et. al [20], which proved the efficacy of the design method. More TO design examples as well as their materializations on the silicon-on-insulator (SOI) material platform appeared in the next few years, e.g. low-loss T-junction waveguide [20], 60° PhCW bend [21][22], double 90° PhCW bends [23] and PhCW-based Y-splitters [24]. In this chapter, we focus on applying frequency-domain TO to the design of slow light couplers for photonic crystal waveguides based on both normal round holes as well as ring-shaped holes in a two-dimensional photonic crystal structure. Both devices were fabricated and large improvements in transmissions are seen in the slow-light region of the transmitted light. The optimization code was designed and written by Jakob S. Jensen in collaboration with Ole Sigmund at MEK, DTU [19].

The results presented in this chapter are published in [77] and [78].

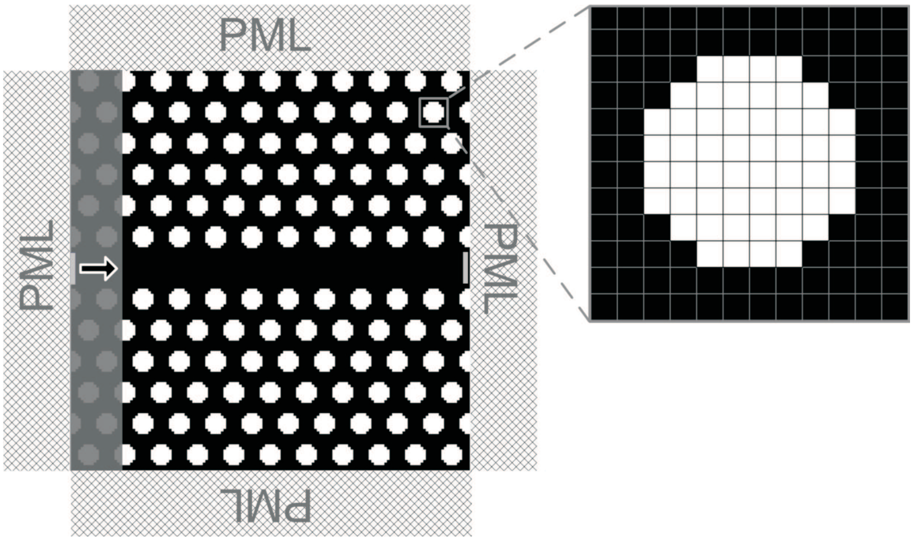


Figure 4.1: The design problem for PhCW-based slow light couplers. The physical domain is a photonic crystal waveguide in a triangular lattice with air holes. It is cladded by perfectly match layers (PML) as absorbing boundary conditions. The wave input is excited at the entrance of the PhCW (as indicated by the arrow) and the output is measured at the exit of the waveguide. The design domain is illustrated as the grey stripe near the entrance of the waveguide, where the slow light modes are being reflected or coupled in, depending on the local geometrical structure.

4.1 Rationale

To design a PhCW-based slow light coupler, the problem is formulated as shown in Fig. 4.1.

From the modeling perspective, let us review the discussions in section 3.2. The governing equation of the E-polarized wave propagation in the form of Helmholtz equation is as follows:

$$\nabla^2 E + \mu\epsilon w^2 E = 0 \tag{4.1}$$

The equation is then implemented by using the finite element (FE) method based on square elements. By assembling the frequency-dependent element matrices into a system matrix $S(\omega)$, we now have a set of linear complex equations:

$$(-w^2\mathbf{u} + iw\mathbf{C} + \mathbf{K})\mathbf{u} = \mathbf{f} \tag{4.2}$$

Here $\mathbf{u}$ is the vector containing the nodal values of field E. Matrix $\mathbf{C}$ accounts for absorbing boundary conditions and artificial material damping which is used to improve smoothness of the optimization problem. Each of the systems of linear equations solves for one time-harmonic wave propagation problem with a specific frequency.

In order to optimize for the higher (or lower) transmission through the waveguide, the time-averaged Poynting vector ($\mathbf{p}$) flowing through the area A is computed by the following equation:

$$\mathbf{p} = \{p_x p_y\}^T = \frac{\omega}{2a} \int_A \mathbb{R}(i(\nabla E)E^*)dA, \tag{4.3}$$

where a is the lattice constant and E^* is the complex conjugate field.

In the following example we show how to formulate an optimization problem when the optimization goal is to maximize the y component of the time-averaged Poying vector p_y in the cell A for a number M of target frequencies $\bar{\omega}_j, j = 1, M$. The optimization objective and bounds can be formulated as:

$$\begin{aligned} &\max_{0 \leq \boldsymbol{\rho} \leq 1} \; C = \sum_{j=1}^{M} p_y(\mathbf{u}_j) \\ &\textit{subject to}: ((-w_j)^2\mathbf{M} + iw_j\mathbf{C} + \mathbf{K})\mathbf{u} = \mathbf{f}(w_j), \;\; j = 1, M, \end{aligned} \tag{4.4}$$

where $\boldsymbol{\rho}$ is the design variable set.

4.2 Design and fabrication of slow light couplers

In this section, two TO design examples are shown to enhance the slow light coupling efficiencies for two different kinds of photonic crystal waveguides.

Small group velocities of light resulting from flat dispersion curves in PhCWs near the cut-off has become an interesting topic in recent years. This is largely due to the fact that the slowed-down light makes PhCWs potential candidates for important applications such as delay lines and optical storage. However, the mismatch of impedances between the PhCW slow light mode and the ridge waveguide mode creates a difficult situation for the light to be coupled in from the ridge waveguide to PhCW, and vice versa. This prevents PhCWs to be efficiently used as slow-light devices in all-optical circuits.

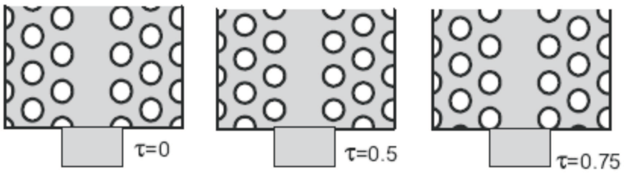


Figure 4.2: Definition of the termination parameter τ.

Vlasov and McNab [79] demonstrated different coupling efficiencies by varying the lattice terminations at the strip/PhC interface. A termination parameter τ was defined by how much the lattice was shifted at the interface (see Fig. 4.2). For their PhC configuration (triangular lattice with radius-pitch ratio (R/a) equal to 0.25), they predicted the best coupling efficiency should occur when $\tau = 0.3$ *or* 0.8, since the photonic surface states originated by the crystal lattice termination are tuned in resonance with the PhCW slow-light mode. The study established a connection between the surface states induced by lattice terminations and the enhancement of the slow light coupling efficiencies. A recipe for improving such coupling efficiencies was proposed by evaluating surface mode from various lattice terminations to find one termination that has the surface mode most in tune with the guided mode of PhCW. However, such an approach can be tedious and inefficient.

In order to test the recipe, we computed the band structures for the W1 PhCW slow light mode as well as surface modes from 8 different termination parameters (see Fig. 4.3). For ease of observation, we only plotted the surface modes that are close to the W1 PhC mode (dotted black) and inside the band gap. No modes higher than normalized frequency $0.26c/a$ or lower than W1 PhC's 11th band (which is the lower bound of the bandgap) are plotted. We notice that the tuned-in termination parameter drifts away from $\tau = 0.3$ *and* 0.8 as the configuration for the lattice changed. Moreover, there is no quantifiable relationship between the surface state frequencies and the termination parameter. This means that many random termination parameters might need to be tested before a good match can be found, which makes it difficult to manually searching for the parameter. Thus, the development of a more general method of manipulating the coupler geometry is of interest. Frequency-domain TO is a good candidate here to find an optimized coupling for a specific lattice configuration, which has no matching surface states to the W1 PhC slow light mode.

4.2.1 PhCW with round holes

The optimization was carried out on PhCWs defined by a line-defect in a triangular photonic crystal lattice of air holes in silicon with pitch (Λ) equal to

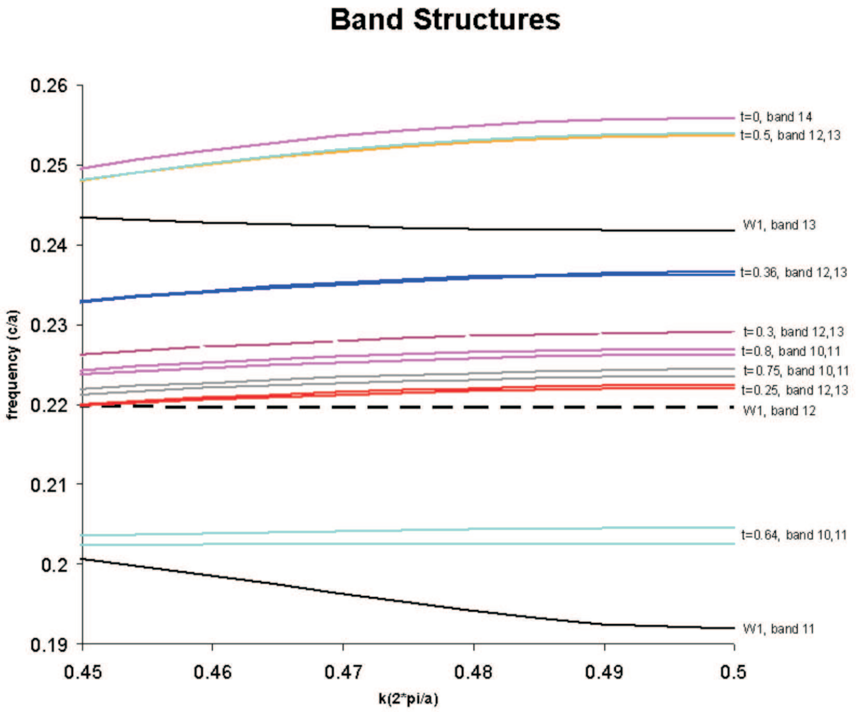


Figure 4.3: Band structures for W1 PhCW mode (dotted black) and surface modes with different termination parameters τ (solid).

$400nm$ and the hole diameter (d) around $260nm$. Fig. 4.4.(a) and Fig. 4.4.(c) illustrate the two different initial configurations at which the optimization begins. Structure (a) has the lattice termination parameter $\tau = 0.5$ if $\tau = 0$ results in a lattice termination cutting through the center of the first row of holes. In structure (c), the termination is shifted by $\Lambda/7$ along the PhCW and thus has $\tau = 0.64$. The design domain, where the dielectric material can be freely redistributed, is set to be a Λ-wide stripe area centered at the original cutting and covering 16 rows of holes in the $\Gamma - M$ direction of the crystal. The target function is to optimize for higher transmission at three frequencies in the slow light regime. The two resulting optimized structures are shown in Fig. 4.4.(b) and Fig. 4.4.(d), respectively for structure (a) and (c).

The structures were fabricated and characterized by Lars Hagedorn Frandsen and Amélie Têtu using e-beam lithography (JEOL-JBX9300FS) and inductively-coupled plasma reactive-ion etching to define the PhCW structure into the $320nm$ top silicon layer of a silicon-on-insulator wafer. The fabricated PhCWs are $12\mu m$ long and connected to tapered ridge waveguides to route light to and from the sample facets.

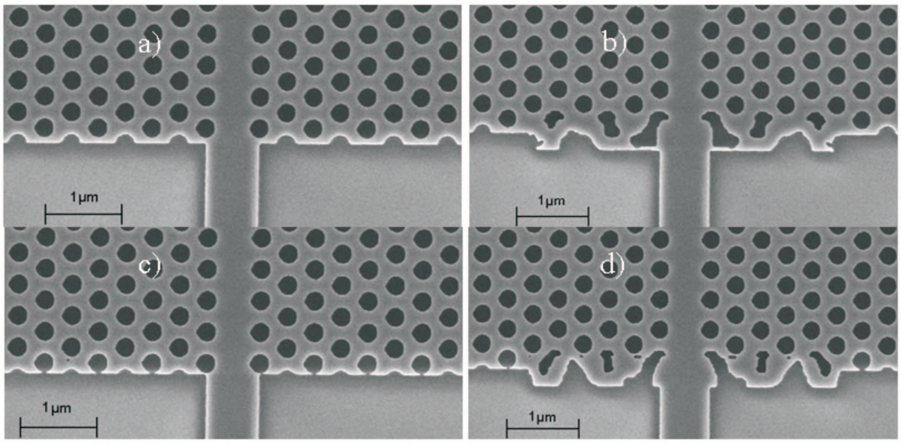


Figure 4.4: Scanning electron micrographs of the un-optimized (a and c) and optimized (b and d) structures: a) $\tau = 0.5$ un-optimized, b) $\tau = 0.5$ optimized, c) $\tau = 0.64$ un-optimized, and d) $\tau = 0.64$ optimized.

Figure 4.5 shows the experimental measurements of the fabricated structures shown in Fig. 4.4. The inset of the figure shows a zoom-in on the slow light regime. As expected from the Finite Difference Time Domain (FDTD) calculations, the spectrum for the un-optimized structure with $\tau = 0.64$ (dotted gray) shows a higher coupling efficiency near the band-edge than that of the structure with $\tau = 0.5$ (dotted black).

Also shown in the figure is the measured transmission for optimized structures starting from waveguides with $\tau = 0.5$ (solid black) and $\tau = 0.64$ (solid gray) terminations. It is clearly seen that the coupling efficiencies for slow light have been improved by $5dB$ and $2dB$ in structures with termination parameters $\tau = 0.5$ and $\tau = 0.64$, respectively, resulting in an improved performance of the PhCW near the band-edge. It is important to notice that the optimizations have converged to approximately the same transmission level in the slow light regime, disregarding the initial configurations proofing the robustness of the method. The optimized waveguide with termination $\tau = 0.5$ is especially noticeable with its unique high transmission in the last 1nm before the cut-off.

The design and characterization show that the topology optimization method can be used to improve the coupling of slow light in and out of PhCWs. The optimized structures show better performance in transmission near the bandgap, thus providing the waveguide with wider and smoother transmission bandwidth.

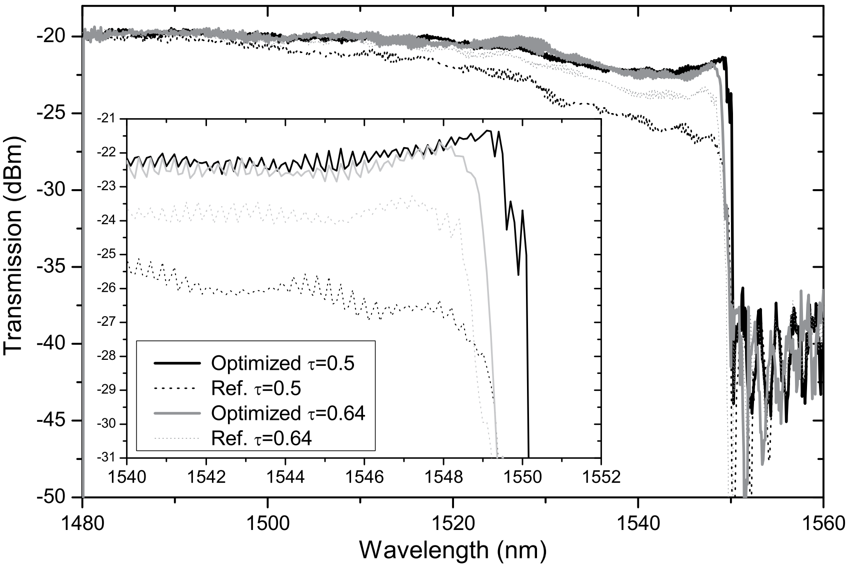


Figure 4.5: Experimental spectra of the optimized waveguides (solid) and their reference (dotted) waveguides for PhCW with round holes.

4.2.2 PhCW with ring-shaped holes

Photonic crystals with ring-shaped holes are dielectric rods each circulated by a round air hole setting in a square or triangular array. It was proposed by Kurt *et al.* [80] and has been shown to have a larger band gap. Here we study the optimization of slow light couplers to PhCW with these ring-shaped holes (RPhCW).

The optimization was performed on an RPhCW consisting of a hexagonal lattice (pitch Λ=405nm) of ring-shaped holes which are defined by their outer ($Ro = 152nm$) and inner ($Ri = 76nm$) radii. The optimization aims at maximizing the wave output of the RPhCW, as illustrated in Fig. 4.6(a). The design domains are Λ-wide strip and 2Λ-wide strip respectively for the two design examples shown in Fig. 4.6(b) and (d). The optimized structures are shown in Fig. 4.6(c) and (e)

Figure 4.7 shows the 2D simulated transmission spectra for the two optimized structures in Figs. 4.6(c) and (e). The calculation is done in the commercial software *Crystal Wave*. To examine the optimized design's tolerance to ring size fluctuations present due to fabrication errors, we performed calculations for the optimized structure with different ring width ($G = R_o - R_i$), $G = 0.16\Lambda$,

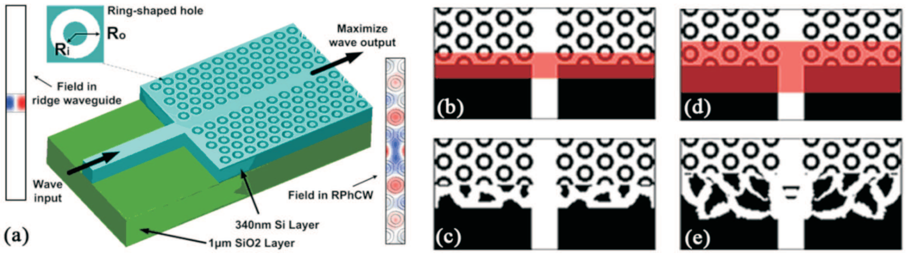


Figure 4.6: (a) Schematic diagram of the RPhCW, (b,d) Un-optimized coupling region with different design domains (red shadows) for topology optimization, (c,e) Optimized slow-light couplers for the different design domains, respectively.

$G = 0.20\Lambda$ and $G = 0.22\Lambda$, while $G = 0.20\Lambda$ is the original ring specification. The bands move to longer wavelengths with decreasing ring width, but in all 3 configurations the transmittances of the TO-optimized RPhCWs are dramatically improved within a $50nm$ bandwidth next to the cut-off. The inset of Fig. 4.7(b) illustrates the coupling improvement in this band width. The maximum coupling enhancement decreases only from 2.4dB to 2.1dB when changing the ring width by 0.06Λ, indicating that the optimized slow-light coupler is quite robust and can tolerate small variations of the ring width.

Both TO-designed couplers were fabricated and characterized by Ph.D. student Minhao Pu. The samples were fabricated using e-beam lithography (JEOL-JBX9300FS) and inductively-coupled plasma reactive-ion etching to define the RPhCW structures into the $340nm$ top silicon layer of a silicon-on-insulator wafer (see Fig. 4.8). The fabricated RPhCWs are $5\mu m$ long and connected to tapered ridge waveguides to route light to and from the sample facets. An RPhCW with standard coupling region was also fabricated as a reference (see Fig. 4.8(a)). For each of those structures, the in-coupling region was fabricated based on the TO designs in Fig. 4.6(c) and (e) while the out-coupling region was the exact mirror reflection of the in-coupler. Light transmission experiments were performed, where light from a tunable laser source ($1520 - 1620nm$) was launched into the sample and collected by an optical spectral analyzer to record the transmission spectrum. The polarization of the input light was adjusted to the quasi-TE mode with a fiber polarization controller.

The measured transmission spectra are shown in Fig. 4.9. It is clear that both optimized structures have higher coupling efficiencies near the band cut-off than the reference structure as we expected. Also shown in the figures are the extracted coupling improvements for the two couplers (see insets in Fig. 4.9). An enhancement in the combined in- and out-coupling of up to $5dB$ is observed in the slow light regime close to the band cut-off of $1600nm$ for both designs. Thus, the experimental results confirmed the predicted improved coupling performance of 2.5dB per coupling of the topology-optimized slow-light interfaces.

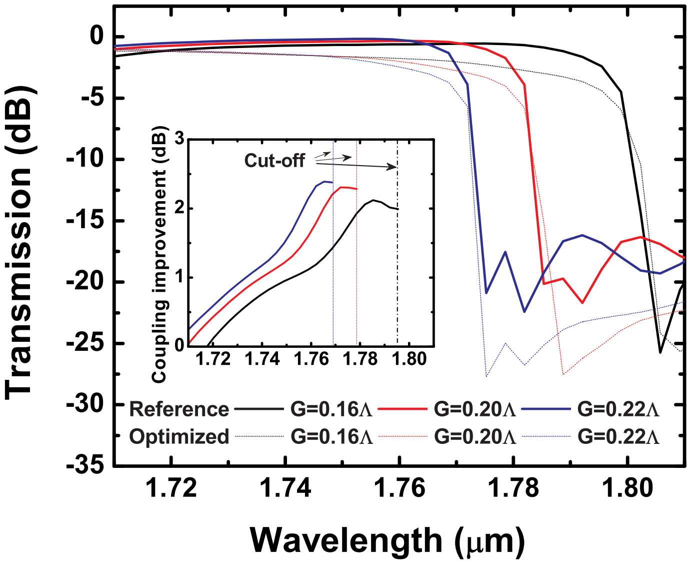


Figure 4.7: 2D simulated transmission spectra for the RPhCW with different ring-gap widths, insets are the calculated coupling improvement for the optimized coupler. Figure courtsey of Minhao Pu.

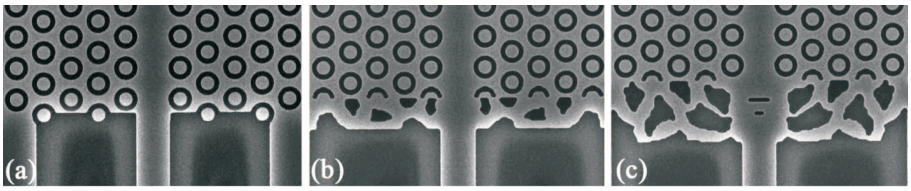


Figure 4.8: Scanning electron micrographs of the un-optimized (a) and two optimized (b,c) structures. Figure courtsey of Minhao Pu.

4.3 Conclusions

In this chapter, we investigated the utilization of FEM-based topology optimization in designing slow light couplers for two different configurations of photonic crystal waveguides. For both structures, TO has shown its capability of improving the slow light coupling efficiency by several dBs, disregarding the initial geometries of the design domain. Moreover, the ring sizes in the RPhCW are perturbed by 0.04 Λ and 0.02 Λ, in both of which cases the optimized structures still see a considerable enhancement of slow light coupling

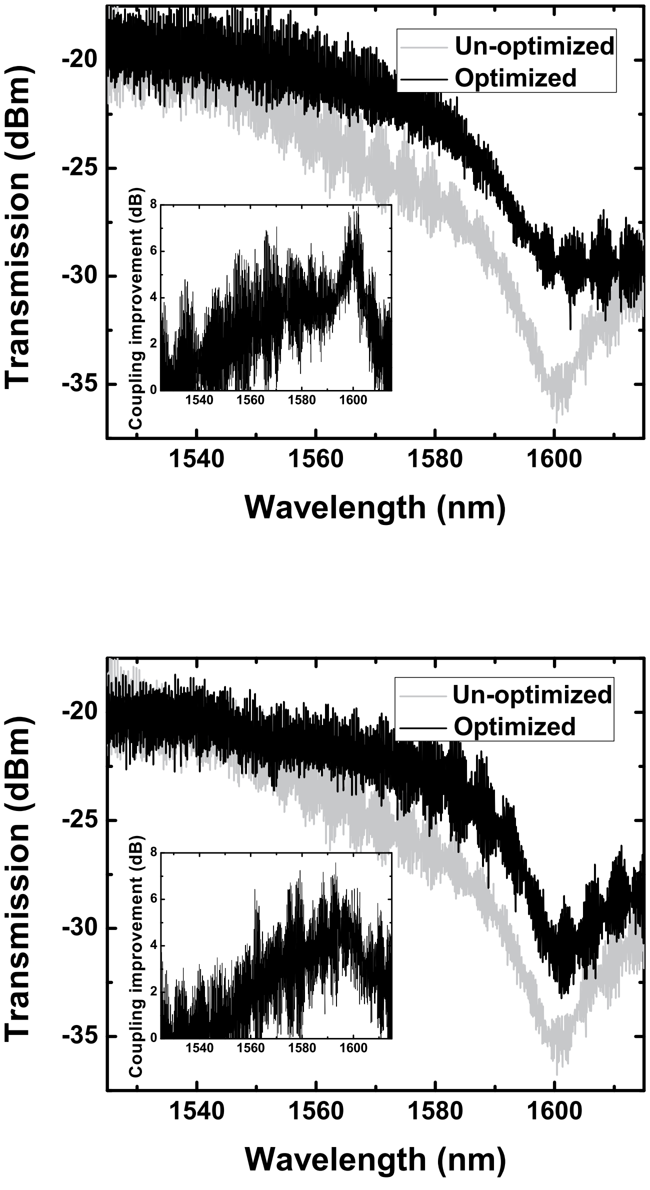


Figure 4.9: Measured transmission spectra for the two optimized couplers in RPhCW. (*a*) Λ-wide design domain, (*b*) 2Λ-wide design domain. Insets are the extracted coupling improvements for the two couplers. The optimized spectra (black) see a clear enhancement of transmission in the slow light regime, compared to the un-optimized spectra (grey). Figure courtsey of Minhao Pu.

efficiency. This proves that the optimized structures are quite robust and are tolerant of small variations from the original structure, an effect inevitable due to fabrication tolerances.

Together with previous design examples in the literature, the frequency-domain TO has proven to be a powerful tool in tuning the device performance when the target frequencies are few. However, in design examples where more frequency samplings need to be evaluated and adjusted, frequency-domain methods face with a computational challenge of solving linear systems for all these frequencies. Time-domain methods, on the other hand, has the capability of solving a wide range of spectrum with one analysis. For the rest of the thesis, we will explore the feasibility and potential of using time-domain topology optimization in designing photonic devices.

CHAPTER 5

1D time-domain topology optimization

In this chapter, we discuss 1D topology optimization based on finite-difference time-domain method. A Bragg grating design case is presented as a benchmark where FDTD-based TO is tested to see whether its solution matches that of theoretical prediction. A more complicated design problem is explored where a grating structure is optimized to be able to transform a Gaussian pulse into a square pulse. Penalization methods are used here to ensure that the final design is strictly black and white while having minimum feature size under control.

Part of the results presented in this chapter is published in [81]

5.1 Problem formulation

The calculation domain is uniformly discretized into plates layered perpendicular to the propagation direction of the light, and is terminated using Mur's absorptive boundary conditions. The material of each plate is determined by design variable ρ_i through linear interpolation:

$$\varepsilon_r^i = \varepsilon_{r_1} + (\varepsilon_{r_2} - \varepsilon_{r_1})\rho_i, \quad (0 \leq \rho_i \leq 1). \tag{5.1}$$

In the above equation ε_r^i is the relative dielectric constant of the i^{th} plate in the design domain, and ε_{r_1} and ε_{r_2} are the relative dielectric constants for the two

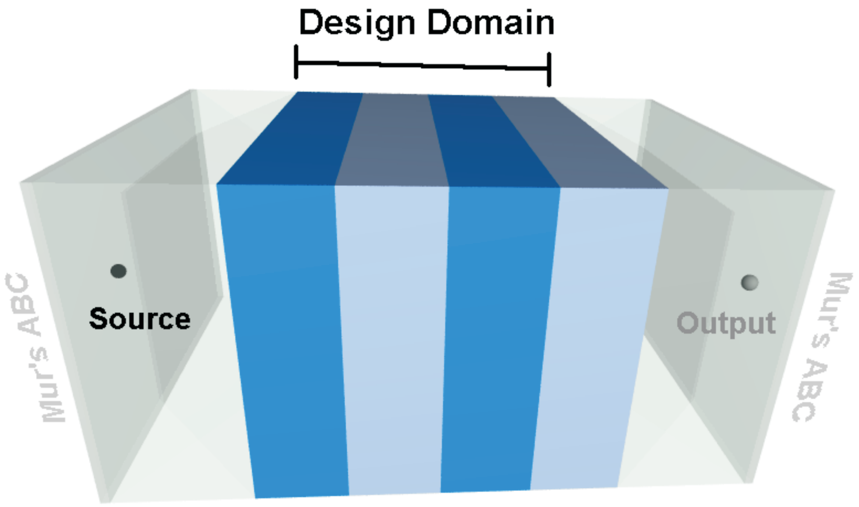


Figure 5.1: 1D grating optimization. The input pulse is excited at the source node (black) and the output pulse at the objective node (grey) is measured and evaluated.

design materials. The goal of the 1D topology optimization is to distribute two dielectric materials (ϵ_r^1 and ϵ_r^2) within the design domain (Fig.5.1) to make a grating structure that fulfills certain transmission functions between the source node and the output node. For example, one can minimize the transmission at certain frequencies to design a special filter, or to alter the temporal shape of the output pulse. Depending on the optimization goal, a specific objective function $F(\rho)$ is prescribed. The gratings are modeled by 1D FDTD method with Mur's ABC terminating the calculation domain.

5.2 Sensitivity analysis

The element-level sensitivities, which are the gradients of the objective function with respect to each of the design variables, are key to the redistribution of the design materials. Here we use the adjoint-variable method [82][83] to efficiently retrieve the sensitivity information using only two system analyses. Only two system analyses are needed for extracting the exact sensitivities for all design variables. The first is the forward analysis yielding the response of the system while the extra adjoint analysis is carried out by swopping the input and output nodes and feeding the system with the adjoint current which depends on the field values from the forward analysis.

The detailed derivation of the sensitivity analysis based on FDTD is presented in Appendix A.1.

5.3 Proof of concept

It is well known that theoretically, Bragg Gratings (BG) are the best structures to minimize the transmission at a specific frequency. We try to use TO aiming at minimizing the transmissions of waves at a single frequency, and expect the resulting optimized structure to be an exact BG with the correct period.

A $1GHz$ harmonic wave was used as the excitation. For BGs with two alternating materials ($\varepsilon_1 = 1$ and $\varepsilon_2 = 2.25$), the period of gratings (Λ) is calculated as below:

$$
\begin{aligned}
L_1 &= \frac{\lambda}{4n_1}, \\
L_2 &= \frac{\lambda}{4n_2}, \\
\Lambda &= L_1 + L_2,
\end{aligned}
\tag{5.2}
$$

where λ is the wavelength of the harmonic wave excitation and n_1 and n_2 are the refractive indices of the two design materials. L_1 and L_2 are the layer thicknesses of material 1 and 2, respectively. BGs with the above layer thicknesses and period can then be constructed to minimize the transmission at the frequency of $1GHz$.

In our problem formulation, the optimization is set to minimize the field energy transmitted through the structure by optimizing design variables $\boldsymbol{\rho}$:

$$
\begin{aligned}
&\min_{0 \le \boldsymbol{\rho} \le 1} \int_T E_o(\boldsymbol{\rho})^2 dt, \\
&s.t. : Maxwell's\ Equations,
\end{aligned}
\tag{5.3}
$$

The topology optimized grating (blue) is seen in Fig. 5.2 in comparison with that of a theoretical 5-layers BGs (red). The resemblance between the two layouts as well as their spectra is clear. Thus we confirm that the optimization reached its theoretical minimum.

A more challenging example was tested to optimize a structure that has multiple transmission channels and cutoff channels. In this specific case, we want 3 transmission dips at $0.8GHz$, $1.0GHz$ and $1.2GHz$, with the two designing materials $\varepsilon_1 = 1$ and $\varepsilon_2 = 2.25$. The objective of the optimization is to not only minimize the transmissions at the 3 dip frequencies, but also to maximize

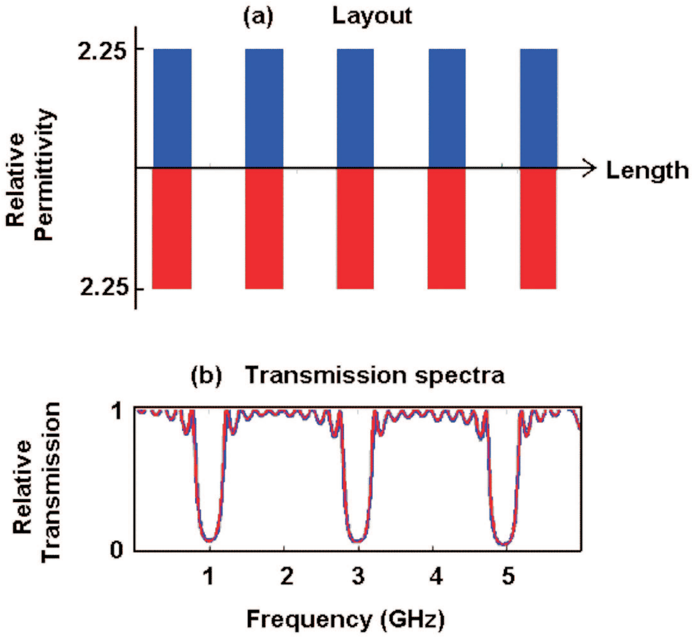


Figure 5.2: Bragg gratings optimization. (a) Layouts of the optimized gratings (blue) and the theoretical Bragg gratings (red). (b) Tranmission spectra of the optimized gratings (blue) and theoretical Bragg gratings (red).

the transmissions at the 2 peak frequencies in between the dips. Thus 2 modellings were needed with different excitations: the first with a superposition of 3 equal-weighted harmonic waves of the dip frequencies; the second with a similar superposition of the peak frequencies. The objective function is thus formulated as to maximize the time-integrated energy transmitted between the peak channels and the dip channels:

$$\begin{aligned} &\min_{0\leq\boldsymbol{\rho}\leq 1} \int_T [E_o^{dip}(\boldsymbol{\rho})^2 - E_o^{peak}(\boldsymbol{\rho})^2]dt, \\ &s.t.: Maxwell's\ Equations, \end{aligned} \tag{5.4}$$

The optimized layout and the corresponding transmission spectrum is depicted below in Fig. 5.3. It can be seen that the optimized structure fulfills the multichannel filtering function we defined earlier.

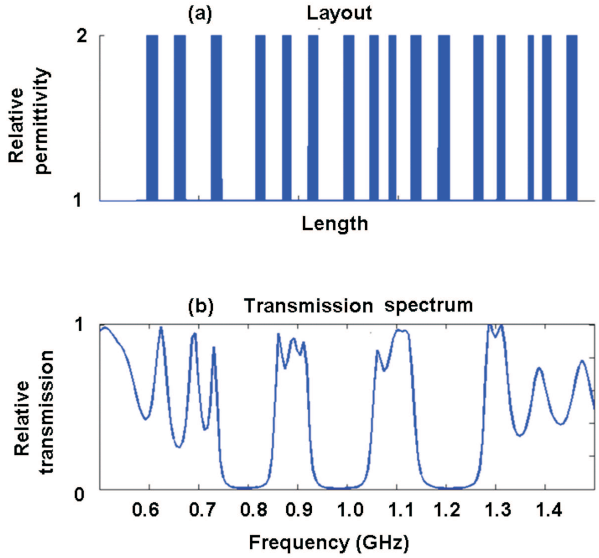


Figure 5.3: The layout and the spectrum of the topology optimized multichannel filter.

5.4 Optimization of 1D pulse-shaping filters

In this section, we investigate using FDTD-based TO to optimize the material distribution in a grating to accomplish pulse filtering functionalities. The difficulties with using the field history directly in the objective function are addressed and a solution is proposed and test. Appropriate filtering methods are used to ensure the fabricability of the final design.

5.4.1 Motivation

Shaping optical pulses into arbitrary wave forms is desirable in many applications including optical communication, nonlinear optics and biomedical imaging [28]. Different methods to design pulse-shaping filters are present today, based on various hardware configurations and design algorithms. Fourier synthesis methods are used by Weiner [28] to design filters by distributing gratings and lenses. While having the benefits of incorporating programmable modules such as liquid-crystal spatial light modulators, this system is bulky, lossy and difficult to integrate into waveguide systems. Fibre gratings have also become good candidates for shaping pulses. Many inverse design algorithms have been introduced [84, 85, 32, 86], among which the inverse scattering layer peeling method (LPM) [87] is the most used. However, LPM is only effective for designing 1D fibre gratings for two reasons: 1. The resulting topology from LPM optimizations are with continuous index variations, which needs an extra synthesis method to be converted to feasible designs for two-phase platforms including thin-films and planar waveguides [88]; 2. So far LPM has not been generalized to design 2D structures. Hence new design methods have to be developed if 2D filters are desired in planar waveguide systems.

5.4.2 Objective function

Different from minimizing the energy flux in the frequency-domain method [19], the objective function (OF), $F(\rho)$, for the time-domain pulse-shaping optimization should aim at minimizing the time-integrated squared difference between the output pulse $E_o(t)$ and the prescribed pulse $g(t)$, namely:

$$F(\rho) = \int_{t=0}^{T_{\max}} [E_o(\rho, t) - g(t)]^2 dt. \tag{5.5}$$

5.4.2.1 Envelope objective function

There are two challenges in using the formulation of OF in Eqn. 5.5. First of all, the optimization will be easily trapped in local minima due to rapid oscillations of the electric field. Consider a case where we want to delay a Gaussian pulse by simply increasing the dielectric constant of the bulk material (see Fig. 5.4a). Figure 5.4b shows the objective function values as a function of τ, which is the distance in time steps between the output pulse and the target pulse. For perfect overlap between the two pulses ($\tau = 400$), the objective function value is 0. For τ different from 400, there are oscillations with local minima for each carrier wave period. The optimization is prone to be trapped at these minima, thus a new objective function needs to be introduced to circumvent this problem.

Instead of optimizing the electric field directly as described in Eqn. 5.5, we propose an envelope OF that optimizes the time-averaged electric field strength:

$$\hat{F} = \int_{t=0}^{T_{\max}} [\frac{\sum\limits_{t' \in N_t} E_o^2(\rho, t')}{N} - \frac{\sum\limits_{t' \in N_t} g^2(t')}{N}] dt, \tag{5.6}$$

where N is the number of time steps in one carrier wave period. N_t is the set of time points having the span of one carrier wave period and is centred at time t. By averaging the squared E field within each period, the envelope is effectively extracted (see Fig. 5.5). The original inverse design problem is now reformulated into a more regularized one. Note, however, that since the

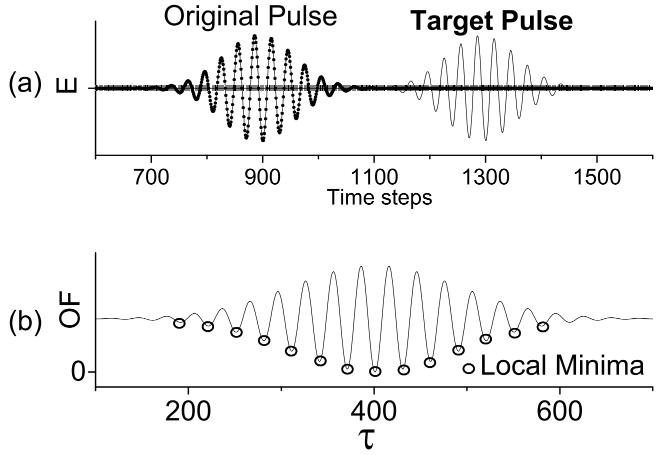


Figure 5.4: A simple design case of delaying a pulse. (a) The electric field of the original (dotted) and target (smooth) pulses. (b) The objective function (OF) values as a function of τ, which is the number of time steps by which the original pulse is moved towards the target pulse.

envelope is being optimized now, the phase of the field will not be in direct control of the optimization anymore.

5.4.2.2 Sensitivity analysis for the envelope objective function

While the system matrix derivative in the sensitivity expression stays the same for the envelope OF, the adjoint current in the backward analysis changed. The adjoint current depends on $\frac{\partial f}{\partial E}$ with f being the merit function in the objective function. In the case of the envelope objective function, f takes the following form:

$$f(t) = [\frac{\sum\limits_{t' \in N_t} E_o^2(t')}{N} - \tilde{g}(t)]^2, \tag{5.7}$$

Here $\tilde{g}(t)$ is the prescribed target pulse that has already taken an averaged form.

Now define:

$$h(t) = \frac{\sum\limits_{t' \in N_t} E_o^2(t')}{N}, \tag{5.8}$$

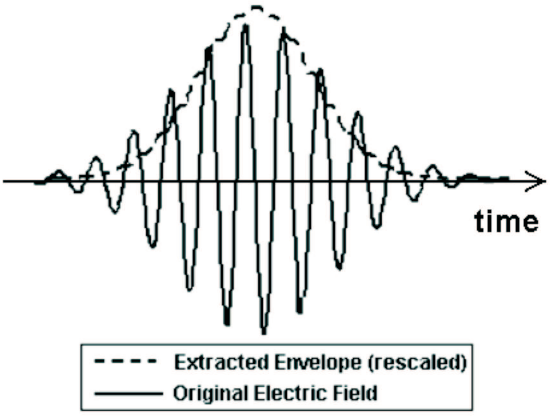


Figure 5.5: Envelope (dotted) extracted from the original electric field (solid) (see Eqn. 5.6). The amplitude of the envelope is scaled down to be comparable qualitatively to the shape of the electric field.

and take $h(t)$ into the derivative of the merit function in Eqn. 5.7 at time t':

$$\frac{\partial f}{\partial E}|_{t'} = \sum_{\tilde{t}=N_{t'}} \frac{\partial f}{\partial h(\tilde{t})} * \frac{\partial h(\tilde{t})}{\partial E(t')}. \tag{5.9}$$

In the above equation, the two derivative terms can be further deduced:

$$\begin{aligned} \frac{\partial f}{\partial h(\tilde{t})} &= 2[h(\tilde{t}) - g(\tilde{t})], \\ \frac{\partial h(\tilde{t})}{\partial E(t')} &= \frac{2}{N} E(t'). \end{aligned} \tag{5.10}$$

Take Eqn. 5.10 back into the derivative of the merit function in Eqn. 5.9, we have:

$$\frac{\partial f}{\partial E}|_{t'} = \frac{1}{N} \sum_{\tilde{t}\in N_{t'}} 4[\frac{\sum_{\bar{\tilde{t}}\in N_{\tilde{t}}} E^2(\bar{\tilde{t}})}{N} - \tilde{g}(\tilde{t})]E(t') \tag{5.11}$$

5.4.2.3 Explicit penalization

The second challenge to an ordinary OF in Fig. 5.5 is to ensure the fabricability of the optimized design. To comply with the gradient-based optimization algorithm, a continuous design variable ρ is used. However, in order to fabricate the device on a two-phase material platform, a black and white design with reasonable minimum feature size has to be guaranteed. The linear interpolation relationship between the local material ε_i and design variable ρ_i calls for extra measures to penalize intermediate values between 0 and 1.

In the case of designing Bragg gratings (see section 5.3), the 0-1 design is naturally enforced without further measures. This is because the transmission is to be minimized, which is accomplished by utilizing the highest index contrast available in the materials. Thus, the algorithm is encouraged to use the materials with the highest and the lowest permittivity available, instead of 'grey' material in between. Unfortunately, this mechanism cannot be applied to the pulse-shaping filter optimizations, where the goal is not to minimize the transmission, but to shape the temporal distribution of the field energy. Based on the above discussions, extra measures need to be taken to enforce the black/white design in our problem.

The popular form of simplified penalization used in standard TO [89] is not feasible here since it is difficult to impose a volume constraint on a scattering problem. Implicit morphological filters employed on densities [90] are usually good candidates for minimum length scale control but not successful in penalizing the intermediate design variables entirely in our case. Although the grey elements can be removed by post processings, the performances of the altered structures may deviate a lot from the thresholded ones. This is because that in $1D$, each element denotes a plane instead of a grain in $2D$. It can be seen that many elements in one plane are coupled and thus one element change forces the others to change altogether. This obviously increase the system's sensitivities to each of these element.

After various trials, it is shown that Borrvall and Petersson's [91] explicit and mesh-independent penalization imposed directly on the objective function offers a good solution to our problem. The new OF will have the following form:

$$\bar{\hat{F}} = \hat{F} + d \sum_i \tilde{\rho}_i (1 - \tilde{\rho}_i), \tag{5.12}$$

where the average density $\tilde{\rho}_i$ is defined as $\tilde{\rho}_i = \sum\limits_{j \in Di} \rho_j / L$. D_i is the vicinity of the i^{th} plate and L is the length of D_i. The averaging of the densities works as a minimum length scale control to avoid generating too fine features in the optimized layout. Clearly the penalization term in Eqn. 5.12 is minimum only if all design variables are either 1 or 0. d is a scaling factor that determines the extent to which intermediate values should be penalized. A too small d would render the penalization ineffective, while a too large d would cause the optimization to deviate from the original goal. Here a continuation method was used where d starts from 5×10^5 and gradually increases through the optimization iterations to shift the optimization focus from pulse shaping to black and white design.

One challenge posed by the problem is that the optimization might fall prey to local minima easily where the pulse shape is still quite far away from ideal. This could be caused by the combination of two reasons. Firstly, the 1D configuration of the layered structure determines that light can only have one path between the source and the output. Any small changes in the layers might drastically change the grating characteristic. This means that as soon as the optimization is at a local minimum, it takes a large incentive for the algorithm to 'climb' uphill again. The second reason is that as the scaling factor d increase rapidly in the end, the optimization gives more preference to structures that brings down the penalization term in Eqn. 5.12, instead of the real objective of pulse shaping. To circumvent this problem, we carry out a set of optimizations instead of one and vary the starting guesses of the design domain in hope for search the whole solution map. Numerical experiments show that by running 50 optimizations based on random starting guesses, a handful of good designs can be obtained for our problem.

The penalization term in Eqn. 5.12 brings an additional sensitivity term which only has an explicit dependance on the design variable ρ. The exact form of the altered sensitivities can be seen in chapter 6 section 6.5.

5.4.2.4 Modified objective function

Combining the above two alterations to the original OF, we now minimize the time-integrated envelope difference between the output and the target pulse plus the density penalization sum by varying the design varible ρ. In the meantime, Maxwell's equations should be obeyed as well as the upper and lower boundaries of the materials used for our system. The problem is thus formulated as:

$$\begin{gathered}\min_{0\leq\boldsymbol{\rho}\leq 1} \bar{\tilde{F}}(\boldsymbol{\rho}),\\ \text{subject to : Maxwell's equations,}\end{gathered} \tag{5.13}$$

5.4.3 Results

The optimization method is tested where an input Gaussian pulse with a full width at half maximum (FWHM) of $26fs$ and a carrier wavelength of $1.55\mu m$ is to be converted to a square pulse in a stack of thin-film layers consisting of Si ($\varepsilon_r = 12.3763$) and SiO_2 ($\varepsilon_r = 2.3339$). Due to the limited spectral range of the input pulse, the filtered square pulse will show overshoot and ringing known as the Gibbs effect. To comply with this mathematical limit, the target pulse is then generated by superimposing five Gaussian pulses with different time delays to mimic a square pulse so that its spectrum fits under that of the input pulse. The resulting target pulse has a FWHM of $90fs$ with its

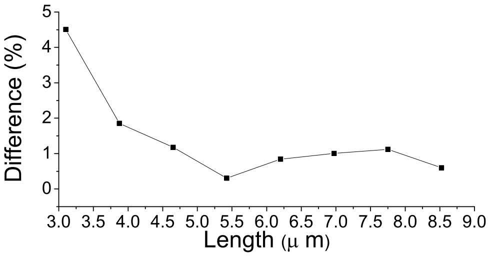


Figure 5.6: The effectiveness of the optimization as a function of design domain size.

amplitude decreased five-fold compared to that of the un-filtered pulse. The design domain is discretized into elements the size of $\lambda/165$, where λ is the wavelength of the carrier wave of the pulses. Design domains with 8 different lengths were optimized, each with 50 randomly distributed starting guesses to increase the chance of finding a solution close to global minimum. For the design domain of $5.42\mu m$ with 578 elements, a single optimization takes 12.6 minutes on an state-of-the-art single processor computer with the Fortran 95 code.

Figure 5.6 shows the optimization results in terms of the integrated difference between the optimized pulse envelope and the target pulse envelope. For each design domain size, a best result was picked among all 50 optimizations of different starting guesses. The optimized pulse approaches the target pulse better as the length of the design domain increases, since more reflectors are allowed in a longer device to tailor the pulse more accurately. This is a clear trend to be seen in the length between $3.10\mu m$ and $5.42\mu m$, with the deviation from the target pulse down to 0.3% at the latter. However, if the design domain is too big, the optimization becomes more difficult due to the many more local minima introduced by the increasing number of design variables. The advantage of longer design domain is then lost, which can be seen after $5.42\mu m$.

The best design is that of $5.42\mu m$, or 3.5 times the wavelength of the carrier wave of the pulse. Fig. 5.7(a) illustrates the optimized pulse (solid black) compared to the target pulse (solid grey) and the original Gaussian pulse (dotted black). The layout of such a pulse-shaping filter is depicted in Fig. 5.7b. With the help of the explicit penalization term in Eqn. 5.12, the thinnest layer in this structure is $65.8nm$ which ensures fabricability.

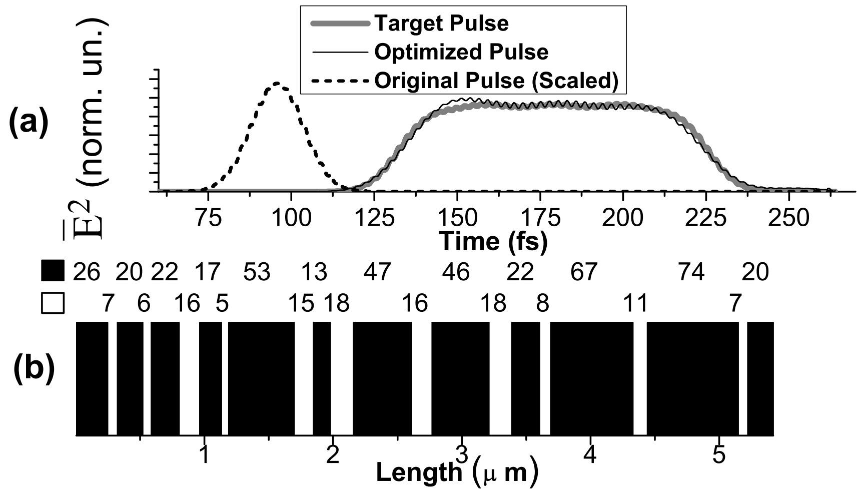


Figure 5.7: Square pulse shaping. (a) The optimized output pulse with scaled amplitude (solid black) vs. the target pulse (solid grey) as well as the scaled input Gaussian pulse (dotted black); (b) The optimized layout of the thin-film pulse-shaping filter with Si (black) and SiO_2 (white). The thickness of each layer is noted in terms of the corresponding number of grid points shown above the topology.

5.5 Conclusions

Topology optimization based on FDTD in one-dimension has been demonstrated in this chapter. A proof of concept in designing Bragg gratings show that TO is indeed capable of finding the theoretically best solutions in simple cases. For more complicated problems where simple rationales of favoring high index contrasts are not available, extra numerical tools need to be applied in order to assure sensible optimization results. Black/white designs and minimum feature size controls are both crucial in producing manufacturable designs. In our experiments, explicit penalization showed its capability of achieving such designs in 1D grating optimization problems. Since 1D structures with the aid of explicit penalization scheme are prone to local minima, optimizations based on various starting guesses are required to ensure a more thorough search in the solution space. The results are promising and feasible designs which fulfills pulse shaping functionalities are generated.

Chapter 6

Minimum lengthscale control and black/white designs

Real-world structures and devices are usually composed of one or more distinct materials which are shaped in an organized way to assume certain desirable properties. For example, steel can be shaped into triangular sub-units to form a truss structure which has a small compliance. Silicon and air can be arranged in a periodic fashion to create a photonic band gap where light within a certain frequency band is prohibited to pass through. In topology optimization, a direct interpretation of these distinct materials would lead to distributed, discrete-valued design problems. These problems are usually very hard to solve, especially considering the large number of design variables common in a topology problem. On the contrary, a convex problem with continuous design variables is possible to solve by using gradient information.

Consider a case where silicon (solid) and air (void) are to be distributed in the design domain to fulfill certain optical functionalities. A linear interpolation can be used to project the design variables (ρ (also known as densities) to the physical domain which contains the actual permittivities (ε):

$$\begin{aligned} &x_i = \varepsilon^0 + (\varepsilon^1 - \varepsilon^0)\rho_i, \\ &0 \leq \rho_i \leq 1, \; i \in N. \end{aligned} \tag{6.1}$$

Here ε_0 and ε_1 are the permittivities of air and silicon respectively and N is the set of all design variables. A zero-valued design variable would render the local permittivity to be that of the air, while a design variable equal to 1 identifies a Si element. Everything in between 0 and 1, or white and black, are the grey elements which are un-physical yet crucial in formulating the design problem into a convex continuous optimization problem.

The existence of intermediate elements in the final design, however, hinder the manufacturability of the optimized topology. The lack of means to interpret these grey material in the manufacturing process means that the theoretical performance predicted for the topology during the modeling process will not hold for the actual device, where only two-phase materials (black/white) are available. Another issue contributing to the difficulties of modeling and manufacturing the optimized designs is the minimum lengthscale of a structure. In an ideal case, a more refined mesh in the optimization would in general produce the same topology as a less refined one, just with more smoothes on the boundaries given rise by finer staircases. For a non-regularized design problem, however, there exist a mesh dependency problem. Small and new features surface when the mesh is refined, thus resulting in a quite different topology compared to a that from a coarse mesh. A direct problem caused by the mesh dependency is that it is difficult to get rid of small features containing only one or two finite elements/grid points. The coarse discretizations within these features give rise to numerical artifacts during the modeling based on finite methods (FEM , FDTD, etc.), and give inaccurate predictions of the device performances. Moreover, the minimum lengthscale also determines whether the topology can be manufactured with decent precision given the fabrication tolerance.

Various techniques have been applied in order to control the black/white composition as well as the minimum lengthscale of the optimized design. In structural optimizations, a material interpolation method called Solid Isotropic Material with Penalization (SIMP) is often used to penalize intermediate densities and eventually steer the design towards black/white topologies. The SIMP model is effective under the volume constraint: $\eta = \frac{V}{V^*} \leq 1$, where V is the aggregated densities over the design domain and $V*$ is the density sum if the whole design domain is filled with solid-phase material. Since the total material is capped, the optimization has to use material wisely so that this constraint is not broken. SIMP penalizes the grey materials to make them less effective in the objective function (e.g. minimum compliance), and thus forcing them out of the final design. However, SIMP does not work well with problems where the sensitivities of the design variables constitute more than one sign. While larger volume of materials are always preferrable in making a mechanical structure strong, the same principle does not hold for designing other physical problems (e.g. heat conduction, optical pulse shaping, etc.). This will be demonstrated in the next section. Filters are a class of popular methods used by TO to control minimum lengthscales. Together with SIMP, density filtering provides mesh independent designs by limiting variations of densities in a close neighborhood [92][93]. Sensitivity filtering also performs well by making a similar filtering approach to the sensitivities instead of the densities [94][95]. An intrinsic characteristic with the above filtering techniques is that they produce grey elements around the boundaries of materials. For mechanical problems, simple thresholding can be used to remove these grey area while still maintaining the optimized device performance to a large extent. For our pulse-shaping examples, however, the light scattering pattern becomes extremely sensitive to the boundary variations from the original topology and the thresholded one. The device performance is drastically changed, rendering the design invalid. More advanced projection methods are developed to further enforce the discreteness of the optimized design, alleviating the problems caused by large grey transition areas [96][90][97].

The optimized topologies as well as their performances are analyzed both qualitatively and quantitatively to give a better understanding of different filter mechanisms in our pulse filtering problem.

6.1 Test problem formulation

A Gaussian pulse with the FWHM of $26fs$ is sent through the device via a ridge waveguide. The design domain is a rectangular region of the size $946nm$ by $1604nm$ Fig. 6.1. It is coupled in and out by ridge waveguides and the whole computational domain is discretized into a uniform grid of the spacing 26.3nm. Silicon and air are to be distributed in the design domain so that the input pulse will be transformed into a target pulse. The target output pulse (grey) in Fig. 6.2(b) is extracted by letting the input pulse propagate through a known test structure in Fig. 6.2(a). Note that even though the test structure from where the target pulse is extracted is an obvious solution candidate, it is not the only solution available. What we try to find is a topology that fits our design criteria, which has black/white topology and maintains minimum lengthscale control, while accomplishing the pulse filtering function. Since the target pulse is the response of the test topology, it ensures that there exists at least one solution. However, the optimized topology is not necessarily equal to this prescribed topology, given the non-uniqueness of the scattering problem.

The above design problem, in its simplest form, takes the following formulation:

$$\begin{gathered}\min_{0\leq\rho\leq1}\ \log\left(\int_t [\bar{E^2}(\rho)-\bar{G^2}]dt\right) - C,\\ \text{subject to}:\ \text{Maxwell's equations}\end{gathered} \tag{6.2}$$

$\bar{E^2}(\rho)$ and $\bar{G^2}$ are the time-averaged envelopes of the output pulse to be optimized and the target pulse, respectively. The reason why we optimize envelope instead of electric field as well as the derivation of sensitivities for the envelope objective function has been described in details in section 5.4.2.1. A logarithm

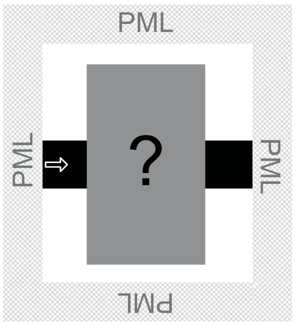
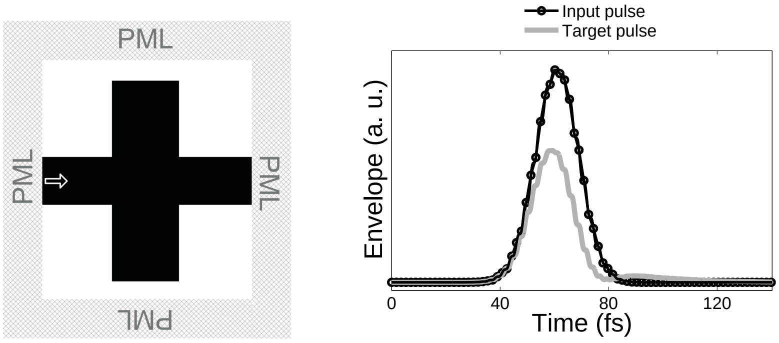


Figure 6.1: Problem formulation for the filter test. The design domain (grey) is coupled by ridge waveguides. The source is excited at the beginning of the waveguide (indicated as the arrow). The output pulse is evaluated at the end of the output ridge waveguide. The calculation domain is terminated by PML layers.

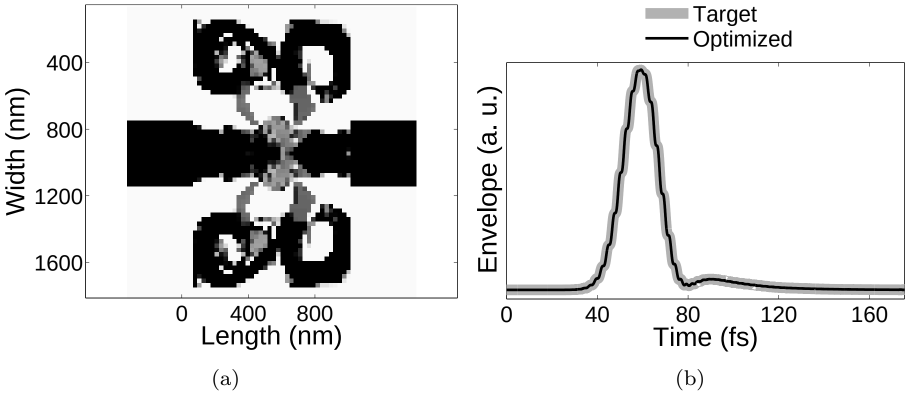


(a) (b)

Figure 6.2: (a) The test structure is coupled in and out by ridge waveguides. The source is excited at the beginning of the waveguide (indicated as the arrow) and the target output pulse is extracted at the end of the output ridge waveguide. (b) The input pulse (black-circled line) is transformed into the target pulse (grey line) after propagating through the test structure.

is taken since the time-integrated square difference between the output pulse is a large number and might cause scaling problems for MMA. For similiar reasons, a constant C is also used here to scale the objective function values close to 1. Naturally, the smaller objective function value, the better fitting between the optimized pulse and its target. Without using any penalization measure nor filters, such an optimization gives a topology in shown in Fig. 6.3.

As illustrated in Fig. 6.3, the optimization finds a solution where the optimized pulse fits the target pulse perfectly. The objective function value is as low as



(a)

(b)

Figure 6.3: Optimized results without penalization or filters. (a) The optimized topology. $M_{nd} = 17.01\%$. (b) The optimized pulse (black) vs. the target pulse (grey). The objective function value is at 0.6056.

0.6056. However, the resulting topology suffers from large area of grey elements. Moreover, there exist many stand-alone elements in the final design, indicating bad minimum lengthscale control. To quantitatively evaluate the severeness of grey areas in the optimized design, a measure of discreteness is used here [90]:

$$M_{nd} = \sum_{i=1}^{N} \frac{4{\rho_i}^*(1-{\rho_i}^*)}{N} \times 100\% \tag{6.3}$$

ρ^* here denotes the filtered design densities, which in the current case equate the original design variables ρ since no filters are used. N is the total number of grid points in the design domain. For a complete grey design where all densities are at 0.5, the measure of discreteness gives a value of 100%. On the other hand, for an absolute black and white design, the measure of discreteness becomes 0%. For the unfiltered design shown above, the M_{nd} value is at 17.01%.

6.2 SIMP

In structural TO, a model called Solid Isotropic Material with Penalization is often used to interpolate the design variables:

$$\begin{aligned} &E_{ijkl,i} = (\rho_i)^p E^0_{ijkl},\ p > 1, \\ &\textstyle\int_\Omega \rho_i d\Omega \leq V,\ 0 \leq \rho_i \leq 1,\ i \in N. \end{aligned} \tag{6.4}$$

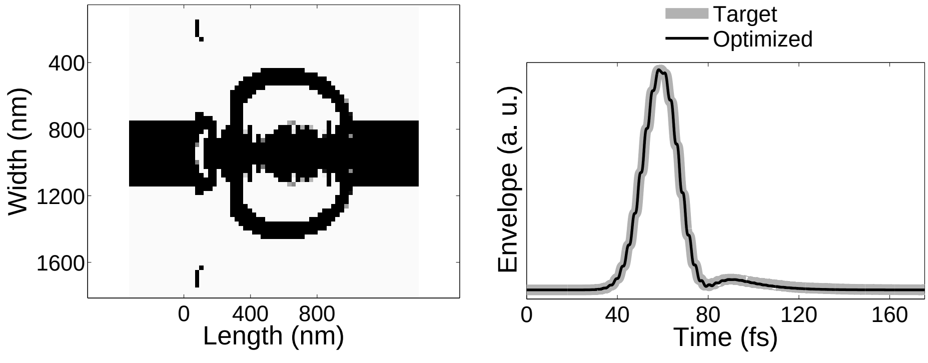
E^0_{ijkl} is the stiffness tensor of a given material. Since the penalization parameter p is larger than 1, the factual stiffness of one element, $E_{ijkl,i}$, is smaller than the one where p is not in play. Thus, the stiffness comes at a more expensive price (more material used) compared to those with design variables either of 0 or 1. With a volume constraint that limits how much solid material is available for the whole structure, the algorithm is bound to choose design variables close to 0 or 1.

Such schemes, however, do not work well with designing photonic devices. While a smaller stiffness usually contributes to weaker mechanical structures, smaller permittivities do not necessarily link to higher nor lower transmissions of light. This is because the improved performance of such devices are based on the distribution of highly localized light patterns introduced by interfaces between two materials. These patterns might just be as well or even better bounded by grey elements than black/white elements. In 1D problems where all waves have to pass through each layer of materials, the volume constraint can hardly force out the grey layers who might be critical in defining the localized modes. The 0s and 1s are no longer more cost-effective than the intermediates elements in this case. Thus, the optimization algorithm will not prefer black/white elements like it does in the mechanical problems or transmission minimizing/maximizing problems. In 2D problems, it is more flexible for the waves to travel via different routes and thus the specific local modes are not as critical as in 1D. There the grey elements can be more effectively penalized by the volume constraint.

Figure 6.4 shows the optimization results of the test problem with SIMP interpolation under no volume constraint while Fig 6.5 is for SIMP under the volume constraint of 0.3 ($\frac{V}{V^*} \leq 0.3$). Without using the volume constraint, the SIMP scheme shows its ability to decrease the grey elements. While under the volume constraint, we observe that the objective function value deteriorates from 4.1290 to 7.5764. A clear discrepancy between the optimized pulse and the target pulse can also be seen. This degrading indicates a strong bound of the volume constraint, i.e. the solid material available is not sufficient to filter the pulse completely. However, even more grey elements exist under the volume constraint ($M_{nd} = 1.7\%$ compared to $M_{nd} = 0.84\%$ without volume constraint). This proves that our prediction that the SIMP/volume constraint combination does not penalize the grey elements efficiently in our problem.

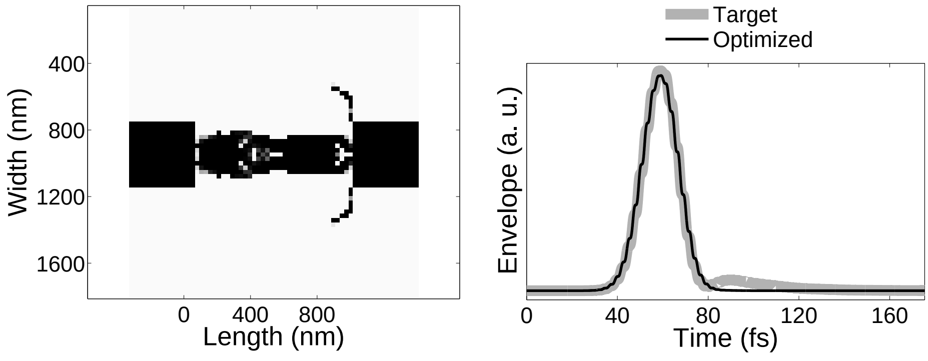
6.3 Density filters

To compute the filtered density for element i, we sum up the weighted densities for all the elements in the neighborhood of i (including i), and divide the sum



(a) (b)

Figure 6.4: Optimized results using SIMP under no volume constraint. (a) The optimized topology. $M_{nd} = 0.84\%$. (b) The optimized pulse (black) vs. the target pulse (grey). The objective function value is at 4.129.



(a) (b)

Figure 6.5: Optimized results using SIMP under a volume constraint of 0.3. (a) The optimized topology. $M_{nd} = 1.7\%$. (b) The optimized pulse (black) vs. the target pulse (grey). The objective function value is at 7.5764.

by the total weights:

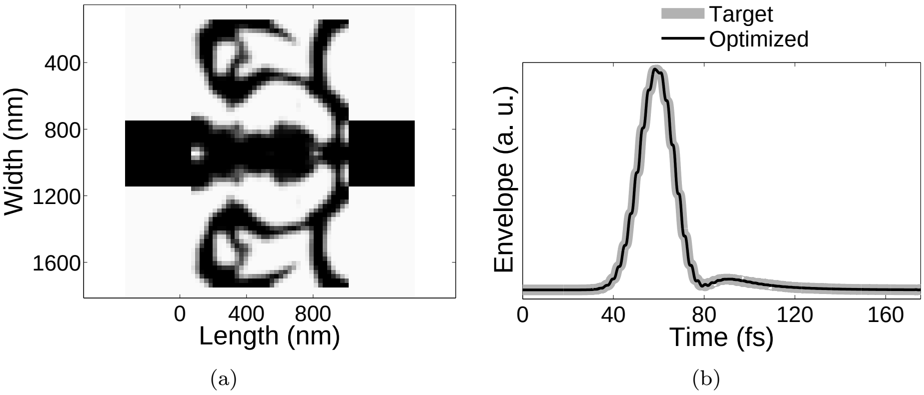
$$\begin{aligned} \tilde{\rho}_i &= \frac{\sum\limits_{j \in N} W_i^j \rho_j}{\sum\limits_{j \in N} W_i^j}, \\ W_i^j &= \left\{ \begin{array}{c} 1 - \frac{D_{ij}}{r}, if\ D_{ij} \leq r \\ 0,\ otherwise \end{array} \right. \end{aligned} \tag{6.5}$$

Here D_ij is the distance between element i and j. If this distance is smaller or equal to the filter size r, the weight of element j in the filtered densities is inversely proportional to D_ij; otherwise, the weight is 0.

The filtered densities $\tilde{\rho}$ are then used in the physical topology on which the analysis is based on. Note that the filtered sensitivities $(\frac{\partial F}{\partial \tilde{\rho}})$ should be adjusted to derive the real sensitivites $\frac{\partial F}{\partial \rho}$ by using a chain rule to take into considerations of the filtered sensitivites in the neiborhood of the center element:

$$\begin{aligned}
\frac{\partial F}{\partial \rho_i} &= \sum_{j \in N} \frac{\partial F}{\partial \tilde{\rho}_j} \frac{\partial \tilde{\rho}_j}{\partial \rho_i}, \\
\frac{\partial \tilde{\rho}_j}{\partial \rho_i} &= \frac{W_i^j}{\sum\limits_{j \in N} W_i^j}, \\
\frac{\partial F}{\partial \rho_i} &= \frac{\sum\limits_{j \in N} W_i^j \frac{\partial F}{\partial \tilde{\rho}_j}}{\sum\limits_{j \in N} W_i^j}, \\
W_i^j &= \begin{cases} 1 - \frac{D_{ij}}{r}, if\ D_{ij} \leq r \\ 0,\ otherwise. \end{cases}
\end{aligned} \tag{6.6}$$

After the filtering, any boundaries between black and white (or grey) in ρ will be smeared out to larger grey regions in $\tilde{\rho}$. Since the grey areas are penalized in SIMP, a black/white topology in ρ will be eventually favored by the optimization to have lesser grey area in the filtered densities $\tilde{\rho}$. Moreover, the resulting structure is more likely to get rid of small geometries since their existence results in larger perimeters with grey transition area, which is bound to be penalized in the SIMP scheme. This mechanism ensure the minimum lengthscale control of the structures. Figure 6.6 shows the optimization results using density filtering with SIMP interpolation under volume constraint 0.7. The topology has much less small features and holes, compared to the results without density filters. The objective function also converges well at the value of 3.0257. However, the filtered densities have a diffusive effect around its edges of the features. Since the filtered densities are the topology we build our modelling on, they are the physical structure we evaluate. Due to the smearing out effect of the density filtering, the grey area will always exist when the filter radius r is larger than 1. In some cases, e.g. some compliance minimization problems, these grey elements can be removed in post-processing using thresholding and the structure still maintains a favorable objective function value. In other cases, e.g. designing optical filters, removing these grey elements results in significant changes in the performance of the optimized structures due to the high sensitivity of the objective function to these elements. An alternative is to gradually decrease the filter radius to 1 during the optimization. However, this



(a) (b)

Figure 6.6: Optimized results using density filters. (a) The optimized topology. $M_{nd} = 27\%$. (b) The optimized pulse (black) vs. the target pulse (grey). The objective function value is at 3.0257.

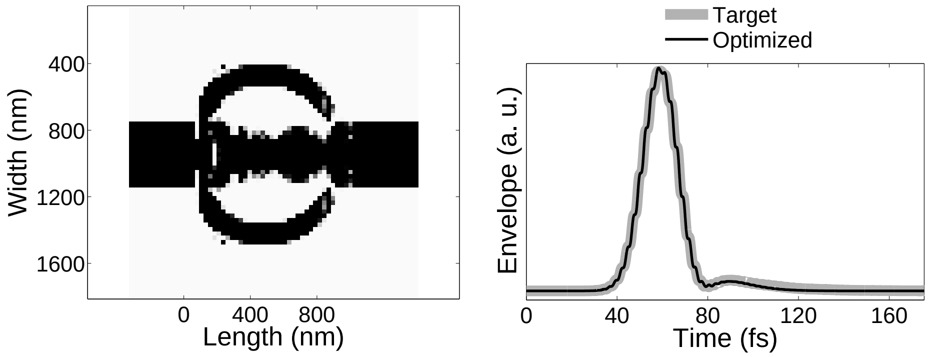
method might reintroduce mesh dependent designs as well as small features.

6.4 Sensitivity filters

Sensitivity filters modify the sensitivities of a certain element by incorporating the neighboring sensitivities:

$$\begin{aligned} \tilde{\rho}_i &= \frac{\sum\limits_{j \in N} W_i^j \rho_j}{\sum\limits_{j \in N} W_i^j}, \\ W_i^j &= \begin{cases} 1 - \frac{D_{ij}}{r}, if\ D_{ij} \leq r \\ 0,\ otherwise \end{cases} \end{aligned} \tag{6.7}$$

As a popular tool in structural optimization, sensitivity filters offer some very desirable qualities. First of all, the resulting topology by using sensitivity filters are mesh-independent and have minimum length scale control by the size of r when used with SIMP interpolation/penalization under a volume constraint. Moreover, the mechanism is inexpensive to implement and does not impose extra constraints in the optimization problem. The optimization results for the test example using the sensitivity filters with $r = 2$ and volume constraint of 0.3 is shown in Fig. 6.7. The same as in density filters, a continuation of filter radii is necessary to get rid all the grey area around the boundaries.



(a) (b)

Figure 6.7: Optimized results using sensitivity filters. (a) The optimized topology. $M_{nd} = 3.18\%$. (b) The optimized pulse (black) vs. the target pulse (grey). The objective function value is at 5.0149.

6.5 Explicit penalization

Borrvall and Petersson first introduced a regularized penalization term on the intermediate densities in 2000 [91]. Different from the density-based filters, this term $G(\rho)$ is explicitly presented in the objective function:

$$\begin{aligned}\bar{F}(\rho) &= F(\rho) + G(\rho),\\ G(\rho) &= d\sum_{e\in N}\tilde{\rho}_e(1-\tilde{\rho}_e).\end{aligned} \tag{6.8}$$

Here $\tilde{\rho}_e$ is the filtered density:

$$\tilde{\rho}_e = \frac{1}{|N_e|}\sum_{i\in N_e}\rho_i, \tag{6.9}$$

where N_e is the neighborhood of element e and $|N_e|$ is the number of elements in the neighborhood.

The penalization term is non-zero only when the filtered density $\tilde{\rho}_e$ is zero or one, hence the intermediate densities will be unfavored during the minimization of the objective function. The filtered densities instead of the original densities are used in order to avoid mesh dependency design as well as small objects in the design. Thus the optimization problem become regularized. In order to ensure the convergence of the original design problem, a continuation method was taken where the penalization term weight d gradually increases from an initial small value. When d is small, the original objective function F becomes

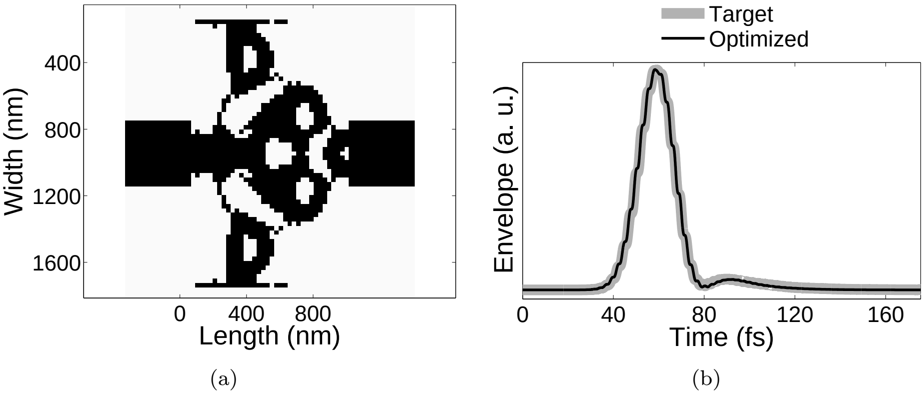


Figure 6.8: Optimized results using explicit penalization. (a) The optimized topology. $M_{nd} = 0.0\%$. (b) The optimized pulse (black) vs. the target pulse (grey). The objective function value is at 4.9536.

the dominant factor in the optimization goal. In later stages, d increases to sufficiently large values so that the penalization of grey element begins.

The sensitivities of the penalization term can be derived by applying the chain rule:

$$\begin{aligned} \frac{\partial G}{\partial \rho_i} &= \sum_{e \in N_i} \frac{\partial G}{\partial \tilde{\rho}_e} \frac{\partial \tilde{\rho}_e}{\partial \rho_i}, \\ \frac{\partial G}{\partial \tilde{\rho}_e} &= d(1 - 2\tilde{\rho}_e), \\ \frac{\partial \tilde{\rho}_e}{\partial \rho_i} &= \frac{1}{|N_e|}. \end{aligned} \tag{6.10}$$

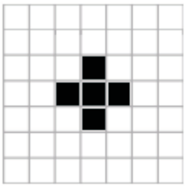
Since the penalization term is explicitly added to the original objective function, the optimization will eventually shift away from the real objective to favor the penalization of grey elements. Experience show that it is especially a big problem in 1D pulse filtering problems. A black/white design is obtained usually at the price of largely distorted pulse shapes. One remedy is to start the optimizations from different random initial topologies in order to search the solution space more thoroughly.

Optimization results on the test example using explicit penalization is illustrated in Fig. 6.8. It can be seen that using explicit penalization does generate completely black and white design without compromising the objective function much in this 2D example. However, thin lines of solid phase persist, indicating less than satisfactory minimum lengthscale control.

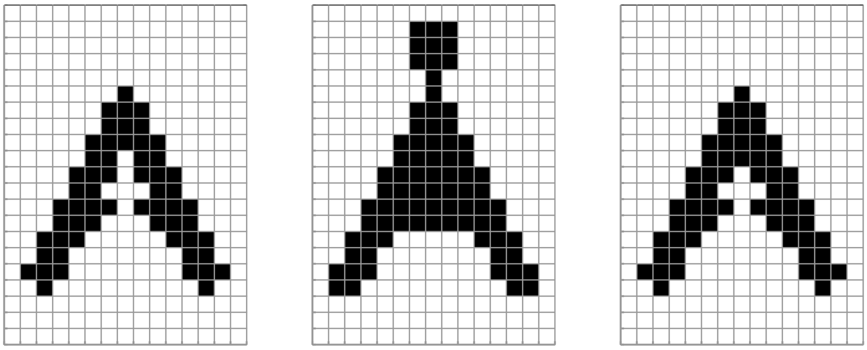
6.6 Modified Heaviside filters

Although density filter helps ensure mesh independency, the averaging effect within the neighborhood of any element makes sharp edges impossible to obtain in the densities. Thus a new scheme is needed to not only give minimum length scale control, but offer black and white topology.

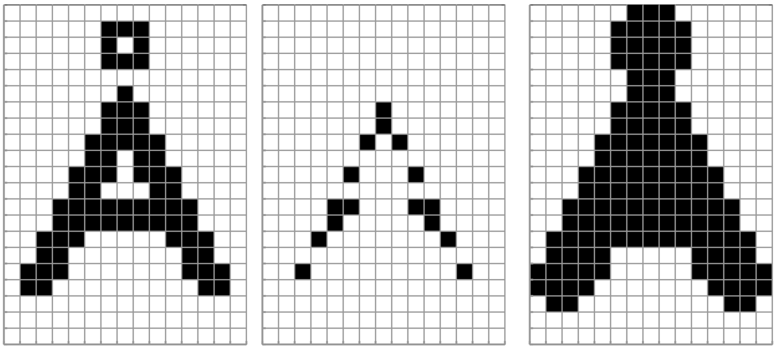
Morphology filters were introduced by Sigmund in [90]. The idea came from morphological filters used in image processing, where erosion and dilation are the basic operators. A structuring element (SE), usually of a couple pixel wide and tall, are used to gauge a binary picture for features. In erosion, the SE is centered at a certain pixel. If any pixel within the area covered by the SE is white, the center pixel is rendered white. A new image is created when all pixels are visited by the SE. A similar procedure goes for dilation, just reverse: if any pixel within the SE-covered area is black, the center pixel is rendered black. In other words, erosion makes the object (black) smaller while the dilation makes it bigger. The erosion-dilation combination (first erosion, then dilation), known as open, removes features smaller than the SE while still keeping the size of other features. The dilation-erosion combination (first dilation, then erosion), known as close, connects features over gaps smaller than the SE, while keeping the other features untouched. In real design problems, we usually want to get rid of both small features and small gaps. One way to achieve this is to take an open-close, which is a combination of the two operators. Figure 6.6 shows how these operators work on a binary image.



(a) Structuring element.



(b) Opened topology. (c) Closed topology. (d) Open-closed topology.



(g) The original topology. (h) Eroded topology. (i) Dilated topology.

Figure 6.7: The effects of morphology filters on a topology.

Such a discrete operator needs to be reformed when working on a continuous grey-scale density field instead of binary field. The Kreisselmeier-Steinhauser formulation was originally used by Guest [96], and later Sigmund [90] suggested a Heaviside formulation for the dilation operator where the local density is approximated as the following:

$$\bar{\rho}_e = 1 - e^{-\beta\tilde{\rho}_e} + \tilde{\rho}_e e^{-\beta} \tag{6.11}$$

In order to ensure convergence, a continuation process is taken where the value of β is gradually increased. When β is small, the Heaviside density filter behaves approximately the same as the mean density filter; when β is a large value, e.g. 500, the Heaviside filter resembles that of a step function (see figure 6.8). Since the filter radius does not need to diminish during the continuation process, the mesh independency as well as the minimum length scale control are able to be obtained. This is of great advantage compared to the density filtering and the sensitivity filtering where the filter size needed to be gradually decreased in order to get rid of the smearing out of geometrical features in the densities. However, the density filtering still impacts the Heaviside operator by yielding a final density value slighted deviated from 0 or 1 when the mean density $\tilde{\rho}_e$ within the structural element area is very close to 0 or 1.

Figure 6.9 shows the optimization results of the test case using modified Heaviside filters described above. The optimized pulse has a good conformation to the target pulse, and the objective function value is at 3.16. Compared to the topology by using density filters and sensitivity filters, the morphological filters are obviously very efficient in converging to black and white designs. The measure of discreteness is as low as 0.89%. It also has a clear advantage over explicit penalization when it comes to minimum lengthscale control. No thin lines are visible in either the solid or void phase.

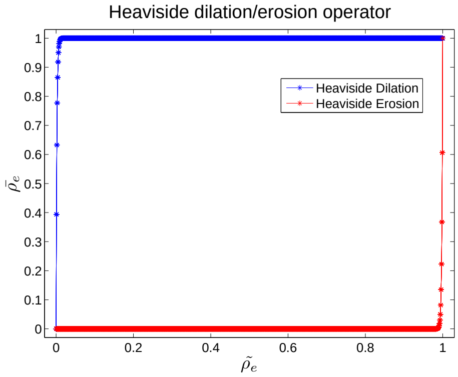
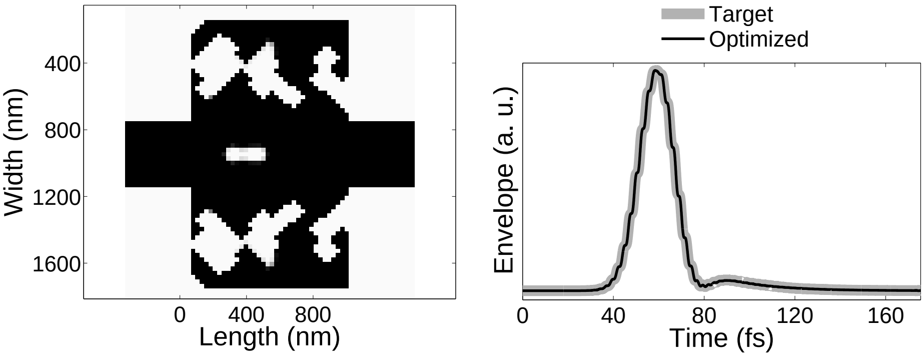


Figure 6.8: Heaviside projection between the physical density ($\bar{\rho}_e$) and the density filtered variable ($\tilde{\rho}_e$) ($\beta = 500$).



(a)

(b)

Figure 6.9: Optimization results using Heaviside morphology filters. (a) The optimized topology. $M_{nd} = 0.89\%$. (b) The optimized pulse (black) vs. the target pulse (grey). The objective function value is at 3.16.

6.7 Conclusions

In order to make the optimized devices manufacturable, the discreteness of the final design and its minimum length control are of the most importance. For optical devices like pulse-shaping filters, high sensitivities to even the smallest changes in the light scatterers determines that traditional control methods suitable for structural optimizations are not effective enough. Projection schemes proved to be a good way to control both issues in our test example. It is based on classic density filtering where a local density is depending on its adjacent neighbors in order to limit the fast variations of densities. However, it does not generate grey transition areas like traditional density filters or sensitivity filters, thus eliminating the need for vast post processing where the device performance can be subject to drastic changes.

CHAPTER 7

2D Time-domain Topology Optimizations of Pulse-shaping Filters

---

Optical pulse shaping involves converting one or more short optical pulses generated by the laser into pulses with pre-determined temporal profiles, as well as pulse trains with specific repetition rates. Current passive pulse-shaping filters are mainly based on distributed system of lenses and gratings [28] as well as long-period fiber-grating (LPG). There are also other proposals of designing pulse-shaping filters. For example, by using arrayed-interferometers, Park et al. suggested using coherence synthesization where a single pulse is replicated, delayed and superimposed to achieve desirable pulse shapes [98]. The duality of space and time was made use of by Palushani et al. where an input sinc pulse propagates through a long single-mode fiber and evolve into a temporal profile similar to its initial spectrum shape, i.e. a square-shape [99].

There are interests as well in integrating the pulse-shaping filters directly into planar waveguide systems. However, available filtering design methods are highly dependable on the platforms where these filters are realized, and none of them can be directly applied to the planar waveguides where the high-index dielectric material can be freely distributed on the two-dimensional ($2D$) platform by using state-of-the-art fabrication methods featured around e-beam lithography. Evolution-inspired optimization methods, e.g. Genetic Algorithms (GA), are of astronomical computational complexities when dealing with free distributions of materials. Thus, Topology Optimization (TO), which has many successes in optimizing free topologies, is considered a great candidate and is investigated here for its efficacy in designing planar waveguide-based pulse-shaping filters.

7.1 The inverse problem

The $2D$ inverse problem for the optimization is posed as Fig. 7.1. The whole calculation domain is invariant in the z direction. It consists of a design domain where dielectric materials are to be distributed within, boundary region where Berenger's Split-field PML absorbing boundary conditions are to be applied, and auxiliary regions including an input and an output ridge waveguide in silicon coupled to the design domain. An input pulse is excited at the source node (blue) in the input waveguide and propagates along the y axis. It is evaluated at the output node (green) and a prescribed pulse shape is expected. The whole calculation domain is uniformly discretized. Since the design domain as well as the input and output ridge waveguides are symmetric about the horizontal middle line of the calculation domain, only half of the space needs to be calculated as well as optimized. The topology is then mirrored to the other half space to obtain a full device layout.

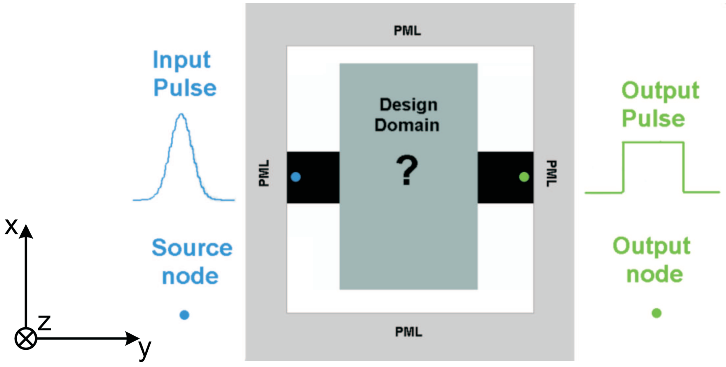


Figure 7.1: The two-dimensional inverse problem.

7.2 Square-pulse filters

Square-shaped optical pulses are widely used in high-speed optical signal processing. For example, such a square pulse can improve the timing-jitter tolerance in ultrahigh-speed optical time domain demultiplexing where the pulse is used as a nonlinear optical switching [100].

In the following example, a Gaussian pulse which has a full-width at half-maximum (FWHM) of $26fs$ and centered at $1577nm$ is to be excited at the source node within the input ridge waveguide. A square target pulse of the width of $136fs$ is prescribed to be the target pulse at the output node of the device. Silicon ($\epsilon_r^d = 12.3763$) and air ($\epsilon_r^0 = 1$) are being distributed within the design domain by the topology optimization in order to fulfill the filtering functionality. The design domain is a $2.367\mu m \times 2.393\mu m$ rectangle region, which is discretized uniformly into 90×91 elements. Each of the square element is of the size $26.3nm \times 26.3nm$.

7.2.1 Original problem formulation

The original problem is formulated as to minimize the squared difference between the extracted output pulse envelope $\bar{E}_x(t)$ and the target pulse envelope $\bar{G}(t)$. The envelope is extracted from the electric field history of the output node by averaging the squared field over each period of the carrier wave frequency (see section 5.4.2.1). Since the pulse is a guided wave propagating along the y-axis through the single mode ridge waveguides and the filtering region, E_x will be the dominating field component over E_y. Thus, in the following text, only E_x will be considered in the evaluation of the output pulse shape and sensitivity analysis.

$$\begin{aligned} &\min_{0 \leq \boldsymbol{\rho} \leq 1} \int\limits_{t=0}^{T_{max}} [\bar{E}_x(t) - \bar{G}(t)]^2 dt, \\ &s.t. : Maxwell's\ Equations, \end{aligned} \tag{7.1}$$

7.2.2 Delay variable

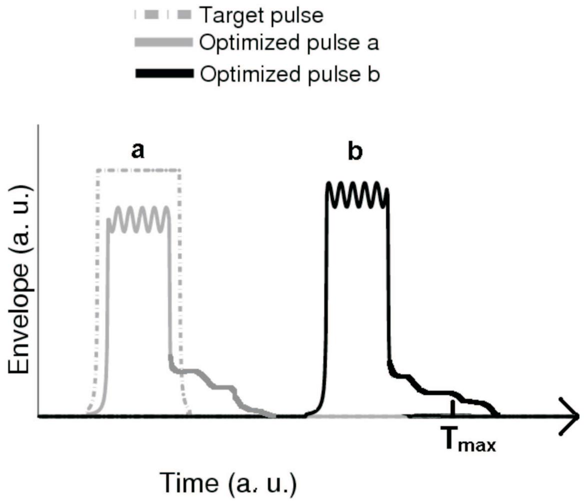
Since the wave packet travels at different group velocities in solid and in void media, the delay of the optimized pulse varies with the changing topologies just as its pulse shape. By prescribing the target pulse to occur at a specific point in the time line, we explicitly fixate the delay, which adds a constraint to how much solid material is allowed to conform to the fixed pulse delay. This hinders the optimization in the sense that the pulse shape is not the only criteria the algorithm has to take into account. See Fig. 7.2(a) for example: in iteration a, the optimized pulse (solid grey) has a worse conformance to the shape of the target pulse (dotted grey) compared to the optimized pulse in iteration b (solid black). However, since the target pulse is fixated on the time line and overlaps more with the pulse in a, this iteration results in a smaller objective function value than iteration b. Thus the algorithm prefers a to b under the false criteria, even though a provides a better topology regarding ideal output pulse shape produced.

In order to improve the flexibility of the optimization, a delay variable τ is introduced (see Fig. 7.2(b)). With the new variable, the target pulse is prescribed as $G(t,\tau)$ where τ determines the delay time of the current target pulse relative to the initial one. The initial delay is extracted by sending the target pulse through the design domain filled with vacuum. Now compare iteration a and b again: it is obvious that the pulse in iteration b has a closer resemblance to its corresponding target pulse (dotted black). The new objective function incorporating the delay variable τ becomes:

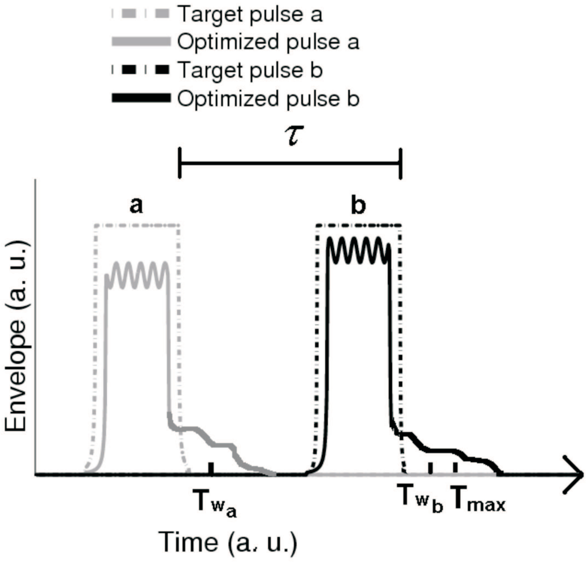
$$F = \int\limits_{t=0}^{T_{\max}} [\bar{E}_x(t) - \bar{G}(t,\tau)]^2 dt, \tag{7.2}$$

where τ is a dimensionless design variable between 0 and 1 that scales the physical delay time linearly between 0 and the maximum delay time D_{max}. The value of D_{max} depends on the size of the design domain and is usually assigned a large number to ensure enough room in the time line to move the target pulse back and forth. When $\tau = 0$, there is no further delay on the original prescribed target pulse; when $\tau = 1$, a maximum delay time D_{max} is imposed on the target pulse.

One problem with using the delay variable is that it tends to favor maximum delay due to the existence of trailing waves. Since the evaluation time terminates at T_{max}, a pulse of larger delay with much of its trailing waves beyond T_{max} is more preferable to the optimization. The wave energy beyond T_{max} is simply ignored instead of contributing to the squared differences between the optimized pulse and the target pulse. To compensate this undesired effect of the delay variable, a flexible evaluation window T_w is introduced. T_w can be determined by extending the target pulse end time by a constant amount of time, and is varied every iteration depending solely on the delay time τ. Since the tail waves fallen inside T_w is always taken into account in the evaluation of the objective function, the optimized pulses are purely merited by their shapes without being affected by the truncation of tail waves. Note that these trailing waves are still part of the evaluation, with their weight kept at the same level throughout the optimization. For the square pulse optimization, for example, the ending time of T_w is set to be 20% of the target square pulse width beyond the target's ending, which favors a pulse with high distinction ratio between the square pulse and its immediate trailing waves. Another benefit of running flexible simulation time in each iteration is that it helps reducing the total computational load. Instead of calculating the full range of time steps within T_{max} in every single iteration, the necessary time range to simulate is


Target pulse
Optimized pulse a
Optimized pulse b
a
b
Envelope (a. u.)
Tmax
Time (a. u.)

(a) Evaluation of two optimization iterations using a fixated target pulse.


Target pulse a
Optimized pulse a
Target pulse b
Optimized pulse b
τ
a
b
Envelope (a. u.)
Twa
Twb Tmax
Time (a. u.)

(b) Evaluation of two optimization iterations using a floating target pulse. An evaluation window is set in respect to the amount of the delay τ.

Figure 7.2: The delay variable and the evaluation window are introduced to improve the flexibility of the prescribed target pulse.

effectively reduced to the size of the evaluation window T_w.

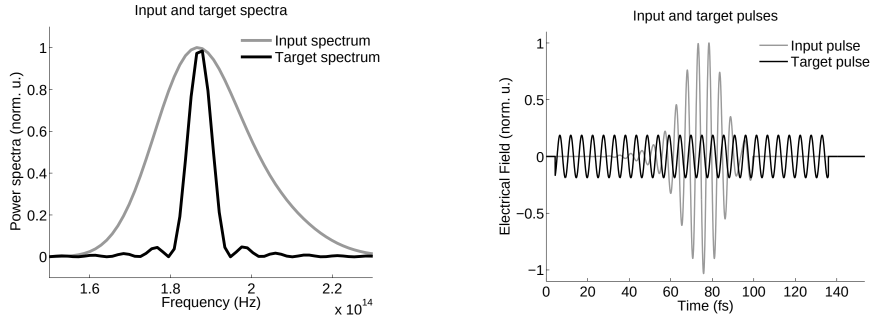
The sensitivities of the delay variable can be calculated by taken a first-order finite difference approximation of the partial derivative of the objective function to τ:

$$\frac{\partial F}{\partial \tau} \approx \frac{\int\limits_{t=0}^{T_w} [\bar{E}_x(t) - \bar{G}(t,\tau+\Delta\tau)]^2 dt - \int\limits_{t=0}^{T_w} [\bar{E}_x(t) - \bar{G}(t,\tau-\Delta\tau)]^2 dt}{2\Delta\tau} \tag{7.3}$$

7.2.3 Transmission efficiencies for the filters

For pulse-shaping filters, low transmission loss is a desirable effect. High transmission losses make it necessary to stage the pulse-shaping filters with amplifiers. The non-linearities associated with optical amplifiers could introduce pulse-shape deform in the input pulse, if the signal is to be amplified prior to entering the filter; or it could directly deteriorate the output pulse if the amplifier is connected after the filter.

Since our pulse filtering is a linear process, the maximum transmitted pulse energy is limited by the spectral shape and amplitude of the input pulse. The spectrum of the optimized pulse has to be confined under the spectrum of the input pulse. The amplitude of the target pulse can be determined by capping its spectrum's peak under that of the input pulse spectrum. Figure 7.3 illustrates the spectrum as well as the electric field of the square-wave target pulse whose amplitude is limited by the power spectrum of the input pulse.



(a) The power spectra for the input and target pulse (normalized by the peak of the input spectrum).

(b) The input pulse (a Gaussian pulse) and the target pulses (harmonic waves modulated by a square function). The pulses' amplitudes are normalized to the peak of the input pulse).

Figure 7.3: Target pulse energy limited by the input pulse spectrum.

7.2.4 Minimum length-scale control and black/white design

Various filters were tested for the 2D optimization problem, and Modified Heaviside filters were chosen (see section 6.6) Although such filtering scheme does not converge well for 1D problems, we find it fitting for the 2D problems. This is mainly due to the increased flexibilities of positioning various lighter scatters in a two-dimensional design domain instead of having to force all the light going through the same scatters as in the 1D problems. The β factor for the filter starts at a small value of 0.2 where the projection between the filter densities $\tilde{\rho}$ and the real physical densities are almost linear. β is doubled every 300 iterations until β is 500.

7.2.5 Results

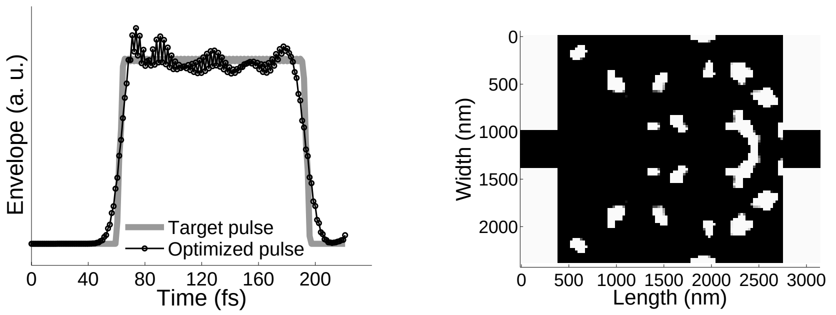
The results for the square-pulse filter optimization where the target pulse amplitude is limited by the power spectrum of the input pulses are shown in Fig. 7.4. The measure of discreteness (see section 6.1 for definition) for the optimized design is 0.79%, which indicates an almost black and white topology. The minimum feature in the topology is an element of the size of 3×4, which corresponds to an area of the size $78.9nm \times 105.2nm$.

The transmission efficiency η is computed as the ratio between the pulse energy of the output pulse E_o and the input pulse E_i:

$$\eta = \frac{\int\limits_{t=0}^{T_{max}} E_o^2(t)dt}{\int\limits_{t=0}^{T_{max}} E_i^2(t)dt} \tag{7.4}$$

The transmission efficiency for the optimized structure is 35.2%. Among the transmitted energy, 96.6% of it is under the main square wave rather than the tail waves. The optimization takes 5505 iterations to converge as β gradually increases to the value of 500. It takes a total time of 28 hours and 31 minutes in *DTU FOTONIK* cluster built on *Dual-Core AMD Opteron(tm) Processor 2218*, with the allocation of $12GB$ memory.

The optimized structure is also tested for input pulses with wavelength shifts to investigate the robustness of the design. The input pulse carrier wavelength is shifted upwards to $1582nm$ and also downwards to $1577nm$ with the same amount of periods as in the original input pulse. The resulting output pulses for the shifted frequencies see small deviations from the output pulse with center



(a) The optimized pulse (black line with circle markers) and the target pulse (solid grey line).

(b) Optimized topology with silicon (black) and air (white).

Figure 7.4: Optimization results for the square-pulse filter with spectral-limited amplitude.

frequency in Fig. 7.5. The object function values are evaluated and compared for the three different frequencies in Table 7.1.

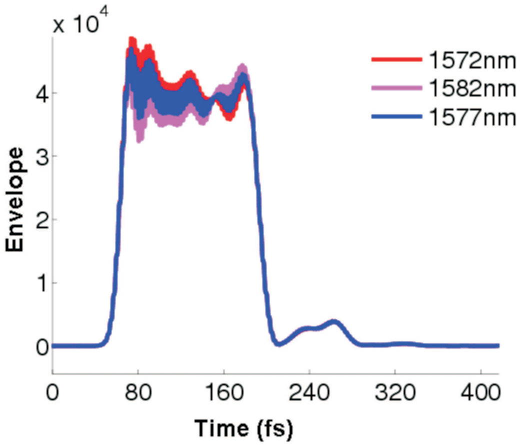


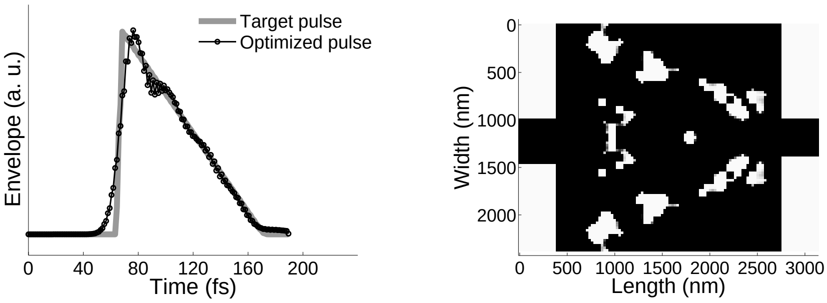
Figure 7.5: Ouput pulses with 3 different input pulses' carrier wavelengths: 1572nm (red), 1577nm (blue) and 1582nm (magenta).

Table 7.1: Objective function value for input pulses with 3 different center wavelengths

Center wavelength (nm)	Objective function (log(F))
1572	29.6668
1577	27.3734
1582	28.9686

7.3 Saw-tooth filters

A saw-tooth pulse is an asymmetric triangular pulse with one of the slopes ascending/descending much more steeply than the other. For our optimization problem, a Gaussian pulse with a FWHM width of $26fs$ centered at $1577nm$ is sent through a rectangular design domain. The size of the design domain, the input/output ridge waveguides as well as the discretization are configured the same way as the previous square pulse example. The target pulse envelope is prescribed to be a saw-tooth pulse with a sharp rising edge and a slowly falling edge with linear slope. The FWHM of the saw-tooth pulse is set to be $50.25fs$. The optimized results are shown in Fig. 7.6. The final design has a highly discrete design with the measure of discreteness at 0.88%. The transmission efficiency for the optimized filter is 63.2%, where 99.7% of the energy is under the saw-tooth wave envelope instead of the tail waves.



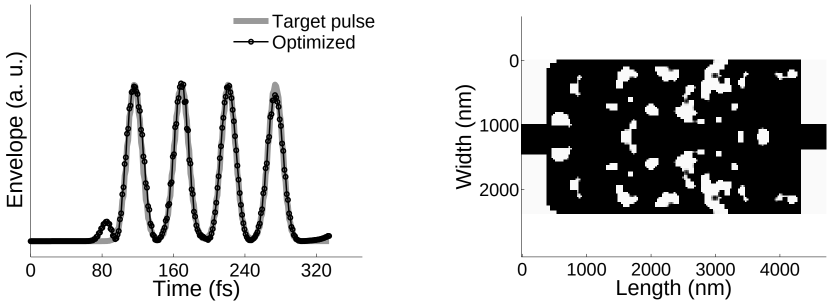
(a) The optimized pulse (black line with circle markers) and the target pulse (solid grey line).

(b) Optimized topology with silicon (black) and air (white).

Figure 7.6: Optimization results for the saw-tooth filter.

7.4 Pulse-splitting filters

The input pulse is a Gaussian pulse with FWHM of $26fs$ centered at $1577nm$. The target is a pulse train constituted of 4 pulses each with $20.4fs$ FWHM, and spaced $52.6fs$ apart. The design domain is a rectangular region with the area of $3.94\mu m \times 2.39\mu m$. The discretized grid spacing is $26.3nm$, which results in a total of 6900 design variables in the half space modeled and optimized. The final design and its performance are illustrated in Fig. 7.7. The topology is almost all black and white, with the measure of discreteness 0.91%.



(a) The optimized pulse (dotted black) and the target pulse (solid grey).

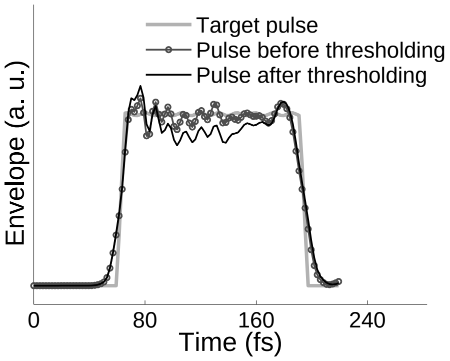
(b) Optimized topology with silicon (black) and air (white).

Figure 7.7: Optimization results for the pulse-splitting filter.

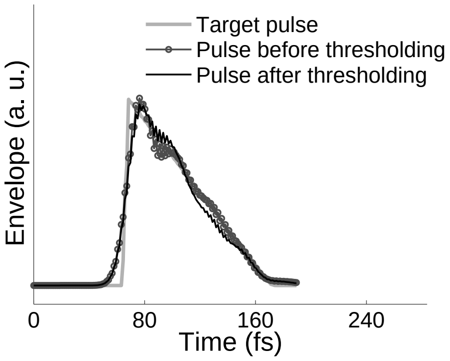
7.5 Thresholded performance

To study the optimized structures more closely, we examine their performances after applying a thresholding procedure on them. In this process, the grey elements still existent in the optimized designs are completely removed. If a grey element has a permittivity larger than 6.6882 (exact midway between silicon and air), it is replaced with a silicon element. Similarly, if a grey element has a permittivity smaller than that, it is treated as an air element instead. The structures are then re-evaluated and compared with the pulse-filtering performance of the original designs, as well as the target pulses (see Fig. 7.8).

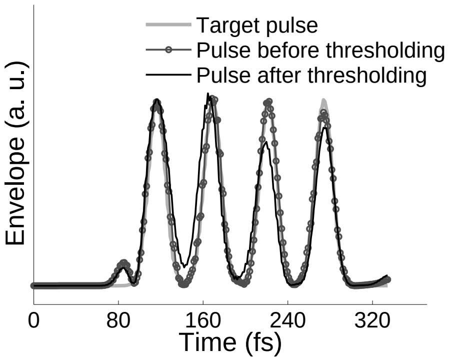
From the comparisons, it is clear that the thresholded structures exhibit qualitatively the same performances as the original designs. However, small distortions do occur. For example, the top of the square pulse experience a slight dip compared to the flatness of the original optimized pulse. The third peak of the pulse train shows a $1.1dB$ drop for the thresholded topology. These changes are due to the highly sensitive nature of these devices to small variations of local scatterers.



(a) Square-pulse filter performance after thresholding.



(b) Saw-tooth filter performance after thresholding.



(c) Pulse-splitting filter performance after thresholding.

Figure 7.8: Comparisons of pulse shaping performances of the thresholded structures and the original designed topology for the square-pulse filter, the saw-tooth filter, and the pulse-splitting filter.

7.6 Conclusions

In this chapter, three different pulse-shaping filters (square wave, saw-tooth wave and pulse-splitting) are optimized using topology optimization based on 2D finite-difference time-domain method. By distributing silicon and air, the optimization successfully generate topologies that fulfill the pulse filtering functionalities. The filtered pulses show good resemblance to the prescribed target pulse, indicating effective optimizations. For square wave filters, it is shown that the optimized structures are quite robust to small variations in the input pulse carrier wavelength. The final designs are highly discrete and almost void of stand-alone elements, showing good minimum lengthscale control.

CHAPTER 8

Conclusions and future work

Conclusions

This thesis exploits topology optimization in both frequency domain and time domain to design nanophotonic devices.

- Slow light couplers

Small group velocities of light due to flat dispersion curve near the band cutoff in photonic crystals waveguides have exhibited the potential in various optical function blocks. However, the impedance mismatch between the ridge waveguide and the photonic crystal waveguide largely prohibits efficient slow light coupling. By using topology optimization based on frequency domain Helmholtz equations, new coupling topologies adjacent to the coupling area are designed and tested. It is demonstrated, both numerically and experimentally, that the optimized structures have improved coupling efficiencies by between 2dB to 5dB, depending on the initial reference structure. For PhCWs based on ring holes, it is also shown that the optimized design has quite robust performance when the ring sizes fluctuate around the original value.

- Pulse-shaping filters based on 1D gratings

1D grating-based square pulse shaping filters are designed using topology optimization based on finite-difference time-domain method. The methodology was tested on the design of Bragg gratings and shows consistency with theoretical predictions. An objective function based on envelope function instead of the field values is designed to tackle the local minima problem in optimizing wave packets. Explicit penalization is used to control the discreteness of the design as well as the minimum lengthscale. With the extra penalization scheme, bad convergence is observed where the optimization focus shifts away from fulfilling the pulse shaping functionality. Multiple optimizations starting from random initial guesses proved to be effective in finding good topologies by searching the solution space more thoroughly.

- Pulse shaping filters based on 2D SOI structures

3 different pulse-shaping functionalities, i.e. square pulse filters, saw-tooth pulse filters and pulse-splitting filters, are optimized by using topology optimization based on 2D finite-difference time-domain method. Apart from geometrical design variables, a new delay variable is introduced to freely position the target pulse as the optimization see fit, giving the optimization more flexibility. Modified Heaviside filters are utilized to give highly discrete designs with reasonable lengthscale controls. The design for square pulse is shown perform robustly under small perturbations to the input pulse carrier wavelength.

Future work

- Real-world device modeling

2D simulations of planar photonic devices have been shown to predict device performance efficiently in the past. However, to better predict the device performance, the modeling should be able to take into account the vertical propagation losses for these devices, including our pulse-shaping filters. Due to the large number of iterations necessary for topology optimizations combined with the high computational cost of full 3D FDTD simuations, alternative measures should be preferred to model those losses instead of a full vectorial simulation.

- Robustness designs

Over- and under-etching are common issues in fabricating planar SOI devices using e-beam lithography. These factors result in the topological deviations of the fabricated device from the original designs, giving rise to performance degrade. To find topologies whose performances are more insensitive to the small variations of material boundaries, a robust design method can be applied to generate more manufacturing tolerant designs [101].

APPENDIX A

Sensitivity analysis for topology optimizations based on finite-difference time-domain method

In this appendix, we derive the sensitivity expressions by using FDTD-based adjoint-variable analysis on structured grids for 1D problems. For 2D problems defined in section 7.1 where E_x becomes the dominant field component, E_z can be ignored in evaluation of the output pulse and the sensitivity analysis. This simplification makes the 1D sensitivity analysis recipe also applicable in the 2D case.

A.1 Sensitivities for 1D problems

A.1.1 Problem formulation

First, let us revisit the 1D problem formulation at the beginning of chapter 5. The goal of the 1D topology optimization is to distribute two dielectric materials (ε_r^1 and ε_r^2) within the design domain (Fig. A.1) to make a grating structure

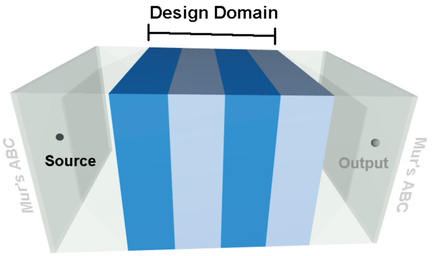


Figure A.1: 1D grating optimization. The input pulse is excited at the source node (black) and the output pulse at the objective node (grey) is measured and evaluated.

that fulfills certain transmission functions between the source node (green) and the output node (red). For example, one can minimize the transmission at certain frequencies to design a special filter, or to alter the temporal shape of the output pulse. Depending on different purposes, a specific objective function $F(\boldsymbol{\rho})$ is prescribed.

The relationship between the material property and the design variable for a specific grid cell is expressed as:

$$\varepsilon_r^i = \varepsilon_r^0 + (\varepsilon_r^d - \varepsilon_r^0)\rho_i, (0 \leq \rho_i \leq 1), \tag{A.1}$$

where

ε_r^i is the relative dielectric constant for element i,
ε_r^0 is the relative dielectric constant for the lower index material,
ε_r^d is the relative dielectric constant for the higher index material, and
ρ^i is the i^{th} design variable.

In the following example, the objective function F of the measured fields is defined as to minimize the local measurement f at the objective node when integrated over the complete analysis time T_{max}:

$$F(E, \boldsymbol{\rho}) = \int_0^{T_{\max}} f(E, \boldsymbol{\rho}) dt. \tag{A.2}$$

The local measurement can be decided upon specific optimization criteria, *e.g.* the transmitted energy.

A.1.2 Definition of sensitivities

The sensitivities of the objective function to design parameter variations are defined by the gradient of F in the design parameter $\boldsymbol{\rho}$ space:

$$\nabla_{\boldsymbol{\rho}} F = \left[\frac{\partial F}{\partial \rho_1}, \frac{\partial F}{\partial \rho_2}, \cdots \frac{\partial F}{\partial \rho_N} \right]. \tag{A.3}$$

It is a row vector with the size of the number of grid cells in the design domain, and can be calculated by two methods: finite difference method and analytical sensitivity method. For the sensitivity of a specific design variable, the derivative is:

$$\frac{\partial F}{\partial \rho_i} \approx \frac{\Delta_i F}{\Delta_i \rho}. \tag{A.4}$$

The variation $\Delta_i F$ in Eqn. A.4 can be expressed as

$$\Delta_i F = \Delta_i^e F + \int_0^{T_{\max}} \left(\frac{\partial f}{\partial E} \cdot \Delta_i E\right) dt, \tag{A.5}$$

where E is the field solution, and the superscript e in $\Delta_i^e F$ denotes the variation related to the explicit dependence on ρ_i. For example, when there is perturbation at the observation node (*i.e.* a small change of the local dielectric constant), f will have a change that is explicitly depending on this perturbation. $\int_0^{T_{\max}} (\frac{\partial f}{\partial \mathrm{E}} \cdot \Delta_i E) dt$ is in correspondence to the implicit variation of the objective function defined in Eqn. A.2 via the field solutions.

By taking Eqn. A.5 into Eqn. A.4, we have the sensitivity of a certain design variable as:

$$\frac{\Delta_i F}{\Delta_i \rho} \approx \frac{\Delta_i^e F}{\Delta_i \rho} + \frac{\int_0^{T_{\max}} (\frac{\partial f}{\partial E} \cdot \Delta_i E) dt}{\Delta_i \rho} \tag{A.6}$$

A.1.3 The finite difference method for calculating sensitivities

The most straight forward way to get the sensitivities is to use the Finite Difference Method, where the material property of every individual element is perturbed to result in a small change in the objective function value. The sensitivity for a specific element in the design domain is defined as:

$$\nabla_{\boldsymbol{\rho}}^{i} F = \frac{F(\rho_i + \Delta\rho_i)}{\Delta\rho_i}, \tag{A.7}$$

where $\Delta\rho_i$ is a very small perturbation around $10^{-7} \sim 10^{-4}$.

The Finite Difference Method is an accurate way to calculate the sensitivities, but for a design domain with N elements, $N+1$ system analyses would have to be carried out to calculate the full vector of $\nabla_{\boldsymbol{\rho}} F$. This makes the method very cumbersome to execute, especially when one single system analysis by $FDTD$ is already computationally heavy.

A.1.4 1D sensitivity analysis by using the adjoint-variable method

The sensitivities can be calculated analytically by using only 2 system analysis, based on the known theory of adjoint-variable analysis [82],[102]. Note that in our case, though the expression for calculating sensitivities is derived analytically, the field values used in the equations are still the numerical results from the $FDTD$ simulation.

A.1.4.1 Derivation of the implicit sensitivity term

Maxwell's coupled equations are mathematically equivalent to the non-coupled second-order wave equation for the E-field:

$$\frac{1}{\mu}\frac{\partial^2 E_z}{\partial x^2} - \varepsilon\frac{\partial^2 E_z}{\partial t^2} = \frac{\partial J_s}{\partial t}. \tag{A.8}$$

After discretization, Eqn. A.8 can be reduced to the linear matrix form:

$$\mathbf{M}\ddot{\boldsymbol{\theta}} + \mathbf{K}\boldsymbol{\theta} = \mathbf{Q}, \tag{A.9}$$

with initial conditions

$$\boldsymbol{\theta}(0) = 0 \; and \; \dot{\boldsymbol{\theta}}(0) = 0, \tag{A.10}$$

where $\boldsymbol{\theta}$ is a row vector of the field values E_z at all elements, and $\dot{\boldsymbol{\theta}}$ and $\ddot{\boldsymbol{\theta}}$

are the first and second time derivatives of $\boldsymbol{\theta}$ respectively. Equation A.9 is a linear matrix form of the wave equation common in the finite element (FEM) formulation. $\mathbf{M}$ and $\mathbf{K}$ are global system matrices assembling the material properties of all the grid cells in the calculation domain, while $\mathbf{Q}$ is the global matrix for depicting the excitation condition. These global matrices are in correspondence to the terms in Eqn. A.8 through the following relations:

$$\begin{gathered}\mathbf{M}\ddot{\boldsymbol{\theta}} \rightarrow -\varepsilon\tfrac{\partial^2 E_z}{\partial t^2},\ \mathbf{M} \text{ is depending on } \varepsilon\ ; \\ \mathbf{K}\theta \rightarrow \tfrac{1}{\mu}\tfrac{\partial^2 E_z}{\partial x^2},\ \mathbf{K} \text{ is depending on } \tfrac{1}{\mu}\ ; \text{ and} \\ \mathbf{Q} \rightarrow \tfrac{\partial J_s}{\partial t}.\end{gathered}$$

In this notation, Eqn. A.5 is rewritten as

$$\Delta_i F = \Delta_i^e F + \int_0^{T_{\max}} (\frac{\partial f}{\partial \boldsymbol{\theta}} \cdot \Delta_i \boldsymbol{\theta}) dt. \tag{A.11}$$

When the system equation Eqn. A.9 is perturbed at the i^{th} design variable by $\Delta\rho_i$, it becomes:

$$(\mathbf{M} + \Delta_i \mathbf{M})(\ddot{\boldsymbol{\theta}} + \Delta_i \ddot{\boldsymbol{\theta}}) + (\mathbf{K} + \Delta_i \mathbf{K})(\boldsymbol{\theta} + \Delta_i \boldsymbol{\theta}) = \mathbf{Q} + \Delta_i \mathbf{Q} \tag{A.12}$$

By expanding equation Eqn. A.12, we have:

$$(\mathbf{M}\ddot{\boldsymbol{\theta}} + \mathbf{K}\boldsymbol{\theta}) + (\Delta_i \mathbf{M}\ddot{\boldsymbol{\theta}} + \Delta_i \mathbf{K}\boldsymbol{\theta}) + (\mathbf{M} + \Delta_i \mathbf{M})\Delta_i \ddot{\boldsymbol{\theta}} + (\mathbf{K} + \Delta_i \mathbf{K})\Delta_i \boldsymbol{\theta} = \mathbf{Q} + \Delta_i \mathbf{Q}. \tag{A.13}$$

Define the residual term R as:

$$\mathbf{R} = \mathbf{M}\ddot{\boldsymbol{\theta}} + \mathbf{K}\boldsymbol{\theta} - \mathbf{Q}, \tag{A.14}$$

and replace the system equation Eqn. A.9 into Eqn. A.13, we have:

$$\tilde{\mathbf{M}}_i \Delta_i \ddot{\boldsymbol{\theta}} + \tilde{\mathbf{K}}_i \Delta_i \boldsymbol{\theta} + \Delta_i \mathbf{R} = 0, \tag{A.15}$$

where

$$\begin{aligned} &\tilde{\mathbf{M}}_i \Delta_i = \mathbf{M} + \Delta_i \mathbf{M}, \\ &\tilde{\mathbf{K}}_i \Delta_i = \mathbf{K} + \Delta_i \mathbf{K},\ \mathit{and} \\ &\Delta_i \mathbf{R} = \Delta_i \mathbf{M}\ddot{\boldsymbol{\theta}} + \Delta_i \mathbf{K}\boldsymbol{\theta} - \Delta_i \mathbf{Q}. \end{aligned}$$

Now let us go back to the sensitivity expression in Eqn. A.11. The implicit dependance of both $\frac{\partial f}{\partial \boldsymbol{\theta}}$ and $\Delta_i \boldsymbol{\theta}$ on the design variable makes it difficult to calculate the sensitivity value in a straightforward way. That is why an auxiliary row vector $\boldsymbol{\lambda}$ is introduced to get rid of the $\Delta_i \boldsymbol{\theta}$ term. $\boldsymbol{\lambda}$ has the same and size as $\boldsymbol{\theta}$ and is called the adjoint variable. Its value is unknown for now, but its solution will be derived later on.

We first pre-multiply Eqn. A.15 with $\boldsymbol{\lambda}$. The product is then integrated over the observation time, just as needed in calculating the objective function:

$$\int_0^{T_{\max}} \boldsymbol{\lambda}_i^T \cdot (\tilde{\mathbf{M}}_i \Delta_i \Delta_i \ddot{\boldsymbol{\theta}} + \tilde{\mathbf{K}}_i \Delta_i \Delta_i \boldsymbol{\theta}) dt = - \int_0^{T_{\max}} \boldsymbol{\lambda}_i^T \cdot \Delta_i \mathbf{R} dt, \tag{A.16}$$

Next we integrate in time the term $\boldsymbol{\lambda}_i^T \cdot \tilde{\mathbf{M}}_i \Delta_i \ddot{\boldsymbol{\theta}}$ by parts twice to remove $\ddot{\boldsymbol{\theta}}$. It results in:

$$\begin{aligned} \boldsymbol{\lambda}_i^T \cdot \tilde{\mathbf{M}}_i \Delta_i \dot{\boldsymbol{\theta}} \Big|_0^{T_{\max}} - \dot{\boldsymbol{\lambda}}_i^T \cdot \tilde{\mathbf{M}}_i \Delta_i \boldsymbol{\theta} \Big|_0^{T_{\max}} + \int_0^{T_{\max}} (\ddot{\boldsymbol{\lambda}}_i^T \tilde{\mathbf{M}}_i + \boldsymbol{\lambda}_i^T \tilde{\mathbf{K}}_i) \cdot \Delta_i \boldsymbol{\theta} \cdot dt \\ = - \int_0^{T_{\max}} \boldsymbol{\lambda}_i^T \cdot \Delta_i \mathbf{R} dt. \end{aligned} \tag{A.17}$$

The first two terms in the above equation are 0 at $t = 0$, since $\dot{\boldsymbol{\theta}}$ and $\ddot{\boldsymbol{\theta}}$ are 0 at $t = 0$ as in Eqn. A.10. We can further eliminate the terms at $t = T_{max}$ too by choosing $\boldsymbol{\lambda}$ to be:

$$\begin{aligned} \boldsymbol{\lambda}_i^T (T_{\max}) &= 0, \\ \dot{\boldsymbol{\lambda}}_i^T (T_{\max}) &= 0. \end{aligned} \tag{A.18}$$

We can also choose values for $\boldsymbol{\lambda}$ to satisfy

$$\ddot{\boldsymbol{\lambda}}_i^T \tilde{\mathbf{M}}_i + \boldsymbol{\lambda}_i^T \tilde{\mathbf{K}}_i = \frac{\partial f}{\partial \boldsymbol{\theta}}, \tag{A.19}$$

so that combined with Eqn. A.18, Eqn. A.17 finally becomes

$$\int_{t=0}^{T_{\max}} \frac{\partial f}{\partial \boldsymbol{\theta}} \Delta_i \boldsymbol{\theta} \cdot dt = -\int_{t=0}^{T_{\max}} \boldsymbol{\lambda}_i^T \cdot \Delta_i \mathbf{R} \cdot dt \tag{A.20}$$

Noticing that the left hand side of the above equation is also the implicit sensitivity term in Eqn. A.5, the sensitivity expression in Eqn. A.6 can now be rewritten as:

$$\frac{\Delta_i F}{\Delta_i \rho} \approx \frac{\Delta_i^e F}{\Delta_i \rho} - \int_0^{T_{\max}} \boldsymbol{\lambda}_i^T \cdot \frac{\Delta_i \mathbf{R}}{\Delta_i \rho}(\boldsymbol{\theta}) dt, \tag{A.21}$$

where the derivative residual is:

$$\frac{\Delta_i \mathbf{R}}{\Delta_i \rho}(\boldsymbol{\theta}) \approx \frac{\Delta_i \mathbf{M} \ddot{\boldsymbol{\theta}}}{\Delta_i \rho} + \frac{\Delta_i \mathbf{K} \boldsymbol{\theta}}{\Delta_i \rho} - \frac{\Delta_i \mathbf{Q}}{\Delta_i \rho} \tag{A.22}$$

Note that the field $\boldsymbol{\theta}$ and its derivatives $\dot{\boldsymbol{\theta}}$ and $\ddot{\boldsymbol{\theta}}$ remain constant during the differentiation of system matrices $\mathbf{M}$ and $\mathbf{K}$. In problems where the source node is not part of the design domain, $\frac{\Delta_i \mathbf{Q}}{\Delta_i \rho}$ to 0.

In the following sections, we are going to derive the detailed calculations of $\boldsymbol{\lambda}$ from the adjoint analysis as well as the derivative residual term $\frac{\Delta_i \mathbf{R}}{\Delta_i \rho}(\boldsymbol{\theta})$, which are both necessary to arrive at the analytical sensitivity values in Eqn. A.21.

A.1.4.2 Derivative residual

In this section, the matrix derivative residual term $\frac{\Delta_i \mathbf{R}}{\Delta_i \rho}(\boldsymbol{\theta})$ is going to be calculated by evaluating $\frac{\Delta_i \mathbf{R}}{\Delta_i \rho}(E_z)$ at every grid cell.

Using central finite difference approximation with spatial interval Δx and time step Δt, the residual for the 1D wave equation in Eqn. A.8 is discretized as:

$$\Delta x^2 D_{xx} E_z - \alpha \cdot D_{tt} E_z - \beta \cdot D_t J_s \tag{A.23}$$

where

$$
\begin{aligned}
&D_t J_s = J_s(t+\Delta t) - J_s(t-\Delta t)\\
&\alpha = \mu_r \varepsilon_r (\frac{\Delta x}{c\Delta t})^2\\
&\beta = \frac{2\mu_0 \mu_r \Delta x^2}{\Delta t}\\
&D_{\xi\xi}E_z = E_z(\xi - \Delta\xi) - 2E_z + E_z(\xi + \Delta\xi),\ \xi = (x,t).
\end{aligned}
$$

Thus, we can have the derivative residual term to the design variable perturbation to be:

$$
\frac{\Delta_i \mathbf{R}}{\Delta_i \rho}(E_z) = \frac{\Delta_i \Delta x^2 D_{xx}E_{z,i} - \Delta_i \alpha \cdot D_{tt}E_{y,i} - \Delta_i \beta \cdot D_t J_s}{\Delta_i \rho} \tag{A.24}
$$

Since E_z remains constant during the matrix derivative approximation, the first increment term in the numerator becomes insensitive to the design parameter ρ and is left as 0. The same applies to the third term, which has no dependence on the material property ε_r^i at all. So Eqn. A.24 simply becomes:

$$
\begin{aligned}
\frac{\Delta_i \mathbf{R}}{\Delta_i \rho}(E_z) &= \frac{-(\Delta_i \alpha)\cdot D_{tt}E_{y,i}}{\Delta_i \rho}\\
&= -\frac{\Delta_i \mu_r \varepsilon_r^i (\frac{\Delta x}{c\Delta t})^2}{\Delta_i \rho}\cdot D_{tt}E_{y,i}\\
&= -\frac{\Delta_i \mu_r [\varepsilon_r^0 + (\varepsilon_r^d - \varepsilon_r^0)\rho_i](\frac{\Delta x}{c\Delta t})^2}{\Delta_i \rho}\cdot D_{tt}E_{y,i}\\
&= -\mu_r(\varepsilon_r^d - \varepsilon_r^0)(\frac{\Delta x}{c\Delta t})^2 \cdot D_{tt}E_{y,i}
\end{aligned} \tag{A.25}
$$

where $E_{y,i}$ is the field history at i^{th} design element.

Equation A.25 shows that the residual derivative to the i^{th} is the product of a constant containing the material parameters ε_r^d, ε_r^0 and μ_r as well as the second time derivative of the field history E_z at grid cell i.

A.1.4.3 The adjoint problem

By recapturing equations Eqn. A.18 and Eqn. A.19, one can see that the adjoint variable $\boldsymbol{\lambda}$ must satisfy:

$$\ddot{\boldsymbol{\lambda}}_i^T \tilde{\mathbf{M}}_i + \boldsymbol{\lambda}_i^T \tilde{\mathbf{K}}_i = \frac{\partial f}{\partial \boldsymbol{\theta}},$$

as well as

$$\boldsymbol{\lambda}_i^T (T_{\max}) = 0 \ \ and \ \ \dot{\boldsymbol{\lambda}}_i^T (T_{\max}) = 0.$$

It is straightforward to see that the above two equations very much resemble the original system equations Eqn. A.9 and Eqn. A.10 for solving second-order wave equation, with 2 major discrepancies:

a) The system matrices $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{K}}$ are perturbed with $\Delta_i \rho$, and there are N different adjoint variable, each for a specific design variable ρ_i;

b) The boundary conditions are set for the terminal at $t = T_{\max}$ instead of at $t = 0$.

In order to calculate the values for $\boldsymbol{\lambda}$, we will first approximate the perturbed system matrices $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{K}}$ with the unperturbed $\mathbf{M}$ and $\mathbf{K}$, thus we will have only one adjoint variable $\boldsymbol{\lambda}$ and the subscript is dropped. It has been shown [83] that given the relative fine mesh and small perturbations, this approximation comes with very good accuracy.

Based on the properties of matrix transpose:

$$A^T B = (B^T A)^T \tag{A.26}$$

we rewrite Eqn. A.19 with the unperturbed matrices:

$$\mathbf{M}^T \ddot{\boldsymbol{\lambda}} + \mathbf{K}^T \boldsymbol{\lambda} = (\frac{\partial f}{\partial \boldsymbol{\rho}})^T \tag{A.27}$$

By comparing the above equation with Eqn. A.9 and bear in mind that $\mathbf{M}$ and $\mathbf{K}$ are symmetric matrices, we can conclude that Eqn. A.27 is also the solution to the same wave equation in Eqn. A.8, except that the $FDTD$ simulation should run backward in time due to the terminal values determined in Eqn. A.18.

Next, we quote the 1D $FDTD$ update equations Eqn. 3.29a and Eqn. 3.29ain section 3.3.3 without magnetic sources:

$$\begin{aligned} H_y\Big|_{i+1/2}^{n+1/2} &= \left(\frac{2\mu_{i+1/2} - {\sigma^*}_{i+1/2}\Delta t}{2\mu_{i+1/2} + {\sigma^*}_{i+1/2}\Delta t}\right) H_y\Big|_{i+\frac{1}{2}}^{n-\frac{1}{2}} \\ &+ \frac{2\Delta t}{2\mu_{i+\frac{1}{2}} + {\sigma^*}_{i+1/2}\Delta t}\left(\frac{E_z\,|_{i+1}^{n} - E_z\,|_{i}^{n}}{\Delta}\right) \end{aligned} \tag{A.28a}$$

$$\begin{aligned} E_z\,|_{i}^{n+1} &= \left(\frac{2\varepsilon_i - \sigma_i\Delta t}{2\varepsilon_i + \sigma_i\Delta t}\right) E_z\,|_{i}^{n} \\ &+ \frac{2\Delta t}{2\varepsilon_i + \sigma_i\Delta t}\left(\frac{H_y\,|_{i+1}^{n} - H_y\,|_{i}^{n}}{\Delta}\right. \\ &\left. - J_{source_y}\,|_{i}^{n+1}\right) \end{aligned} \tag{A.28b}$$

Based on the above equations, we solve for $H_y\Big|_{i+\frac{1}{2}}^{n-\frac{1}{2}}$ and $E_z\,|_i^n$ respectively instead of $H_y\Big|_{i+1/2}^{n+1/2}$ and $E_z\,|_i^{n+1}$, and also change the sign of the lossy term σ and σ^* since the computation now runs backward in time. This give us the backward update equations:

$$\begin{aligned} \hat{H}_y\Big|_{i+1/2}^{n-1/2} &= \left(\frac{2\mu_{i+1/2} - {\sigma^*}_{i+1/2}\Delta t}{2\mu_{i+1/2} + {\sigma^*}_{i+1/2}\Delta t}\right) \hat{H}_y\Big|_{i+\frac{1}{2}}^{n+\frac{1}{2}} \\ &- \frac{2\Delta t}{2\mu_{i+\frac{1}{2}} + {\sigma^*}_{i+1/2}\Delta t}\left(\frac{\hat{E}_z\,|_{i+1}^{n} - \hat{E}_z\,|_{i}^{n}}{\Delta}\right) \end{aligned} \tag{A.29a}$$

$$\begin{aligned} \hat{E}_z\,|_{i}^{n} &= \left(\frac{2\varepsilon_i - \sigma_i\Delta t}{2\varepsilon_i + \sigma_i\Delta t}\right) \hat{E}_z\,|_{i}^{n+1} \\ &- \frac{2\Delta t}{2\varepsilon_i + \sigma_i\Delta t}\left(\frac{\hat{H}_y\,|_{i+1}^{n} - \hat{H}_y\,|_{i}^{n}}{\Delta}\right. \\ &\left. + \hat{J}_{source_y}\,|_{i}^{n+1}\right) \end{aligned} \tag{A.29b}$$

The comparison of the above two systems of update equations shows that we can use the original $FDTD$ update scheme to solve for $(\hat{H}, -\hat{E})$ in the backward analysis, starting from the terminal values at $t = T_{\max}$ and running backward to $t = 0$. The excitation for the backward scheme is the adjoint current $\hat{J}_S$. Its values will be determined in the following section.

A.1.4.4 The adjoint current

The adjoint current is computed according to the correspondence between $\frac{\partial f}{\partial \boldsymbol{\theta}}$ in the adjoint problem definition Eqn. A.19 and $\beta \cdot D_t \hat{J}_S$ in the discretized wave equations of Eqn. A.23 at every time step and grid cell:

$$\beta \cdot D_t \hat{J}_S = \frac{\partial f}{\partial E_z}. \tag{A.30}$$

From the above equation we can see that the adjoint current depends on the definition of local measurement f for individual optimization problems. For the simplest case, we determine f to be:

$$f = \frac{E^2_{z,objective}}{2}, \tag{A.31}$$

which means the goal of the optimization is to minimize the time-integrated transmitted energy at the objective node.

By expanding Eqn. A.30 at grid cell i and time n, we have

$$\beta \cdot (\hat{J}^{n+1}_{S,i} - \hat{J}^{n}_{S,i}) = \frac{\partial f}{\partial E_z} |_{i,n} \tag{A.32}$$

Take Eqn. A.31 into the above equation, it becomes:

$$\begin{aligned}\beta \cdot (\hat{J}^{n+1}_{S,i} - \hat{J}^{n}_{S,i}) &= \frac{\partial (E^2_{z,objective}/2)}{\partial E_z} |_{i,n} \\ &= \left\{ \begin{array}{ll} E^n_{z,objective}, & i = objective\ node \\ 0, & i \neq objective\ node. \end{array} \right\}\end{aligned} \tag{A.33}$$

Since the RHS of the above equation is only nonzero at the objective node, it means the adjoint current $\hat{J}_{S,i}$ will only be excited at the object node in the backward analysis:

$$\begin{aligned}\beta \cdot (\hat{J}^{n+1}_{S,objective} - \hat{J}^{n}_{S,objective}) &= E^n_{z,objective} \\ \hat{J}^{n+1}_{S,objective} &= \hat{J}^{n}_{S,objective} + \frac{E^n_{z,objective}}{\beta}, \\ where\ \hat{J}^{0}_{S,objective} &= 0.\end{aligned} \tag{A.34}$$

The above adjoint current update equation shows that the adjoint current value at a certain time step depends on its own value at the previous time step as well as the forward analysis field solutions at the objective node, at the corresponding time.

A.1.4.5 Implementation of sensitivity analysis using the adjoint-variable method

By taking Eqn. A.25 as well as the simulated adjoint variable $\boldsymbol{\lambda}$ into the sensitivity calculation in Eqn. A.21, we can now determine the sensitivity for the i^{th} design variable to be:

$$\frac{\Delta_i F}{\Delta_i \rho} \approx \frac{\Delta_i^e F}{\Delta_i \rho} + \sum_{n=1}^{n_{\max}} \mu_r(\varepsilon_r^d - \varepsilon_r^0)(\frac{\Delta x}{c\Delta t})^2 \cdot D_{tt} E_{z,i}^n \cdot \lambda_i^n \tag{A.35}$$

In most cases where the i^{th} element is not the objective node, the first term in the RHS of the above equation is 0 since the local measurement does not change explicitly so long as the objective node material property stays the same.

A flow chart streamlining the procedures described above is illustrated below:

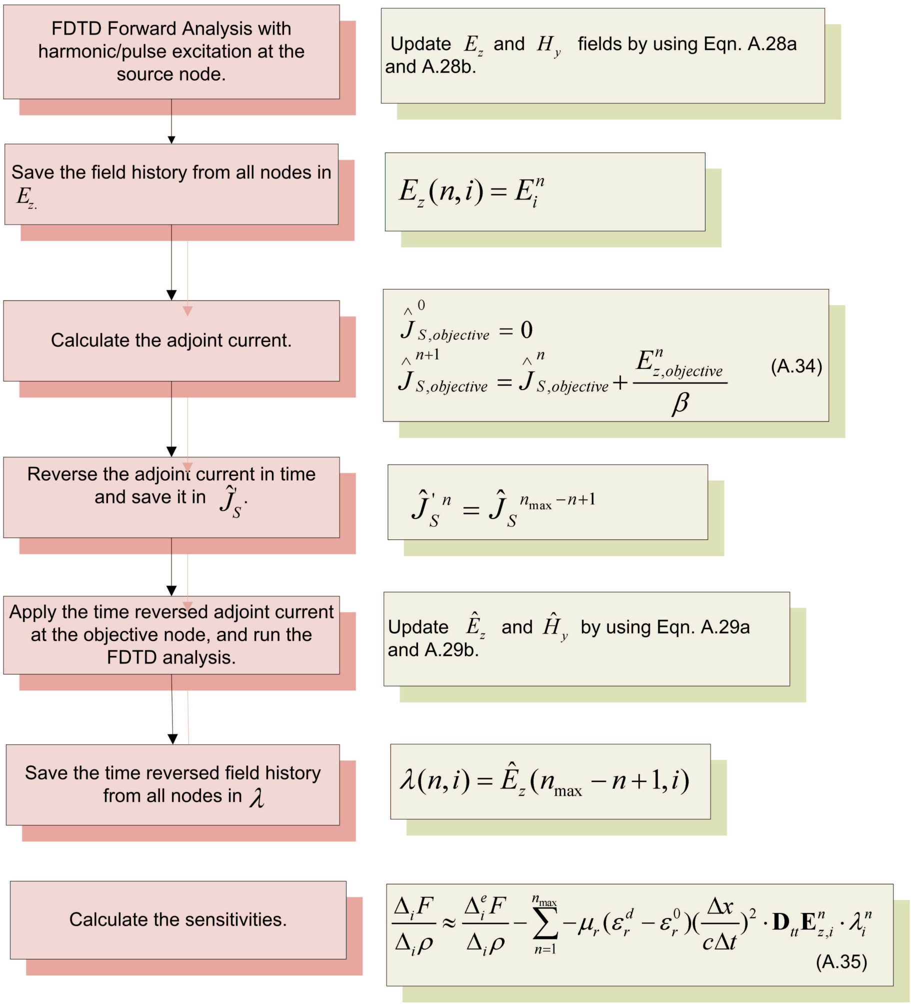

FDTD Forward Analysis with harmonic/pulse excitation at the source node.
Update E_z and H_y fields by using Eqn. A.28a and A.28b.
Save the field history from all nodes in E_z.
$E_z(n,i) = E_i^n$
Calculate the adjoint current.
$\hat{J}^{0}_{S,objective} = 0$
$\hat{J}^{n+1}_{S,objective} = \hat{J}^{n}_{S,objective} + \frac{E^n_{z,objective}}{\beta}$
(A.34)
Reverse the adjoint current in time and save it in $\hat{J}_S'$.
$\hat{J}_S'^{\,n} = \hat{J}_S^{\,n_{\max}-n+1}$
Apply the time reversed adjoint current at the objective node, and run the FDTD analysis.
Update $\hat{E}_z$ and $\hat{H}_y$ by using Eqn. A.29a and A.29b.
Save the time reversed field history from all nodes in λ
$\lambda(n,i) = \hat{E}_z(n_{\max} - n + 1, i)$
Calculate the sensitivities.
$\frac{\Delta_i F}{\Delta_i \rho} \approx \frac{\Delta_i^e F}{\Delta_i \rho} - \sum_{n=1}^{n_{\max}} -\mu_r(\varepsilon_r^d - \varepsilon_r^0)(\frac{\Delta x}{c\Delta t})^2 \cdot \mathbf{D}_{tt}\mathbf{E}^n_{z,i} \cdot \lambda_i^n$
(A.35)

Figure A.2: Flow chart for analytical sensitivity calculation using AVM.

Bibliography

[1] G. Roelkens, D. V. Thourhout, R. Baets, R. Nötzel, and M. Smit. Laser emission and photodetection in an InP/InGaAsP layer integrated on and coupled to a Silicon-on-Insulator waveguide circuit. *Optics Express*, 14(18):8154–8159, 2006.

[2] A. W. Fang, H. Park, O. Cohen, R. Jones, M. J. Paniccia, and J. E. Bowers. Electrically pumped hybrid AlGaInAs-silicon evanescent laser. *Optics Express*, 14(20):9203–9210, 2006.

[3] E. Yablonovitch. Inhibited spontaneous emission in solid-state physics and electronics. *Physics Review Letters*, 58(20):2059–2062, 1987.

[4] S. John. Strong localization of photons in certain disordered dielectric superlattices. *Physical Review Letters*, 58(23):2486, 1987.

[5] C. Hafner, C. Xudong, J. Smajic, and R. Vahldieck. Efficient procedures for the optimization of defects in photonic crystal structures. *Journal of the Optical Society of America A*, 24(4):1177–1188, 2007.

[6] A. Chutinan and S. Noda. Waveguides and waveguide bends in two-dimensional photonic crystal slabs. *Physical Review B*, 62(7):4488, 2000.

[7] M. Palamaru and P. Lalanne. Photonic crystal waveguides: Out-of-plane losses and adiabatic modal conversion. *Applied Physics Letters*, 78(11):1466, 2001.

[8] A. Talneau, P. Lalanne, M. Agio, and C. M. Soukoulis. Low-reflection photonic-crystal taper for efficient coupling between guide sections of arbitrary widths. *Optics Letters*, 27(17):1522–1524, 2002.

[9] I. Ntakis, P. Pottier, and R. M. D. L. Rue. Optimization of transmission properties of two-dimensional photonic crystal channel waveguide bends

through local lattice deformation. *Journal of Applied Physics*, 96(1):12 – 18, 2004.

[10] A. Chutinan, M. Okano, and S. Noda. Wider bandwidth with high transmission through waveguide bends in two-dimensional photonic crystal slabs. *Applied Physics Letters*, 80(10):1698, 2002.

[11] S. Olivier, H. Benisty, M. Rattier, C. Weisbuch, M. Qiu, A. Karlsson, C. J. M. Smith, R. Houdré, and U. Oesterle. Resonant and nonresonant transmission through waveguide bends in a planar photonic crystal. *Applied Physics Letters*, 79(16):2514–2516, 2001.

[12] S. Fan, S. G. Johnson, J. D. Joannopoulos, C. Manolatou, and H. A. Haus. Waveguide branches in photonic crystals. *Journal of the Optical Society of America B*, 18(2):162–165, 2001.

[13] J. Moosburger, M. Kamp, A. Forchel, S. Olivier, H. Benisty, C. Weisbuch, and U. Oesterle. Enhanced transmission through photonic-crystal-based bent waveguides by bend engineering. *Applied Physics Letters*, 79(22):3579 – 3581, 2001.

[14] Y. Zhang and B. Li. Ultracompact waveguide bends with simple topology in two-dimensional photonic crystal slabs for optical communication wavelengths. *Optics Letters*, 32(7):787–789, 2007.

[15] J. Smajic, C. Hafner, and D. Erni. Design and optimization of an achromatic photonic crystal bend. *Optics Express*, 11(12):1378–1384, 2003.

[16] W. J. Kim and J. D. O'Brien. Optimization of a two-dimensional photonic-crystal waveguide branch by simulated annealing and the finite-element method. *Journal of the Optical Society of America B*, 21(2):289–295, 2004.

[17] J. Smajic, C. Hafner, and D. Erni. Optimization of photonic crystal structures. *Journal of the Optical Society of America A*, 21(11):2223–2232, 2004.

[18] D. Erni, D. Wiesmann, M. Spühler, S. Hunziker, E. Moreno, B. Oswald, J. Fröhlich, and C. Hafner. Application of evolutionary optimization algorithms in computational optics. *ACES Journal: Special Issue on Genetic Algorithms*, 15, 2000.

[19] J. S. Jensen and O. Sigmund. Systematic design of photonic crystal structures using topology optimization: Low-loss waveguide bends. *Applied Physics Letters*, 84(12):2022–2024, 2004.

[20] P. Borel, A. Harpøth, L. Frandsen, M. Kristensen, P. Shi, J. Jensen, and O. Sigmund. Topology optimization and fabrication of photonic crystal structures. *Optics Express*, 12:1996–2001, 2004.

[21] L. Frandsen, A. Harpøth, P. Borel, M. Kristensen, J. Jensen, and O. Sigmund. Broadband photonic crystal waveguide 60° bend obtained utilizing topology optimization. *Optics Express*, 12(24):5916–5921, 2004.

[22] A. V. Lavrinenko, A. Têtu, L. H. Frandsen, J. Fage-Pedersen, and P. I. Borel. Optimization of photonic crystal 60° waveguide bends for broadband and slow-light transmission. *Applied Physics B.*, 87(1):53–56, 2007.

[23] J. Jensen, O. Sigmund, L. Frandsen, P. Borel, A. Harpoth, and M. Kristensen. Topology design and fabrication of an efficient double 90° photonic crystal waveguide bend. *Photonics Technology Letters, IEEE*, 17:1202–1204, 2005.

[24] P. Borel, L. Frandsen, A. Harpoth, M. Kristensen, J. Jensen, and O. Sigmund. Topology optimised broadband photonic crystal y-splitter. *Electronics Letters*, 41(2):69–71, 2005.

[25] O. Sigmund and K. Hougaard. Geometric properties of optimal photonic crystals. *Physical Review Letters*, 100(15):153904–4, 2008.

[26] R. Matzen, J. S. Jensen, and O. Sigmund. Topology optimization for transient response of photonic crystal structures. *Journal of the Optical Society of America B*, 27(10):2040–2050, 2010.

[27] T. Nomura, K. Sato, K. Taguchi, T. Kashiwa, and S. Nishiwaki. Structural topology optimization for the design of broadband dielectric resonator antennas using the finite difference time domain technique. *International Journal for Numerical Methods in Engineering*, 71(11):1261–1296, 2007.

[28] A. M. Weiner. Femtosecond pulse shaping using spatial light modulators. *Review of Scientific Instruments*, 71:1929–1960, 2000.

[29] N. F. Scherer, A. J. Ruggiero, M. Du, and G. R. Fleming. Time resolved dynamics of isolated molecular systems studied with phase-locked femtosecond pulse pairs. *The Journal of Chemical Physics*, 93(1):856, 1990.

[30] P. Petropoulos, M. Ibsen, A. D. Ellis, and D. J. Richardson. Rectangular pulse generation based on pulse reshaping using a SuperstructuredFiber bragg grating. *Journal of Lightwave Technology*, 19(5):746, 2001.

[31] R. Feced, M. N. Zervas, and M. A. Muriel. An efficient inverse scattering algorithm for the design of nonuniform fiber bragg gratings. *IEEE Journal of Quantum Electronics*, 35(8):1105–1115, 1999.

[32] J. Skaar, L. Wang, and T. Erdogan. On the synthesis of fiber bragg gratings by layer peeling. *IEEE Journal of Quantum Electronics*, 37(2):165–173, 2001.

[33] J. Holland. *Adaptation in natural and artificial systems.* MIT Press, 1975.

[34] D. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley Longman, 1989.

[35] R. C. E. J. Kennedy. Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks*, pages 1942–1948, 1995.

[36] T. Bäck, F. Hoffmeister, and H. Schwefel. A survey of evolution strategies. *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 2–9, 1991.

[37] S. Kirkpatrick. Optimization by simulated annealing: Quantitative studies. *Journal of Statistical Physics*, 34:975–986, 1984.

[38] M. P. Bendsøe and N. Kikuchi. Generating optimal topologies in structural design using a homogenization method. *Computer Methods in Applied Mechanics and Engineering*, 71(2):197–224, 1988.

[39] M. Bendsøe and O. Sigmund. *Topology Optimization: Theory, Methods and Applications.* Springer, 2003.

[40] E. J. Haug and J. S. Arora. Design sensitivity analysis of elastic mechanical systems. *Computer Methods in Applied Mechanics and Engineering*, 15(1):35–62, 1978.

[41] O. Sigmund. Design of multiphysics actuators using topology optimization - part i: One-material structures. *Computer Methods in Applied Mechanics and Engineering*, 190(49-50):6577–6604, 2001.

[42] W. R. Frei, D. A. Tortorelli, and H. T. Johnson. Topology optimization of a photonic crystal waveguide termination to maximize directional emission. *Applied Physics Letters*, 86(11):111114–111114–3, 2005.

[43] J. Dahl, J. Jensen, and O. Sigmund. Topology optimization for transient wave propagation problems in one dimension. *Structural and Multidisciplinary Optimization*, 36(6):585–595, 2008.

[44] J. S. Jensen and O. Sigmund. Topology optimization for nanophotonics. *Laser Photonics Review.* To appear in the March issue. DOI 10.1002/lpor.201000014.

[45] E. Haug, K. Choi, and V. Komkov. *Design sensitivity analysis of structural systems.* Academic Press, Orlando, FL, 1986.

[46] K. Svanberg. The method of moving asymptotes - a new method for structural optimization. *International Journal for Numerical Methods in Engineering*, 24(2):359–373, 1987.

[47] J. Goh, I. Fushman, D. Englund, and J. Vučković. Genetic optimization of photonic bandgap structures. *Optics Express*, 15(13):8218–8230, 2007.

[48] J. C. Maxwell. On physical lines of force. *Philosophical Magazine*, 21(4), 1861.

[49] J. Wang. Generalised moment methods in electromagnetics. *Microwaves, Antennas and Propagation, IEE Proceedings H*, 137(2):127–132, 1990.

[50] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *Journal of Computational Physics*, 73(2):325–348, 1987.

[51] A. A. Ergin, B. Shanker, and E. Michielssen. Fast evaluation of Three-Dimensional transient wave fields using diagonal translation operators. *Journal of Computational Physics*, 146(1):157–180, 1998.

[52] K. S. Yee. Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Transaction on Antennas and Propagation*, 14(3):302–307, 1966.

[53] A. Taflove and M. Brodwin. Numerical solution of Steady-State electromagnetic scattering problems using the Time-Dependent maxwell's equations. *IEEE Transaction on Microwave Theory and Techniques*, 23(8):623–630, 1975.

[54] A. Taflov and S. Hagness. *Computational Electrodynamics*. Artech House, 2000.

[55] J. L. Volakis, A. Chatterjee, and L. C. Kempel. *Finite Element Method Electromagnetics: Antennas, Microwave Circuits, and Scattering Applications*. Wiley-IEEE Press, 1998.

[56] J. Jin. *The Finite Element Method in Electromagnetics*. Wiley-IEEE Press, 2 edition, 2002.

[57] A. Cangellaris, C. Lin, and K. Mei. Point-matched time domain finite element methods for electromagnetic radiation and scattering. *IEEE Transactions on Antennas and Propagation*, 35(10):1160–1173, 1987.

[58] J. Lee, R. Lee, and A. Cangellaris. Time-domain finite-element methods. *IEEE Transactions onAntennas and Propagation*, 45(3):430–442, 1997.

[59] V. Shankar, A. H. Mohammadian, and W. F. Hall. A time-domain, finite-volume treatment for the maxwell equations. *Electromagnetics*, 10:127–145, 1990.

[60] N. K. Madsen and R. W. Ziolkowski. A three-dimensional modified finite volume technique for maxwell's equations. *Electromagnetics*, 10:147–161, 1990.

[61] A. C. Cangellaris and D. B. Wright. Analysis of the numerical error caused by the stair-stepped approximation of a conducting boundary in fdtd simulations of electromagnetic phenomena. *IEEE Transactions on Antennas and Propagation*, 39:1518–1525, 1991.

[62] S. Benhassine, W. Carpes, and L. Pichon. Comparison of mass lumping techniques for solving the 3D maxwell's equations in the time domain. *IEEE Transactions on Magnetics*, 36(4):1548–1552, 2000.

[63] A. Fisher, R. Rieben, G. Rodrigue, and D. White. A generalized mass lumping technique for vector finite-element solutions of the time-dependent maxwell equations. *IEEE Transactions on Antennas and Propagation*, 53(9):2900–2910, 2005.

[64] D. Riley and C. Turner. VOLMAX: a solid-model-based, transient volumetric maxwell solver using hybrid grids. *IEEE Antennas and Propagation Magazine*, 39(1):20–33, 1997.

[65] T. Rylander and A. Bondeson. Stable FEM-FDTD hybrid method for maxwell's equations. *Computer Physics Communications*, 125(1-3):75–82, 2000.

[66] U. Andersson. *Time-domain methods for the Maxwell equations.* PhD thesis, Kungl Tekniksa Högskolan, 2001.

[67] S. Zivanovic, K. Yee, and K. Mei. A subgridding method for the time-domain finite-difference method to solve maxwell's equations. *IEEE Transactions on Microwave Theory and Techniques*, 39(3):471–479, 1991.

[68] T. W. P. Thoma. A consistent subgridding scheme for the finite difference time domain method. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, 9(5):359–374, 1996.

[69] J. S. Jensen and O. Sigmund. Topology optimization of photonic crystal structures: a high-bandwidth low-loss t-junction waveguide. *Journal of the Optical Society of America B*, 22(6):1191–1198, 2005.

[70] A. Bossavit and I. Mayergoyz. Edge-elements for scattering problems. *IEEE Transactions on Magnetics*, 25(4):2816–2821, 1989.

[71] R. D. Cook, D. S. Malkus, M. E. Plesha, and R. J. Witt. *Concepts and Applications of Finite Element Analysis, 4th Edition.* Wiley, 4th edition, 2001.

[72] F. Zhen, Z. Chen, and J. Zhang. Toward the development of a three-dimensional unconditionally stable finite-difference time-domain method. *IEEE Transactions on Microwave Theory and Techniques*, 48(9):1550–1558, 2000.

[73] J. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114(2):185–200, 1994.

[74] D. S. Katz, E. T. Thiele, and A. Taflove. Validation and extension to three dimensions of the berenger PML absorbing boundary condition for FD-TD meshes. *IEEE Microwave and Guided Wave Letters*, 4(8):268–270, 2002.

[75] Z. Sacks, D. Kingsland, R. Lee, and J. Lee. A perfectly matched anisotropic absorber for use as an absorbing boundary condition. *IEEE Transactions on Antennas and Propagation*, 43(12):1460–1463, 1995.

[76] S. Gedney. An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices. *IEEE Transactions on Antennas and Propagation*, 44(12):1630–1639, 1996.

[77] L. Yang, A. Lavrinenko, L. Frandsen, P. Borel, A. Tetu, and J. Fage-Pedersen. Topology optimisation of slow light coupling to photonic crystal waveguides. *Electronics Letters*, 43(17):923–924, 2007.

[78] M. Pu, L. Yang, L. H. Fradsen, J. S. Jensen, O. Sigmund, H. Ou, K. Yvind, and J. M. Hvam. Topology-optimized slow-light couplers for ring-shaped photonic crystal waveguide. *National Fiber Optic Engineers Conference, OSA Technical Digest (CD) (Optical Society of America, 2010)*, paper JWA30, 2010.

[79] Y. Vlasov and S. McNab. Coupling into the slow light mode in slab-type photonic crystal waveguides. *Optics Letters*, 31(1):50–52, 2006.

[80] H. Kurt and D. Citrin. Annular photonic crystals. *Optics Express*, 13:10316–10326, 2005.

[81] L. Yang, O. Sigmund, A. V. Lavrinenko, and J. Hvam. Design of one-dimensional optical pulse-shaping filters by time-domain topology optimization. *Applied Physics Letters*, 95(26):261101 – 261101–3, 2009.

[82] E. J. Haug, K. K. Choi, and V. Komkov. *Design Sensitivity Analysis of Structural Systems*. Orlando, FL: Academic, 1986.

[83] N. K. Nikolova, H. W. Tam, and M. H. Bakr. Sensitivity analysis with the FDTD method on structured grids. *IEEE Transactions on Microwave Theory and Techniques*, 52(4):1207–1216, 2004.

[84] J. Skaar, B. Sahlgren, P. Fonjallaz, H. S. y, and R. Stubbe. High-reflectivity fiber-optic bandpass filter designed by use of the iterative solution to the Gel'fand-Levitan-Marchenko equations. *Optics Letters*, 23(12):933–935, 1998.

[85] J. Skaar and K. M. Risvik. A genetic algorithm for the inverse problem in synthesis of fiber gratings. *Journal of Lightwave Technology*, 16(10):1928, 1998.

[86] N. Plougmann and M. Kristensen. Efficient iterative technique for designing bragg gratings. *Optics Letters*, 29(1):23–25, 2004.

[87] R. Feced, M. N. Zervas, and M. A. Muriel. An efficient inverse scattering algorithm for the design of nonuniform fiber bragg gratings. *IEEE Journal of Quantum Electronics*, 35(8):1105–1115, 1999.

[88] J. Skaar, L. Wang, and T. Erdogan. Synthesis of thick optical Thin-Film filters with a Layer-Peeling Inverse-Scattering algorithm. *Applied Optics*, 40(13):2183–2189, 2001.

[89] M. P. Bendsøe and O. Sigmund. *Topology Optimization: Theory, Methods and Applications.* Springer, Berlin, 2003.

[90] O. Sigmund. Morphology-based black and white filters for topology optimization. *Structural and Multidisciplinary Optimization*, 33(4-5):401–424, 2007.

[91] T. Borrvall and J. Petersson. Topology optimization using regularized intermediate density control. *Computer Methods in Applied Mechanics and Engineering*, 190(37-38):4911 – 4928, 2001.

[92] T. E. Bruns and D. A. Tortorelli. Topology optimization of non-linear elastic structures and compliant mechanisms. *Computer Methods in Applied Mechanics and Engineering*, 190(26-27):3443–3459, 2001.

[93] B. Bourdin. Filters in topology optimization. *International Journal for Numerical Methods in Engineering*, 50(9):2143–2158, 2001.

[94] O. Sigmund. *Design of material structures using topology optimization.* Technical University of Denmark, 1994.

[95] O. Sigmund. On the design of compliant mechanisms using topology optimization. *Mechanics of Structures and Machines*, 25(4):493–524, 1997.

[96] J. K. Guest, J. H. Prévost, and T. Belytschko. Achieving minimum length scale in topology optimization using nodal design variables and projection functions. *International Journal for Numerical Methods in Engineering*, 61(2):238–254, 2004.

[97] J. K. Guest. Topology optimization with multiple phase projection. *Computer Methods in Applied Mechanics and Engineering*, 199(1-4):123–135, 2009.

[98] Y. Park, M. H. Asghari, T.-J. Ahn, and J. Azaña. Transform-limited picosecond pulse shaping based on temporal coherence synthesization. *Optics Express*, 15(15):9584–9599, 2007.

[99] E. Palushani, L. Oxenlowe, M. Galili, H. Mulvad, A. Clausen, and P. Jeppesen. Flat-Top pulse generation by the optical fourier transform technique for ultrahigh speed signal processing. *IEEE Journal of Quantum Electronics*, 45(11):1317–1324, 2009.

[100] L. Oxenlowe, M. Galili, R. Slavfk, H. Mulvad, A. Clausen, Y. Park, J. Azana, and P. Jeppesen. Timing jitter tolerant 640 gb/s demultiplexing using a long-period fibre grating-based flat-top pulse shaper. *IET Seminar Digests*, 2007(1):452, 2007.

[101] F. Wang, B. S. Lazarov, and O. Sigmund. On projection methods, convergence and robust formulations in topology optimization. *Structural and Multidisciplinary Optimization*, 2010. DOI 10.1007/s00158-010-0602-y.

[102] A. D. Belegundu and T. R. Chandrupatla. *Optimization Concepts and Applications in Engineering.* Upper Saddle River, NJ: Prentice-Hall, 1999.