Topology Optimization for Transient Wave Propagation Problems

Matzen, René; Sigmund, Ole; Jensen, Jakob Søndergaard

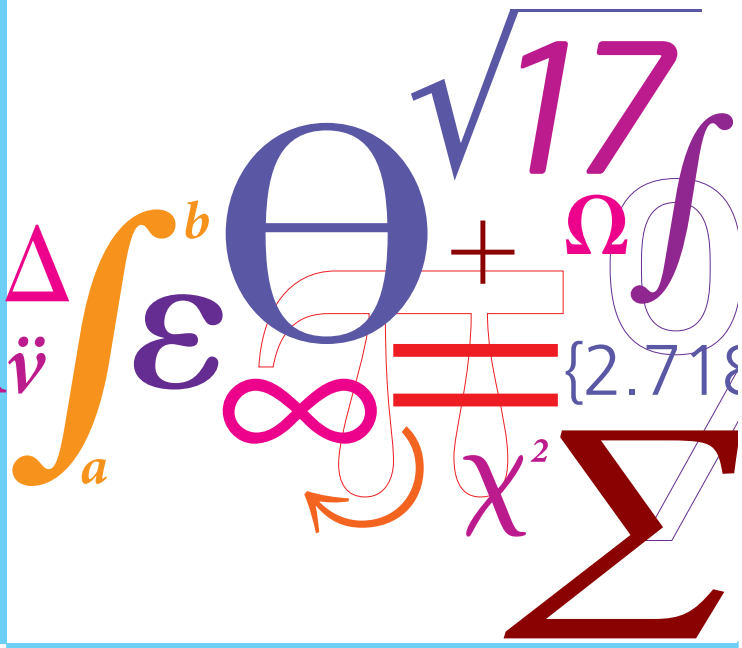
Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

Topology Optimization for Transient Wave Propagation Problems

PhD Thesis

$$(EIv'')'' = q - \rho A\ddot{v}$$



René Matzen
DCAMM Special Repport No. S130
March 2011

DTU Mechanical Engineering
Department of Mechanical Engineering

Topology Optimization for Transient Wave Propagation Problems

René Matzen
Department of Mechanical Engineering
Technical University of Denmark

A thesis submitted for the degree of
Philosophiae Doctor (Ph.D.)

2011 March

DEPT. OF MECHANICAL ENGINEERING
Solid Mechanics

TECHNICAL UNIVERSITY OF DENMARK

Title of the thesis:
Topology Optimization for Transient Wave Propagation Problems

Author:
René Matzen
E-mail: rmat@mek.dtu.dk

Supervisors:
Ole Sigmund, Professor, Dr. techn.
E-mail: sigmund@mek.dtu.dk

Jakob Søndergaard Jensen, Associate Professor
E-mail: jsj@mek.dtu.dk

Address:
Department of Mechanical Engineering, Solid Mechanics
Technical University of Denmark
Nils Koppels Allé, Building 404
DK-2800 Kgs. Lyngby, Denmark

ISBN 978-87-90416-57-7

Preface

This thesis is submitted in partial fulfillment of the requirements for obtaining the Ph.D. degree in mechanical engineering at the Technical University of Denmark (DTU). The Ph.D. project was carried out at the Department of Mechanical Engineering (MEK), Solid Mechanics, at DTU in the period April 1st 2008 - March 31st 2011. Supervisors on the project were Professor Dr. techn. Ole Sigmund and Associate Professor Ph.D. Jakob Søndergaard Jensen from the department.

I am very grateful to my two supervisors for their inspiring support and for always taking their time for fruitful and valuable discussions of the work and the results during the project, in a cosy and truly inspiring atmosphere. I thank Associate Professor Kurt Maute for hosting me during my 3 months external research stay at the Department of Aerospace Science, Colorado University, Boulder, USA autumn 2010. I truly appreciate new academic input in a friendly and welcoming atmosphere.

I would like to express my gratitude to postdoc Boyan S. Lazarov from MEK for his assistance in setting up the parallel communication strategy in my FORTRAN code used to obtain the results during the project. The support from Mathias Stolpe from the Department of Mathematics DTU for providing me necessary insight into various optimization algorithms has also been invaluable.

I am thankful to Krister Svanberg at Royal Institute of Technology, Sweden, who has granted me permission to use the MMA/GCMMA FORTRAN code.

The Ph.D. has been funded by the Eurohorcs/ESF European Young Investigator Award (EURYI, www.esf.org/euryi) through the grant "Synthesis and topology optimization of optomechanical systems", M.B. Richters foundation, and DTU. The support from Danish Center for Scientific Computing is also gratefully acknowledged.

I finally wish to thank my colleagues in the TopOpt Group and at the section of Solid Mechanics for creating a cheerful and inspiring working environment.

Kgs. Lyngby, March 31st 2011.

René Matzen

Resumé (in Danish)

Topologioptimering anvendt på transiente bølgeudbredelsesproblemer

Studiet af elastiske og optiske bølger sammen med intensiv materialeforskning har revolutioneret hverdags- samt avanceret teknologi i det sidste århundrede. Derfor er det vigtigt at fortsætte det undersøgende arbejde for at forbedre såvel eksisterende som ny innovativ teknologi, ved at designe nye materialer og deres layout.

Afhandlingen præsenterer en generel ramme for anvendelsen af topologioptimering til design af materialelayouts for transiente bølgeudbredelsesproblemer. I modsætning til det høje niveau for modellering i frekvensdomænet, er topologioptimering i tidsdomænet stadig i en opbygningsproces. Et generelt optimeringsproblem formuleres med en objekt-funktion, der kan være felt-, hastigheds- og accelerationsafhængig, ligesom den også kan afhænge af filtrerede signaler [P3]. De analytiske designgradienter findes ved hjælp af *adjoint metoden*.

Bølgeudbredelseproblemer foregår oftest i åbne domæner. Det numeriske simuleringsdomæne bør derfor afgrænses ved at indføre absorberende randbetingelser. Til dette formål er udviklet en *finite element* formulering baseret på en *perfekt matchede lag* metode for tidsafhængige skalare og elastiske bølgeudbredelsesproblemer [P2], [P4]. For at reducere de beregningsmæssige omkostninger i forbindelse med gradientbaseret optimering i tidsdomænet, introduceres parallel databehandling til de mest tids- og hukommelseskrævende dele af den iterative optimeringsprocedure.

Ved hjælp af den opstillede model og den udviklede software betragtes to slags optimeringsproblemer, der involverer optiske bølger i *nano-fotoniske* komponenter. Først optimeres en *optisk taper* og et *notch filter* ved energimaksimering [P2], dernæst to slags strukturer, der henholdsvis omformer pulse [P3] og understøtter ikke-dispersivt langsomt lys [P5].

Abstract

The study of elastic and optical waves together with intensive material research has revolutionized everyday as well as cutting edge technology in very tangible ways within the last century. Therefore it is important to continue the investigative work towards improving existing as well as innovate new technology, by designing new materials and their layout.

The thesis presents a general framework for applying *topology optimization* in the design of material layouts for transient wave propagation problems. In contrast to the high level of modeling in the frequency domain, time domain topology optimization is still in its infancy. A generic optimization problem is formulated with an objective function that can be field, velocity, and acceleration dependent, as well as it can accommodate the dependency of filtered signals essential in signal shape optimization [P3]. The analytical design gradients are derived by use of the *adjoint variable method*.

Many wave propagation problems are open-region problems, i.e. the outer boundaries of the modeling domain must be reflection-less. The thesis contains new and independent developments within *perfectly matched layer* techniques for scalar as well as for vectorial elastic wave propagation problems using *finite element* analysis [P2], [P4]. The concept is implemented in a parallel computing code that includes efficient techniques for performing gradient based topology optimization.

Using the developed computational framework the thesis considers four optimization problems from *nano-photonics*: First, an *optical taper* [P1] and a *notch filter* [P2] – both optimized by energy maximization. The last two cases demonstrate pulse shaping and delay in one [P3] and two [P5] dimensions. Whereas the test problem in [P3] is rather academic, the example considered in [P5] optimizes structures that accommodate non-dispersive slow light, with important applications for optical buffering devices.

Publications

The following publications are part of the thesis

[P1] R. Matzen, J. S. Jensen, and O. Sigmund, Transient topology optimization of two-dimensional elastic wave propagation. *Proceedings of the Ninth International Conference on Computational Structures Technology*, B.H.V. Topping, M. Papadrakakis, (Editors), Civil-Comp Press, Stirlingshire, United Kingdom, paper 75, 2008.

[P2] R. Matzen, J. S. Jensen and O. Sigmund, Topology optimization for transient response of photonic crystal structures. *J. Opt. Soc. Am. B*, 27(10):2040-2050, October 2010.

[P3] B. Lazarov, R. Matzen, and Y. Elesin. Topology optimization of pulse shaping filters using the Hilbert transform envelope extraction. *Structural and Multidisciplinary Optimization*, January 2011. Accepted for publication.

[P4] R. Matzen. An efficient finite element time-domain formulation for the elastic second-order wave equation: A non-split complex frequency shifted convolutional PML. *Int. J. Numer. Meth. Engng.*, 2011. DOI: 10.1002/nme.3205.

[P5] R. Matzen, J. S. Jensen and O. Sigmund. Topology optimization of slow light structures in the time domain. Submitted.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

True technological revolutions and innovations in history have evolved from a curiosity to comprehend the physical properties of materials. The discovery of new materials that the Earth provides has triggered a desire amongst engineers to tinker with their properties, hoping to produce substances with even better, e.g. *mechanical* and/or *optical* properties.

The study of elastic and optical waves has been vital to the engineering of new materials. Combined with intensive research in materials, manufacturing and design techniques, the discovery and understanding of new wave types with different features has led to the fabrication of structures down to nano-scale. This has resulted in an endless list of significant applications ranging from sonar, ultrasound scanning, to lasers engineering, high-speed computing, spectroscopy, fiber-optic cables, and integrated optics – just to mention a few – among which some have revolutionized the telecommunication industry. In the design process of new material layouts and structures one could use *Edisonian* approaches, or, more systematically *shape* and *geometry optimization*, or optimization techniques relying on free distribution of material such as *topology optimization* (Bendsøe and Sigmund, 2004).

The aim of this thesis is to study transient propagation of elastic and optical waves and to develop a general framework for optimizing structures subjected to these types of waves, by using the methodology of topology optimization. Topology optimization is a gradient-based iterative optimization technique that has proven very efficient in the design of mechanical structures subjected to static and dynamic loading (Bendsøe and Sigmund, 2004), and for optimizing optical devices subjected to time-harmonic waves (Jensen and Sigmund, 2011). Very recently, topology optimization has been applied to one-dimensional (1D) elastic transient wave propagation problems based on the *finite element time domain*

(FETD) method (Dahl et al., 2008; Jensen, 2009). Also, based on the *finite difference time domain* (FDTD) technique, topology optimization has been used to design 1D optical filters (Yang et al., 2009) as well as three-dimensional (3D) dielectric antennas (Nomura et al., 2007).

Here, we extend the recently developed 1D finite element based framework to elastic and optical problems in two dimensions (2D). Going from 1D to 2D is not straightforward because waves are suddenly allowed to propagate freely within the plane. This imposes requirements on the numerical techniques used to render the boundaries of the computational transparent in the transient simulations of open-region problems. In a FETD framework, efficient mesh truncation techniques can be quite tricky to implement. Furthermore, gradient-based time domain optimization of problems in general involving a large number of design variables appear to be a challenge with respect to computing efforts.

We apply the developed framework to design optical material structures. The first two problems optimize the performance of an optical taper and a *photonic crystal* (PhC) notch filter respectively, and the third problem optimizes structures that slow down the speed of light. All simulations are performed by our own-developed FE code implemented in FORTRAN. It supports parallel computing to lower the computational burden associated with time domain optimization and thereby allows us to optimize within a reasonable time frame for large problems.

1.2 Thesis structure

This thesis serves as a general introduction to the work that has been carried out during the Ph.D. study. Apart from an introduction to the employed theory it also provides an overview and, where it has been found necessary, a detailed description of the developed methods and obtained results presented in the five publications [P1]-[P5].

We start out in chapter 2 by introducing the governing equations for elastic and optical waves, from which we derive the FE model used in the numerical analysis of the inhomogeneous material hosts throughout the thesis. To treat open-region problems efficiently we develop a FE formulation based on the perfectly matched layer technique.

In chapter 3, topology optimization is introduced to transient wave propagation problems. Since time domain topology optimization is still relatively unexplored, we will formulate a generic optimization problem that is believed to cover a wide range of topology optimization problems for transient responses. We derive the sensitivity analysis and subsequently present our thoughts in the process of parallelizing the topology optimization procedure.

In chapter 4, we present three typical time domain optimization problems.

The first two problems consider maximization of the energy response, used in the design of an optical taper and a notch filter. For the third problem we formulate a pulse shaping and delaying strategy to optimize slow light devices.

Finally, in chapter 5, we will state some concluding remarks.

Chapter 2

Time domain modeling of propagating waves

In this chapter time domain modeling of elastic and optical waves is introduced. First, we present the governing equations on continuous form and formulate the boundary value problem (section 2.1-2.2). The latter is subsequently discretized in space by the finite element method and in time by finite differences (section 2.3). To allow for arbitrary propagation characteristics and directions inside the truncated modeling domain we introduce an efficient absorbing boundary technique (section 2.4). Then, we explain how to parallelize the finite element solver (section 2.4), and finally the computation of transmission and reflection spectra is addressed (section 2.7).

2.1 Elastic waves

An elastic wave is a class of mechanical wave that propagates in elastic or viscoelastic materials. As long as the a material is not stressed in tension or compression beyond its elastic limit, it is the restoring forces between particles, combined with inertia of particles, that lead to the oscillatory motions of the medium after vibratory disturbances, and thereby accommodates a propagating wave.

In the present work we consider propagation of elastic waves in 2D structures. The medium inside the structures is assumed to occupy a composite of regions of homogenous linear elastic material as a function of the plane position vector $\boldsymbol{r} = (x_1, x_2)^T$ (with $(\cdot)^T$ denoting the transverse of a vector) in the solution domain Ω_{S} with boundary $\partial\Omega_{\mathrm{S}} = \Gamma_{\mathrm{S}}$. Combining the equations of motion with the constitutive law and the linear kinematic conditions, governs the propagation of elastic waves in anisotropic media by the following system of equations

(Achenbach, 1975)

$$(\textit{Eq. of motion}) \qquad \nabla \cdot \boldsymbol{\sigma} + \rho \boldsymbol{p} = \rho \frac{\partial^2 \boldsymbol{u}}{\partial t^2} \tag{2.1a}$$

$$(\textit{Constitutive law}) \qquad \boldsymbol{\sigma} = \boldsymbol{C} : \boldsymbol{\varepsilon} \tag{2.1b}$$

$$(\textit{Strain-displ.}) \qquad \boldsymbol{\varepsilon} = \tfrac{1}{2}\left[\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T\right] \tag{2.1c}$$

where $\boldsymbol{u} = (u_1, u_2)^T$ and $\boldsymbol{p} = (p_1, p_2)^T$ are the displacement and body force vectors, respectively. Further, ρ is the mass density, $\boldsymbol{\sigma}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{C}$ are stress, strain and constitutive tensors, respectively, and $\nabla = (\partial/\partial x_1, \partial/\partial x_2)^T$ is the divergence operator. For 2D isotropic elastic media the constitutive tensor entries are determined by

$$C_{ijkl} = \lambda \delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{2.2}$$

with μ and λ denoting the Lamé coefficients, and where δ_{ij} is Kronecker's delta. The above set of equations (2.1)-(2.2) can model three fundamental types of elastic waves: Shear (S-) waves with oscillations transverse to the propagation direction. Pressure (P-) waves, whose polarization is always longitudinal, and surface (Rayleigh) waves. Where S- and P-waves are body waves, surface waves can, as the name also indicates, only exist on the surface with an exponential decay into the bulk. For example, sound propagates through air as P-waves, seismic waves can propagate as S-, P- and surface waves, and water waves propagate as surface waves.

2.2 Optical waves

Optical waves were probably described for the first time in 1705 when Isaac Newton released *Opticks* about the nature of light and colors and the diffraction of light based on phenomenological observations.

It took more than 150 years, before J.C. Maxwell (1865) combined a set of equations into the *Maxwell's equations* that form the basis theory of classical electrodynamics and optics in the modeling the propagation of electromagnetic waves. In contrast to elastic waves, electromagnetic waves can both travel in matter and vacuum. Here, we will consider the propagation of light in isotropic dielectric media containing no free charges and currents, for which the macroscopic behavior of light is generally governed by

$$\nabla \times \boldsymbol{H} = \varepsilon_0 \varepsilon_r \frac{\partial \boldsymbol{E}}{\partial t} \tag{2.3a}$$

$$\nabla \times \boldsymbol{E} = -\mu_0 \frac{\partial \boldsymbol{H}}{\partial t} \tag{2.3b}$$

where $\boldsymbol{E} = (E_1, E_2, E_3)^T$ and $\boldsymbol{H} = (H_1, H_2, H_3)^T$ are the electric and magnetic fields, respectively, ε_r is the relative permittivity, and ε_0 and μ_0 is vacuum permittivity and permeability, respectively, that can be combined to yield the vacuum speed of light $c = 1/\sqrt{\varepsilon_0 \mu_0}$.

Throughout the thesis we will assume wave propagation inside 2D structures with infinite extension in the third dimension. Consequently, any modes are invariant under reflection through the plane of propagation and this mirror symmetry allows us to uncouple the modes into two distinct polarizations (Jackson, 1999). Transverse-electric (TE) modes have an in-plane electric vector field, and out-of-plane scalar magnetic field component. Transverse-magnetic (TM) modes have the reverse. If we combine equations (2.3a) and (2.3b) into one for either $\boldsymbol{H}$ or $\boldsymbol{E}$, it can be shown that the scalar fields in the solution domain Ω_{S} are governed by the generic scalar wave equation

$$\frac{\partial}{\partial x_1}\left(A\frac{\partial u}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(A\frac{\partial u}{\partial x_2}\right) - B\frac{\partial^2 u}{\partial t^2} = 0 \tag{2.4}$$

where $u = H_3$, $A = 1/\varepsilon_r$ and $B = c^{-2}$ for TE modes, and $u = E_3$, $A = 1$ and $B = \varepsilon c^{-2}$ for TM modes.

As a remark it should be emphasized that the scalar wave equation (2.4) can be applied in the modeling of a certain class of elastic waves and acoustic waves with the right choice of material parameters for A and B. In the elastic case the solution will be displacement, and pressure in the acoustic case. In order to find a solution to the wave equation (2.4) we need to combine it with a set of boundary conditions on Γ_{S}. By means of the Robin boundary condition, given by

$$\boldsymbol{n} \cdot (A\nabla u) + \sqrt{AB}\frac{\partial u}{\partial t} = 2\sqrt{AB}\frac{\partial u_0}{\partial t} \tag{2.5}$$

with $\boldsymbol{n} = (n_1, n_2)^T$ denoting the outward unit vector normal to the boundary, we can specify sections from which the waves should be excited with amplitude u_0 and/or sections through which normal incident waves only ought to be transmitted without reflections. The last condition is known as the *first-order absorbing boundary condition*.

2.3 Finite element time domain formulation

Throughout the thesis we will in general deal with inhomogeneous materials for which no analytical solution exists to the elastic and scalar wave equations. To obtain a numerical solution the finite element (FE) method is employed (Hughes, 2000; Zienkiewicz et al., 2005; Jin and Riley, 2009).

The FE method holds an unsurpassed capability for modeling complex geometries and materials, and in the time domain it facilitates a simple switch between an unconditionally or a conditionally stable formulation. Unconditional stability is very important when parts of the geometry needs to be finely resolved to avoid the limiting correlation between the smallest element size and the largest possible time step dictated by the CFL[1] condition (Courant et al., 1928).

The involved geometries in the present work allow for the use of square elements to discretize the computational domain. From an optimization perspective this is desirable since the topology of the optimized design is not known in advance. Furthermore, the numerical advantages of using these elements comprise high stability and efficiency in the time stepping and they are well-suited for parallel computing. It should be stressed that for regular grids the finite difference time domain (FDTD) method (Taflove and Hagness, 2005) can be equally used for both electromagnetic and elastic waves. However, the choice of the finite element time domain (FETD) method relies on our experience with this numerical technique.

For simplicity, we will only present the finite element analysis for the scalar wave equation (2.4) combined with the first-order absorbing boundary condition in equation (2.5). More efficient absorbing boundary conditions are treated in details in section 2.4.

2.3.1 Spatial discretization

To solve the approximate boundary value problem defined by equations (2.4) and (2.5), we multiply equation (2.4) by an appropriate testing function T, integrate over the 2D solution domain Ω_S with boundary $\partial\Omega_\mathrm{S} = \Gamma_\mathrm{S}$ and obtain the weak form representation

$$\iint_{\Omega_\mathrm{S}} \left[\nabla T \cdot A\nabla u + TB\frac{\partial^2 u}{\partial t^2}\right] \mathrm{d}\Omega + \int_{\Gamma_\mathrm{S}} \left[T\sqrt{AB}\frac{\partial u}{\partial t} - TU\right] \mathrm{d}\Gamma = 0 \tag{2.6}$$

where $U = 2\sqrt{AB}\partial u_0/\partial t$ is the boundary source on $\Gamma_\mathrm{inc} \subseteq \Gamma_\mathrm{S}$. The weak-form derivation can be found in various books about the finite element method, see e.g. Zienkiewicz et al. (2005); Jin and Riley (2009).

To seek the finite element solution of equation (2.6) we subdivide the solution domain into M square finite elements with edge size $h = \Delta x_1 = \Delta x_2$, and by expanding the considered scalar field $u(\boldsymbol{r}, t)$ at any point within each element using nodal bilinear basis functions, it can be expressed as

$$u(\boldsymbol{r}, t) \approx \sum_{k=1}^{4} N_i(\boldsymbol{r}) u_k^e(t) \tag{2.7}$$

[1] Courant-Friedrichs-Lewy

where u_k^e denotes the value of u at node kth of element e, and N_k is the basis function corresponding to the kth node. Substitution of equation (2.7) into (2.6) and using the same N_k as the testing function T yields the system of second-order ordinary differential equation

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{f} \tag{2.8}$$

where $\dot{(\cdot)} = \mathrm{d}/\mathrm{d}t$, and $\ddot{(\cdot)} = \mathrm{d}^2/\mathrm{d}t^2$. $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ represent sparse, symmetric, and positive definite matrices that are assembled from their respective element-level constituent square 4×4 matrices $\mathbf{M}^e$, $\mathbf{C}^e$, and $\mathbf{K}^e$ whose entries are given by

$$M_{ij}^e = \iint_{\Omega^e} BN_iN_j\mathrm{d}\Omega \tag{2.9a}$$

$$C_{ij}^e = \int_{\Gamma^e} \sqrt{AB}N_iN_j\mathrm{d}\Gamma \tag{2.9b}$$

$$K_{ij}^e = \iint_{\Omega^e} A\nabla N_i \cdot \nabla N_j\mathrm{d}\Omega \tag{2.9c}$$

Additionally, $\mathbf{u} = [u_1, u_2, \ldots, u_{N_{\text{node}}}]^T$ where N_{node} denotes the total number of nodes, and the entries of the element contribution to the excitation vector $\mathbf{f} = \sum \mathbf{f}^e$ are given by

$$f_i^e = \int_{\Gamma^e \subseteq \Gamma_{\text{inc}}} N_iU\mathrm{d}\Gamma \tag{2.10}$$

For the elastic wave equation (2.1) the principle of *virtual work* is used to obtain the weak-form representation. The resulting discrete equation is identical to that in equation (2.8), albeit with element square matrices of size 8×8 instead. Further, the scalar entries in equations (2.9) transform into sub-matrix entries, and the vector ∇N_k converts to the so-called *strain-displacement* matrix, see e.g. (Zienkiewicz et al., 2005) or [P4].

2.3.2 Time discretization and integration

We approximate the first- and second-order time derivatives in the FE discretized wave equation (2.8) by the differencing formulae derived from the Newmark-beta time integration scheme (Newmark, 1959), i.e.

$$\dot{\mathbf{u}} \approx \frac{\mathbf{u}^{n+1} - \mathbf{u}^{n-1}}{2\Delta t} \tag{2.11}$$

$$\ddot{\mathbf{u}} \approx \frac{\mathbf{u}^{n+1} - 2\mathbf{u}^n + \mathbf{u}^{n-1}}{\Delta t^2} \tag{2.12}$$

where Δt is the time-step. The non-derivatives are approximated using a weighted average

$$\mathbf{u} \approx \beta \mathbf{u}^{n+1} + (1-2\beta)\mathbf{u}^{n} + \beta \mathbf{u}^{n-1} \tag{2.13}$$

$$\mathbf{f} \approx \beta \mathbf{f}^{n+1} + (1-2\beta)\mathbf{f}^{n} + \beta \mathbf{f}^{n-1} \tag{2.14}$$

where β is a parameter that takes a value between 0 and 1. If we substitute equations (2.11)-(2.14) into (2.8) the time marching of the elastic field becomes

$$\left(\frac{1}{\Delta t^2}\mathbf{M} + \frac{1}{2\Delta t}\mathbf{C} + \beta \mathbf{K}\right)\mathbf{u}^{n+1} = \left(\frac{2}{\Delta t^2}\mathbf{M} - (1-2\beta)\mathbf{K}\right)\mathbf{u}^{n}$$
$$-\left(\frac{1}{\Delta t^2}\mathbf{M} - \frac{1}{2\Delta t}\mathbf{C} + \beta \mathbf{K}\right)\mathbf{u}^{n-1} + \beta \mathbf{f}^{n+1} + (1-2\beta)\mathbf{f}^{n} + \beta \mathbf{f}^{n-1} \tag{2.15}$$

When $\beta = 0$, the formula reduces to explicit time integration and the time-step is constrained by the spatial discretization through the CFL condition. In addition, if we apply the mass lumping rule $\mathbf{D}^{e}_{ii} = \sum_j \mathbf{M}^{e}_{ij}$, $\mathbf{D}_{ij} = 0$ for $i \neq j$, the time marching can be performed without solving a matrix system at each time step. Conversely, unconditionally stable time integration is obtained for $\beta \geq 1/4$ with second-order accuracy preserved (Chilton and Lee, 2007). In this case, the choice of Δt depends solely on the accuracy with which the temporal variation of the field needs to be reproduced. However, the price for this assurance is solving a matrix system at each time step. For large models, direct solution methods are not feasible anymore to complete this task, and iterative methods are instead required.

For the scalar wave equation an efficient integration technique tailored to increase the dispersion error accuracy to fourth-order of low-order square finite elements is employed. The accuracy of this integration scheme is achieved by modified integration rules for the mass and stiffness matrix proposed by Yue and Guddati (2005). The efficiency is obtained by rendering the mass matrix $\mathbf{M}$ diagonal when inversion is needed, and in the case of multiplication a linear combination of the consistent and diagonal mass matrix is used, which leaves the integration explicit. The semi-discrete formula of equation (2.8) for this scheme is written as

$$\mathbf{D}\dot{\mathbf{u}}^{n+1/2} = \mathbf{M}_{\delta}\mathbf{v}^{n-1/2} \tag{2.16a}$$

$$\mathbf{D}\dot{\mathbf{v}}^{n} = \mathbf{f}^{n} - \mathbf{K}\mathbf{u}^{n} \tag{2.16b}$$

where $\mathbf{M}_{\delta} = (1-\delta)\mathbf{D} + \delta\mathbf{M}$. The optimal combination factor is $\delta = (\tau^2 - 1)/2$ to obtain the fourth-order accuracy, where $\tau = \Delta t_c v/\Delta x$ with Δt_c denoting the critical time step that follows the CFL condition and v the phase velocity. Half-step approximations are used for central time differences, i.e. $\dot{\mathbf{u}}^{n+1/2} = (\mathbf{u}^{n} - \mathbf{u}^{n-1})/\Delta t$ and $\dot{\mathbf{v}}^{n} = (\mathbf{u}^{n+1/2} - \mathbf{u}^{n-1/2})/\Delta t$.

2.4 Absorbing boundary conditions

Solving open-region elastodynamic and electromagnetic problems implies that the infinite region exterior to the radiating/scattering object has to be truncated with some type of artificial boundary, such that the computational domain becomes limited. Hence, a boundary condition that absorbs waves, leaving the computational domain independent of direction and frequency, needs to be introduced. Such a condition serves, as far as possible, as a transparent boundary that should yield perfect transmission of the scattered/radiated field and thereby minimize the non-physical reflections from the boundary. Several absorbing boundary condition (ABC) techniques have been developed to complete this task. It comprises non-local conditions (i.e. an exact representation of the infinite medium) which are difficult to implement, and local conditions, such as the first-order absorbing boundary condition in equation (2.5), that only exhibit good performance under tailored circumstances (frequency dependent, specific angles of incidence). The introduction of the *perfectly matched layer* (PML) in 1994 revolutionized absorbing boundaries for wave equations because it was designed to efficiently absorb outgoing electromagnetic waves regardless of their propagation characteristics and directions.

It is highly recommendable to use an efficient ABC such as the PML in topology optimization of wave propagation problems since distribution of material intrinsically leads to material heterogeneity and thus propagation characteristics and directions not known a priori.

In the following section 2.6 the PML technique is introduced to the time domain version of the scalar wave equation and to the second order elastic wave equation. Both approaches have been developed as a part of the present work since they, to the authors' knowledge, do not exist for the second order wave equation in the form we use and discretize it by the FE method. For further details consult [P2], [P4].

2.5 Perfectly matched layer

The concept of the transparent perfectly matched interface, proposed by Berenger (1994) to electromagnetic waves, is an interface between two half-spaces, one of which is dispersive, i.e. lossy, and this is known as the perfectly matched layer. By construction, the PML is a local ABC and renders the interface reflectionless in the continuous case, while outwardly propagating waves are attenuated exponentially independent of frequency, angle of incidence and polarization.

The key concept of the PML interpretation is a coordinate transformation in which the spatial variables are mapped independently onto the complex space by

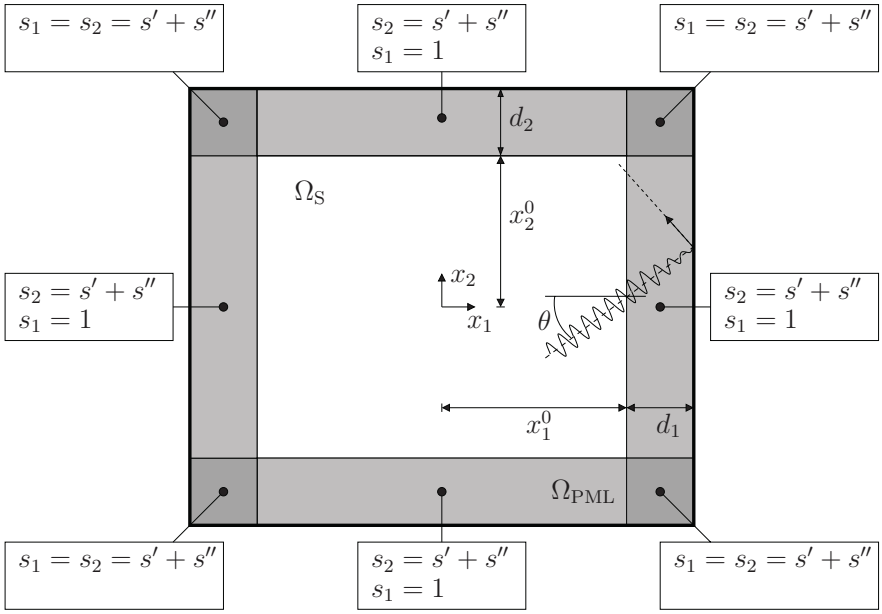


Figure 2.1: Computational domain - truncated by PML layer. A penetrating wave with angle θ is depicted.

a complex stretching function. In particular, this mapping replaces propagating waves with exponentially decaying waves as soon as the waves pass the PML interface located at $x_1 = x_1^0$ along the spatial direction x_1, see figure 2.1. The attenuation occurs over the PML layer width d_1, and continues after the remainder of the wave has been reflected from the PML boundary at $x_1 = x_1^0 + d_1$.

In the 2D PML region Ω_{PML} the original spatial coordinate variables x_i $(i = 1, 2)$ are replaced in each distinct spatial direction by a stretched coordinate, that is defined as (Chew and Weedon, 1994)

$$x_i' = \int_0^{x_i} s_i(\tilde{x}_i, \omega) \mathrm{d}\tilde{x}_i, \quad i = 1, 2 \tag{2.17}$$

The apostrophe $(\cdot)'$ indicates the stretched version of the subtended function, ω is the angular frequency, and s_i $(i = 1, 2)$ are the complex frequency shifted stretched-coordinate metrics proposed by Kuzuoglu and Mittra (1996)

$$s_i(x_i, \omega) = \kappa_i(x_i) + \frac{\beta_i(x_i)}{\alpha_i(x_i) + j\omega}, \quad i = 1, 2 \tag{2.18}$$

where j denotes the imaginary unit. In equation (2.18) $\beta_i \geq 0$ $(i = 1, 2)$ is a coordinate-wise real function that controls the attenuation of the propagating waves. The coordinate-wise real functions $\kappa_i \geq 1$ and $\alpha_i \geq 0$ $(i = 1, 2)$ serve to enhance attenuation of evanescent and near-grazing waves.

Due to the fact that the PML represents a dispersive and lossy medium it is well-suited for implementation in the finite element frequency domain (FEFD) method. However, the FETD implementation becomes more involved since the given frequency dependency in the stretching function in equation (2.18) transforms into convolution in the time domain for the second order wave equation.

2.5.1 Scalar wave equation [P2]

The basic concept of the PML is introduced to the 2D scalar wave equation (2.4) by considering its frequency domain counterpart in the stretched coordinate metrics. It can be shown that based on the original coordinates, the field inside the PML satisfies [P4]

$$\frac{\partial}{\partial x_1}\left(\frac{s_2}{s_1}A\frac{\partial \widehat{u}}{\partial x_1}\right)+\frac{\partial}{\partial x_2}\left(\frac{s_1}{s_2}A\frac{\partial \widehat{u}}{\partial x_2}\right)+\omega^2 s_1 s_2 B\widehat{u}=0 \tag{2.19}$$

where the caret $\widehat{(\cdot)}$ designates the frequency counterpart of the subtended function. If we, for simplicity, assume that propagating waves only need to be attenuated, i.e. $\alpha_i = 0$ and $\kappa_i = 1$ $(i = 1, 2)$, the time domain version of equation (2.19), by using the inverse Fourier transform, becomes [P2]

$$\frac{\partial}{\partial x_1}\left(\mathcal{L}_{A,1}(t)A\frac{\partial u}{\partial x_1}\right)+\frac{\partial}{\partial x_2}\left(\mathcal{L}_{A,2}(t)A\frac{\partial u}{\partial x_2}\right)-\mathcal{L}_B(t)Bu=0 \tag{2.20}$$

In equation (2.20) the operator $\mathcal{L}_1(t)$ is given by

$$\mathcal{L}_B(t)=\frac{\partial^2}{\partial t^2}+(\beta_1+\beta_2)\frac{\partial}{\partial t}+\beta_1\beta_2 \tag{2.21a}$$

and the operator $\mathcal{L}_{A,i}(t)$ is given by

$$\mathcal{L}_{A,i}(t)=1-a_i e^{-\beta_i t}\bar{u}(t)*, \quad i=1,2 \tag{2.21b}$$

with $a_1 = \beta_1 - \beta_2$ and $a_2 = \beta_2 - \beta_1$. In equation (2.21b), $\bar{u}(t)$ denotes the Heaviside step function and $*$ stands for temporal convolution. Note that if $\beta_i = 0$ $(i = 1, 2)$ the governing equation (2.20) is also valid inside the physical simulation domain Ω_{S}. There exists different strategies for choosing β_i and the choice depends on the physical origin and type of the waves considered. For electromagnetic waves see e.g. (Koshiba et al., 2001) and [P2]. For elastic waves consult [P4].

The dispersive and lossy material behavior in the PML enters the governing equation through the operators $\mathcal{L}_{A,i}(t)$ $(i = 1, 2)$ and $\mathcal{L}_B(t)$ and thereby modifies the weak-form representation in equation (2.6). This gives rise to a dissipative term, modification of the stiffness matrix, as well as a temporal convolution term

in the FE discretized wave equation (2.8) [P2]. Normally, the computational efforts associated with the evaluation of discrete convolution are high, since integration of the entire solution history is needed at each time step. However, due to the exponential form of the convolution kernel in equation (2.21b), it can be evaluated recursively (Luebbers and Hunsberger, 1992). This only requires information about the field from the previous time step. Two different recursive convolution techniques are presented in [P2] and [P4], and the latter is recommended since it demonstrates the highest stability properties for long time simulations.

2.5.2 Elastic wave equation [P4]

Compared to the scalar wave equation, the derivation of the time domain version of the elastic wave equation in the PML is more complex. In essence, the introduction of the stretched coordinate metric modifies the isotropic constitutive tensor in equation (2.2) to become anisotropic, and thus alternates the PML interpretation from initially being based on complex coordinate stretching to an anisotropic material model.

In the present work a novel convolutional PML formulation for elastic waves is developed that results in a FETD algorithm of high computational efficiency. It offers a simple switch between implicit and explicit time integration exhibiting remarkable long time stability behavior. We leave the derivation of the governing equations to [P4], and instead provide an example illustrating the efficiency of the suggested method.

We have chosen an experiment that addresses the issue when an evanescent wave field is generated in the PML layer as a result of propagating waves at near-grazing incidence. We launch an explosive vertical directional point-source inside a 2D semi-infinite elastic half-space consisting of a homogeneous material

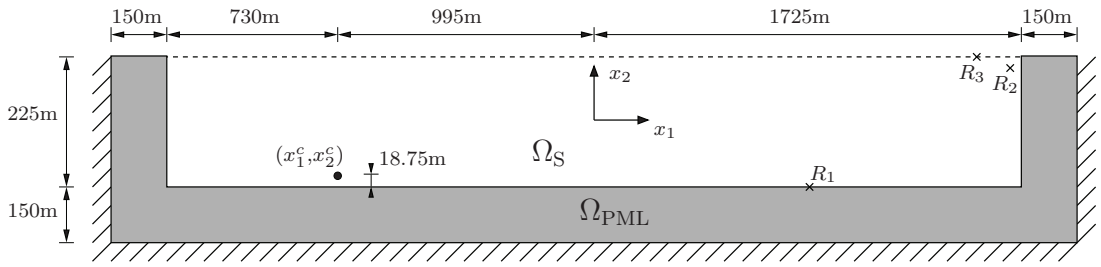


Figure 2.2: PML-truncated elastic half-space - subjected to a vertically downward point-source located at (x_1^c, x_2^c) driven by an explosive Ricker wavelet. $R1$, $R2$, and $R3$ indicate the location of the receivers at which the wave response is recorded.

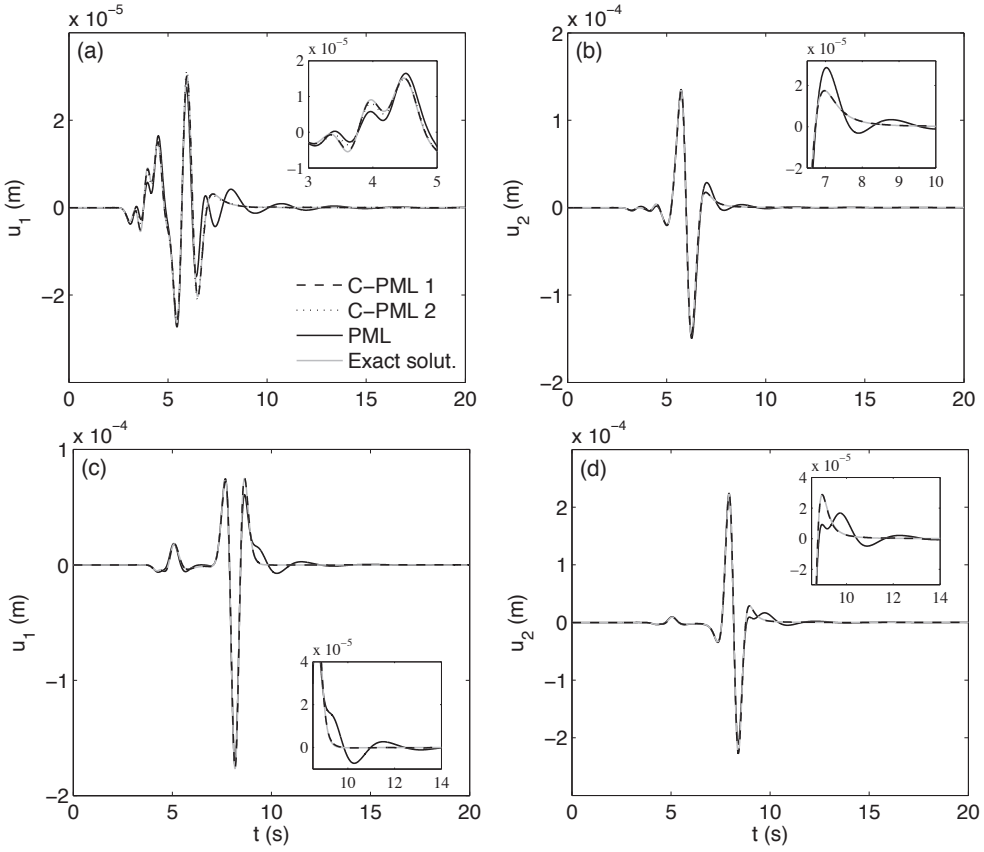


Figure 2.3: Transient responses - CFS-PML, standard PML, and analytical transient response comparison of the horizontal u_1 (left) and the vertical u_2 (right) at three receiving positions: **(a)-(b)** R_1, and **(c)-(d)** R_3 located at (862.5m, -112.5m), and (1650.0m, 112.5m), respectively. Note, that the response based on two C-PML models are recorded at R_1: one with (C-PML 1) evanescent scaling active, i.e. $\kappa_i > 1$, and one without (C-PML 2).

distribution and with a free surface. Since the host consists of homogeneous material the medium is not contaminated with refracted, reflected and diffracted waves, and it also means that an analytical solution exists. The FORTRAN code EX2DDIR of Berg et al. (1994) has been used to compute the exact analytical solution of the responses $\boldsymbol{u} = (u_1, u_2)^T$.

The material properties of the elastic isotropic half-plane considered are: density $\rho = 2000\text{kg/m}^3$, longitudinal wave velocity $c_p \approx 774.6\text{m/s}$, and shear wave velocity $c_s = c_p/2$ in a state of plane strain. The explosive point-source is driven by a *Ricker wavelet* with a dominant frequency $f_c = 1\text{Hz}$.

To accommodate the scenario of waves at near-grazing incidence we have reduced the semi-infinite domain with a free surface to a 3450m × 225m elongated

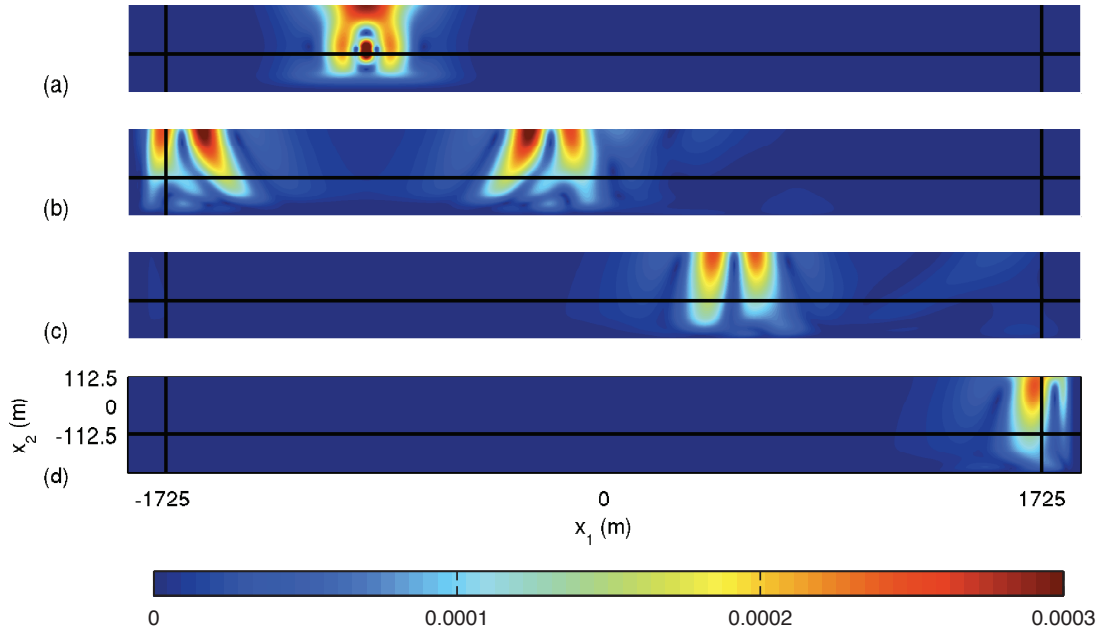


Figure 2.4: Snapshots of the displacement field magnitude - defined as the norm $||u|| = \sqrt{u_1^2 + u_2^2}$ at time steps: **(a)** 1.0s, **(b)** 3.0s, **(c)** 5.0s, and **(d)** 8.5s. The waves are driven by a Ricker wavelet source located at (x_1^c, x_2^c) = (-937.50m, -93.75m).

domain truncated by a 150m thick PML layer, as illustrated in figure 2.2. Both regions are discretized by square elements with edge size $\Delta x_1 = \Delta x_2 = 4.6875$m and we run the simulation for 20s with a mass lumped explicit time integration scheme. At the chosen source frequency the size of the computational domain in vertical direction is comparable to one wavelength. The Ricker wavelet source is located near the bottom-PML, i.e. at (-937.50m, -93.75m), and we have allocated three receivers at (-862.5m, -112.5m), (-1706.25m, 93.75m), and (-1650.0m, 112.5m) to track u_1 and u_2 responses. In figure 2.3 the response records display perfect agreement with the analytical solution when $\beta_i > 1$ and $\alpha_i > 0$ $(i = 1, 2)$ are activated in the stretching function (2.18). If they both remain inactive or only $\alpha_i > 0$ $(i = 1, 2)$ is active the pressure wave will be affected by energy reentering the domain as spurious reflected evanescent waves, see the inset of figure 2.3a.

The near-bottom PML location of the source implies that surface waves form as soon as pressure and shear waves have interacted with the surface. If we consider the snapshots of the displacement magnitude at different time steps in figure 2.4, and simultaneously compare figure 2.3a, -c to 2.3b, -d, respectively, we notice that Rayleigh waves will dominate. The location of the source also gives

rise to an inclined wave vector away from the surface at 3.0s as illustrated in 2.4b. The direction of the wave vector eventually becomes parallel to the free surface, see 2.4c-d, and the propagating waves are efficiently absorbed in the PML.

To this end it should be emphasized that the high stability and efficiency are achieved by imposing homogeneous Direchlet conditions on the outer boundary (Cohen, 2002). Alternatively, the PML could be backed by ABCs, such as the first-order ABC, to enhance the performance even further. The total reflection then becomes $R_{\mathrm{PML-ABC}} = R_{\mathrm{PML}} \cdot R_{\mathrm{ABC}}$ (Jin and Riley, 2009).

2.6 Parallelizing the finite element solver

Time domain analysis of optical waves often requires hundreds of thousands of time steps. Since we will use iterative gradient-based optimization techniques that typically require several hundred optimization steps to reach convergence, all computational efforts will be predominantly related to solving the state field in the evaluation of the objective function and its design gradients. Consequently, parallelization of the state field solver will induce significant speedup in the optimization process.

In the parallelization of the explicit time integration scheme we employ a *non-blocking* strategy to overlap computation and all interprocessor communications. The algorithmic primitives of the strategy, that follow the pseudo-code in table 2.1, provide communication robustness and efficiency (Krysl and Bittnar, 2001), and it can be implemented by using the *message passing interface* (MPI) library (Snir et al., 1995). The explicit time integration scheme in equation (2.16) used for the scalar wave equation requires two half-step evaluations to proceed to the next time step $t + \Delta t$. This means that data synchronization needs to be performed twice. The explicit version of the Newmark-beta scheme in equation (2.15) can suffice with one synchronization step. However, this gain in efficiency sacrifices the dispersion accuracy in the time stepping.

Table 2.1: Pseudo code.

Non-blocking communication algorithm
1. Compute RHS of equation (2.15), (2.16) for boundary elements
2. Post receive requests
3. Send boundary information from step 1 to the buffer
4. Compute RHS of equation (2.15), (2.16) for interior elements
5. Wait for receive requests completion
6. Check that messages have been sent
7. Update state field

To fully exploit the advantages of non-blocking communication, load-balancing is crucial and should therefore allow all processing units to complete their portion of the work simultaneously, at the next global synchronization point. Load-balancing is the primary concern in the mesh partitioning. Secondly comes minimization of communication requirements between the processors. Due to fast *infiniband* network connections on the market today the influence hereof diminishes. For the wave propagation problems presented in this thesis all geometries are represented by regular square elements for which reason it is straightforward to partition the mesh in load-balance without using commercial software such as METIS (Karypis and Kumar, 2009).

2.7 Transmission/reflection spectra

It is common to map the performance of a scattering object by the frequency response of the transmitted and reflected power flux. Instead of computing the flux for each frequency component short pulse excitation in a single time domain simulation offers an efficient way to compute a broad spectrum response. The power through a surface S is determined by the surface integral of the time averaged Poynting vector $\overline{\boldsymbol{P}}^{\omega}(\boldsymbol{x})$

$$P(\omega) = \int_S \boldsymbol{n} \cdot \overline{\boldsymbol{P}}^{\omega}(\boldsymbol{x}) \mathrm{d}S \tag{2.22}$$

where ω is the angular frequency and $\boldsymbol{n}$ is the unit normal vector. The entries of the time averaged Poynting vector for elastic waves is, in tensor notation, expressed by (Royer and Dieulesaint, 2000)

$$\overline{P}_j^{\omega}(\boldsymbol{x}) = \frac{1}{2}\mathrm{Re}\left[-\sigma_{ij}^{\omega}(\boldsymbol{x})^* \dot{u}_i^{\omega}(\boldsymbol{x})\right] \tag{2.23a}$$

for $j = 1, 2, 3$. For electromagnetic waves the Poynting vector is determined by (Jackson, 1999)

$$\overline{\boldsymbol{P}}^{\omega}(\boldsymbol{x}) = \frac{1}{2}\mathrm{Re}\left[\boldsymbol{E}^{\omega}(\boldsymbol{x})^* \times \boldsymbol{H}^{\omega}(\boldsymbol{x})\right] \tag{2.23b}$$

with $(\cdot)^*$ designating the complex conjugate, and $(\cdot)^{\omega}$ indicating the frequency dependency. For a short input pulse it might be tempting to compute the power integral $P(t)$ of the instantaneous Poynting vector, and subsequently employ the Fourier transform to obtain the frequency counterpart. This is incorrect, however, since we need the frequency response of the power flux cf. equations (2.23) and this is not a linear function of fields. The correct approach is to first compute the frequency dependent fields by accumulating the Fourier transformed field

contribution for every point in the flux plane at each time step while progressing in time. The Fourier transform is approximated by

$$h(\omega) = \frac{1}{\sqrt{2\pi}} \sum_n e^{j\omega} h(n\Delta t)\Delta t \approx \frac{1}{\sqrt{2\pi}} \int_0^T e^{j\omega t} h(t) \mathrm{d}t \tag{2.24}$$

Once time stepping is completed, the power flux $P(\omega)$ can then be computed by equation (2.22). To resolve the necessary details in the response, it is important to bear in mind that the desired frequency resolution of the response should be less or equal to the intrinsic resolution $\Delta f = 1/T$ with T denoting the total simulation time.

Transmission is defined by the ratio of the transmitted power $P_t(\omega)$ and the incident power $P_{\text{in}}(\omega)$, i.e. $|T(\omega)|^2 \equiv P_t(\omega)/P_{\text{in}}(\omega)$. In practice, to take numerical artifacts into account, such as numerical dispersion, the transmitted and incident power are determined through the same surface with and without the scattering object, respectively. Thus, two simulations are required to construct the transmission spectrum.

The reflection spectrum is slightly more tricky to compute. The reflected power cannot simply be evaluated by the surface integral of the Poynting vector in backward direction, since this contains both reflected and incident power. Essentially, the Fourier transformed incident field(s) obtained at the surface without the scatterer are subtracted from the registered field. Hence, for electromagnetic waves the reflected Poynting vector is given by

$$\overline{\boldsymbol{P}}_r^{\omega}(\boldsymbol{x}) = \frac{1}{2}\mathrm{Re}\Big\{ [\boldsymbol{E}^{\omega}(\boldsymbol{x}) - \boldsymbol{E}_{\text{in}}^{\omega}(\boldsymbol{x})]^* \times [\boldsymbol{H}^{\omega}(\boldsymbol{x}) - \boldsymbol{H}_{\text{in}}^{\omega}(\boldsymbol{x})] \Big\} \tag{2.25}$$

Again, two simulations are needed, one without the scatterer in which we need to store the Fourier transformed fields in each spatial integration point located on the surface S, and subsequently one with the scatterer in which we subtract the incident field components. The reflection spectrum is then computed by $|R(\omega)|^2 \equiv P_r(\omega)/P_{\text{in}}(\omega)$.

Chapter 3

Topology optimization for transient wave propagation problems

Topology optimization is an iterative optimization procedure that seized the day some 20 years ago when Bendsøe and Kikuchi (1988) introduced the method as a homogenization technique to design continuum structures for minimum compliance. Shortly after, Bendsøe (1989) proposed the SIMP[1] model to obtain structures with well-defined material phases, and his pioneering work has established the basis for a structural optimization methodology that has grown in popularity ever since both in academia and in the industry as a preferred design tool. For some classes of problems commercial software exists e.g. provided by Altair Engineering and FE-Design.

As design method topology optimization seeks to meet prescribed target design response functions, by distributing material freely within a design space, for a given sets of loads and boundary conditions. With this method one may perceive the optimized geometry as a gray-scale image, in which each pixel indicates the density of a material parameter, such as Young's modulus in solid mechanics. Hence, the free pixel-like distribution of material densities in which white and black pixels render void and solid material, respectively, does not restrict the optimized design to attain any particular geometry with fixed boundaries. This inevitably leaves topology optimization as a serious alternative to size and shape optimization methods, and not least Edisonian approaches, in terms of improving the performance of a given structure.

Since its introduction, the method has been subject to extensive research activity which has broadened the areas of applications significantly. This com-

[1] Solid Isotropic Material with Penalization

prises the design of compliant mechanism (Sigmund, 1997; Bruns and Tortorelli, 2001; Pedersen et al., 2001; Li et al., 2004), the crashworthiness of automotive structures subjected to transient loads (Pedersen, 2004; Forsberg and Nilsson, 2007), material with extremal properties (Sigmund, 1994, 1997, 2000; Gibiansky and Sigmund, 2000), and time-harmonic dynamic problems by eigenvalue optimization for free vibrations (Diaz and Kikuchi, 1992; Pedersen, 2000) as well as optimization of the structural dynamic compliance for single and multiple operating frequencies (Ma et al., 1995; Du and Olhoff, 2007).

Topology optimization has also been applied to problems from other (multi)-physical branches such as Stokes flow problems (Gersborg-Hansen et al., 2005; Borrvall and Petersson, 2003), microfluidic mixers (Andreasen et al., 2009), MEMS[1] devices (Larsen et al., 1997; Sigmund, 2001a,b), fluid-structure interaction problems (Andreasen and Sigmund, 2010; Kreissl et al., 2010) and a wide range of wave propagation problems in acoustics and electromagnetism that will be reviewed in the following section 3.1. A comprehensive and detailed overview of the theory, method, and applications of topology optimization is provided by Bendsøe and Sigmund (2004).

3.1 Topology optimization applied to wave propagation problems

Around the beginning of this century Cox and Dobson (1999, 2000) introduced for the first time a material distribution technique to maximize band gaps of 2D photonic crystal (PhC) structures. The idea was later adopted and generalized in a pioneering work by Sigmund and Jensen (2003) that applied the framework of classical topology optimization to elastic wave propagation problems in the systematic design of 2D *phononic* band gap materials and structures. Their study has formed a firm basis for applying topology optimization to all sorts of wave propagation problems, an ongoing research activity that receives a great deal of attention today.

Shortly after its introduction to 2D elastic wave propagation problems, topology optimization has been applied to the design of phononic crystal plate structures (Halkjær et al., 2005) for elastic bending waves, and the optimized designs have also been experimentally verified by Halkjær et al. (2006). For the same type of waves vibration suppressing and energy transporting, non-periodic plate structures have been obtained by Larsen et al. (2009). Different studies for in-plane shear and pressure waves have applied topology optimization to design structures exhibiting minimal absorption and transmittance (Jensen, 2007b) as well

[1] Micro Electro-Mechanical Systems

as non-linear tunable elastic structures with waveguide functionalities (Evgrafov et al., 2008a). The usefulness of the design freedom associated with topology optimization has also been demonstrated on three-dimensional (3D) phononic band gap filter and waveguide devices that support the propagation of surface-acoustic waves (Rupp et al., 2007), and to design structures that accommodate maximal eigenfrequency separation (Jensen and Pedersen, 2006). The performance of acoustic devices ranging from an acoustic horn (Wadbro and Berggren, 2006; Wadbro et al., 2010), sound barriers (Duhring et al., 2008) to devices with structural (Yoon et al., 2007) and optical (Duhring, 2009) interaction included in the analysis has been improved drastically by utilizing topology optimization.

Simultaneously to the studies involving elastic waves, electromagnetic wave propagation problems have been considered. In the microwave regime, an efficient skin-depth resolving methodology for distributing conducting material has been developed by Aage et al. (2010) to optimize energy harvesting metallic devices. Erentok and Sigmund (2011) use a different approach to design 3D sub-wavelength antennas by topology optimization. Prior to these recent studies, topology optimized patch antennas were obtained by Kiziltas et al. (2003, 2004). Optical wave propagation design problems, in particular involving photonic crystal (PhC) structures, have recently been treated extensively by topology optimization to improve the functionality of nano-photonic devices (Jensen and Sigmund, 2011). The research activity within this field encompasses optimization of various PhC devices for E- and H-polarized light, such as geometrical properties for mirrors (Sigmund and Hougaard, 2008), bends (Jensen and Sigmund, 2004; Tsuji et al., 2006; Watanabe et al., 2007), splitters (Jensen and Sigmund, 2005; Watanabe et al., 2006), and waveguide terminations (Frei et al., 2005) with minimal losses. Amalgamation of these devices should ultimately accommodate the full control of light in integrated optical circuits, and experimental verifications of the topology optimized devices have disclosed remarkable improvements in their optical performance (Borel et al., 2004, 2005). In addition to that, topology optimization has proven very useful in tailoring dispersion properties of optical fibers (Riishede and Sigmund, 2008), and PhC waveguides with enhanced slow light behavior initially formulated by (Stainko and Sigmund, 2007) and very recently extended by Wang et al. (2011) to include robustness against manufacturability uncertainties. The design of metamaterials has also been addressed by Sigmund (2009); Diaz and Sigmund (2010) as well as the design of grating couplers for surface plasmons (Andkjær et al., 2010).

Common for all the above mentioned studies, reviewed comprehensively by Jensen and Sigmund (2011), is their frequency domain analysis of time-harmonic waves. Essentially, it implies that one analysis is associated with one driving frequency. A frequency range can be covered by multiple analyses or alternatively by Padé approximations (Jensen, 2007a). Conversely, time domain optimiza-

tion can be employed to handle broadband signals by short pulse excitation with a single analysis cycle. In addition, it also allows for signal shaping and the treatment of nonlinear material resulting in local frequency permutations. Due to these promising features topology optimization of transient wave propagation problems has recently attracted increasing awareness. Reported in the literature to this date, are topology optimization of one-dimensional filters and pulse modulators (Dahl et al., 2008; Jensen, 2009) based on the FETD method, and, conducted by FDTD analyses, optimization of 3D broadband antennas (Nomura et al., 2007) as well as 1D signal shaping devices (Yang et al., 2009). Compared to frequency domain approaches, topology optimization of transient wave propagation problems has clearly received very little attention. General time domain topology optimization has been applied to design structures under dynamics loads by Min et al. (1999), transient heat conduction problems by Turteltaub (2001), and thermally actuated compliant mechanisms by Li et al. (2004).

In the present work, the method of topology optimization is extended to 2D transient wave propagation problems analyzed by the FETD method. Two problems belonging to the optical regime are considered, with the first one formulated as a maximum stored energy optimization problem, and the second one recast as a signal shaping problem. This chapter contains an introduction to the continuous design parametrization (section 3.2), followed by a formulation of a generic time domain optimization problem (section 3.3) and the associated sensitivity[1] analysis (section 3.4). The last part (section 3.5) describes how the computing effort of time domain optimization can be lowered by parallel computing.

3.2 Design parametrization and material interpolation

Topology optimization is here applied to the design of structures subjected to transient waves. A simulation domain Ω_{S} with the boundary $\Gamma_{\mathrm{S}} = \partial\Omega_{\mathrm{S}}$ is excited either by a radiative source located at $\mathbf{x}_{\mathrm{r}} \in \Omega_{\mathrm{S}}$ or an incoming plane wave specified on the boundary $\Gamma_{\mathrm{inc}} \subseteq \Gamma_{\mathrm{S}}$. The aim is then to find a material distribution consisting of air and solid in a prescribed design domain Ω_{D} such that given design response functions are optimized/satisfied in an evaluation domain Ω_{E}, see figure 3.1.

For the design parametrization each element in the design domain is linked to a unique design variable ξ_e for $e \in \Omega_{\mathrm{D}}$ that controls the material phases, such

[1] Sensitivity is here used as the synonym for first-order gradients.

that

$$\xi_e = \begin{cases} 0, & \text{if air} \\ 1, & \text{if solid material} \end{cases} \tag{3.1}$$

Hence, the optimization is intrinsically rendered discrete to obtain pure material phases. To facilitate the use of gradient-based optimization algorithms, and thereby ease the optimization process significantly, the discrete element variables are translated into continuous (density) equivalents satisfying $0 \leq \rho_e \leq 1$ for $e \in \Omega_{\mathrm{D}}$. For convenience, we assemble them in the global design vector $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_M)^T$ with M denoting the number of design variables.

It has been demonstrated for transmittance type problems, that density variables entering a material interpolation scheme linearly, can be successfully adopted to represent material properties of the specific element in wave propagation problems (Sigmund and Jensen, 2003)

$$A(\rho_e) = A_1 + \rho_e(A_2 - A_1) \tag{3.2}$$

Here, A_1 and A_2 denote pure air and solid phases, respectively, and A designates the distributed material property. For optical waves it concerns permittivity, and for elastic waves mass density and Young's modulus (or Lamé coefficients). However, the continuous design parametrization suffers from the drawback that it allows for "intermediate" or "gray" densities to appear in the optimized designs, corresponding to neither of the desired material phases. From a structural point of view the manufacturer needs optimized designs with well-defined boundaries. Hence, in order to steer the optimization process towards $0-1$ designs some sort of penalization is required. Here, if necessary, we use a technique that introduces

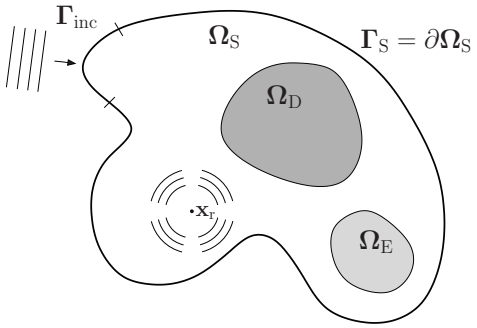


Figure 3.1: Design problem diagram - the simulation domain Ω_{S} contains a design region Ω_{D} and a design response evaluation region Ω_{E}. The structure is subjected to transient propagating waves excited either by a radiative source located at $\mathbf{x}_{\mathrm{r}}$ or an incoming plane wave at $\Gamma_{\mathrm{inc}} \subseteq \Gamma$. On the outer boundary Γ absorbing boundary conditions are imposed.

artificial damping ("Pamping") in intermediate material regions to remove the gray elements from here (Jensen and Sigmund, 2005).

The minimum length scale of the design features is controlled using a density filter that operates on the design as follows (Bourdin, 2001; Bruns and Tortorelli, 2001)

$$\widetilde{\rho}_e = \frac{\sum_{i \in N_e} w(\mathbf{x}_i) v_i \rho_i}{\sum_{i \in N_e} w(\mathbf{x}_i) v_i} \tag{3.3}$$

where N_e is the neighborhood of element e specified by the elements that have centers within a given filter radius R of the center element e. Further, $w(\mathbf{x}_i)$ is the cone-shaped weighting function given by $w(\mathbf{x}_i) = R - ||\mathbf{x}_i - \mathbf{x}_e||$, and v_i denotes the volume of element i in the neighborhood with spatial (center) location $\mathbf{x}_i$.

After density filtering we remove gray elements in the transition region between solid and void by single or dual projection of the material phase with the minimum length scale preserved (Guest, 2009; Sigmund, 2007), or by a single projection with no minimum length scale conserved (Xu et al., 2010). Even though the latter eliminates the minimum length scale, it has proven useful in obtaining well-defined optimized design (Wang et al., 2011).

3.3 Formulating the optimization problem

The motivation for using topology optimization is to optimize an objective subjected to a number of constraints by distributing material freely in a prescribed design domain Ω_{D}.

For time domain optimization in general, a given design functional G can be a function of signal displacement $\mathbf{u}$, velocity $\dot{\mathbf{u}}$, and acceleration $\ddot{\mathbf{u}}$. We also incorporate the option for using various signal processing convolution operations on $\mathbf{u}$ such as the Hilbert transform, high- and low-pass filters, etc., which we, for convenience, designate $\widehat{\mathbf{u}}$.[1] Based on this we define an objective function as

$$\Phi(\boldsymbol{\rho}) = \int_0^T G(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \widehat{\mathbf{u}}, \boldsymbol{\rho}) \mathrm{d}t \tag{3.4}$$

where T is the termination time of the transient simulation, and $\mathbf{u} \equiv \mathbf{u}(\boldsymbol{\rho}, t)$ where the design dependence is implicit, and ρ is the design vector. The problems considered in this thesis only involve at most two constraints $g_0(\boldsymbol{\rho})$ and $g_1(\boldsymbol{\rho})$, of which $g_1(\boldsymbol{\rho})$ is assumed to have a similar form as the design response in equation

[1] Filter convolution is defined by $\widehat{\mathbf{u}}(t) = \int_{-\infty}^{\infty} \mathbf{u}(s) H(t-s) \mathrm{d}s$ where $H(t)$ is the filter kernel.

(3.4). Thus recast on a generic form, a time domain optimization problem can be formulated as

$$\begin{aligned} \min_{\boldsymbol{\rho}\in\mathbf{R}^M} \quad & \Phi(\boldsymbol{\rho}) \\ \text{s.t.:} \quad & \text{Governing equation (2.8)} \\ & g_0(\boldsymbol{\rho}) = V(\boldsymbol{\rho})/V^* - 1 \leq 0 \\ & g_1(\boldsymbol{\rho}) \leq 0 \\ & 0 \leq \rho_e \leq 1, \qquad e \in \Omega_{\mathrm{D}} \end{aligned} \tag{3.5}$$

where $g_0(\boldsymbol{\rho})$ is the so-called volume constraint in which V is the volume occupied by present material, and V^* is the upper bound on the allowable material usage in the design domain Ω_{D}. In wave propagation problems, unless strictly imposed, the volume constraint is not necessarily active. However, when the optimization problems involve several local minima the optimized design will be strongly dependent on the choice of V^*.

We will consider two types of design functionals G can attain: ① *stored energy* in a prescribed output domain Ω_{E}, and ② *signal shaping* at multiple points in Ω_{E}.

The optimization problem in equation (3.5) is solved by the globally convergent method of moving asymptotes (GCMMA) proposed and implemented in FORTRAN 77 by Svanberg (2002), hereafter referred to as the *optimizer*. This is a gradient-based method that can solve non-linear mathematical programming problems. The algorithm employs design information from previous optimization steps and current gradient information based on which it efficiently carries out the design update satisfying global convergence of the objective and constraint functions. The GCMMA is efficient for a large number of design variables and constraints which is typically required in topology optimization problems.

3.4 Sensitivity analysis

For the continuous optimization problem we need to compute the design sensitivities of the objective and the constraint functions in order to facilitate the use of gradient-based optimization algorithms. Essentially, the sensitivities contain information on how much a function will change upon an infinitesimal design variable perturbation.

For a large number of design variables the adjoint variable method (AVM) (Michaleris et al., 1994; Tortorelli et al., 1990, 1991; Tortorelli and Michaleris, 1994; Cardoso and Arora, 1992; Tsay and Arora, 1990; Arora and Haug, 1979) offers a clever and computationally efficient alternative to the direct differentiation method (DDM) (Greene and Haftka, 1991; Hsieh and Arora, 1985, 1984).

In particular, AVM evaluates the sensitivities with computational costs independent of the number of design variables by introducing an auxiliary problem that is always linear.[1] Sensitivity computation is therefore usually comparable to the cost of solving for the state field once. However, the benefits of AVM are only useful when the objective and constraint functions are recast on integral form.

The explicit design sensitivity expression is derived by considering the design response defined in equation (3.4). To make the approach general we proceed by expressing the sensitivities in terms of the residual for the equation governing the state field as a function of the displacement $\mathbf{u}$, velocity $\dot{\mathbf{u}}$ and accelerations $\ddot{\mathbf{u}}$ and the design $\boldsymbol{\rho}$, i.e. $\mathbf{r} \equiv \mathbf{r}(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \boldsymbol{\rho})$. Based on equation (3.4) we form the augmented design response

$$\Phi_{\mathrm{A}} = \int_0^T G(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \widehat{\mathbf{u}}, \boldsymbol{\rho}) \mathrm{d}t + \int_0^T \boldsymbol{\lambda}^T \mathbf{r}(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \boldsymbol{\rho}) \mathrm{d}t \tag{3.6}$$

where the adjoint (Lagrange multiplier) vector $\boldsymbol{\lambda} \equiv \boldsymbol{\lambda}(\boldsymbol{\rho}, t)$ depends implicitly on the design and explicitly on time. The augmented design response is always equal to the original design response in equation (3.6), since $\mathbf{r} = \mathbf{0}$ is ensured from the primal transient analysis when solving equation (2.8). This means that their associated sensitivities must be identical, i.e. $\mathrm{d}\Phi/\mathrm{d}\boldsymbol{\rho} = \mathrm{d}\Phi_A/\mathrm{d}\boldsymbol{\rho}$. By a clever choice of $\boldsymbol{\lambda}$, which can be chosen freely since $\mathbf{r} = \mathbf{0}$, the sensitivity analysis can be considerably simplified as shown in the following.[2]

Differentiation of equation (3.6) with respect to each element design variable ρ_e for $e \in \Omega_{\mathrm{D}}$ by the chain rule, and subsequently utilization of integration by parts yields

$$\begin{aligned}
\frac{\mathrm{d}\Phi_{\mathrm{A}}}{\mathrm{d}\rho_e} = & \left[\left(\frac{\partial F}{\partial \ddot{\mathbf{u}}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}} \right) \frac{\mathrm{d}\dot{\mathbf{u}}}{\mathrm{d}\rho_e} \right]_0^T \\
& + \left[\left(\frac{\partial G}{\partial \dot{\mathbf{u}}} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial G}{\partial \ddot{\mathbf{u}}} - \dot{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}} \right) \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}\rho_e} \right]_0^T \\
& + \int_0^T \left(\ddot{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}} - \dot{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial \mathbf{u}} - \frac{\partial G}{\partial \mathbf{u}} + \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial G}{\partial \dot{\mathbf{u}}} - \frac{\mathrm{d}^2}{\mathrm{d}t^2} \frac{\partial G}{\partial \ddot{\mathbf{u}}} - (-1)^p \widehat{\frac{\partial G}{\partial \widehat{\mathbf{u}}}} \right) \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}\rho_e} \mathrm{d}t \\
& + \int_0^T \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial \rho_e} \mathrm{d}t + \int_0^T \frac{\partial \boldsymbol{\lambda}^T}{\partial \rho_e} \mathbf{r} \mathrm{d}t
\end{aligned} \tag{3.7}$$

where $p = \{0, 1\}$, depending on the ordinate symmetry of the kernel function in the filter convolution operations. When the kernel holds even symmetry $p = 0$,

[1] This is independent of what type the primal state problem is: linear or non-linear.

[2] In the non-linear case the accuracy of the sensitivities obtained from AVM depends on the precision with which we solve for $\mathbf{r} = \mathbf{0}$.

and odd $p = 1$. The implicit system derivatives $\mathrm{d}\mathbf{u}/\mathrm{d}\rho_e$ and $\mathrm{d}\dot{\mathbf{u}}/\mathrm{d}\rho_e$ are annihilated from the sensitivity expression in equation (3.7) by selecting the appropriate $\boldsymbol{\lambda}$. If we assume that the field is not design dependent at $t = 0$, i.e. $\mathbf{u}(0) = \dot{\mathbf{u}}(0) = \mathbf{0}$, the adjoint response is obtained by solving the following terminal value problem

$$\ddot{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}} - \dot{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial \mathbf{u}} = \frac{\partial G}{\partial \mathbf{u}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial G}{\partial \dot{\mathbf{u}}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial G}{\partial \ddot{\mathbf{u}}} + (-1)^p \widehat{\frac{\partial G}{\partial \widehat{\mathbf{u}}}} \tag{3.8a}$$

$$\boldsymbol{\lambda}(T)^T \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}(T) = -\frac{\partial G}{\partial \ddot{\mathbf{u}}}(T) \tag{3.8b}$$

$$-\dot{\boldsymbol{\lambda}}(T)^T \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}(T) + \boldsymbol{\lambda}(T)^T \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}}(T) = -\frac{\partial G}{\partial \dot{\mathbf{u}}}(T) + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial G}{\partial \ddot{\mathbf{u}}}(T) \tag{3.8c}$$

Clearly, the terminal value problem above is always linear with respect to $\boldsymbol{\lambda}$. Once the adjoint response has been found the sensitivities are subsequently computed by the time integral

$$\frac{\mathrm{d}\Phi}{\mathrm{d}\rho_e} = \int_0^T \left(\frac{\partial G}{\partial \rho_e} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial \rho_e} \right) \mathrm{d}t \tag{3.9}$$

that can be evaluated by using the trapezoidal integration rule.

3.4.1 The linear wave equation

Throughout the thesis we consider linear wave propagation problems for which the residual is given by

$$\mathbf{r}(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \boldsymbol{\rho}) = \mathbf{f} - (\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u}) \tag{3.10}$$

where the system matrices $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are assumed design dependent, time-invariant, and symmetric. If we additionally introduce the time shift $t = T - \tau$ and solve for the residual in equation (3.10), the adjoint terminal value problem, based on $\bar{\boldsymbol{\lambda}}(\tau)$, simplifies to

$$\mathbf{M}\ddot{\bar{\boldsymbol{\lambda}}} + \mathbf{C}\dot{\bar{\boldsymbol{\lambda}}} + \mathbf{K}\bar{\boldsymbol{\lambda}} = \mathbf{f}_a(T - \tau) \tag{3.11a}$$

$$\bar{\boldsymbol{\lambda}}(0)^T \mathbf{M} = -\frac{\partial F}{\partial \ddot{\mathbf{u}}}(T) \tag{3.11b}$$

$$\dot{\bar{\boldsymbol{\lambda}}}(0)^T \mathbf{M} + \bar{\boldsymbol{\lambda}}(0)^T \mathbf{C} = -\frac{\partial G}{\partial \dot{\mathbf{u}}}(T) + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial G}{\partial \ddot{\mathbf{u}}}(T) \tag{3.11c}$$

with the adjoint load

$$\mathbf{f}_a(t) = \left(\frac{\partial G}{\partial \mathbf{u}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial G}{\partial \dot{\mathbf{u}}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial G}{\partial \ddot{\mathbf{u}}} + (-1)^p \widehat{\frac{\partial G}{\partial \widehat{\mathbf{u}}}} \right)^T \tag{3.11d}$$

since $\mathrm{d}/\mathrm{d}t = -\mathrm{d}/\mathrm{d}\tau$. Thus, the original terminal value problem has been transformed into an initial value problem. We recover the original adjoint variable by substituting $\boldsymbol{\lambda}(T-\tau) = \bar{\boldsymbol{\lambda}}(\tau)$, and the sensitivities are subsequently evaluated by equation (3.9).

In the situation where the design functional G only depends on the state field $\mathbf{u}$ and $\hat{\mathbf{u}}$ the initial conditions in equations (3.11b)-(3.11c) reduce to $\boldsymbol{\lambda}(0) = \dot{\boldsymbol{\lambda}}(0) = \mathbf{0}$. Nonetheless, the result in equation (3.11a) is an adjoint problem whose form is identical to that of the primal transient analysis in equation (2.8), albeit with the adjoint load, cf. equation (3.11d), acting as the excitation term. Hence, the same integration scheme can be used to find the adjoint response.

The evaluation of the sensitivities by AVM proceeds as follows: After the transient primal analysis for the state field is concluded, the adjoint response $\boldsymbol{\lambda}$ is computed at the exact same time steps by reusing the exact same ABC setup from the primal analysis (Chung et al., 2000). While integrating the adjoint response in time, the integrand in equation (3.9) is evaluated and the contribution to the sensitivities is added by using the trapezoidal integration rule. Consequently, when time marching of the adjoint equation (3.11a) is completed the sensitivities are obtained. This process requires storage of $\mathbf{u}$ from which, when needed, $\ddot{\mathbf{u}}$ and $\dot{\mathbf{u}}$ are computed by equations (2.11)-(2.12).

If density filtering techniques have been used to regularize the topology optimization problem, the sensitivities are based on the filtered design variable $\widetilde{\rho}_e$. However, the design update needs the originally intended sensitivities $\mathrm{d}\Phi/\mathrm{d}\rho_e$. We recover those from $\partial\Phi/\partial\widetilde{\rho}_e$ by using the chain-rule

$$\frac{\mathrm{d}\Phi}{\mathrm{d}\rho_e} = \sum_{i \in N_e} \frac{\partial \Phi}{\partial \widetilde{\rho}_i}\frac{\mathrm{d}\widetilde{\rho}_i}{\mathrm{d}\rho_e} \tag{3.12a}$$

with

$$\frac{\mathrm{d}\widetilde{\rho}_i}{\mathrm{d}\rho_e} = \frac{w(\mathbf{x}_e)v_e}{\sum_{j \in N_i} w(\mathbf{x}_j)v_j} \tag{3.12b}$$

for the standard density filter in equation (3.3).

3.4.2 Localizing functions

In practical problems point-wise constraints and the need for controlling constraints and the objective in a localized time interval very often occur. On integral form the point-wise constraint can be incorporated by introducing the Dirac delta function (Tortorelli and Haber, 1989). We suggest that a time interval $[T_1, T] \subseteq [0, T]$ can be specified by the Heaviside step function $\bar{u}(t - T_1)$ as an alternative to splitting the time integral in equation (3.4) into $\int_{T_1}^{T} = \int_0^T - \int_0^{T_1}$ which requires two adjoint analyses. As a result, the design response function becomes

$$\Phi(\boldsymbol{\rho}) = \int_0^T G(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \widehat{\mathbf{u}}, \boldsymbol{\rho}) L(t) \mathrm{d}t \tag{3.13}$$

where $L(t)$ denotes the localizing function which modifies the adjoint load in equation (3.11d) to $\mathbf{f}_a(t)L(t)$.

3.5 Parallel computing applied to topology optimization

The iterative nature of the topology optimization method implies the necessity for repeating the state field and the sensitivity analysis (e.g. by the FE method) to succeed each design update until convergence is reached. The flowchart of the optimization procedure appear in figure 3.2. For wave propagation prob-

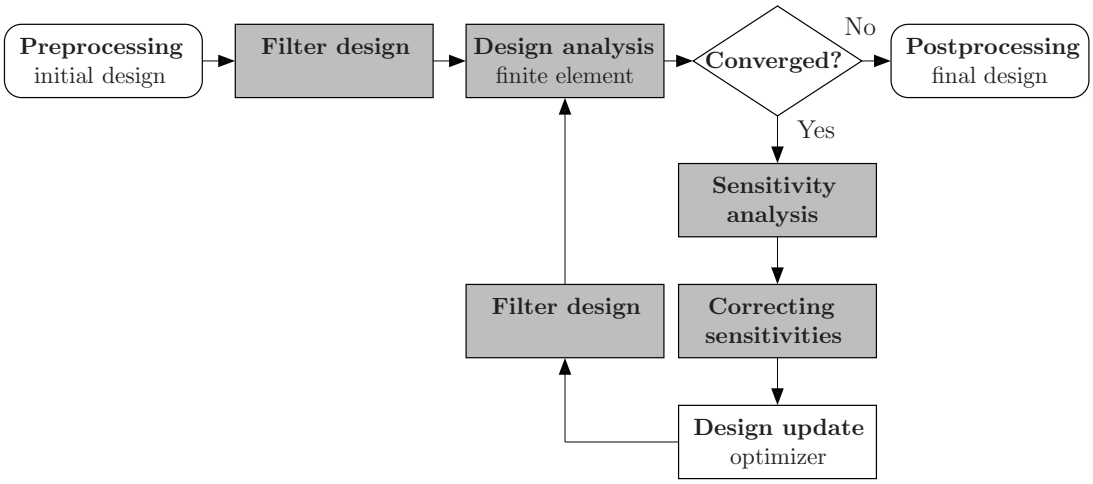


Figure 3.2: Optimization flowchart - for the design problems encompassed by the generic formulation in equation (3.5). The gray boxes indicate parallelized steps in the optimization process.

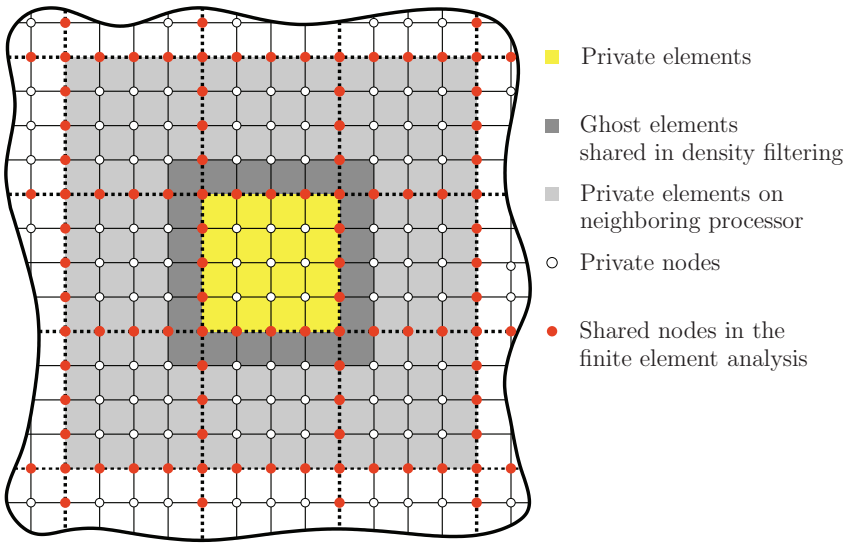


Figure 3.3: Mesh decomposition diagram - based on a node-cut partitioning through the red nodes. The center processor (containing yellow elements) with one element thick ghost layer share boundary nodes with eight adjacent processors. The ghost layer is shared in the "Correcting sensitivity" step in the optimization flowchart c.f. figure 3.2.

lems it is not uncommon that the optimization process requires several hundreds of optimization steps in order to converge. We saw in the previous section 3.4 that sensitivity evaluation in time domain optimization is associated with high demands on memory storage, since the field from the primal forward analysis needs to be stored in order to be reused in the adjoint load in equation (3.11d) and in the computation of the sensitivities by the time integral in equation (3.9). Additionally, to allow for detailed structural features in the optimized designs high element resolution is required. On this basis topology optimization inherently calls for the use of parallel computing, and recently it has been applied to topology optimization of different types of steady-state problems, see Borrvall and Petersson (2001); Kim et al. (2004b); Mahdavi et al. (2006); Evgrafov et al. (2008b); Aage and Lazarov (2011).

The intrinsic element-based density concept behind topology optimization renders the optimization method a natural subject to a parallel computing strategy that relies on a node-cut partitioning of the mesh. As depicted in figure 3.3 this partition method means cuts through finite element edges and faces connected by the (red) border nodes. Sharing of nodes amongst the processors requires very little duplication of computing effort. Advantageously, data locality is thereby retained, and it aligns with the diagonal nature of the mass matrix in our explicit time integration scheme permitting the FE equation for the nodal degrees of free-

dom to be solved independently of one another. The node-cut mesh partitioning holds superior candidacy since all computing efforts in our case, i.e. explicit time integration of the FE equations, density filtering, and sensitivity analysis, are associated with element based calculations (Krysl and Bittnar, 2001; Danielson and Namburu, 1998).

When parallelizing the topology optimization procedure there are three aspects to consider: density filtering, the state field solver, and the optimizer. According to Amdahl's Law (Amdahl, 1967) speedup that scales linearly with the number of processors is only achieved for a large number of processors if all three functions are parallelized. However, we only parallelize the first two, still yielding linear scaling for a maximum number of 288 processors available. The design is updated on the *root* processor to which sensitivity information from each contributing processor is sent. The updated design is subsequently broadcasted back to the relevant processors.

Parallelization of the density filtering process follows naturally as a result of node-cut partitioning of the mesh. However, to facilitate the use of density filters we need to add a ghost layer to each partition, see figure 3.3. The thickness of the layer corresponds to the filter radius measured in an element number metric. Here, it is important to stress that the elements in the ghost layer become shared with the neighboring processors in the "Correcting sensitivity" (chain rule) step of the optimization flowchart, cf. figure 3.2, due to the implications of the chain rule in equation (3.12b) when the original sensitivities $\mathrm{d}\Phi/\mathrm{d}\rho_e$ for $e \in \Omega_\mathrm{D}$ have to be retrieved. Hence, global synchronization of the processing units is required here.

Chapter 4

Designing optical devices for applications

In this chapter three optical components are designed by the use of topology optimization. The first problem optimizes an optical taper (4.2). This is followed by optimization of a photonic crystal notch filter (4.3), and the final design problem of the chapter considers optimization of slow light devices (4.4). However, we start out by introducing the concept of *photonic crystals*.

4.1 Two-dimensional photonic crystals

Photonic crystals (PhC) are periodic optical nano-structures that are designed to control and manipulate the motion of *photons* (light modes) in a number of desirable ways (John, 1987; Yablonovitch, 1987). They serve as the optical analogue to the periodicity of a semiconductor crystal that can manipulate the motion of *electrons* on a microscopic level if atoms and molecules are replaced by macroscopic media with differing dielectric constants of high contrast. Similar to how a periodic potential of semiconducting material can inhibit the motion of electrons, one can, by a periodic distribution of dielectric, construct photonic crystals with *photonic band gaps* (PBGs), preventing light from propagating in certain directions with specified frequencies. Conversely, a photonic crystal can, if properly designed, accommodate *propagating modes* in PhC waveguides (WGs), and it can also support resonant *localized modes* inside microcavities (MCs).

The advantage of PhCs is their ability to provide strong light-matter interaction while keeping the size of the optical components greatly reduced. Amalgamation of the PBG mirror, PhC-WG, and PhC-MC as the basic components has for instance been utilized to design various PhC devices such as filters, bends, and splitters for various applications (Joannopoulos et al., 2008). Additionally,

the unique tunable dispersion properties of PhCs allow for the generation of slow light, leaving a PhC as an obvious candidate for engineering in the realization of on-chip integration of slow-light devices (Baba, 2008).

Even though 3D PhCs exist, we here limit the study to 2D crystals. This implies that the PhC is periodic along two of its axes and homogeneous along the third axis. In the inset of figure 4.1, a *supercell* sample consisting of a triangular lattice of air holes embedded in dielectric material with a line defect is shown, and we imagine the specimen to be infinitely tall. For certain geometrical configurations this crystal can have a PBG in the x_1x_2 plane, comprised by the light gray rectangular region of the band diagram in figure 4.1. Inside the band gap guided modes can live if defects are introduced in the perfect lattice.

As described in section 2.2, electromagnetic waves can be separated into two distinct polarizations in the 2D case; TE and TM modes, for which the band structures can be fundamentally different. Here, we consider TE modes, because their band gaps are favored in triangular lattices compared to TM band gaps. They are governed by the scalar wave equation (2.4) with $u = H_3$, $A = 1/\varepsilon_r$ and $B = c^{-2}$. Because we only model the transverse fields it is adequate to use nodal-based finite elements in order to satisfy the electromagnetic field continuity equations (Jin and Riley, 2009).

Since we consider macroscopic systems containing dielectric material we utilize

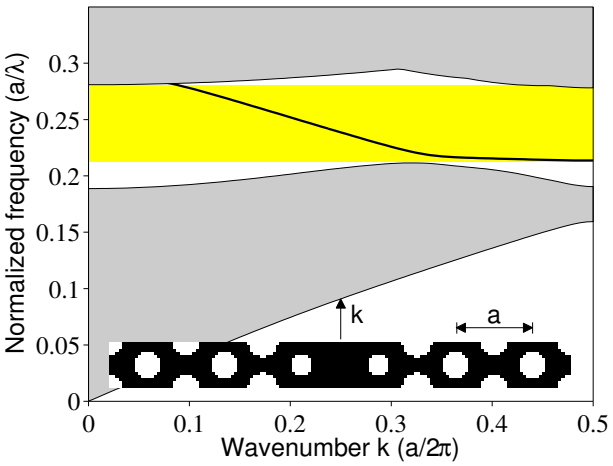


Figure 4.1: The projected band structure - of the line defect (inset) formed by filling a row of air holes in a perfect crystal plotted versus the wave vector component k along the waveguide axis. The extended modes in the crystal become continuum regions (gray), whereas inside the band gap (light gray rectangular region) a defect band is introduced. The pitch/lattice constant of the photonic crystal is denoted a.

that no fundamental *length scale* exists by introducing $(x_1', x_2') = (x_1, x_2)/a$ to the scalar wave equation (2.4) where a is a characteristic length, typically the *pitch* (or lattice constant) of the PhC (Joannopoulos et al., 2008). This will lead to the scaled time $t' = tc/a$, and scaled angular frequency $\omega' = \omega a/c$ whereby the master equation becomes *scale invariant*. In the remainder of this chapter the scaled parameters are used, unless otherwise stated, and it is straightforward to switch between these and the non-scaled parameters as soon as a is known.

4.2 Topology optimization of an optical taper [P1]

The optical taper is an essential component when several optical functions have to be integrated into one chip. This is e.g. the case if two waveguides have to be seamlessly joined.

A possible application could be to guide light from an external dielectric ridge waveguide (WG) into a PhC-WG, which will be considered in section 4.4. Here, however, we start out by optimizing a tapering device with dielectric contrast $\varepsilon_r = 11.56$ coupling two *monomodal*[1] ridge waveguides with the setup given in figure 4.2a. The goal is to distribute dielectric material inside the design domain Ω_D such that waveguide tapering from 2 µm to 0.5 µm over a 0.5 µm distance is achieved without backscattering effects by maximizing the output energy. This leads to the following optimization formulation

$$\begin{aligned} \min_{\boldsymbol{\rho}\in\mathbf{R}^M} \quad & \Phi(\boldsymbol{\rho}) = -\int_0^T \sum_{e\in\Gamma_\mathrm{out}} \mathbf{u}^T\mathbf{Q}^e\mathbf{u}\mathrm{d}t \\ \text{s.t.:} \quad & \text{Governing equation (2.8)} \\ & 0 \le \rho_e \le 1, \qquad e \in \Omega_\mathrm{D} \end{aligned} \tag{4.1}$$

where $\mathbf{u}$ holds the discrete values of the transverse magnetic field, H_3, and $\mathbf{Q}^e$ is a diagonal matrix with non-zero entries for the degrees of freedom at Γ_out.

For the incoming waves we use a Gaussian distributed wave packet with dominant frequency $\omega_0 a/2\pi c = 0.37$ and a *full width at half maximum* (FWHM) bandwidth equal to 0.2. For the spatial and time discretization $\Delta x_1 = \Delta x_2 = a/20$ and $\Delta t = 0.9\Delta x_1/c$, respectively, have been used.

The initial material distribution is not based on prior knowledge of potential well-performing geometries, such as the parabolic taper (Luyssaert, 2005). The optimization is started from a homogenous material distribution corresponding

[1] A monomodal waveguide only supports one guided mode.

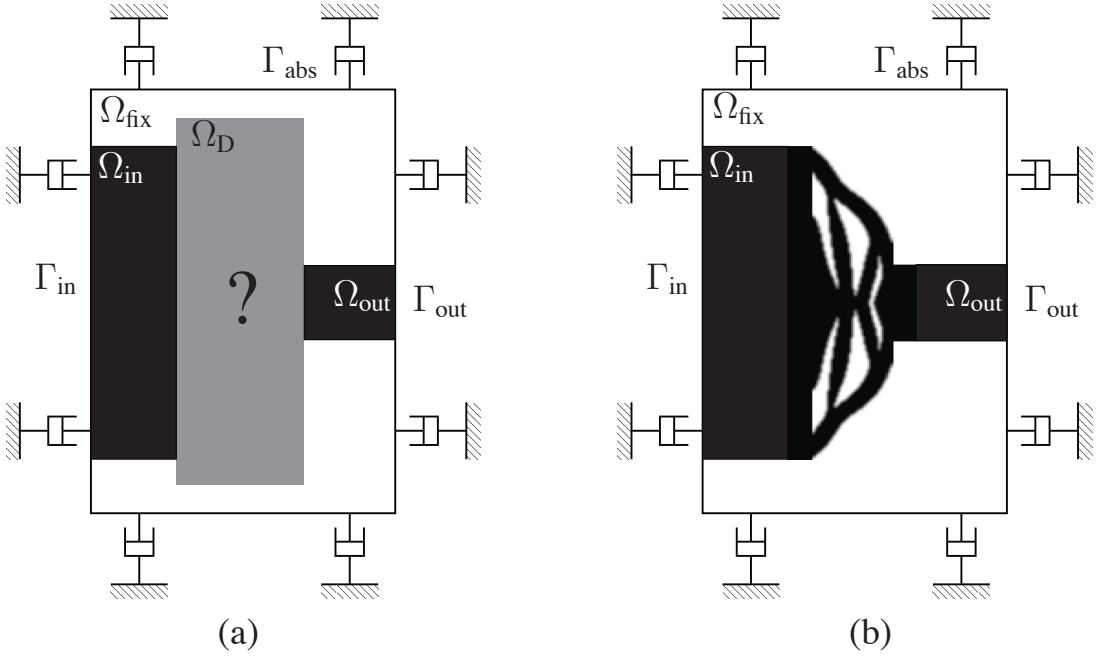


Figure 4.2: The optical taper - (a) The design setup with Ω_{D} as the design domain of length 0.5 µm. The monomodal wave is specified at Γ_{in} in the inlet waveguide Ω_{in} of width 2 µm, and exits in the outlet waveguide of Ω_{out} of width 0.5 µm. Energy is maximized at the ouput port Γ_{out}. **(b)** Optimized taper design.

to $\rho_e = 0.5$ for $e \in \Omega_{\mathrm{D}}$. In the region with inactive elements, the in- and outlet regions consist of dielectric material, the remainder is free space. We have introduced a minimum length scale by the standard density filter with radius $R = 2.5\Delta x_1$ cf. equation (3.3).

The result of the optimization is the taper design in figure 4.2b that exhibits almost zero-backscattering. Hence, if we compare this taper design to e.g. a linear taper, which is a commonly used geometry for optical tapers, the performance has been significantly improved, for further details see [P1].

4.3 Topology optimization of a photonic crystal notch filter [P2]

In signal processing, a *band-stop* or *band-rejection* filter is a frequency-selective device that discriminates among the various frequency components of an input signal. By proper design of the filter we can create a device that passes most frequencies unaltered and simultaneously attenuates the input signal within a specified frequency range to very low level.

Among the class of band-stop filters, the notch filter is characterized by having a narrow stop band. Optical notch filters are essential devices in *Raman spectroscopy*. Here, they are used to remove the incoming laser signal from the detected signal such that Raman shifted satellite peaks, with energy closely related but several orders of magnitude weaker than that of the laser, are not eclipsed by the stray light and the molecule informations of the given subject under investigation can be clearly discerned (Somerville et al., 2010). Furthermore, notch filters are standard components in digital live sound reproduction systems and in instrument amplifiers used to trim or hinder signal feedback without affecting the rest of the transmission spectrum.

We will realize the notch filter by using a 2D PhC composed of a triangular lattice of air holes in the high-contrast dielectric material GaAs, in which we introduce a line defect waveguide and a point defect cavity. By using the hole radius $r/a = 0.35$ the PhC supports a complete TE band gap in the normalized frequency range $a/\lambda = 0.21 - 0.33$, where λ denotes the wavelength in vacuum. The dielectric contrast between GaAs and air is $\varepsilon_r = 11.4$ for wavelengths around 1.5 μm. A schematic drawing of the filter is shown in figure 4.3. Based on this rather abstract diagram of the energy flow inside the system, we will use *temporal coupled-mode theory* to predict quantitative features about the filter functionality, and subsequently employ these predictions as inspiration to formulating the optimization problem.

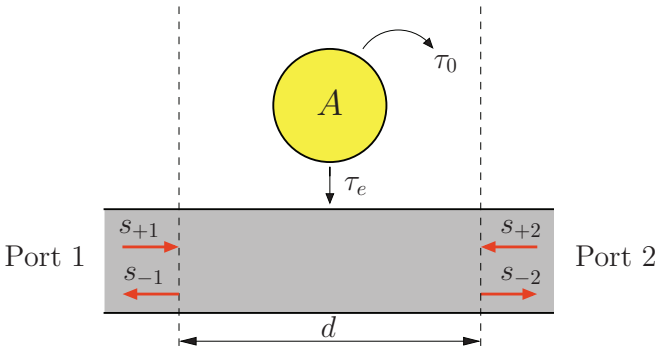


Figure 4.3: PhC filter feature diagram - a single-mode waveguide with input/output s_{+1}/s_{-1} and s_{+2}/s_{-2} at port 1 and 2, respectively; and a resonant single-mode of field amplitude A (thus with stores energy $|A|^2$) and frequency ω_0 coupled to the waveguide with lifetime τ_e, and with loss lifetime τ_0 to the PhC. The dashed lines indicate the reference planes.

4.3.1 Temporal coupled-mode theory and the Q factor

The main idea behind *temporal coupled-mode theory* (CMT) proposed by Haus (1984) is to describe the coupling between a set of distinct idealized components, i.e. the cavity and the waveguide in our case, and thereby make analytical predictions about the physical features of the coupled system in figure 4.3. The key assumption in CMT is slow energy exchange between the involved components, that will result in a weak coupling. For the present system we guarantee that by surrounding the cavity with a sufficient number of PhC layers.

As a result of the PBG the light can only live inside the cavity and in the waveguide, for which reason the remainder of the PhC is neglected in the analysis. The standing mode in the cavity, whose amplitude A has time dependence $\exp(-j\omega_0 t)$, has a resonant frequency ω_0. In the waveguide, the amplitudes of the incoming (outgoing waves) are denoted by s_{+1} (s_{-1}) and s_{+2} (s_{-2}). Thus, s_{+1} and s_{-2} are amplitudes in the forward direction, whereas s_{-1} and s_{+2} are amplitudes in the backward direction. Associated power is determined by the absolute value of the amplitudes squared.

In the present system the cavity has two loss mechanisms: The mode decays with lifetime τ_e and τ_0 into the waveguide and the surrounding PhC, respectively. The net lifetime is then given by $1/\tau = 1/\tau_e + 1/\tau_0$. Input energy from s_{+1} in the waveguide can potentially couple into the cavity, be reflected back to s_{-1}, or transmitted to s_{-2} (the opposite scenario with an incoming wave from s_{+2} also exists). Consequently, the CMT equations governing the amplitude evolution of the cavity mode in time are given by (Haus, 1984)

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \left(-j\omega_0 - \frac{1}{\tau_0} - \frac{1}{\tau_e}\right) A + \kappa_1 s_{+1} + \kappa_2 s_{+2} \tag{4.2}$$

$$s_{-1} = e^{-j\gamma d}(s_{+2} - \kappa_2^* A) \tag{4.3}$$

$$s_{-2} = e^{-j\gamma d}(s_{+1} - \kappa_1^* A) \tag{4.4}$$

where j denotes the imaginary unit. In equation (4.2), $1/\tau_0$ is the decay rate due to losses, $1/\tau_e$ is the rate of decay into the waveguide, κ_1 and κ_2 are the input coupling coefficients associated with forward and backward propagating modes. In equations (4.3)-(4.4), γ is the propagation constant of the waveguide mode, d is the distance between in- and output ports, and $(\cdot)^*$ stands for complex conjugation. Based on power conservation, see Manolatou et al. (1999), we obtain the coupling equations

$$\kappa_i = \sqrt{\frac{1}{\tau_{e_i}}} e^{j\theta_i}, \quad i = 1, 2 \tag{4.5}$$

with $1/\tau_{e_1}$ and $1/\tau_{e_2}$ defined as the decay rates into the waveguide in the forward

and backward direction, respectively, satisfying $1/\tau_{e_1} + 1/\tau_{e_2} = 2/\tau_e$.[1] Further, θ_i are the respective phases which have been necessary to introduce, in order to meet a potential phase mismatch between the cavity and waveguide mode, and the reference planes (shown as dashed lines in figure 4.3). Assuming an incoming wave at port 1 (i.e. $s_{+2} = 0$) whose associated field s_{+1} has a $e^{-j\omega t}$ time dependence, the field everywhere in the system, due to linear material behavior, must also oscillate equivalently (i.e. $\mathrm{d}A/\mathrm{d}t = -j\omega A$). If we define the transmitted frequency response as $T(\omega) \equiv s_{-2}/s_{+1}$, then at steady state, by substituting equation (4.2) into (4.4), we obtain

$$\frac{s_{-2}}{s_{+1}} \equiv T(\omega) = -e^{-j\gamma d}\left(1 - \frac{\frac{1}{\tau_{e_1}}}{-j(\omega-\omega_0) + \frac{1}{\tau_0} + \frac{1}{\tau_e}}\right) \tag{4.6}$$

Instead of the decay rates, it is more convenient to measure the performance of the PhC notch filter based on the dimensionless *quality factor*, and by definition the *unloaded* and the *external* quality factors relate to the decay rates by $Q_0 = \omega_0\tau_0/2$ and $Q_e = \omega_0\tau_e/2$, respectively. The *loaded* (total) in-plane Q factor of the system is determined by

$$\frac{1}{Q_{\text{in}}} = \frac{1}{Q_e} + \frac{1}{Q_0} \tag{4.7}$$

In terms of the Q factors, the filter transmission spectrum is given by

$$|T(\omega)|^2 = 1 - \frac{\frac{1}{4Q_e^2}\left(1 + 2\frac{Q_e}{Q_0}\right)}{\left(\frac{\omega-\omega_0}{\omega_0}\right)^2 + \frac{1}{4Q_{\text{in}}^2}} \tag{4.8}$$

This response will produce a Lorentzian peak at the resonance frequency ω_0 with the fractional width $\Delta\omega/\omega_0$ at half maximum given by $1/Q_{\text{in}}$. At resonance, i.e. $\omega = \omega_0$, it simplifies to

$$|T(\omega_0)|^2 = \frac{Q_{\text{in}}^2}{Q_0^2} \approx \frac{Q_e^2}{Q_0^2} + O([Q_e/Q_0]^3) \tag{4.9}$$

where the rightmost expression is for the usual case $Q_0 >> Q_e$, i.e. when the energy decay rate from the isolated cavity is several orders of magnitude smaller than the decay rate from the cavity into the waveguide (Kim et al., 2004a). We learn from equation (4.9) that $Q_{\text{in}} \approx Q_e$ which means that the energy decay inside the cavity is only determined by the coupling power between the waveguide

[1] Since the cavity mode consists of a pure standing wave it decays equally into backward and forward propagating waveguide modes, with no net power flux in either direction inside the cavity.

and the cavity. It leaves Q_e as the limiting factor on the total performance Q_{in} of the PhC notch filter. Another intriguing property in this physical limit is maintenance of strong coupling (i.e. very low transmission), even upon considerable improvement of Q_e.

Similarly for the reflected frequency response defined by $R(\omega) = s_{-1}/s_{+1}$, we find that the reflected energy at resonance is given by

$$|R(\omega_0)|^2 = \frac{1}{(1+\frac{Q_e}{Q_0})^2} \approx 1 - 2\frac{Q_e}{Q_0} + O([Q_e/Q_0]^3) \tag{4.10}$$

if $Q_0 >> Q_e$. This shows us that energy is conserved at resonance, i.e. $|R(\omega_0)|^2 + |T(\omega_0)|^2 = 1$, where transmitted energy drops to zero and all light is reflected back towards the waveguide entry.

To measure the coupling performance of the notch filter, one may record the transient decay of the stored energy U_0, reached at time $T_{\max}$ inside the cavity given by (Jackson, 1999)

$$U(t) = U_0 e^{-\omega_0 t/Q_{\text{in}}} \tag{4.11}$$

from which Q_{in} can be computed. We previously learned about the present system that $Q_e \approx Q_{\text{in}}$, c.f. equation (4.9). Hence, once Q_{in} is known Q_e can be directly obtained. The functionality of a notch filter is improved if the energy decay to the waveguide is delayed as much as possible. This corresponds to narrowing the Lorentzian peak in the transmission spectrum. Therefore, the goal is to enhance the external quality factor Q_e.

4.3.2 Formulation and method

The purpose of the presented method is to distribute air and dielectric material in the design domain Ω_{D} around the cavity such that the stored energy in the cavity core region Ω_{E} is maximized within the decaying regime of the transient energy response. Including the Heaviside step function in the objective allows us to specify the elapsed time T_1, after which the energy should be maximized. By choosing T_1 large enough the optimization should favor cavity designs with lower decay rate, and the choice of T_1 is expressed in multiples of $T_{\max}$.

The optimization problem with the setup given in figure 4.4 can be formulated as

$$\begin{aligned} \min_{\boldsymbol{\rho}\in\mathbf{R}^M} \quad & \Phi(\boldsymbol{\rho}) = -\log_{10}\left[\int_0^T \sum_{e\in\Omega_{\text{E}}} \mathbf{u}^T\mathbf{Q}^e\mathbf{u}\bar{u}(t-T_1)\mathrm{d}t\right] \\ \text{s.t.:} \quad & \text{Governing equation (2.8)} \\ & g_0(\boldsymbol{\rho}) = V(\boldsymbol{\rho})/V^* - 1 \leq 0 \\ & 0 \leq \rho_e \leq 1, \qquad e \in \Omega_{\text{D}} \end{aligned} \tag{4.12}$$

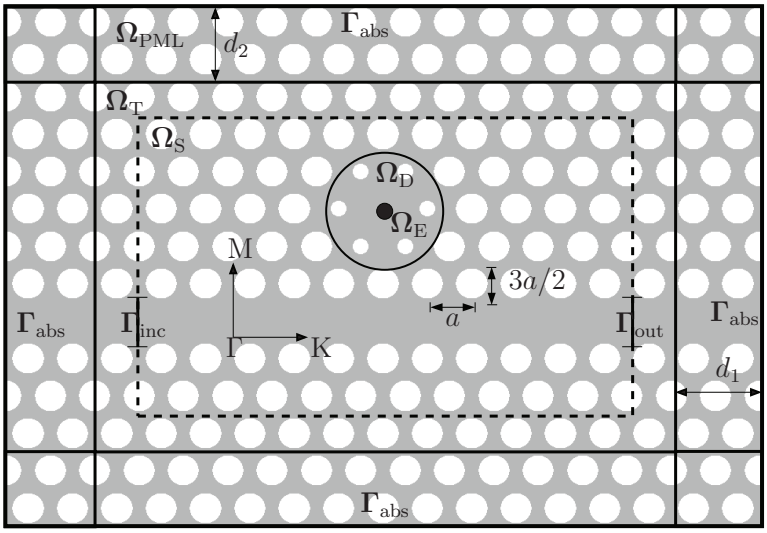


Figure 4.4: Notch filter optimization setup - $\Gamma - \mathrm{K}$ directional PhC-WG coupled to a PhC-MC. The solution domain Ω_{S} and the transition domain Ω_{T} are truncated by the PML region Ω_{PML}. The circle encloses design domain Ω_{D}. In Ω_{E} the energy is maximized. The PhC is built by blocks of size $a/2 \times \sqrt{3}a/2$, with $\Omega_{\mathrm{T}} \cup \Omega_{\mathrm{S}}$ containing 30×14 building blocks, and Ω_{PML} elongated 24 and 4 blocks on both sides in direction $\Gamma - \mathrm{K}$ and $\Gamma - \mathrm{M}$, respectively.

where $\mathbf{u}$ holds the discrete values of the transverse magnetic field, H_3, and the entries of the element matrix $\mathbf{Q}^e$ are determined by equation (2.9a). The logarithm has been introduced to the objective function to ensure better numerical scaling. We obtain a continuous design parametrization through a linear interpolation of $1/\varepsilon_r$ between the two candidate materials, c.f. equation (3.2). The objective integral in equation (4.12) is smoothen to avoid the introduction of any unnecessary numerical resonance artifacts in the adjoint load. The strategy can be found in [P2].

As illustrated in figure 4.4 we include the nearest neighboring holes to the cavity in the design domain Ω_{D}, since it has been demonstrated that alternating their geometry can improve the Q factor significantly (Painter et al., 1999; Kim et al., 2004a). The field inside the device is generated by the Gaussian modulated source

$$g(t) = e^{-(t-t_0)^2/T_0^2} \sin[2\pi\omega_0(t-t_0)] \tag{4.13}$$

where ω_0 denotes the dominant angular frequency. The $1/e$-intensity half width of the Gaussian distribution is set to $T_0 = 180a/c$ centered around $t_0 = 500a/c$. These settings produce a pulse with FWHM that, in the band gap range, corresponds to approximately $60 - 100$ wave periods (or wavelengths in vacuum). For the spatial and time discretization we use $\Delta x_1 = \Delta x_2 = a/14$ and $\Delta t = 0.9\Delta x_1/c$,

respectively.

The solution of the resonant type optimization problem in equation (4.12) suffers from strong non-uniqueness that will lead to multiple local minima. Some of them are results of the possible existence of degenerate modes (Joannopoulos et al., 2008), others arise from local resonance effects yielding poor performance away from the target frequency (Jensen and Sigmund, 2005). Resonance type optimization problems are inherently sensitive because very small design changes in a tangle of local minima can easily deteriorate the physical performance at resonance. Therefore, to prevent the optimization algorithm from roaming around among different local minima in the beginning of the optimization process, we have found it necessary to employ a globally convergent algorithm, such as the GCMMA.

In order to avoid any non-manufacturable geometry features in the optimized designs we combine the standard density filter in equation (3.3) with two inversely related projection schemes, one linked to each material phase to control the length scale of void and dielectric simultaneously. This type of multiphase projection will lead to intermediate design variables in the transition regions between void and solid. We penalize this trend by introducing artificial mass proportional damping ("Pamping") in gray elements (Jensen and Sigmund, 2005).

The design process is executed in a sequence of three consecutive steps: First, based on a trial and error radius variation of nearest neighboring holes to the cavity, we find a well-performing geometry. This serves as an initial guess in the optimization of the isolated cavity, and if the performance is improved, the optimized cavity is subsequently used to start the optimization of the coupled device.

4.3.3 Results

In the search for a cavity geometry that supports a high Q mode as initial guess, we have found the geometry in figure 4.5a. It pulls down a monopole H_3 mode from the air band at the frequency $\omega_0 a/2\pi c = 0.3030$ with $Q_0 = 2.2 \times 10^5$. We retrieve the corresponding mode in figure 4.5c by the inverse Fourier transform. For the numerical simulations the solution domain Ω_{S} is surrounded by a two cell sizes thick PML.

The design region Ω_{D} used for the coupled system in figure 4.4 is also adopted for the isolated cavity. In the cavity center we excite the monopole mode by a point source with a temporal variation given in equation (4.13), and run the simulation for $T = 100,000\Delta t$. The objective has been evaluated in a circular region with radius $3a/14$ in the cavity center. It should be emphasized that the choice of T_1 in the objective function is not critical here, since the isolated cavity has only a single loss mechanism. To enhance manufacturability we impose

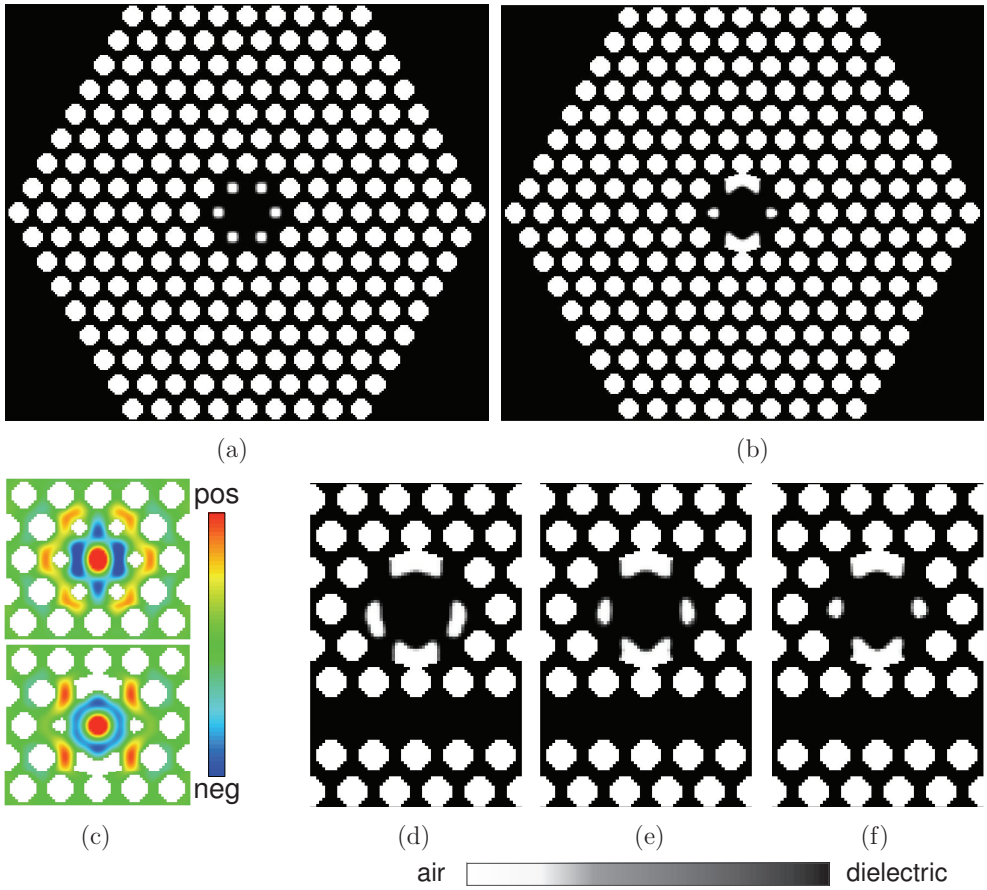


Figure 4.5: Monopole cavity - isolated cavity with **(a)** initial geometry and **(b)** optimized geometry, and **(c)** corresponding mode distribution. **(d)-(f)** coupled system designs for **(d)** $T_1 = 0$, **(e)** $T_1 = T_{\max}$, and **(f)** $T_1 = 4T_{\max}$.

vertical and horizontal symmetry conditions through the cavity center, and do not constrain the amount of dielectric material, i.e. $V^* = \sum_e v_e$ with v_e denoting the element volume.

The optimized design reached after 311 design iterations appears in figure 4.5b. As expected it does not include any structural features that violate the minimum length scale prescribed by the density filter with radius $R = 2.5a/14$, and the blending region between air and dielectric material corresponds to the thickness of one finite element. The Q factor of the associated mode is computed to be $Q_0 = 4.8 \times 10^5$. Clearly, the initial design as well as the mode pattern in figure 4.5 have been altered significantly to achieve the doubling of the Q factor. An overview of all calculated Q factors can be found in table 4.1 below.

In practical applications of the PhC notch filter the initial geometry and corresponding mode cannot be chosen uncritically. The influence of *leaky modes*

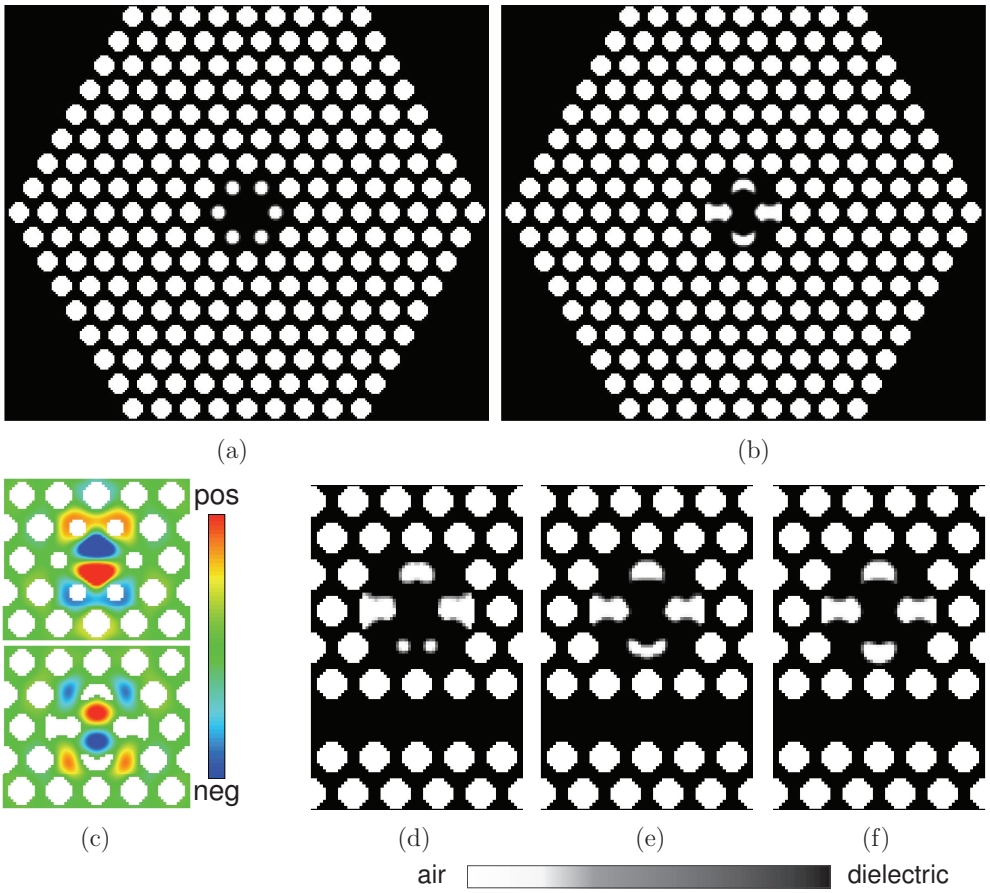


Figure 4.6: Dipole cavity - isolated cavity with **(a)** initial geometry and **(b)** optimized geometry, and **(c)** corresponding mode distribution. **(d)-(f)** coupled system designs for **(d)** $T_1 = 0$, **(e)** $T_1 = T_{\max}$, and **(f)** $T_1 = 4T_{\max}$.

above the light line in the dispersion diagram (see e.g. the band diagram in figure 4.10) needs to be taken into consideration. Unlike the monopole mode, we have identified a dipole mode at the frequency $\omega_0 a/2\pi c = 0.2480$ that is not leaky. Due to the change of mode it has been necessary to increase the radius of the evaluation domain for the objective function to $6a/14$.

The system is now excited by two point sources located in the eye of the dipole with opposite signs. In contrast to the optimization for the monopole mode, it has now been necessary to constrain the amount of dielectric material to $V^* = 0.8 \sum_e v_e$ to avoid ending up in local minima bearing degenerate modes. The optimized geometry reached after 592 design iterations is displayed in figure 4.6b. Compared to the mode in figure 4.6c for the initial geometry with $Q_0 = 9.9 \times 10^4$, the Q factor of the optimized cavity mode is slightly improved to $Q_0 = 11 \times 10^4$. Similar to the improved monopole mode, the topological features

of the optimized dipole design prevent the mode energy from escaping into the PhC above and below the cavity.

For the coupled system in figure 4.4 we have considered both modes despite the limited practical application of a notch filter supporting the monopole mode. Now, we increase the design freedom by imposing only a vertical symmetry restriction through the cavity center. In a distance sufficiently away from the coupling region we have excited the system from a line source with the same temporal variation as before and run the simulations for $T = 120,000\Delta t$. Since the coupled system accommodates multiple loss mechanisms, the localizing time T_1 in the objective cannot be chosen freely with respect to optimizing $Q_{\text{in}}(\approx Q_e)$. Here, we have chosen the three cases $T_1 = 0$, $T_1 = T_{\max}$, and $T_1 = 4T_{\max}$ to investigate the significance. We compute Q_{in} by utilizing that ω_0/Q equals the FWHM of resonant shape in the transmitted power $|T|^2$ for the coupled device.

The optimized designs are displayed in figure 4.5d-f and 4.6d-f, respectively. We first notice that all the designs comply with the minimum length scale introduced by the density filter. For the monopole mode the optimization has caused minor redistribution of the dielectric material to adjust the cavity mode to the coupling. According to table 4.1, none of the obtained Q_{in} for the optimized coupled system for the monopole mode have been improved compared to how the optimized isolated cavity performs in the coupling. However, the results in table 4.1 indicate, that for higher values of T_1 above the $T_{\max}$-limit will increase Q_{in} and thereby delay the energy decay. Conversely, for $T_1 < T_{\max}$ the energy will, once it has coupled into the cavity, decay rapidly back into the waveguide, yielding a strong interaction between the waveguide and the cavity. This behavior suggests that achieving high Q devices in principle counteracts strong coupling. Using the analytical expression for the transmission in equation (4.9) for the obtained Q factors leads to the result that all monopole notch filter designs exhibit practically zero transmission at resonance [P2]. It confirms what the CMT analysis predicts, that strong coupling is retained even after radical improvement of Q_{in}.

Table 4.1: Q[10^3] Factors for Coupled System Configurations.

	Monopole			Dipole		
	Q_{in}	Q_e	Q_0	Q_{in}	Q_e	Q_0
Initial cavity	16	17.3	220	2.1	2.1	99
Optimized cavity	60	68.3	474	4.1	4.2	110
$T_1 = 0$	2.1	2.1	139	2.0	2.0	90
$T_1 = T_{\max}$	7	7.3	168	3.3	3.4	100
$T_1 = 4T_{\max}$	16	17.2	239	6.0	6.2	140

Consequently through specification of T_1, the optimization formulation makes it possible to control the counteracting relation between strong coupling and high Q performance. One should also expect that optimization of the isolated cavity will result in the highest Q_0. According to table 4.1 this is also the case for the monopole mode. However, the dipole mode favors the largest Q_0 for the unsymmetrical cavity design, obtained for the coupled system when $T_1 = 4T_{\max}$. We believe this to be a result of the imposed two-planes symmetry conditions for the isolated cavity optimization, or alternatively a consequence of the enforced

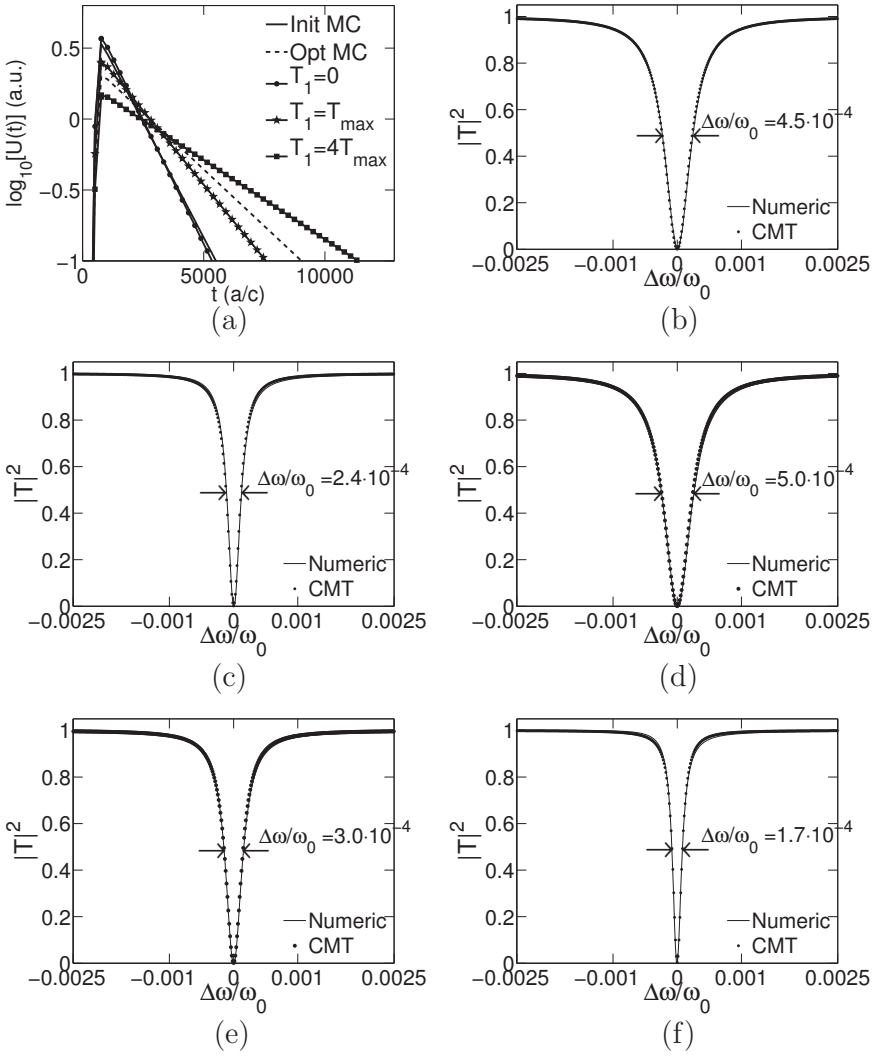


Figure 4.7: Dipole notch filter performance - **(a)** logarithmic envelope of the stored energy decay $U(t)$. Transmission spectrum for **(b)** initial design, **(c)** isolated cavity, coupled system **(d)** $T_1 = 0$, **(e)** $T_1 = T_{\max}$, and **(f)** $T_1 = 4T_{\max}$.

volume constraint that could lock the optimization inside a local minimum.

The transmission spectra in figure 4.7 show that the original and the optimized cavity geometries for the dipole mode in figure 4.6 exhibit almost zero transmission. Also, the transmitted energy in the vicinity of the resonance shows almost perfect agreement with the prediction from the CMT analysis in equation 4.8. The isolated cavity optimization only improves Q_{in} by a factor of two. In contrast to the monopole mode, optimizing for $T_1 = 4T_{\max}$ further increases Q_{in} by 50% as a result of minor design changes (compare figure 4.6b and -f), yielding a total improvement of 185%. The design result for $T_1 = 0$ in figure 4.6d shows substantial material redistribution in the interaction region in order to enhance coupling. However, the envelopes of the energy decay in figure 4.7a only display marginal changes among the various designs in the coupling performance. As expected the H_3-field plot in figure 4.8 of the best dipole candidate ($T_1 = 4T_{\max}$) shows nearly zero transmission at resonance.

We should finally mention, that the use of 2D PhC of infinite height omits the influence of the out-of-plane energy losses which constitute a limiting factor with regards to improving the total Q factor in PhC slabs. Thus, improving the in-plane Q does not necessarily imply the same trend for the out-of-plane Q (Kim et al., 2004a). Furthermore, we cannot take the third dimension into account by the effective index method, since it holds limited accuracy for high index-contrast structures or near cut-off. However, the experience from previous studies (Borel et al., 2004; Frandsen et al., 2004; Borel et al., 2005; Jensen et al., 2005), is that 2D optimized designs in general yield good behavior in 3D as well.

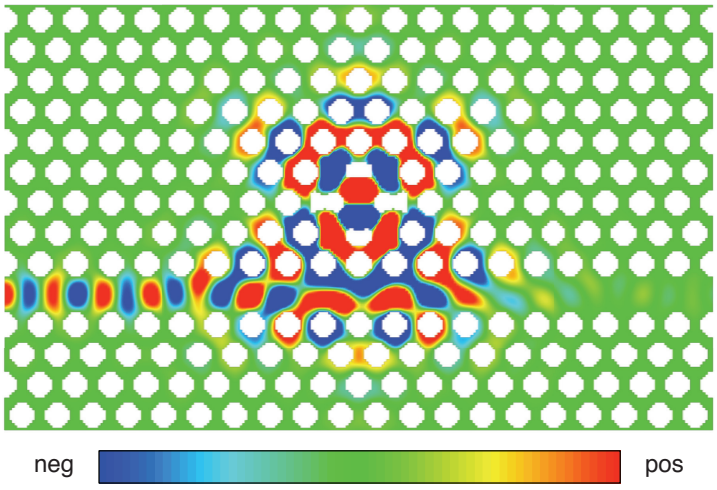


Figure 4.8: H_3 field pattern for dipole notch - the optimized candidate for $T_1 = 4T_{\max}$, and the material distribution is shown with 0.6 threshold. (Almost) Zero transmission is seen at the waveguide outlet.

4.4 Topology optimization of slow light devices [P3], [P5]

It is well-known that the speed of light is an upper bound for waves conveying matter, energy or information (Diener, 1996). However, after the discovery of the slow-light phenomenon as a result of material dispersion, it seems that no lower limit for the group velocity at which light can travel exists (Hau et al., 1999; Kash et al., 1999). Since then, this promising technology has attracted much attention, as the application prospects are numerous, e.g. in future optical networks and information processing systems. To mention a few remarkable properties, slow-light offers the opportunity for accurate time-domain processing of optical signals, low power consumption in optical switching devices (Baba, 2008), optical buffering (Tucker et al., 2005), and enhancement of weak linear as well as nonlinear light-matter interaction processes (Eggleton et al., 1999; Soljacic et al., 2002; Soljacic and Joannopoulos, 2004; Krauss, 2007).

In the present work we realize the slow light device by exploiting the unique

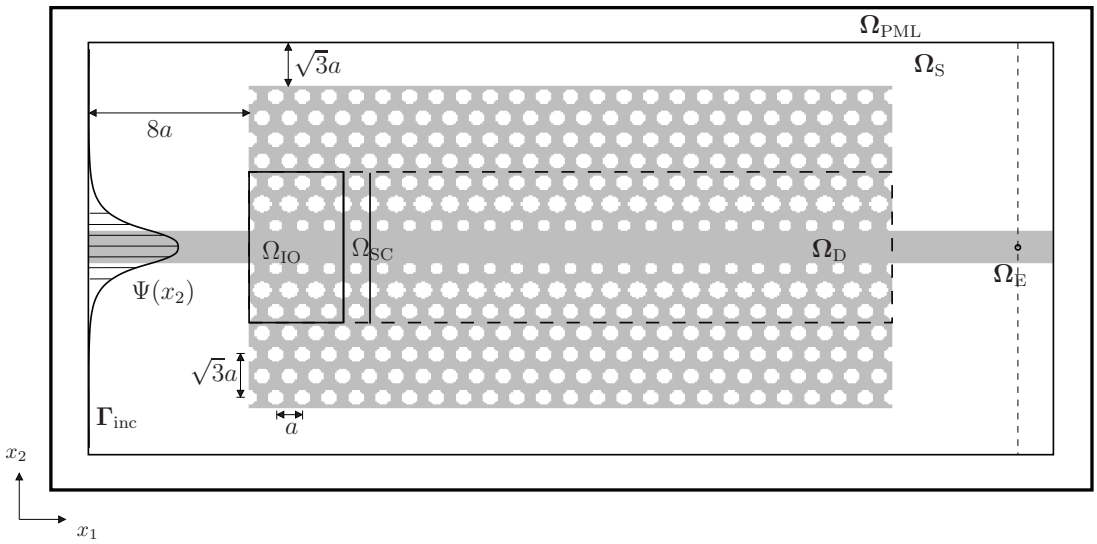


Figure 4.9: Slow light optimization setup - Ridge wave guide optimization setup. The computational domain contains a solution region Ω_{S}, PML region Ω_{PML}, and design region Ω_{D} encapsulated by the dashed boundary. The 'active' design set consists of x_2 mirrored in-/outlet region Ω_{IO} and a supercell Ω_{SC} that are both mapped to Ω_{D}. Whenever an optimized design is presented we only show the material distribution in Ω_{IO} and Ω_{SC}. The waveguide mode is excited on Γ_{inc} with an analytically given amplitude profile $\Psi(x_2)$. The objective is measured in point(s) Ω_{E}.

tunable dispersion properties of the W1 PhC waveguide (PhC-WG) in figure 4.9, where a single row of holes has been removed. As illustrated, the waveguide has in- and outlet connections to straight ridge waveguides. Equivalent to the notch filter from the previous section, the surrounding crystal is formed by perforated air holes in the dielectric material GaAs, but now with the hole radius altered to $r/a = 0.30$. This yields a band gap in the normalized frequency $a/\lambda = 0.21-0.30$, see figure 4.1. We will assume that the pitch of the lattice is 370nm.

The goal of the presented method is to improve the performance of slow light devices by the use of topology optimization. We start from a W1 waveguide and formulate an optimization problem that delays the pulse as the objective while its shape is conserved through a relaxed signal shaping constraint. Essentially, this leaves us with an optimization problem in which the in-/outlet region is designed to remove the possible mismatch between the ridge and the PhC waveguide mode, while the interior material distribution is changed to delay the probe pulse as much as possible. It is shown that allowing for distortion of the pulse shape yields designs with an enhanced slow down effect. We subsequently employ a pulse shaping formulation to minimize the dispersion of the resulting probe signal for the maximized time delay.

4.4.1 Slow light structures characteristics

There are different ways to determine if a (PhC) structure for a given geometry or material distribution exhibits slow light behavior. A common method is to study the dispersion relation $\omega(k)$ from which the group velocity can be computed by

$$v_g \equiv \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{c}{n_g} \tag{4.14}$$

where k is the wave number. In the rightmost expression the dimensionless *group index* n_g is introduced as an alternative measure for the slow light behavior, and we can obtain that by numerical differentiation, once $\omega(k)$ is known.

We have performed an eigenvalue analysis to obtain the dispersion characteristics for the supercell structure in figure 4.10a. In conjunction with the dispersion diagram in the left part of figure 4.10b, the group index versus normalized frequency plot in figure 4.10c reveals that inside the band gap $a/\lambda = 0.21-0.30$, the waveguide mode exhibits (semi)-slow light behavior with very little group velocity dispersion at group velocity $v_g \approx c/(40\pm10\%)$. The useful bandwidth of the structure (i.e. the range over which the group index remains constant within $\pm10\%$) is $B_{\mathrm{FWHM}} \approx 12$nm centered around the normalized frequency $a/\lambda = 0.2163$. Consequently, the given supercell structure serves as a suitable starting guess the for optimization process. The geometry has been found via a parameter search in

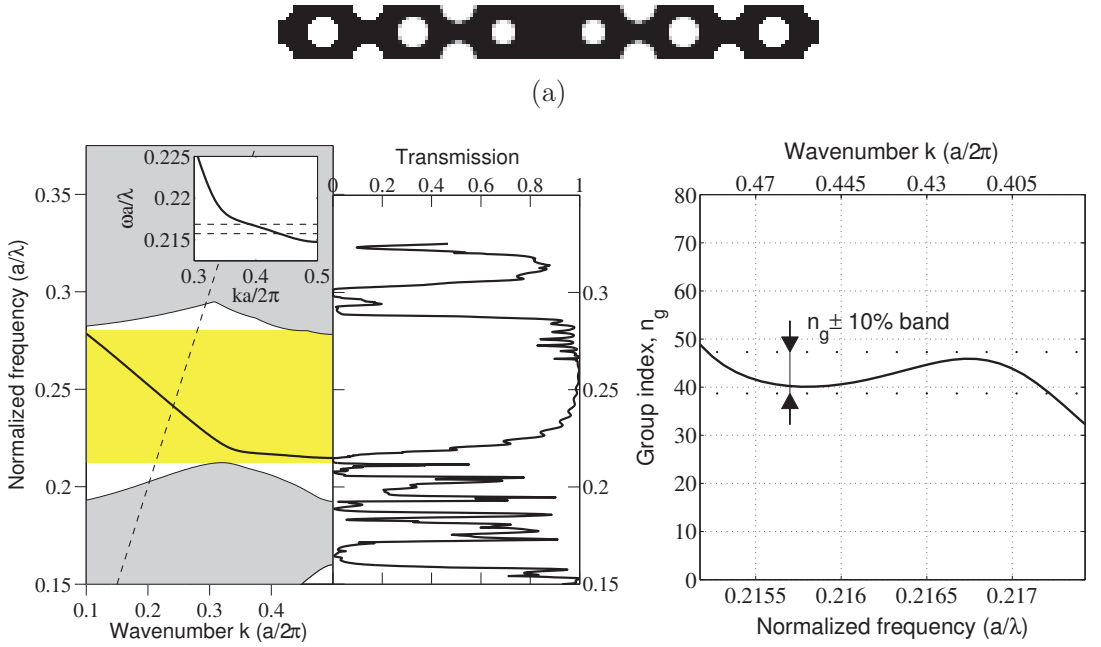


Figure 4.10: Slow light start design - **(a)** The rotated supercell used in the eigenvalue computation to obtain the **(b)** band diagram. This shows the normalized frequencies versus normalized wave vectors for an even (solid) and odd (omitted here) PhC-WG mode in the band gap, and the dashed line indicates the light line above which leaky modes live. The inset of the graph magnifies the dispersion properties of the guided mode where the dotted lines confine a region with low dispersion. **(c)** Group index n_g versus normalized frequency.

which the radius of the three nearest neighboring holes to the waveguide has been varied (Frandsen et al., 2006; Schulz et al., 2010).

For time domain simulations it is not that straightforward to retrieve the dispersion characteristics as they appear in the band diagram (Oskooi et al., 2010). Another reliable way to detect slow light behavior is to analyze the transmission spectrum, which we construct by exciting the inlet ridge waveguide with a short pulse. We obtain the transmission spectrum by integrating the frequency dependent Poynting flux over a vertical line located after the PhC-WG. Since the PhC-WG sample has a PhC mirror mechanism: the PhC-WG and the non-optimized in- and outlet, it functions as a so-called *Fabry-Pérot* (F-P) cavity. This leads to F-P fringes in the transmitted energy, and the group velocity can be determined by (Letartre et al., 2001).

$$v_g = c\frac{2L_c\Delta\lambda}{\lambda^2} \tag{4.15}$$

where $\Delta\lambda$ is the free spectral range between adjacent F-P peaks, and L_c is the length of the (closed) cavity. The transmission spectrum in the right part of figure 4.10b reveals F-P oscillations in the normalized frequency $a/\lambda = 0.21 - 0.25$, thus insinuating slow light behavior. Furthermore, the spectrum shows an unambiguous correlation between the eigenvalue analysis and our FETD data.

Where the above techniques require frequency information, the *time-of-flight* (ToF) method, see e.g. Inoue et al. (2002); Jacobsen et al. (2005), relies on time observations of the pulse delay η as it traverses through the PhC-WG to compute the group velocity by

$$v_g = \frac{\eta}{L} \tag{4.16}$$

where L is the length of the PhC-WG. For the device in figure 4.9 the time delay η can be determined by recording the probe pulse at a given point in the exit ridge waveguide with and without PhC-WG. For a precise measure we extract the envelope of the probe pulses and define the delay as the temporal distance between the envelope peaks. It is important to be sufficiently downstream in the ridge waveguide to avoid potential near-field disturbances from surface modes living along the in-/outlet interface.

The definition in equation 4.16 gives us a quantitative measure, namely the delay η, which can be adopted in the formulation of a time domain optimization problem.

4.4.2 Formulating a pulse shaping and delaying problem

The optimization problem with the setup given in figure 4.9 is based on pulse shaping of the probe signal envelope $r(t)$ registered at Ω_{E} in the ridge waveguide. The *Hilbert* transform is used to retrieve the envelope

$$r(t) = \sqrt{u(t)^2 + \hat{u}(t)^2} \tag{4.17}$$

where $u(t)$ is the signal, and $\hat{u}(t)$ is its corresponding Hilbert transform. The aim of the optimization is to obtain signals at multiple registration points with an envelope following a prescribed envelope, obtained from the same input signals propagating in a straight ridge waveguide. The pulse conforming measure is defined as [P3]

$$f = \frac{\displaystyle\int_0^T \sum_{i\in\Omega_{\mathrm{E}}} [r_i(t) - \alpha p_i(t-\eta)]^2 \mathrm{d}t}{\displaystyle\int_0^T \sum_{i\in\Omega_{\mathrm{E}}} \alpha^2 p_i(t)^2 \mathrm{d}t} \tag{4.18}$$

where $p_i(t)$ is the prescribed envelope function delayed η and scaled by α, and T is the termination time of the transient simulation. To render a physically

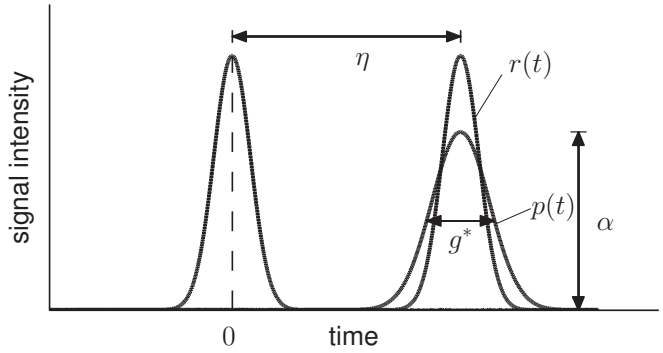


Figure 4.11: Pulse delay strategy - for pulses with a group delay 0 and η. The parameter η is introduced as a design variable. The g^*-parameter controls the temporal spreading of the output pulse envelope. The α-parameter mainly specifies the transmitted energy, i.e. the amplitude decrease.

interpretable measure, we have normalized with respect to the prescribed pulse envelope area(s).

To formulate a pulse delaying strategy, we introduce $\eta' = \eta/s$ as a design variable with s denoting a scaling factor. The shape of the pulse is controlled by treating the design response measure in equation (4.18) as a constraint. The objective is to maximize the delay, which we formulate as

$$\begin{aligned} \min_{\boldsymbol{\rho}\in\mathbf{R}^M,\eta'\in\mathbf{R}} \quad & -\eta' \\ \text{s.t.:} \quad & \text{Governing equation (2.8)} \\ & g_1(\boldsymbol{\rho},\eta') = f(\boldsymbol{\rho},\eta')/g^* - 1 < 0 \\ & 0 \le \rho_e \le 1, \qquad e \in \Omega_{\mathrm{D}} \\ & 0 \le \eta' \le 1 \end{aligned} \tag{4.19}$$

where $0 < g^* \le 1$ has been introduced to allow for broadening and distortion of the delayed pulse in figure 4.11 after traversing the PhC-WG. The maximum allowable distortion is obtained for $g^* = 1$ (i.e. the initial pulse shape has been completely destroyed) and is gradually diminished when $g^* \to 0$ (i.e. the output pulse shape is perfect). The gradients of g_1 with respect to the structurally related design variables are found through the adjoint sensitivity analysis method presented in chapter 3. The sensitivity of g_1 with respect to the time related design variable η' is given by

$$\frac{\partial g_1}{\partial \eta'} = \frac{\displaystyle\int_0^T \sum_{i\in\Omega_{\mathrm{E}}} -2\left[r_i(t) - \alpha p_i(t - s\eta')\right] \alpha \frac{\partial p_i(t-s\eta')}{\partial \eta'} \mathrm{d}t}{\displaystyle g^* \int_0^T \sum_{i\in\Omega_{\mathrm{E}}} \alpha^2 p_i(t)^2 \mathrm{d}t} \tag{4.20}$$

where the change of variable $\Pi = t - s\eta'$ yields $\partial p_i/\partial \eta' = -s\partial p_i/\partial \Pi$ which we approximate by finite differences. It is important to stress that the delay variable $\eta = s\eta'$ is several orders of magnitudes larger compared to the structurally related design variables that vary between 0 and 1. Hence, the sensitivity expression in equation (4.20) has to be scaled adequately through s in order to gain a well-posed optimization problem. A strategy for choosing the scaling parameter is $0 \leq \eta/s \leq 3$ yielding $\eta'_{\max} = 3$.

Once a material distribution is obtained for a maximum delay and a given allowable distortion g^*, we will start from the resulting design and use the pulse shaping response function in equation (4.18) as the objective, now with a fixed η, to minimize the pulse distortion. The optimization problem is now formulated as

$$\begin{aligned} \min_{\boldsymbol{\rho}\in\mathbf{R}^M} \quad & f(\boldsymbol{\rho}, \eta') \\ \text{s.t.:} \quad & \text{Governing equation (2.8)} \\ & 0 \leq \rho_e \leq 1, \qquad e \in \Omega_{\mathrm{D}} \end{aligned} \tag{4.21}$$

Hence, the optimization is completed in two consecutive steps: ① a *pulse delay* step, followed by ② a *pulse restoring* step. The pulse shaping constraint in the delay formulation is applied in a single point in the ridge waveguide center sufficiently downstream. To achieve perfect transmission for the output signal in the subsequent restoring step, we apply the pulse shaping objective in multiple points along a vertical wall, such that the ridge waveguide mode profile will be inherited automatically in the registered output signal.

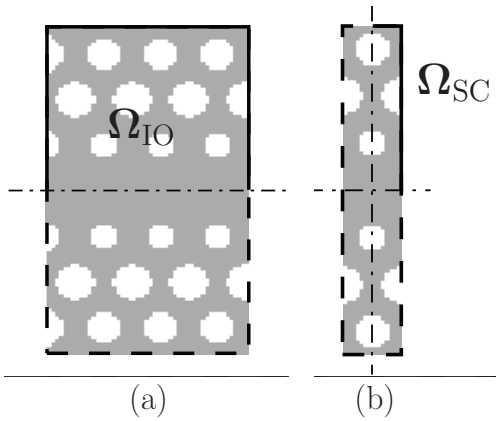


Figure 4.12: Symmetry conditions for the 'active' design set - x_1 symmetry is imposed for Ω_{IO}, and x_1-x_2 symmetry is imposed for Ω_{SC}. This reduces the number of active design variables further.

4.4.3 Method and results

Similar to the optimization of the notch filter, the purpose is again to distribute air and dielectric material in the design domain Ω_{D}, such that the probe pulse at Ω_{E} is delayed as much as possible. As illustrated in figure 4.9, the design domain includes, in addition to the waveguide, the three nearest neighboring rows of holes on both sides of the waveguide. To retain a certain periodicity and symmetry of the final design, we only allow for design freedom in the x_1 mirrored part of in-/outlet region Ω_{IO} as well as in the x_1-x_2 mirrored part of the supercell Ω_{SC} distributed periodically downstream over the lattice distance $17a$ (see the symmetry details in figure 4.12). Even though the actual design update is carried out only in a small part of the design domain, we still need to compute the sensitivities in the entire design domain, and subsequently map and sum those properly for the 'active' design variables. In figure 4.9 the active design set $\Omega_{\mathrm{IO}} \cup \Omega_{\mathrm{SC}}$ is highlighted in relation to the entire design domain Ω_{D}. It should be combined with figure 4.12 to understand the symmetry conditions.

The geometrical restrictions will reduce the optimization design space whereby potentially well-performing designs are disregarded because nothing dictates that a symmetric and periodic structure should be the best performing. However, by limiting the optimization to the in-/outlet region and the supercell, we can directly determine the dispersion characteristics of the resulting structure by a simple eigenvalue analysis of the supercell. The periodicity also enhances manufacturability, that is further boosted by standard density filtering of the design variables with filter radius $R = 2.5\Delta x$ combined with material phase projection (Sigmund, 2007). Even though the applied projection function does not preserve the minimum structural length scale for some settings (Wang et al., 2011), it results in well-defined designs, see e.g. figure 4.13a, -b, -c and 4.15b. It should be emphasized, that whenever an optimized design is presented we only illustrate the in-/outlet region together with the supercell structure.

The structure is excited by a ridge waveguide mode along the entire vertical interface separating the solution domain and the PML region. The mode profile is analytically known and the temporal variation of the source follows the Gaussian modulated distribution in equation (4.13) with $\omega_0 a/2\pi c = 0.2163$ ($\sim$ 1710nm wavelength) as dominant frequency. We have chosen to optimize for a 6nm FWHM bandwidth which requires an $1/e$-intensity $T_0 = 700a/c$ ($\sim$ 1ps) of the Gaussian envelope centered around $t_0 = 2000a/c$. This choice is less than the useful 12nm bandwidth of the initial design.[1] For the spatial discretization, the computational domain is composed of square elements with edge size $\Delta x_1 = \Delta x_2 = a/14$. We run the simulations for $T = 8000a/c$ ($\sim$ 10ps) with a

[1] Conversion from FWHM bandwidth to Gaussian $1/e$ width is given by $T_0 = \sqrt{4\ln 2}/(\pi B_{\mathrm{FWHM}})$

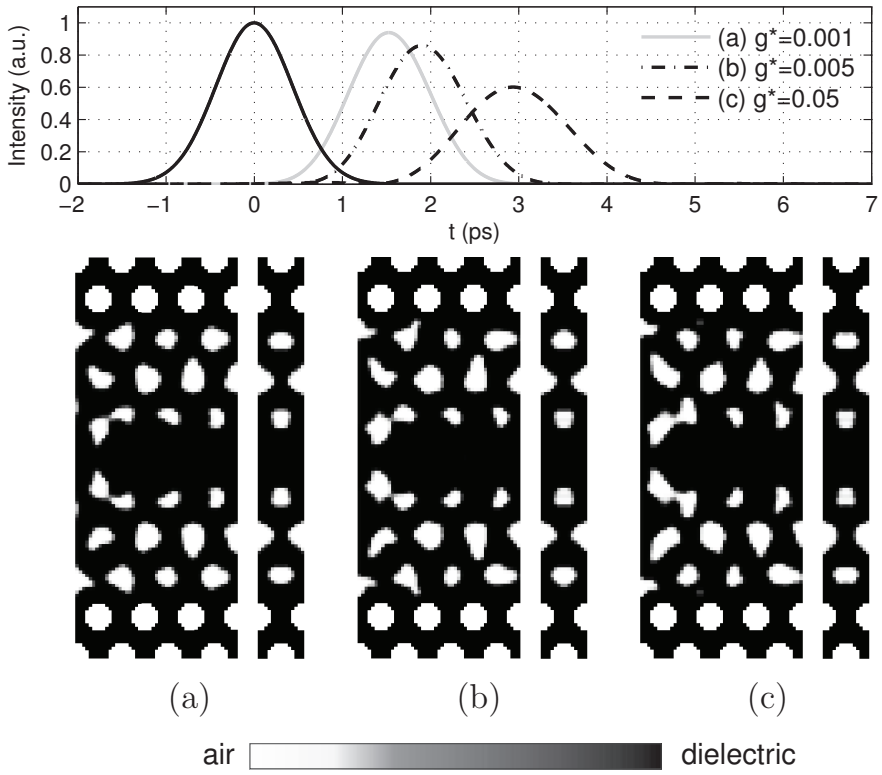


Figure 4.13: Pulse delayed slow light designs - TOP: Transient intensity response of the ridge waveguide pulse peaking at $t = 0$ps, and time delayed pulse for $g^* = 0.001$, 0.005, and 0.05 peaking at $t = 1.5$ps, 2.0ps, and 3.0ps, respectively. In/-outlet and supercell design for **(a)** $g^* = 0.001$, **(b)** $g^* = 0.005$, and **(c)** $g^* = 0.05$.

time step $\Delta t = 0.90\Delta x_1/c$.

The first part of the optimization has been dedicated to investigate the design as well as the time delay response to three different values for the relaxation parameter $g^* = 0.001$, 0.005, and 0.05 of the prescribed pulse envelope area in the pulse shaping constraint. The designs and time delays appear in figure 4.13. For the lowest relaxation factor $g^* = 0.001$ the pulse is delayed 1.5ps which doubles to 3ps for $g^* = 0.05$. However, the price for the largest delay is a considerable broadening of the pulse and amplitude distortion which corresponds to an unfavorable reduction of the useful bandwidth of the device. In practice heavy distortion implies that some of the information carried by the pulse has been lost, and for signal processing devices it is important to preserve the pulse shape to a certain degree. Further, the peak intensity reduction of a pulse makes it less effective for driving nonlinear effects.

Based on an eigenvalue analysis of the optimized supercell structures the group index versus normalized frequency plot in figure 4.14 confirms that the maximized

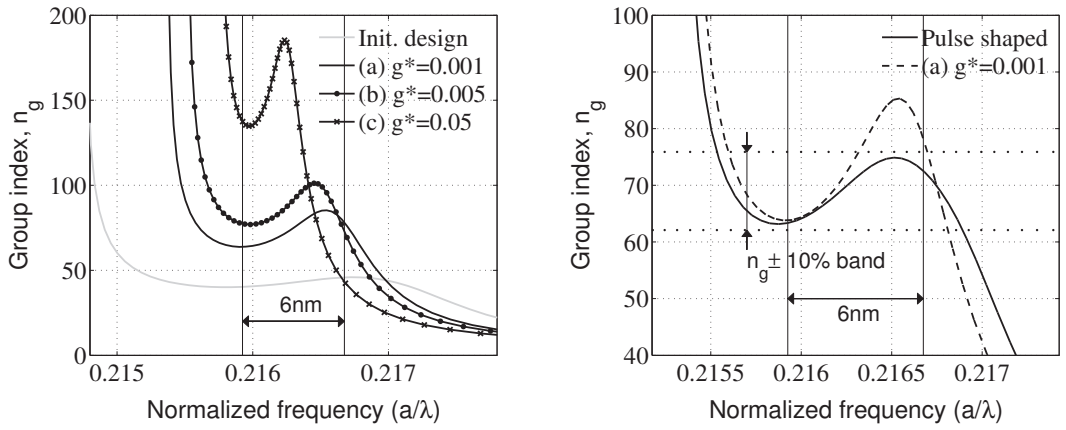


Figure 4.14: Group index versus normalized frequency - (a) the optimized design obtained from the delay formulation in equation (4.19) for three different values for the pulse relaxation: $g^* = 0.001$, 0.005, and 0.05. **(b)** A comparison between the pulse delayed design ($g^* = 0.001$) and the pulse delayed-shaped design.

delays give rise to increased group indices. However, the bandwidth in which the optimized devices can be considered useful for slow light purposes is simultaneously decreased. The enhanced time delays induce very strong wavelength dependence of the group index. In particular, we see that the 'flat-band' region for the initial guess evolves into an s-shaped kink whose center is moving towards northwest, causing bandwidth shrinkage. For the largest relaxation value the average group index is $n_g \approx 160$, for which, however, the 10% criterion is not satisfied. This points to the fact, that a bandwidth extension is only obtained at the cost of the group index. Frankly, this detrimental effect is inherently associated to slow light systems, leaving the designer of such systems with a difficult task.

All the optimized designs have reached convergence[1] within 300 optimization iterations. In addition to changing the supercell layout, the geometry of the in-/outlet region has been altered to lower the interface penalty losses, when coupling the access waveguide mode into the PhC-WG. For real applications it is important to terminate the PhC properly, since coupling losses, together with backscattering losses, as a result of fabrication disorders, will degrade the functionality of the slow light device. Intriguingly, a comparison of the optimized supercell structures in figure 4.13a, -b, -c and the initial supercell layout in figure 4.10a reveals that the optimization has reduced the size of the air inclusions. We consider this a logical action in order to reduce backscattering losses as they originate from the

[1]For the convergence criteria we use that the relative change of the objective function should be less 10^{-4}.

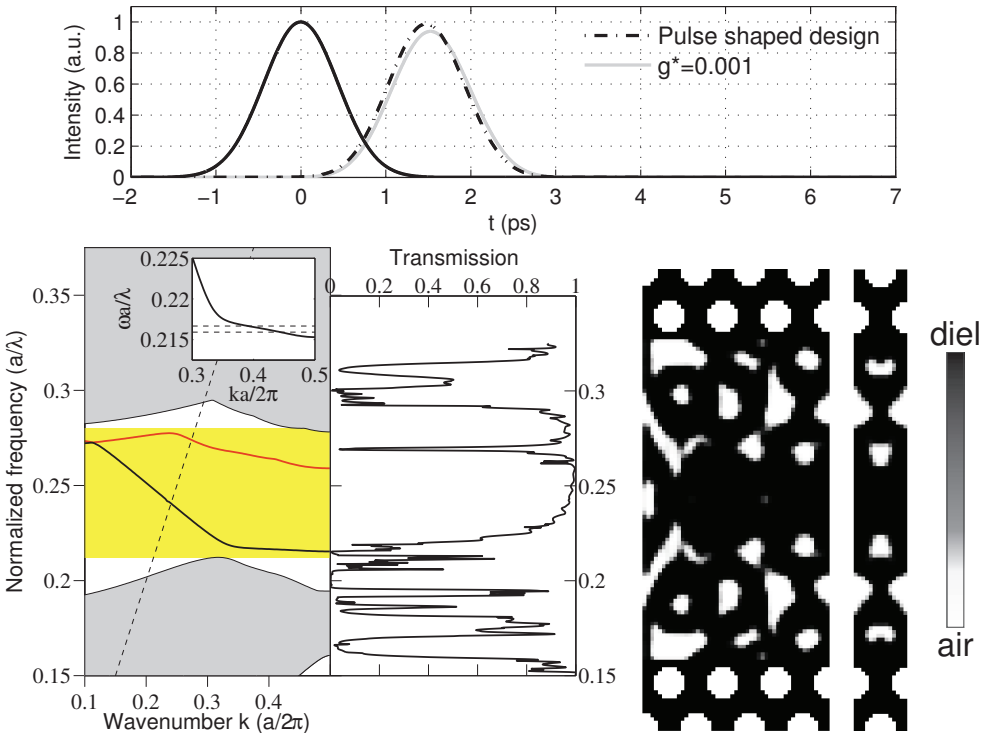


Figure 4.15: Pulse delayed-shaped slow light design - TOP: Transient intensity response of the initial envelope peaking at $t = 0$ps, and the optimized envelope peaking at $t = 1.5$ps with and without pulse shaping. **(a)** Band diagram for the optimized supercell structure and the transmission spectrum for the structure of finite length without optimized in-/outlets. **(b)** Optimized in-/outlet and supercell structure.

overlap of the optical modes and the hole surfaces (Schulz et al., 2010).

In figure 4.14b it is seen that even for the lowest relaxation factor $g^* = 0.001$ we obtain a supercell design with a group index variation around $n_g \approx 70$ that is not confined to the allowable $\pm 10\%$ band for the specified 6nm bandwidth. To tweak this unsatisfied behavior we have executed the *pulse restoring* step with the design in figure 4.13a as initial design and the maximized delay as the fixed time delay in the pulse shaping objective. Whereas the delay formulation typically reaches convergence within 300 optimization iterations (it naturally depends on the relaxation), the pulse shape formulation has required 1710 iterations for converging to the design in figure 4.15b. We believe the reason for this rather high number of iterations lies in the fact that the optimization has two concerns; while modification of the supercell structure will result in a change of the waveguide

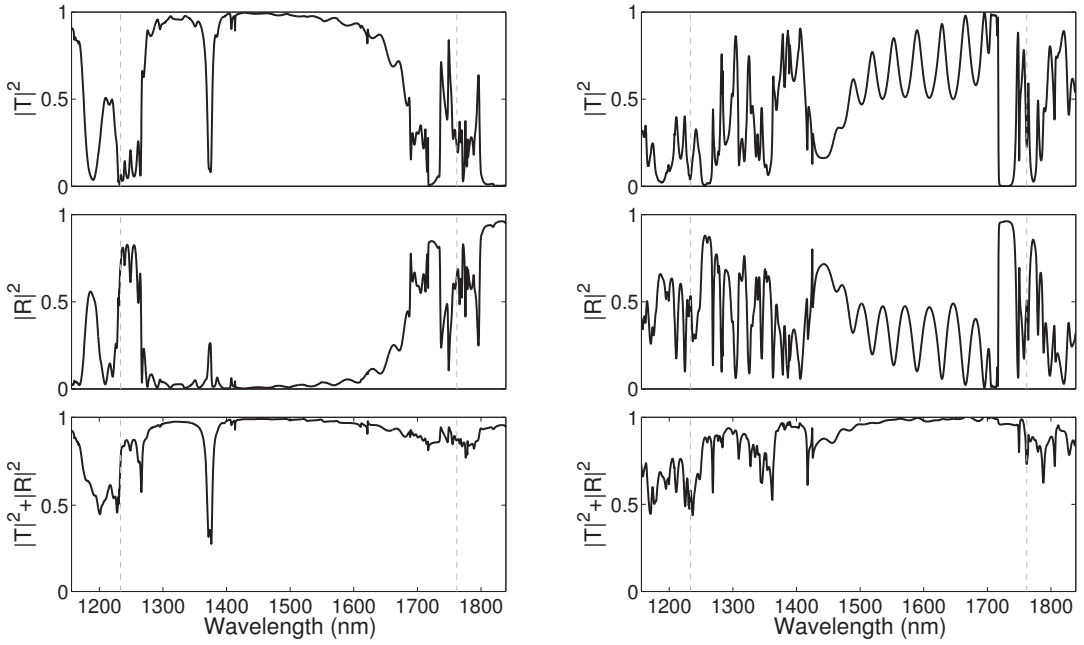


Figure 4.16: Transmission ($|T|^2$), reflection ($|R|^2$) and energy balance ($|T|^2+|R|^2$) spectrum for the optimized slow light device - (a) consisting of the supercell structure in figure 4.15b without optimized in-/outlets. **(b)** The structure with optimized in-/outlets in figure 4.18a.

mode, the in-/outlet needs to be simultaneously adjusted to eliminate coupling losses. To meet and balance both regards, the optimization will then progress in small increments.

As it appears in figure 4.14b, the optimized design yields a group index variation that satisfies the $\pm 10\%$ criterion. It is obvious to pose the question why the pulse shaping optimization strategy does not result in a complete flat-band region. Possible explanations could be that the optimization has ended up in a local minimum or that a flat-band region is not physically obtainable. We should also remember that as designers we are competing against the intrinsic dispersion property of the PhC-WG, that the group index can only be increased at the cost of the operational bandwidth. It is also noted in figure 4.14b that the group index versus normalized frequency curve is shifted slightly to lower frequencies compared to the start design. Based on a comparison of the supercell designs in figure 4.13a and 4.15b showing that the amount of dielectric material has increased moderately, the horizontal shift of group index curve is expected, since the frequency of the guided mode scales by $1/\sqrt{\varepsilon}$ in a medium of dielectric

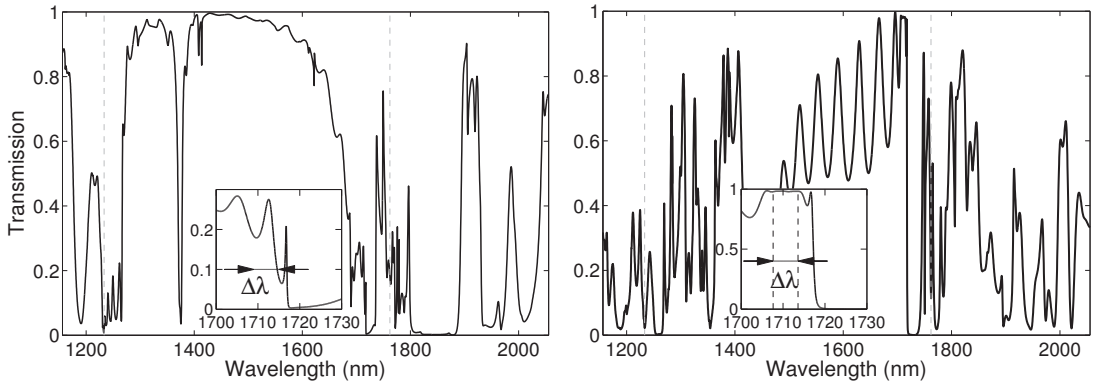


Figure 4.17: Transmission spectrum for the optimized slow light device - (a) consisting of the supercell structure in figure 4.15b without optimized in-/outlets. The inset shows Fabry-Pérot oscillations near the upper band gap edge ($\lambda \approx 1716$nm) separated by $\Delta\lambda = 5$nm yielding $v_g \approx c/65$. **(b)** The structure with optimized in-/outlets in figure 4.18a.

constant ε (Wang et al., 2011).

The transmission, reflection[1] and energy balance spectrum for the finite waveguide structure assembled from the supercell design in figure 4.15b without the optimized in-/outlet regions are presented in figure 4.16a. They show clearly defined Fabry-Pérot (F-P) fringes with decreased spacing towards the slow light wavelength region. In the slow light region region near the band edge ($\lambda = 1716$nm) the wavelength spacing is measured to $\Delta\lambda = 5$nm that for an effective F-P cavity length $L = 13a$ yields $v_g \approx c/65$, see figure 4.17a. The transmission spectrum also contains a dip at $\lambda = 1375$nm with almost zero transmittance. To explain this we need to consider the band diagram for the optimized supercell in figure 4.15a, showing that a guided mode (the upper curve in the rectangular band gap region) has been pulled down from the air band. This mode has an odd profile and, apart from living in the waveguide, it can also live in the upper/lower part of the design domain between the waveguide and the surrounding PhC. A comparison of transmittance against the band diagram in figure 4.15a shows that anti-crossing of the even (black curve) and the odd (red curve) modes at $\lambda \approx 1375$ gives rise to a mode gap inside the band gap, which explains the dip in transmission (Notomi et al., 2001).

Finally, figure 4.16b and 4.17b show that almost perfect transmission, i.e. $|T|^2 \approx 0.98$, is achieved in the wavelength range specified by the source, if we

[1]The reflected energy is recorded at a vertical line in front of the PhC-WG by using equation (2.25) from which the energy balance can be computed.

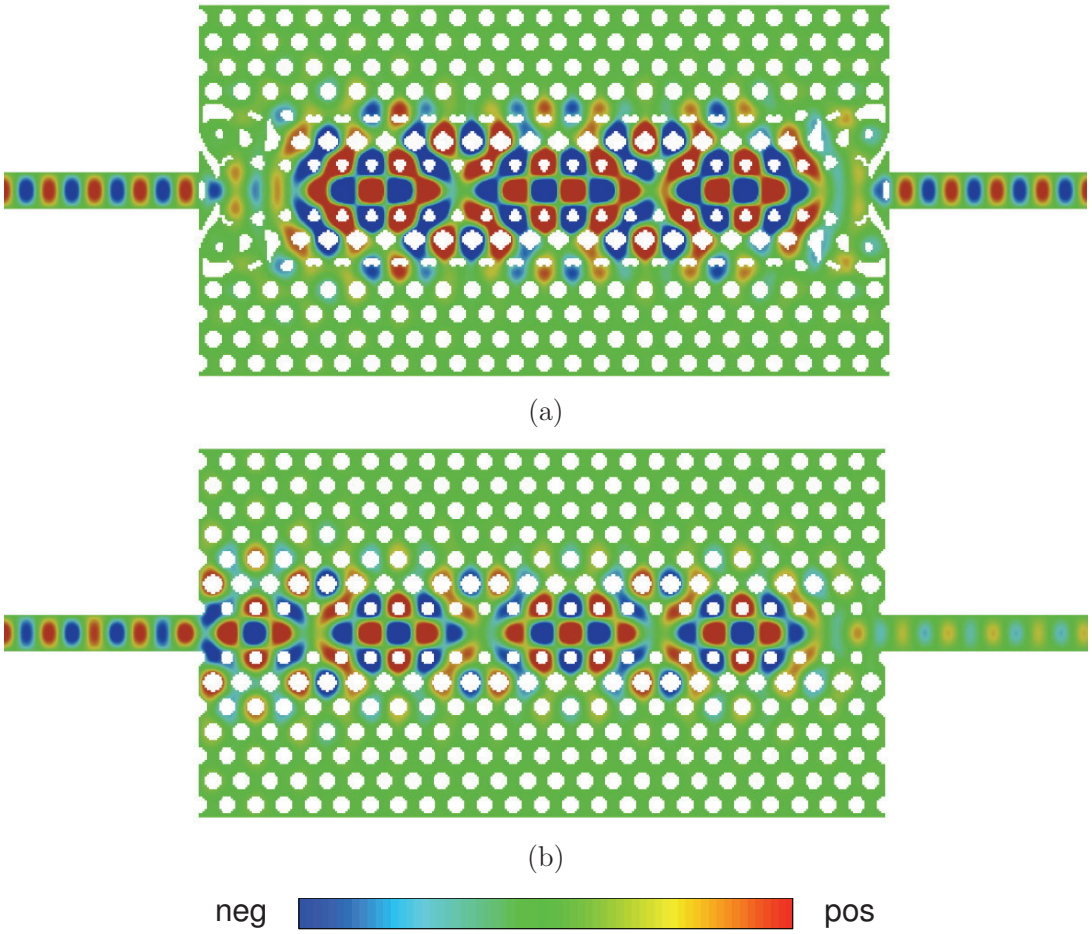


Figure 4.18: H_3 **field pattern for slow light device** - **(a)** The optimized structure first by delaying and subsequently shaping the pulse. **(b)** The initial geometry. The material distribution is shown with 0.6 threshold

include the optimized in-/outlet design. The associated H_3 field pattern in figure 4.18a for the dominant source wavelength also confirms perfect transmission compared to the poor performing start design with a field pattern given in figure 4.18b.

Chapter 5

Concluding remarks

5.1 Conclusions

Since the introduction of topology optimization to mechanical problems in 1989, the popularity of the method has almost exploded. It has become a subject of intensive research, not only for mechanical problems, but also for a wide range problems from other (multi)-physical branches such as fluid dynamics, acoustics, electromagnetism, MEMS and many others.

Around the beginning of the century, a firm basis for applying topology optimization to wave propagation problems was established. The list of problems that have been attacked until today is endless and it covers all possible types of waves, i.e. electromagnetic, acoustic, and elastic waves. However, most of the contributions consider time-harmonic waves, and there are only few that optimize for transient waves.

The purpose of the present thesis has been to establish a general framework for applying topology optimization to 2D transient wave propagation problems. To treat waves with arbitrary directions and propagation characteristics we have developed a *finite element time domain* formulation based on the efficient *perfectly matched layer* technique [P2], [P4]. A generic optimization problem is formulated which we believe will cover a wide range of transient wave propagation problems. We have discussed how to apply parallel computing to lower the computational costs associated with gradient-based time domain optimization.

The developed framework has been applied to three problems involving the design of optical material structures. The first problem optimizes an optical taper coupling two monomodal waveguides by maximizing the energy in the output waveguide [P1]. The second problem optimizes a notch filter with respect to narrowing the filter stop band. We have formulated that as a maximization problem of the stored energy in the decaying regime [P2]. Finally, we have improved the slow light performance of a *photonic crystal waveguide* by using a delaying strat-

egy for the objective, while a pulse shaping constraint serves to control the pulse distortion [P3], [P5].

5.2 Future work

Future research should be directed towards:

- Incorporating elasto-optical wave interaction in order to benefit from the resulting change in refractive index.
- Solving advanced nonlinear optical problems, e.g. by including the instantaneous Kerr-nonlinearity.
- Implementing a robust formulation that includes material uncertainties.
- Solving large-scale 3D problems in order to e.g. include out-of-plane losses in the numerical analysis.

References

Aage, N. and Lazarov, B. (2011). Parallel multiphysics topology optimization using the method of moving asymptotes. *Submitted*. 32

Aage, N., Mortensen, N.A. and Sigmund, O. (2010). Topology optimization of metallic devices for microwave applications. *International Journal For Numerical Methods In Engineering*, 83(2):228–248. 23

Achenbach, J. (1975). *Wave Propagation in Elastic Solids*. Amsterdam, Holanda : North-Holland. 6

Amdahl, G.M. (1967). Validity of the single processor approach to achieving large scale computing capabilities. In *Proceedings of the April 18-20, 1967, spring joint computer conference*, AFIPS '67 (Spring), pages 483–485. ACM, New York, NY, USA. 33

Andkjær, J., Nishiwaki, S., Nomura, T. and Sigmund, O. (2010). Topology optimization of grating couplers for the efficient excitation of surface plasmons. *Journal of the Optical Society of America B-optical Physics*, 27(9):1828–1832. 23

Andreasen, C. and Sigmund, O. (2010). Saturated poroelastic actuators generated by topology optimization. *Structural and Multidisciplinary Optimization*, 2:1–14. 22

Andreasen, C.S., Gersborg, A.R. and Sigmund, O. (2009). Topology optimization of microfluidic mixers. *International Journal for Numerical Methods in Fluids*, 61(5):498–513. 22

Arora, J.S. and Haug, E.J. (1979). Methods of design sensitivity analysis in structural optimization. *Aiaa Journal*, 17(9):970–974. 27

Baba, T. (2008). Slow light in photonic crystals. *Nature Photonics*, 2(8):465–473. 36, 50

Bendsøe, M.P. (1989). Optimal shape design as a material distribution problem. *Structural and Multidisciplinary Optimization*, 1(4):193–202. 21

Bendsøe, M.P. and Kikuchi, N. (1988). Generating optimal topologies in structural design using a homogenization method. *Computer Methods In Applied Mechanics and Engineering*, 71(2):197–224. 21

Bendsøe, M. and Sigmund, O. (2004). *Topology Optimization; Theory, Methods, and Applications*. Springer Verlag Berlin Heiddelberg New York, 2nd edition edition. 1, 22

Berenger, J.P. (1994). A perfectly matched layer for the absorption of electromagnetic-waves. *Journal of Computational Physics*, 114(2):185–200. 11

Berg, P., If, F., Nielsen, P. and Skovgaard, O. (1994). Analytical reference solutions. In K. Helbig, editor, *Modeling the Earth for Oil Exploration*, pages 421–427. Pergamon, Brussels, Belgium. 15

Borel, P.I., Frandsen, L.H., Harpoth, A., Kristensen, M., Jensen, J.S. and Sigmund, O. (2005). Topology optimised broadband photonic crystal y-splitter. *Electronics Letters*, 41(2):69–71. 23, 49

Borel, P.I., Harpoth, A., Frandsen, L.H., Kristensen, M., Shi, P., Jensen, J.S. and Sigmund, O. (2004). Topology optimization and fabrication of photonic crystal structures. *Optics Express*, 12(9):1996–2001. 23, 49

Borrvall, T. and Petersson, J. (2001). Large-scale topology optimization in 3d using parallel computing. *Computer Methods In Applied Mechanics and Engineering*, 190(46-47):6201–6229. 32

Borrvall, T. and Petersson, J. (2003). Topology optimization of fluids in stokes flow. *International Journal For Numerical Methods In Fluids*, 41(1):77–107. 22

Bourdin, B. (2001). Filters in topology optimization. *International Journal for Numerical Methods in Engineering*, 50(9):2143–2158. 26

Bruns, T.E. and Tortorelli, D.A. (2001). Topology optimization of non-linear elastic structures and compliant mechanisms. *Computer Methods In Applied Mechanics and Engineering*, 190(26-27):3443–3459. 22, 26

Cardoso, J.B. and Arora, J.S. (1992). Design sensitivity analysis of nonlinear dynamic-response of structural and mechanical systems. *Structural Optimization*, 4(1):37–46. 27

Chew, W.C. and Weedon, W.H. (1994). A 3d perfectly matched medium from modified maxwells equations with stretched coordinates. *Microwave and Optical Technology Letters*, 7(13):599–604. 12

Chilton, R. and Lee, R. (2007). The discrete origin of fetd-newmark late time instability, and a correction scheme. *Journal of Computational Physics*, 224(2):1293–1306. 10

Chung, Y.S., Cheon, C., Park, I.H. and Hahn, S.Y. (2000). Optimal shape design of microwave device using fdtd and design sensitivity analysis. *IEEE Transactions On Microwave Theory and Techniques*, 48(12):2289–2296. 30

Cohen, G.C. (2002). *Higher-Order Numerical Methods for Transient Wave Equations*. Springer-Verlag Heidelberg, New York. 17

Courant, R., Friedrichs, K. and Lewy, H. (1928). Über die partiellen Differenzengleichungen der mathematischen Physik. *Mathematische Annalen*, 100(1):32–74. 8

Cox, S.J. and Dobson, D.C. (1999). Maximizing band gaps in two-dimensional photonic crystals. *SIAM Journal on Applied Mathematics*, 59(6):2108–2120. 22

Cox, S.J. and Dobson, D.C. (2000). Band structure optimization of two-dimensional photonic crystals in h-polarization. *Journal of Computational Physics*, 158(2):214–224. 22

Dahl, J., Jensen, J.S. and Sigmund, O. (2008). Topology optimization for transient wave propagation problems in one dimension design of filters and pulse modulators. *Structural and Multidisciplinary Optimization*, 36(6):585–595. 2, 24

Danielson, K.T. and Namburu, R.R. (1998). Nonlinear dynamic finite element analysis on parallel computers using fortran 90 and mpi. *Advances In Engineering Software*, 29(3-6):179–186. 33

Diaz, A.R. and Kikuchi, N. (1992). Solutions to shape and topology eigenvalue optimization problems using a homogenization method. *International Journal For Numerical Methods In Engineering*, 35(7):1487–1502. 22

Diaz, A.R. and Sigmund, O. (2010). A topology optimization method for design of negative permeability metamaterials. *Structural and Multidisciplinary Optimization*, 41(2):163–177. 23

Diener, G. (1996). Superluminal group velocities and information transfer. *Physics Letters A*, 223(5):327–331. 50

Du, J. and Olhoff, N. (2007). Minimization of sound radiation from vibrating bimaterial structures using topology optimization. *Structural and Multidisciplinary Optimization*, 33(4):305–321. 22

Duhring, M.B. (2009). Design of acousto-optical devices by topology optimization. In *Proceedings of WCSMO-8, 8th World Congress on Structural Multidisciplinary Optimization*. LNEC, Lisbon, Portugal. 23

Duhring, M.B., Jensen, J.S. and Sigmund, O. (2008). Acoustic design by topology optimization. *Journal of Sound and Vibration*, 317(3-5):557–575. 23

Eggleton, B.J., de Sterke, C.M. and Slusher, R.E. (1999). Bragg solitons in the nonlinear schrodinger limit: experiment and theory. *Journal Of The Optical Society Of America B-Optical Physics*, 16(4):587–599. 50

Erentok, A. and Sigmund, O. (2011). Topology optimization of sub-wavelength antennas. *Antennas and Propagation, IEEE Transactions on*, 59(1):58 –69. 23

Evgrafov, A., Rupp, C.J., Dunn, M.L. and Maute, K. (2008a). Optimal synthesis of tunable elastic wave-guides. *Computer Methods In Applied Mechanics and Engineering*, 198(2):292–301. 23

Evgrafov, A., Rupp, C.J., Maute, K. and Dunn, M.L. (2008b). Large-scale parallel topology optimization using a dual-primal substructuring solver. *Structural and Multidisciplinary Optimization*, 36(4):329–345. 32

Forsberg, J. and Nilsson, L. (2007). Topology optimization in crashworthiness design. *Structural and Multidisciplinary Optimization*, 33(1):1–12. 22

Frandsen, L.H., Harpoth, A., Borel, P.I., Kristensen, M., Jensen, J.S. and Sigmund, O. (2004). Broadband photonic crystal waveguide 60 degrees bend obtained utilizing topology optimization. *Optics Express*, 12(24):5916–5921. 49

Frandsen, L.H., Lavrinenko, A.V., Fage-Pedersen, J. and Borel, P.I. (2006). Photonic crystal waveguides with semi-slow light and tailored dispersion properties. *Optics Express*, 14(20):9444–9450. 52

Frei, W.R., Tortorelli, D.A. and Johnson, H.T. (2005). Topology optimization of a photonic crystal waveguide termination to maximize directional emission. *Applied Physics Letters*, 86(11):111114. 23

Gersborg-Hansen, A., Sigmund, O. and Haber, R. (2005). Topology optimization of channel flow problems. *Structural and Multidisciplinary Optimization*, 30(3):181–192. 22

Gibiansky, L.V. and Sigmund, O. (2000). Multiphase composites with extremal bulk modulus. *Journal of the Mechanics and Physics of Solids*, 48(3):461–498. 22

Greene, W.H. and Haftka, R.T. (1991). Computational aspects of sensitivity calculations in linear transient structural-analysis. *Structural Optimization*, 3(3):176–201. 27

Guest, J.K. (2009). Topology optimization with multiple phase projection. *Computer Methods In Applied Mechanics and Engineering*, 199:123–135. 26

Halkjær, S., Sigmund, O. and Jensen, J. (2006). Maximizing band gaps in plate structures. *Structural and Multidisciplinary Optimization*, 32(4):263–275. 22

Halkjær, S., Sigmund, O. and Jensen, J.S. (2005). Inverse design of phononic crystals by topology optimization. *Zeitschrift fr Kristallographie*, 220(9-10):895–905. 22

Hau, L.V., Harris, S.E., Dutton, Z. and Behroozi, C.H. (1999). Light speed reduction to 17 metres per second in an ultracold atomic gas. *Nature*, 397(6720):594–598. 50

Haus, H.A. (1984). *Waves and Fields in Optoelectronics*. Prentice-Hall: Englewood Cliffs, NJ. 40

Hsieh, C.C. and Arora, J.S. (1984). Design sensitivity analysis and optimization of dynamic-response. *Computer Methods In Applied Mechanics and Engineering*, 43(2):195–219. 27

Hsieh, C.C. and Arora, J.S. (1985). Structural design sensitivity analysis with general boundary-conditions - dynamic problem. *International Journal For Numerical Methods In Engineering*, 21(2):267–283. 27

Hughes, T.J.R. (2000). *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover: New York. 7

Inoue, K., Kawai, N., Sugimoto, Y., Carlsson, N., Ikeda, N. and Asakawa, K. (2002). Observation of small group velocity in two-dimensional algaas-based photonic crystal slabs. *Phys. Rev. B*, 65(12):121308. 53

Jackson, J.D. (1999). *Classical Electrodynamics*. Wiley. 7, 18, 42

Jacobsen, R.S., Lavrinenko, A.V., Frandsen, L.H., Peucheret, C., Zsigri, B., Moulin, G., Fage-Pedersen, J. and Borel, P.I. (2005). Direct experimental and numerical determination of extremely high group indices in photonic crystal waveguides. *Optics Express*, 13:7861–7871. 53

Jensen, J.S. (2007a). Topology optimization of dynamics problems with pade approximants. *International Journal For Numerical Methods In Engineering*, 72(13):1605–1630. 23

Jensen, J.S. (2007b). Topology optimization problems for reflection and dissipation of elastic waves. *Journal of Sound and Vibration*, 301(1-2):319–340. 22

Jensen, J.S. (2009). Space-time topology optimization for one-dimensional wave propagation. *Computer Methods In Applied Mechanics and Engineering*, 198(5-8):705–715. 2, 24

Jensen, J.S. and Pedersen, N.L. (2006). On maximal eigenfrequency separation in two-material structures: the 1d and 2d scalar cases. *Journal of Sound and Vibration*, 289(4-5):967–986. 23

Jensen, J.S. and Sigmund, O. (2004). Systematic design of photonic crystal structures using topology optimization: Low-loss waveguide bends. *Applied Physics Letters*, 84(12):2022–2024. 23

Jensen, J.S. and Sigmund, O. (2005). Topology optimization of photonic crystal structures: a high-bandwidth low-loss t-junction waveguide. *Journal of the Optical Society of America B-optical Physics*, 22(6):1191–1198. 23, 26, 44

Jensen, J.S. and Sigmund, O. (2011). Topology optimization for nano-photonics. *Laser & Photonics Reviews*, 5(2):308–321. 1, 23

Jensen, J.S., Sigmund, O., Frandsen, L.H., Borel, P.I., Harpoth, A. and Kristensen, M. (2005). Topology design and fabrication of an efficient double 90(circle) photonic crystal waveguide bend. *IEEE Photonics Technology Letters*, 17(6):1202–1204. 49

Jin, J.M. and Riley, D. (2009). *Finite Element Analysis of Antennas and Arrays*. John Wiley & Sons, Inc. 7, 8, 17, 36

Joannopoulos, J., Johnson, S., Winn, J. and Meade, R. (2008). *Photonic crystals : molding the flow of light*. Princeton University Press, Princeton. 35, 37, 44

John, S. (1987). Strong localization of photons in certain disordered dielectric superlattices. *Physical Review Letters*, 58(23):2486–2489. 35

Karypis, G. and Kumar, G. (2009). Metis: Unstructured graph partitioning and sparse matrix ordering system, version 4.0. http://www.cs.umn.edu/ metis. 18

Kash, M.M., Sautenkov, V.A., Zibrov, A.S., Hollberg, L., Welch, G.R., Lukin, M.D., Rostovtsev, Y., Fry, E.S. and Scully, M.O. (1999). Ultraslow group velocity and enhanced nonlinear optical effects in a coherently driven hot atomic gas. *Physical Review Letters*, 82(26):5229–5232. 50

Kim, G.H., Lee, Y.H., Shinya, A. and Notomi, M. (2004a). Coupling of small, low-loss hexapole mode with photonic crystal slab waveguide mode. *Opt. Express*, 12(26):6624–6631. 41, 43, 49

Kim, T.S., Kim, J.E. and Kim, Y.Y. (2004b). Parallelized structural topology optimization for eigenvalue problems. *International Journal of Solids and Structures*, 41(9-10):2623–2641. 32

Kiziltas, G., Kikuchi, N., Volakis, J. and Halloran, J. (2004). Topology optimization of dielectric substrates for filters and antennas using simp. *Archives of Computational Methods in Engineering*, 11:355–388. 23

Kiziltas, G., Psychoudakis, D., Volakis, J. and Kikuchi, N. (2003). Topology design optimization of dielectric substrates for bandwidth improvement of a patch antenna. *Antennas and Propagation, IEEE Transactions on*, 51(10):2732 – 2743. 23

Koshiba, M., Tsuji, Y. and Sasaki, S. (2001). High-performance absorbing boundary conditions for photonic crystal waveguide simulations. *Microwave and Wireless Components Letters, IEEE*, 11(4):152 –154. 13

Krauss, T.F. (2007). Slow light in photonic crystal waveguides. *Journal of Physics D-applied Physics*, 40(9):2666–2670. 50

Kreissl, S., Pingen, G., Evgrafov, A. and Maute, K. (2010). Topology optimization of flexible micro-fluidic devices. *Structural and Multidisciplinary Optimization*, 42(4):495–516. 22

Krysl, P. and Bittnar, Z. (2001). Parallel explicit finite element solid dynamics with domain decomposition and message passing: dual partitioning scalability. *Computers & Structures*, 79(3):345–360. 17, 33

Kuzuoglu, M. and Mittra, R. (1996). Frequency dependence of the constitutive parameters of causal perfectly matched anisotropic absorbers. *Ieee Microwave and Guided Wave Letters*, 6(12):447–449. 12

Larsen, A., Laksafoss, B., Jensen, J. and Sigmund, O. (2009). Topological material layout in plates for vibration suppression and wave propagation control. *Structural and Multidisciplinary Optimization*, 37(6):585–594. 22

Larsen, U., Signund, O. and Bouwsta, S. (1997). Design and fabrication of compliant micromechanisms and structures with negative poisson's ratio. *Microelectromechanical Systems, Journal of*, 6(2):99 –106. 22

Letartre, X., Seassal, C., Grillet, C., Rojo-Romeo, P., Viktorovitch, P., d'Yerville, M.L., Cassagne, D. and Jouanin, C. (2001). Group velocity and propagation losses measurement in a single-line photonic-crystal waveguide on inp membranes. *Applied Physics Letters*, 79(15):2312–2314. 52

Li, Y., Saitou, K. and Kikuchi, N. (2004). Topology optimization of thermally actuated compliant mechanisms considering time-transient effect. *Finite Elements in Analysis and Design*, 40(11):1317–1331. 22, 24

Luebbers, R.J. and Hunsberger, F. (1992). Fdtd for nth-order dispersive media. *Ieee Transactions On Antennas and Propagation*, 40(11):1297–1301. 14

Luyssaert, B. (2005). *Compact Planar Waveguide Spot-Size Converters in Silicon-on-Insulator*. Ph.D. thesis, Universiteit Gent. 37

Ma, Z.D., Kikuchi, N. and Cheng, H.C. (1995). Topological design for vibrating structures. *Computer Methods in Applied Mechanics and Engineering*, 121(1-4):259–280. 22

Mahdavi, A., Balaji, R., Frecker, M. and Mockensturm, E.M. (2006). Topology optimization of 2d continua for minimum compliance using parallel computing. *Structural and Multidisciplinary Optimization*, 32(2):121–132. 32

Manolatou, C., Khan, M., Fan, S., Villeneuve, P., Haus, H. and Joannopoulos, J. (1999). Coupling of modes analysis of resonant channel add-drop filters. *Quantum Electronics, IEEE Journal*, 35(9):1322–1331. 40

Maxwell, J.C. (1865). A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155:459–512. 6

Michaleris, P., Tortorelli, D.A. and Vidal, C.A. (1994). Tangent operators and design sensitivity formulations for transient nonlinear coupled problems with applications to elastoplasticity. *International Journal For Numerical Methods In Engineering*, 37(14):2471–2499. 27

Min, S., Kikuchi, N., Park, Y.C., Kim, S. and Chang, S. (1999). Optimal topology design of structures under dynamic loads. *Structural and Multidisciplinary Optimization*, 17:208–218. 24

Newmark, N. (1959). A method of computation for structural dynamics. *ASCE Journal of the Engineering Mechanics Division*, 85(3):67–94. 9

Nomura, T., Sato, K., Taguchi, K., Kashiwa, T. and Nishiwaki, S. (2007). 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. 2, 24

Notomi, M., Yamada, K., Shinya, A., Takahashi, J., Takahashi, C. and Yokohama, I. (2001). Extremely large group-velocity dispersion of line-defect waveguides in photonic crystal slabs. *Physical Review Letters*, 87(25). 61

Oskooi, A.F., Roundy, D., Ibanescu, M., Bermel, P., Joannopoulos, J. and Johnson, S.G. (2010). Meep: A flexible free-software package for electromagnetic simulations by the fdtd method. *Computer Physics Communications*, 181(3):687–702. 52

Painter, O., Lee, R.K., Scherer, A., Yariv, A., O'Brien, J.D., Dapkus, P.D. and Kim, I. (1999). Two-dimensional photonic band-gap defect mode laser. *Science*, 284(5421):1819–1821. 43

Pedersen, C.B.W. (2004). Crashworthiness design of transient frame structures using topology optimization. *Computer Methods In Applied Mechanics and Engineering*, 193(6-8):653–678. 22

Pedersen, C.B.W., Buhl, T. and Sigmund, O. (2001). Topology synthesis of large-displacement compliant mechanisms. *International Journal for Numerical Methods in Engineering*, 50(12):2683–2705. 22

Pedersen, N.L. (2000). Maximization of eigenvalues using topology optimization. *Structural and Multidisciplinary Optimization*, 20(1):2–11. 22

Riishede, J. and Sigmund, O. (2008). Inverse design of dispersion compensating optical fiber using topology optimization. *J. Opt. Soc. Am. B*, 25(1):88–97. 23

Royer, D. and Dieulesaint, E. (2000). *Elastic Waves in Solids I - Free abd Guided Propagation*. Springer, New York. 18

Rupp, C., Evgrafov, A., Maute, K. and Dunn, M. (2007). Design of phononic materials/structures for surface wave devices using topology optimization. *Structural and Multidisciplinary Optimization*, 34(2):111–121. 23

Schulz, S.A., O'Faolain, L., Beggs, D.M., White, T.P., Melloni, A. and Krauss, T.F. (2010). Dispersion engineered slow light in photonic crystals: a comparison. *Journal of Optics*, 12(10). 52, 59

Sigmund, O. (1994). *Design of material structures using topology optimization*. Ph.D. thesis, Department of Solid Mechanics, Technical University of Denmark. 22

Sigmund, O. (1997). On the design of compliant mechanisms using topology optimization. *Mechanics of Structures and Machines*, 25(4):493–524. 22

Sigmund, O. (2000). A new class of extremal composites. *Journal of the Mechanics and Physics of Solids*, 48(2):397–428. 22

Sigmund, O. (2001a). Design of multiphysics actuators using topology optimization - part i: One-material structures. *Computer Methods In Applied Mechanics and Engineering*, 190(49-50):6577–6604. 22

Sigmund, O. (2001b). Design of multiphysics actuators using topology optimization - part ii: Two-material structures. *Computer Methods In Applied Mechanics and Engineering*, 190(49-50):6605–6627. 22

Sigmund, O. (2007). Morphology-based black and white filters for topology optimization. *Structural and Multidisciplinary Optimization*, 33(4-5):401–424. 26, 56

Sigmund, O. (2009). *IUTAM Bookseries*, chapter Systematic Design of Metamaterials by Topology Optimization, pages 151–159. Springer Netherlands. 23

Sigmund, O. and Hougaard, K. (2008). Geometric properties of optimal photonic crystals. *Physical Review Letters*, 100(15):153904. 23

Sigmund, O. and Jensen, J.S. (2003). Systematic design of phononic band-gap materials and structures by topology optimization. *Philosophical Transactions of the Royal Society of London Series A-mathematical Physical and Engineering Sciences*, 361(1806):1001–1019. 22, 25

Snir, M., Otto, S., Huss-Lederman, S., Walker, D. and Dongarra, J. (1995). *MPI: The Complete Reference*. MIT Press Cambridge. 17

Soljacic, M. and Joannopoulos, J.D. (2004). Enhancement of nonlinear effects using photonic crystals. *Nature Materials*, 3(4):211–219. 50

Soljacic, M., Johnson, S.G., Fan, S.H., Ibanescu, M., Ippen, E. and Joannopoulos, J.D. (2002). Photonic-crystal slow-light enhancement of nonlinear phase sensitivity. *Journal Of The Optical Society Of America B-Optical Physics*, 19(9):2052–2059. 50

Somerville, W.R.C., Le Ru, E.C., Northcote, P.T. and Etchegoin, P.G. (2010). High performance raman spectroscopy with simple optical components. *American Journal of Physics*, 78(7):671–677. 39

Stainko, R. and Sigmund, O. (2007). Tailoring dispersion properties of photonic crystal waveguides by topology optimization. *Waves In Random and Complex Media*, 17(4):477–489. 23

Svanberg, K. (2002). A class of globally convergent optimization methods based on conservative convex separable approximations. *SIAM J. Optim.*, 12(2):555–573. 27

Taflove, A. and Hagness, S. (2005). *Computational Electrodynamics: The Finite-Difference Time-domain Method*. Artech House, Inc., third edition edition. 8

Tortorelli, D.A. and Haber, R.B. (1989). First-order design sensitivities for transient conduction problems by an adjoint method. *International Journal for Numerical Methods in Engineering*, 28(4):733–752. 31

Tortorelli, D.A., Haber, R.B. and Lu, S.C.Y. (1991). Adjoint sensitivity analysis for nonlinear dynamic thermoelastic systems. *Aiaa Journal*, 29(2):253–263. 27

Tortorelli, D.A., Lu, S.C.Y. and Haber, R.B. (1990). Design sensitivity analysis for elastodynamic systems. *Mechanics Of Structures And Machines*, 18(1):77–106. 27

Tortorelli, D.A. and Michaleris, P. (1994). Design sensitivity analysis: Overview and review. *Inverse Problems in Engineering*, 1(1):71–105. 27

Tsay, J.J. and Arora, J.S. (1990). Nonlinear structural design sensitivity analysis for path dependent problems.1. general-theory. *Computer Methods In Applied Mechanics And Engineering*, 81(2):183–208. 27

Tsuji, Y., Hirayama, K., Nomura, T., Sato, K. and Nishiwaki, S. (2006). Design of optical circuit devices based on topology optimization. *Ieee Photonics Technology Letters*, 18(5-8):850–852. 23

Tucker, R.S., Ku, P.C. and Chang-Hasnain, C.J. (2005). Slow-light optical buffers: Capabilities and fundamental limitations. *Journal Of Lightwave Technology*, 23(12):4046–4066. 50

Turteltaub, S. (2001). Optimal material properties for transient problems. *Structural and Multidisciplinary Optimization*, 22:157–166. 24

Wadbro, E. and Berggren, M. (2006). Topology optimization of an acoustic horn. *Computer Methods In Applied Mechanics and Engineering*, 196(1-3):420–436. 23

Wadbro, E., Udawalpola, R. and Berggren, M. (2010). Shape and topology optimization of an acoustic horn-lens combination. *Journal of Computational and Applied Mathematics*, 234(6):1781–1787. 23

Wang, F., Jensen, J.S. and Sigmund, O. (2011). Robust topology optimization of photonic crystal waveguides with tailored dispersion properties. *J. Opt. Soc. Am. B*, 28(3):387–397. 23, 26, 56, 61

Watanabe, Y., Ikeda, N., Sugimoto, Y., Takata, Y., Kitagawa, Y., Mizutani, A., Ozaki, N. and Asakawa, K. (2007). Topology optimization of waveguide bends with wide, flat bandwidth in air-bridge-type photonic crystal slabs. *Journal of Applied Physics*, 101(11). 23

Watanabe, Y., Sugimoto, Y., Ikeda, N., Ozaki, N., Mizutani, A., Takata, Y., Kitagawa, Y. and Asakawa, K. (2006). Broadband waveguide intersection with low crosstalk in two-dimensional photonic crystal circuits by using topology optimization. *Opt. Express*, 14(20):9502–9507. 23

Xu, S., Cai, Y. and Cheng, G. (2010). Volume preserving nonlinear density filter based on heaviside functions. *Structural and Multidisciplinary Optimization*, 41:495–505. 26

Yablonovitch, E. (1987). Inhibited spontaneous emission in solid-state physics and electronics. *Physical Review Letters*, 58(20):2059–2062. 35

Yang, L.R., Lavrinenko, A.V., Hvam, J.M. and Sigmund, O. (2009). Design of one-dimensional optical pulse-shaping filters by time-domain topology optimization. *Applied Physics Letters*, 95(26):261101. 2, 24

Yoon, G., Jensen, J. and Sigmund, O. (2007). Topology optimization for acoustic-structure interaction problems using a mixed finite element formulation. *International Journal For Numerical Methods In Engineering*, 70:1049–1075. 23

Yue, B. and Guddati, M.N. (2005). Dispersion-reducing finite elements for transient acoustics. *Journal of the Acoustical Society of America*, 118(4):2132–2141. 10

Zienkiewicz, O., Taylor, R. and Zhu, J. (2005). *The Finite Element Method: Its Basis and Fundamentals*. Elsevier Butterworth-Heinemann, London, 6th edition. 7, 8, 9

Errata [P2]

There are a couple typing errors in some of the formulas in reference [P2]. All the produced results have, however, been obtained by using the correct formulas. Hence, all figure data is correct and remains unaffected.

Equation (1) has a sign error and should read:

$$\mu\mathcal{L}_1(t)H_z - \left[\frac{\partial}{\partial x}\left(\frac{\mathcal{L}_{2,x}(t)}{\varepsilon}\frac{\partial}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\mathcal{L}_{2,y}(t)}{\varepsilon}\frac{\partial}{\partial y}\right)\right] H_z = -\frac{\partial J_{\mathrm{B},z}}{\partial t}$$

In the 'Finite-Element Implementation'-section the permittivity is missing in the 'stiffness'-related term in the weak-form representation in equation (7). It should read:

$$\iiint_V \left[\mu T_z \mathcal{L}_1(t) H_z(t) + \varepsilon^{-1}\nabla T_z \cdot \tilde{\nabla} H_z(t) + T_z \frac{\partial J_{\mathrm{B},z}}{\partial t}\right] \mathrm{d}V + \iint_S \left[Y_c T_z \frac{\partial H_z}{\partial t} + T_z U_z\right] \mathrm{d}S = 0$$

In the 'Sensitivity Analysis'-section the manipulated sensitivity expression in equation (21) needs a ∂-sign in second term of the first line, such that the expression should read:

$$\begin{aligned}\frac{\partial \Phi}{\partial x_e} &= \left[\boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}}\frac{\partial \mathbf{u}}{x_e} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}\frac{\partial \dot{\mathbf{u}}}{\partial x_e} - \dot{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}\frac{\partial \mathbf{u}}{\partial x_e}\right]_0^T \\ &+ \int_0^T \left(\frac{\partial F}{\partial \dot{\mathbf{u}}}\frac{\partial \dot{\mathbf{u}}}{\partial x_e} + \frac{\partial F}{\partial \ddot{\mathbf{u}}}\frac{\partial \ddot{\mathbf{u}}}{\partial x_e} + \frac{\partial F}{\partial x_e}\right) \mathrm{d}t \\ &+ \int_0^T \left(\frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}\ddot{\boldsymbol{\lambda}} - \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}}\dot{\boldsymbol{\lambda}} + \frac{\partial \mathbf{r}}{\partial \mathbf{u}}\boldsymbol{\lambda} - \frac{\partial F}{\partial \mathbf{u}}\right)\frac{\partial \mathbf{u}}{\partial x_e}\mathrm{d}t \\ &+ \int_0^T \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial x_e}\mathrm{d}t\end{aligned}$$

After equation (30) $\tau = 180c/a$ should instead read $\tau = 180a/c$. The theoretical prediction for the transmission at resonance in equation (32) needs a squared-sign in the second term in the denominator, i.e.

$$|T(\omega)|^2 = 1 - \frac{\frac{1}{4Q_\mathrm{e}^2}\left(1 + 2\frac{Q_\mathrm{e}}{Q_0}\right)}{\left(\frac{\omega-\omega_0}{\omega_0}\right)^2 + \frac{1}{4Q_\mathrm{e}^2}\left(1 + \frac{Q_\mathrm{e}}{Q_0}\right)^2}$$

Finally in Fig. 8 and Fig. 10, ω_r should be replaced by ω_0 to denote the resonance frequency.

Errata [P4]

There are two typing errors in reference [P4]. All the produced results have, however, been obtained by using the correct formulas. Hence, all figure data is correct and remains unaffected.

Equation (17c) has an index error and should read:

$$\mathbf{C}' = \mathbf{C}_0 + \mathcal{L}_1(t)\mathbf{C}_1 + \mathcal{L}_2(t)\mathbf{C}_2$$

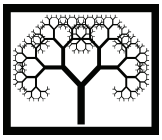
Equation (28a) has a sign error and should read:

$$\widetilde{\mathbf{h}}_i^e = \sum_j \left(\mathbf{M}_{1,ij}^e - \frac{\beta_1^2}{P_{3,1}} \mathbf{K}_{1,ij}^e \right) \mathbf{u}_{1,j}^+ + \left(\mathbf{M}_{2,ij}^e - \frac{\beta_2^2}{P_{3,2}} \mathbf{K}_{2,ij}^e \right) \mathbf{u}_{2,j}^+$$

$$\widetilde{\mathbf{g}}_i^e = \sum_j \mathbf{K}_{2,ij}^e P_{4,1} \mathbf{u}_{1,j}^{++} + \mathbf{K}_{1,ij}^e P_{4,2} \mathbf{u}_{2,j}^{++}$$

[P1]

Paper 75



Proceedings of the Ninth International Conference
on Computational Structures Technology,
B.H.V. Topping and M. Papadrakakis, (Editors),
Civil-Comp Press, Stirlingshire, Scotland

Transient Topology Optimization of Two-Dimensional Elastic Wave Propagation

R. Matzen, J.S. Jensen and O. Sigmund
Solid Mechanics, Department of Mechanical Engineering
Technical University of Denmark, Lyngby, Denmark

Abstract

A tapering device coupling two monomodal waveguides is designed with the topology-optimization method based on transient wave propagation. The gradient-based optimization technique is applied to predict the material distribution in the tapering area such that the squared output displacement (a measure for transmission) in the taper is maximized. High transmission in a large frequency range is gained by use of incident wave packets. To avoid nondiscrete properties in the design domain a density filtering technique is employed.

Keywords: topology optimization, transient analysis, waveguide, the Helmholtz equation,wave packet, density filter.

1 Introduction

Recently, the topology optimization method has been applied to the design of structures and materials subject to harmonic wave propagation problems in the frequency domain. The focus has been put on developing periodic band gap structures in phononic and photonic applications (Sigmund & Jensen 2003 [14]). In light of the former Halkjær et al. 2005 [4] maximize band gaps for bending waves in elastic plate structures. The reverse problem of the transmission-type has also been investigated. Low-loss waveguide bends and a T-junction waveguide for electromagnetic waves are designed by Jensen & Sigmund in [6] and [7], respectively, using the procedure of topology optimization.

A disadvantage of using harmonic excitation is the need for several response analyses at different frequencies to optimize for larger bandwidths. However, topology optimization based on frequency responses represented by Padé approximants is demonstrated to be a very efficient way to obtain accurate approximations over wide fre-

quency ranges (Jensen & Sigmund 2005 [7]; Jensen 2007 [8]).

Another way to excite a broad frequency range is to model wave propagation in the time domain by selecting a proper input pulse. This only requires one analysis. In the framework of topology optimization this has just very recently been employed in antenna design by Nomura et al. (2007) [11], using the finite difference time difference (FDTD) method. Another very recent work by Dahl et al. (2007) [2] presents a topology optimization strategy for design of transient response of 1-D elastic and optic waveguides, where the finite element time difference (FETD) method is used.

Even though the sensitivity analysis for transient problems has existed for two decades (Haug & Arora 1978 [5]), topology optimization of mechanical problems subject to transient dynamic loadings have only received attention within the last few years. For instance, topology optimization of transport vehicles for crashworthiness modeled by frame structures was carried out by Pedersen (2004) [12]. Turteltaub (2005) optimized the performance of a two-phase composite under transient dynamic loading.

In this paper, the topology optimization method is applied to transient wave propagation problems in 2-D elastic media. The aim is to design a tapering device coupling two monomodal waveguides that improves the transmission of a linear taper. The modeling of elastic waves in 2-D media can be divided into two decoupled considerations: in-plane transverse and longitudinal waves and out-of-plane transverse waves (acoustic mode), respectively. This paper focuses on the latter. The wave propagation is modeled by the Helmholtz equation, which is discretized using the FETD method. The formulation can readily be applied to optical wave propagation in dielectric media by changing the material parameters. In order to model the wave propagation properly from a practical point of view we introduce standard absorbing boundary condition for normal incidence.

The paper is conducted in the following way. In Section 2 the Helmholtz equation is first presented in the continuous and in the corresponding finite element form. Subsequently, design variables and material interpolation scheme are established leading to the setup of the optimization problem. The sensitivity analysis is then carried out and the implementation is finally illuminated. In Section 3, the method is applied to design a tapering device for a large bandwidth.

2 Topology optimization of 2-D elastic structures subject to wave propagation

Out-of-plane wave propagation in a non-homogenous media is governed in the time domain by

$$\nabla \cdot A(\mathbf{x})\nabla u(\mathbf{x}) - B(\mathbf{x})\ddot{u}(\mathbf{x}) = f(t), \qquad \text{in } \Omega \tag{1}$$

where $\ddot{(\,)} = \partial^2/\partial t^2$. In eq. (1) $u(\mathbf{x})$ is the unknown field in the plane $\mathbf{x} = (x, y)$, $A(\mathbf{x})$ and $B(\mathbf{x})$ are the spatial-dependent material parameters, $f(t)$ is the time-dependent

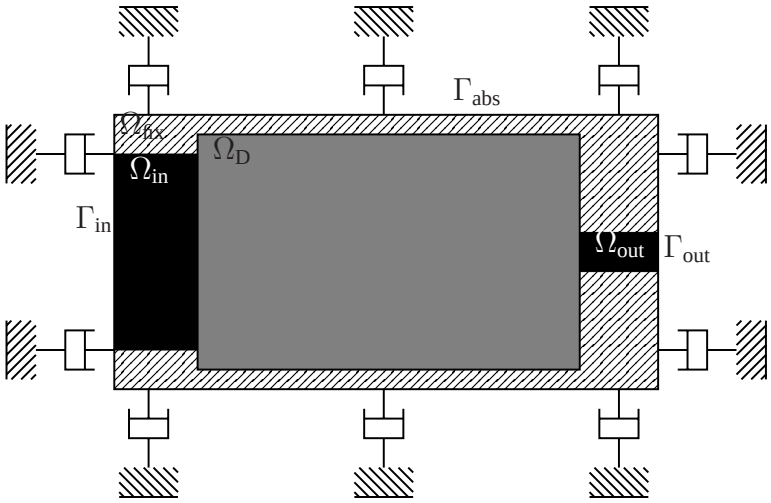


Figure 1: Computational domain Ω decomposed in four subregions: Ω_D is the design domain, Ω_{fix} contains fixed elements, Ω_{in} and Ω_{out} are in- and outlet regions, respectively with dielectric material. Absorbing boundaries are specified on $\partial\Omega = \Gamma_{abs}$. A wave input is provided on Γ_{in}.

dynamic loading and Ω designates the modeling domain. For elastic waves $A = \mu(\mathbf{x})$ and $B = \rho(\mathbf{x})$, where μ is the shear modulus and ρ the density. The optimized design found by topology optimization is a distribution of phase 1 and phase 2 materials characterized by (A_1, B_1) and (A_2, B_2), respectively.

It is convenient to introduce the following dimensionless parameters and variables

$$\tilde{A} = A_1^{-1}A, \quad \tilde{B} = B_1^{-1}B, \quad \tilde{\mathbf{x}} = a^{-1}\mathbf{x}, \quad \tilde{t} = a^{-1}ct, \quad \tilde{\omega} = ac^{-1}\omega \tag{2}$$

where c is the wave speed for phase 1 material and a is a characteristic length. Consequently, eq. (1) can be reused simply by replacing real physical quantities with the dimensionless ones. The Helmholtz equation does also apply for optical waves in dielectric media having transverse electric (TE) and transverse magnetic (TM) polarized fields perpendicular to one another.

When designing the tapering device we focus on the TE-polarization resulting in $\tilde{A} = \varepsilon_r^{-1}$ and $\tilde{B} = 1$, where ε_r is the dielectric constants. Phase 1 material is chosen to be air with a dielectric constant of unity, phase 2 material to be silicon with $\varepsilon_r = 11.56$, the speed of light is $c = 3 \cdot 10^8 \mathrm{m/s}$. To simulate traveling waves we impose the standard absorbing boundary conditions for normal incidence (eq. Krenk & Kirkegaard 2001 [9])

$$\frac{\partial u}{\partial n} + \frac{1}{c}\frac{\partial u}{\partial t} = 0 \tag{3}$$

where n is length the outward-pointing normal vector and c is the wave speed in light.

A Galerkin finite-element procedure is applied to discretize the Helmholtz equation, yielding the system of second order ordinary differential equations

$$[\mathbf{M}]\{\ddot{\mathbf{d}}\} + [\mathbf{C}]\{\dot{\mathbf{d}}\} + [\mathbf{K}]\{\mathbf{d}\} = \{\mathbf{f}\} \tag{4}$$

where $\{\mathbf{d}\}$ is a vector of discretized nodal values of $u(\mathbf{x})$. On element level the nodal values are interpolated as $u(\mathbf{x}) = [\mathbf{N}(\mathbf{x})]\{\mathbf{d}\}$, where $[\mathbf{N}(\mathbf{x})]_{1\times 4}$ is the bilinear shape function column vector. This leads to the following form of the finite element-matrices

$$\begin{aligned}
[\mathbf{M}] &= \sum_{e\in\Omega} \tilde{B}_e[\mathbf{M}^e], & [\mathbf{M}^e] &= \int_{\Omega^e} [\mathbf{N}]^T[\mathbf{N}]\mathrm{d}\Omega \\
[\mathbf{C}] &= \sum_{e\in\Gamma_{\text{abs}}} \sqrt{\tilde{A}_e\tilde{B}_e}[\mathbf{C}^e], & [\mathbf{C}^e] &= \oint_{\Gamma^e} [\mathbf{N}]^T[\mathbf{N}]\mathrm{d}\Gamma \\
[\mathbf{K}] &= \sum_{e\in\Omega} \tilde{A}_e[\mathbf{K}^e], & [\mathbf{K}^e] &= \int_{\Omega^e} \frac{\partial[\mathbf{N}]^T}{\partial x}\frac{\partial[\mathbf{N}]}{\partial x}\mathrm{d}\Omega + \int_{\Omega^e} \frac{\partial[\mathbf{N}]^T}{\partial y}\frac{\partial[\mathbf{N}]}{\partial y}\mathrm{d}\Omega \\
\{\mathbf{f}\} &= \sum_{e\in\Gamma_{\text{in}}} \sqrt{\tilde{A}_e\tilde{B}_e}\{\mathbf{f}^e\}, & \{\mathbf{f}^e\} &= \oint_{\Gamma^e} [\mathbf{N}]^T\mathrm{d}\Gamma
\end{aligned}$$

where $\tilde{A}_e$ and $\tilde{B}_e$ are assumed element-wise constant. The discrete form in eq. (4) is progressed in time using an explicit time stepping scheme. To obtain a very quick time stepping, we use mass lumping in terms of the row sum method, given by $[\mathbf{M}_{ii}] = \sum_j [\mathbf{M}_{ij}]$. This result in a rather insignificant trade-off in accuracy.

2.1 Design variables and material interpolation

We let the design domain Ω_{D} be a subregion of the modeling domain Ω (see Fig. 1). To each finite element within Ω_{D} one design variable x_e is assigned, that either takes the value of zero, corresponding to air, and one, corresponding to silicon; i.e. $x_e \in \{0, 1\}$.

Unfortunately this discrete optimization problem is hard to solve, in particular for large scale problems. Instead we utilize the SIMP interpolation scheme (Bendsøe & Sigmund 2004 [1]) in which the material parameters vary linearly between the parameters for the phase 1 and phase 2 materials, i.e.

$$\begin{aligned}
A_e &= A_1 + x_e(A_1 - A_2) \\
B_e &= B_1 + x_e(B_1 - B_2)
\end{aligned} \tag{5}$$

This facilitates the use of gradient-based algorithms to find an optimized design. Unlike in the discrete problem porous material is now able to appear. However, several methods have been developed to suppress this tendency. Prominent methods are e.g. *pamping* proposed by Jensen & Sigmund (2005), sensitivity filtering and different density filtering techniques proposed by Sigmund (2007). Here, a density filter with a linearly varying weighting function is applied. For further details, consult [15].

2.2 The optimization problem and sensitivity analysis

The objective is to maximize the transmission in the taper. A suitable measure is the squared output displacement on Γ_{out} evaluated as $F(\mathbf{d}, x_e) = \int_0^T H(\mathbf{d}, x_e)\mathrm{d}t$ where $H(\mathbf{d}, x_e) = -\{\mathbf{d}\}^T[\mathbf{L}]\{\mathbf{d}\}$. Here T is the fixed final time and $[\mathbf{L}]$ is a zero matrix with 1 in the diagonal entries corresponding to the DOFs of the object node. In the standard formulation the optimization problem takes the form

$$\begin{aligned} \min_{x_e} \quad & F(\mathbf{d}, x_e) \\ & 0 \leq \rho_e \leq 1, \qquad\qquad e \in \Omega_{\text{D}} \end{aligned}$$

This optimization problem is solved with the mathematical programming tool, the Method of Moving Asymptotes (Svanberg 1987 [16]), in combination with analytical derived sensitivities of the objective function.

Since we are dealing with a large number of design and state variables, the derivation of the analytical sensitivities is based on the adjoint method (Haug & Arora 1978 [5]). Without going into details, the sensitivities are given by

$$\frac{\partial F}{\partial x_e} = -\int_0^T \{\boldsymbol{\lambda}_e\}^T \left(\frac{\partial [\mathbf{K}_e]}{\partial x_e}\{\mathbf{d}_e\} + \frac{\partial [\mathbf{C}_e]}{\partial x_e}\{\dot{\mathbf{d}}_e\} + \frac{\partial [\mathbf{M}_e]}{\partial x_e}\{\ddot{\mathbf{d}}_e\} \right) \mathrm{d}t \tag{6}$$

where the adjoint operator $\boldsymbol{\lambda}$ is obtained from the pseudo backward initial-value-problem in time variable $\tau \equiv T - t$

$$[\mathbf{M}]\{\ddot{\bar{\boldsymbol{\lambda}}}\} + [\mathbf{Z}]\{\dot{\bar{\boldsymbol{\lambda}}}\} + [\mathbf{K}]\{\bar{\boldsymbol{\lambda}}\} = \left\{ \frac{\partial H}{\partial \{\mathbf{d}\}}(T - \tau) \right\}^T \tag{7}$$

with initial conditions $\{\bar{\boldsymbol{\lambda}}(0)\} = 0$ and $\{\dot{\bar{\boldsymbol{\lambda}}}(T)\} = 0$, and where $0 < \tau < T$. The adjoint operator is finally obtained by $\boldsymbol{\lambda}(t) = \bar{\boldsymbol{\lambda}}(T - \tau)$. Evidently, this backward IVP is the same as the state problem in eq. 4, albeit with another forcing term $\{\partial H/\partial\{\mathbf{d}\}\}^T = -2[\mathbf{L}]\{\mathbf{d}\}$. Hence we can use the same time stepping procedure.

2.3 Algorithm

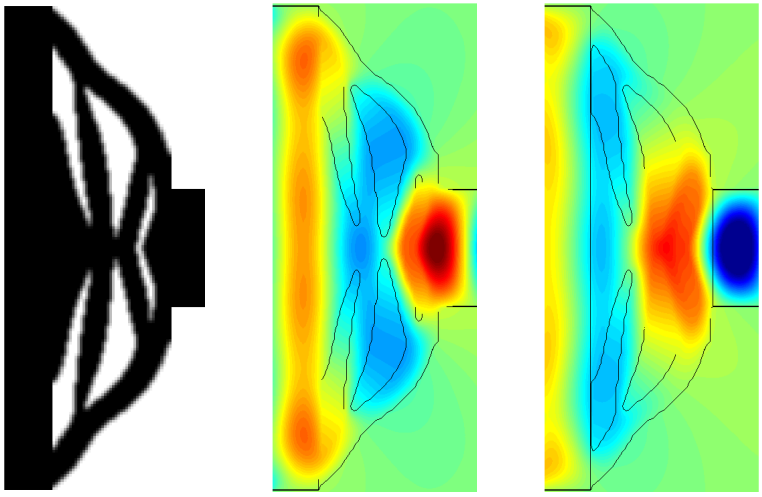
The algorithm has been implemented in MATLAB 2007b. The flowchart for the optimization algorithm is outlined in Table 1.

3 Design of tapering device

The coupling between waveguides with widths of 2μm and 0.5μm will now be studied at the wavelength of 1.55μm. The optimization is repeated for various lengths L of the coupling region in the range of 0.5μm to 3μm. To ensure an adequate representation of the given wavelength, we strive to retain a resolution in the design domain corresponding to at least 20 elements per wavelength in silicon. As input waves we use a

i.	Initialize mesh, load scenario
ii.	Build neighbourhood N_e and weighting function $w(\mathbf{x}_e)$ for $e \in \Omega_{\text{D}}$
iii.	Initialize design variables $\boldsymbol{\rho}$, counter loop=0, change$_{\boldsymbol{\rho}} = 1$ and rel. change$_F = 1$
iv.	do while change$_{\boldsymbol{\rho}} > 0.01$ or rel. change$_F = 10^{-7}$
v.	loop = loop +1
vi.	Filter densities $\boldsymbol{\rho} \rightarrow \hat{\boldsymbol{\rho}}$
vii.	Solve FE problem based on filtered densities $\hat{\boldsymbol{\rho}}$
viii.	Compute sensitivities $\frac{\partial F}{\partial \hat{\rho}_e}$ based on $\hat{\boldsymbol{\rho}}$
ix.	Update sensitivities in a consistent way, i.e. $\frac{\partial F}{\partial \tilde{\rho}_e} \rightarrow \frac{\partial F}{\partial \rho_e}$
x.	Update design variable $\boldsymbol{\rho}_{\text{new}}$ using MMA
xi.	Compute change$_{\boldsymbol{\rho}} = \|\|\boldsymbol{\rho} - \boldsymbol{\rho}_{\text{new}}\|\|_{\text{inf}}$ and rel.change$_F = \|F(\hat{\boldsymbol{\rho}}_{\text{new}}) - F(\hat{\boldsymbol{\rho}})\|/F(\hat{\boldsymbol{\rho}})$
xii.	end do
xiii.	post processing using the filtered densities $\hat{\boldsymbol{\rho}}$

Table 1: Flowchart for the optimization algorithm.



(a) Optimized design (b) Field at $t = 17.1$fs (c) Field at $t = 19.1$fs

Figure 2: (a) Optimized taper of length $L = 0.5\mu$m for coupling between 2μm and 0.5μm. (b)-(c) depict the displacement field where the contours correspond to material $\rho_e > 0.90$. The normalized objective function value is $F_{\text{norm}} = 3.35$. Filter with $R = 2h$ has been used. Transmission efficiency is $I_{\text{eff}} = 0.95$. $ndv = 4752$ and there are approximately 40 elements pr. wavelength. $\Omega = [0, 0.9] \times [0, 2.5]$ is discretized in 80×112 elements. $\Omega_{\text{D}} = [0, 0.5] \times [0, 2.4]$. Note that the entire modeling domain is not displayed.

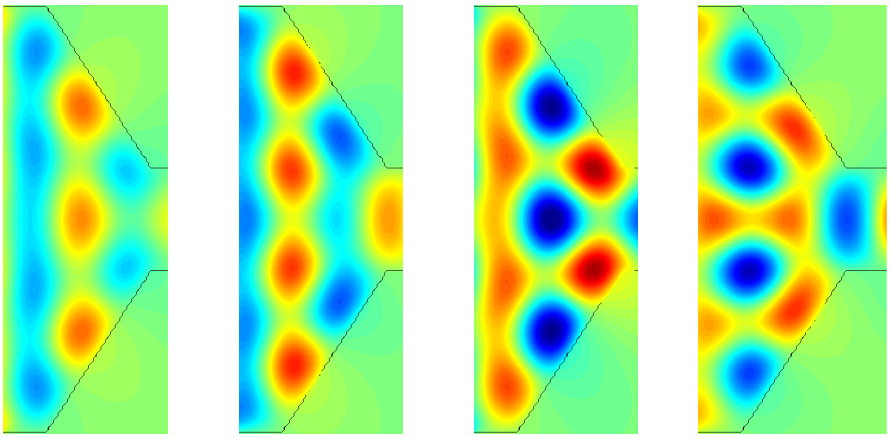


Figure 3: Displacement fields for taper of length $L = 0.5\mu$m at different times. From left to right the normalized times t correspond to 13.1fs, 15.1fs, 17.1fs and 19.1fs.

Gaussian wave packet with normalized center frequency $f_c = 0.37$ (i.e. wavelength $\Lambda = 1.55\mu$m), fractional bandwidth $bw = 0.2$ and threshold of -60dB. It is used due to its wider frequency (and thereby wavelength) distribution compared to harmonic excitation. The characteristic length is chosen to $a = 0.58\mu$m. Basically, the choice of physical data is inspired from the work done by Luyssaert [10].

Initially we optimize without filtering the interpolation densities. Then, dependent on the amount of gray material in the converged design, the optimization is repeated for increasing filter sizes such that the converged design (hopefully) will steer towards structurally interpretable black and white designs. Some specially selected (not necessarily the best) designs are presented in Fig. 2, 5 and 6 based on the physical filtered densities and together with the displacement field. The transmittance of all designs is found in Fig. 4a. It turns out, that the bias of applying a density filter with a linear weighting function is a rather insignificant trade-off in a numerical decrease of the objective function value. Even though a density filtering technique is utilized, the optimization problem is difficult to solve. This means that a large number of optimization steps are required to satisfy the convergence criteria.

The initial material distribution uses no prior knowledge of potential well-performing geometry, such as the parabolic taper [10]. In all cases a homogenous material distribution corresponding to $x_e = 0.5$ for $e \in \Omega_{\mathrm{D}}$ is used as initial guess. In the region with inactive elements, the in- and outlet are made of silicon; the remainder is vacuum.

The optimized design for the taper of length $L = 0.5$ is, according to Fig. 2a, topologically well-defined and simple. The outer form is modified slightly to be more curved compared to the linear taper in Fig. 3. However, the interior is fundamentally different. Since material is freely distributed when using the SIMP approach, the optimization algorithm is allowed to generate so-called *cavities* within the taper. Unlike parameter based optimization using a genetic algorithm, this is a major advantage of topology optimization. Comparing the displacement field before and after the optimization (cf. Fig. 3 and Fig. 2b-c) indicates that these cavities apparently have great

impact on reducing the reflection in the taper. Referring to Fig. 4 the transmission efficiency improves from approximately 30% to 95%.

Independent of taper length, it is seen in Fig. 2a, 5a and 6a that the formation of cavities is often seen in the optimized designs. Their size and geometry change drastically for increasing taper length, and they result in various interfaces between high and low refractive index material within the taper. It is not unlikely, that the appearance of these interfaces accompanied by the reduced amount of reflection, will result in strong resonant behavior. However, the transmission spectrum in Fig. 4b for the optimized taper of length 0.5μm does not indicate strong fluctuations which is a typical sign of resonance. Instead it shows a relative large frequency (wavelength) range within the wave packet threshold in which the efficiency remains high and rather flat. The author finds it puzzling that the maximum value of 92% in the transmission spectrum does not correspond to 95% computed by the time-averaged Poynting vector [13]. However, a reasonable explanation could be that the waves in the relative short in- and outlet regions are not fully propagating. The contribution from the radiated power is therefore missing in the energy estimate. So an extension of these regions in the modeling will assure the waves to be fully propagating and thus a correct energy estimate.

The complexity of the optimized design and the amount of intermediate material grow for increasing length. Hence the need for filters and other penalization methods, such as *pamping*, is essential to avoid nondiscrete design variables x_e. The author strongly believes that implementation of pamping would remove or replace gray material based on the findings in [7] and [15]. Further, a density filter with a Modified Heaviside weight function, proposed by Sigmund (2007) [15], also turns out to be very useful in wave transmission problems. Reversely, it is also worth considering the opportunity, that intermediate material might be favorable in design problems of the transmission type; evidently for increasing transmission distances of the taper. From

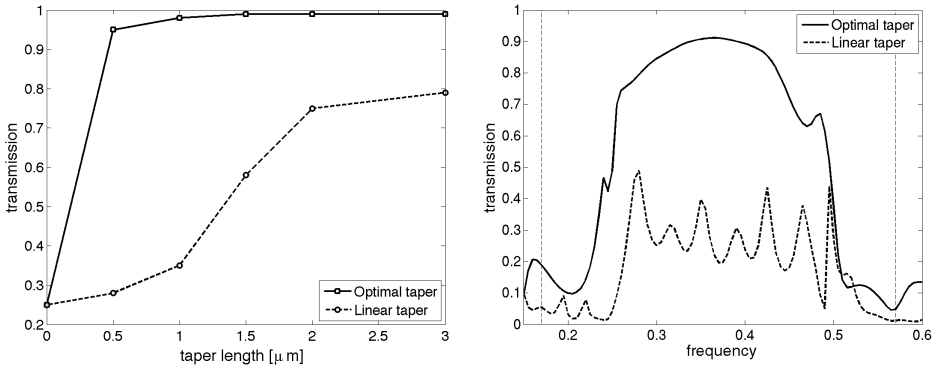
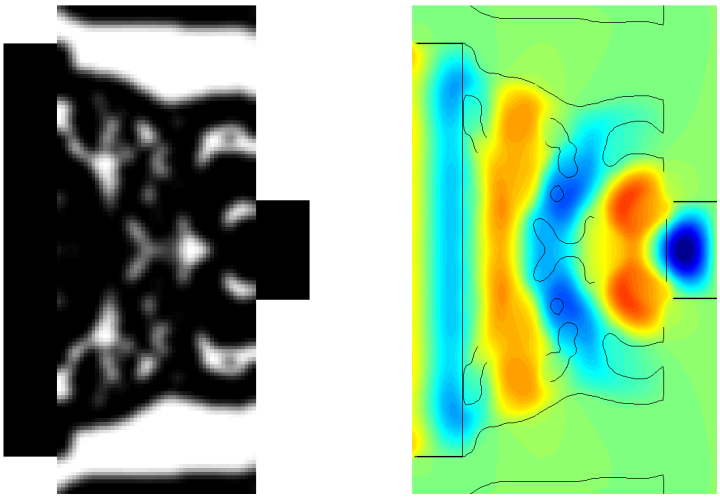
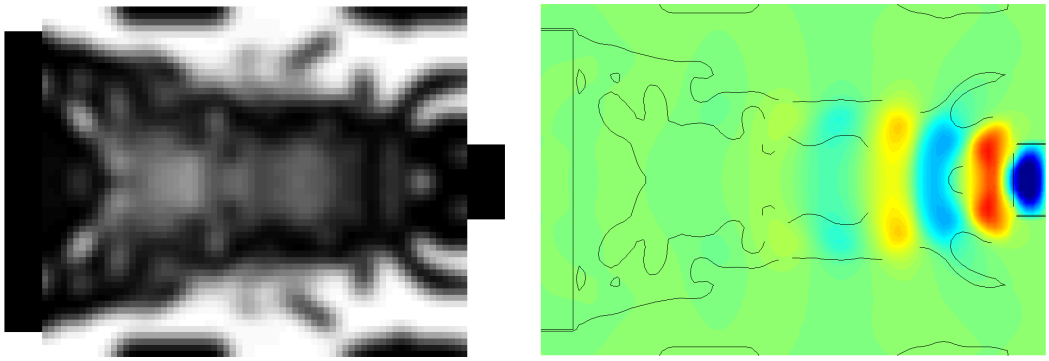


Figure 4: Left: Transmission efficiencies for optimized tapers of various length. **Right:** Transmission spectrum of the taper in Fig. 2. The *dotted lines* indicate the -60dB wave packet threshold. The transmission spectrum has been generated by a fast fourier transform of the instantaneous radiated power fraction from the in- and outlet.



(a) Optimized design (b) Field at $t = 18.8$fs

Figure 5: (a) Optimized taper of length $L = 1.0\mu$m for coupling between 2μm and 0.5μm. (b) depicts the displacement field where the contours correspond to material $\rho_e > 0.80$. The normalized objective function is $F_{\text{norm}} = 3.56$. Filter with $R = 3h$ is employed. Transmission efficiency is $I_{\text{eff}} = 0.98$. $ndv = 3328$ and the resolution is approximately 24 elements pr. wavelength. $\Omega = [0, 1.5] \times [0, 2.5]$ is discretized in 80×67 elements. $\Omega_{\text{D}} = [0.25, 1.25] \times [0, 2.4]$.



(a) Optimized design for $R = 3h$ (b) Field plot at $t = 30.4$fs.

Figure 6: (a) Optimized taper of length $L = 3.0\mu$m for coupling between 2μm and 0.5μm. (b) depicts the displacement field where the contours correspond to material $\rho_e > 0.80$. The normalized objective function value is $F_{\text{norm}} = 3.83$. Filter with $R = 3h$ is employed. Transmission efficiency is $I_{\text{eff}} = 0.99$. $ndv = 1940$ and there are approximately 14 elements pr. wavelength. $\Omega = [0, 3.5] \times [0, 2.5]$ is discretized in 80×29 elements. $\Omega_{\text{D}} = [0.25, 3.25] \times [0, 2.4]$.

a manufacturing point of view this is not desirable. However, since the optimization problem is non-convex, we know that other, potentially more applicable local maxima (optimized designs) may exist.

Material not in direct contact with the main part of the taper is inclined to emerge for increasing taper length. This is believed to serve as Bragg gratings that prevent waves from escaping the domain. Even for resolution refinement of the design domain the basic topology is retained. Another intriguing observation is that the form of the outer contour for the taper of length 3μm approaches the linear form and thereby endeavors adiabatic transmission.

It is interesting that the topology optimization method is able to improve the performance of the taper. From a practical point of view the most well-defined and still very well-performing design proposal (95% transmission) is fortunately gained for the shortest taper length. For instance, joining two waveguides seamlessly in a chip with several optical functions integrated namely requires that components like tapers become sufficiently small without compromising their performance.

4 Conclusions

In this paper the topology optimization method is employed to maximize the transmission in a tapering device by freely distributing silicon in the design domain. The method has revealed its capability to reliably predict optimized (optimal) topologies. The study has been carried out for various taper lengths. The topologically most well-defined design is obtained for the shortest length of 0.5μm. This has improved the transmission of the taper from approximately 30% of the linear taper to 95%.

The longer the taper gets, the more significant becomes the appearance of non-discrete interpolation densities. Therefore filtering of the densities plays an important role to obtain black and white designs. In general, problems of the transmission-type are likely to steer towards designs with gray regions. The tendency in all designs is formation of cavities within the taper. These cavities seem to be important in order to enhance transmission and they do not give rise to resonant phenomena. The transmission for longer tapers is also improved considerably.

Future work includes implementation of the Modified Heaviside filter and *pamping*. A way to improve the designs is to use the transmission in terms of the Poynting vector as objective function instead of the minimum displacement objective function. This is due to the fact that the former correctly estimates the transmitted energy by the flux, whereas the energy estimate in the latter corresponds to elastic energy. Lastly, more efficient absorbing boundaries that the ones for normal incidence should be used.

References

[1] Bendsøe, M. P. & Sigmund, O. (2004) *Topology Optimization; Theory, Methods, and Applications*. Springer Verlag Berlin Heiddelberg New York, 2nd edition.

[2] Dahl, J., Sigmund, O. & Jensen, J.S. (2007) *Topology optimization for transient wave propagation problems in one dimension : Design of filters and pulse modulators.* Struct. Multidisc. Optim. (2007). Vol. 33:401-424. Springer-Verlag 2007

[3] Halkjær, S., Sigmund, O., Jensen, J.S. (2005) *Inverse design of phononic crystals by topology optimization*. Z Kristallogr 220(9-10):895-905

[4] Halkjær, S., Sigmund, O., Jensen, J.S. (2006) *Maximizing band gaps in plate structures*. Struct Multidisc Optim 32(4):263-275

[5] Haug, E. & Arora, J. (1978) *Design sensitivity analysis of elastic mechanical systems*. Comput Methods Appl Mech Eng 15(1):35-62

[6] Jensen, J.S. & Sigmund, O. (2004). *Systematic design of photonic crystal structures using topology optimization: low-loss waveguide bends*. Appl Phys Lett 84(12):2022-2024

[7] Jensen, J. S. & Sigmund. O. (2005) *Topology optimization of photonic crystal structures: a high-bandwidth low-loss T-junction waveguide*. J. Opt. Soc. Am. B. Vol. 22, No. 6.

[8] Jensen, J.S. (2007). *Topology optimization of dynamics problems with Padé approximants*. Int J Numer Methods Eng doi:10.1002/nme.2065

[9] Krenk, S. & Kirkegaard. P.H. (2001) *Local Tensor Radiation Conditions for Elastic Waves* Journal of Sound and Vibration (2001), 247(5): 875-896.

[10] Luyssart, B. (2005). *Compact Planar Waveguide Spot-Size Converters in Silicon-on-Insulator*. Faculteit Ingenieurswetenschappen, Universiteit Gent.

[11] Nomura T., Sato, K., Taguchi K., Kashiwa T., Nishiwaki S. (2007) *Structural topology optimization for the design of broadband dielectric resonator antennas using the finite difference time domain technique*. Int J Numer Methods Eng 71(11):1261-1296.

[12] Pedersen, C. B. W. (2004). *Crashworthiness design of transient frame structures using topology optimization*. Computer Methods in Applied Mechanics and Engineering.

[13] Royer, D. & Dieulesaint, E. (2000). *Elastic Waves in Solids I - Free and Guided Propagation*. Springer Verlag Berlin Heiddelberg New York.

[14] Sigmund, O. & Jensen, J.S. (2003). *Systematic design of phononic band-gap materials and structures by topology optimization*. Phil. Trans. R. Soc. Lond. A (2003) 361:1001-1019.

[15] Sigmund. O. (2007). *Morphology-based black and white filters for topology optimization*. Struct. Multidisc. Optim. (2007). Vol. 33:401-424. Springer-Verlag 2007

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

[17] Turteltaub, S. (2005) *Optimal non-homogenous composites for dynamic loading*. Struct Multtidisc Optim 30(2):101-112

[P2]

Topology optimization for transient response of photonic crystal structures

René Matzen,* Jakob S. Jensen, and Ole Sigmund

Department of Mechanical Engineering, Technical University of Denmark, Nils Koppels Allé, Building 404, 2800 Kgs. Lyngby, Denmark
**Corresponding author: rmat@mek.dtu.dk*

Received June 22, 2010; accepted August 4, 2010;
posted August 13, 2010 (Doc. ID 130543); published September 16, 2010

An optimization scheme based on topology optimization for transient response of photonic crystal structures is developed. The system response is obtained by a finite-element time-domain analysis employing perfectly matched layers as an absorbing boundary condition. As an example a waveguide-side-coupled microcavity is designed. The gradient-based optimization technique is applied to redistribute the material inside the microcavity such that the Q factors of a monopole and a dipole mode are improved by 375% and 285%, respectively, while maintaining strong coupling. This is obtained by maximizing the stored energy inside the microcavity in the decaying regime of the transient response. Manufacturable designs are achieved by filtering techniques capable of controlling minimum length scales of the design features.
OCIS codes: 000.4430, 230.5750, 230.7390.

1. INTRODUCTION

In the field of information technology light plays a progressively important role as an information conveyor. One of the keys to future information technology is the realization of large-scale optical integrated circuits. Herein photonic crystals (PhCs) are believed to be the leading platforms. This is due to their unique capability of exhibiting very strong light-matter interaction while keeping the size of the optical components greatly reduced [1,2]. It is the photonic bandgaps of PhCs that yield this interaction in which wave propagation at certain wavelengths is prohibited. The introduction of PhC mirrors, waveguides (WGs), and resonant microcavitites (MCs) as the three basic PhC components has led to an increased attention on the design of optical devices within the last decade that may accomplish the complete control of light propagation. Amalgamation of these components has been utilized to designing various PhC devices such as filters, bends, and splitters for various applications [3].

In the engineering of PhC devices, the use of inverse problem techniques has recently been introduced as potential designing tools to replace previous approaches [4,5]. In this paper we demonstrate how topology optimization based on transient analysis can be used to tweak the performance of a two-dimensional (2D) PhC filtering device, consisting of a PhC-WG-side-coupled PhC-MC. The aim is to achieve strong coupling between the PhC-WG and -MC while maintaining a high Q factor for the filtered mode residing inside the MC.

The methodology behind topology optimization has previously disclosed its usefulness in the design of various 2D PhC-WG components, e.g., a 90° bend [6], a T-junction [7], and a termination [8]. Other relevant and efficient PhC component designs have been obtained and experimentally verified in [9–12]. The common goal for these optimization examples is the maximization of power transmission of either transverse electric (TE-) or transverse magnetic (TM)-polarized waves at multiple frequencies. The optimized components reveal good performance and agreement with experiments. The basis for the computational model is the finite-element frequency-domain method. A consequence of using frequency-domain methods is that wideband optimization requires multiple frequency analyses.

In the present paper the finite-element time-domain (FETD) method will be employed. Time-domain methods have the advantage of computing the response of a linear system at many frequencies with a single time-domain analysis. This idea has been used for antenna design using the finite-difference time-domain method [13], one-dimensional filter and pulse modulator designs [14,15], and simultaneous space-time optimization [16] in the setting of the FETD method. Additionally, time-domain methods can accommodate strongly nonlinear or active (time-varying) media, whereas frequency methods have difficulties with these physical regimes because the frequency is no longer preserved. Two of the major challenges of the FETD method are the computational cost associated with the computation of the sensitivities and the implementation of efficient absorbing boundary conditions (ABCs), such as the perfectly matched layer (PML) [17]. To the authors' knowledge, a topology optimization scheme based on the FETD method using PMLs as ABCs has not been reported before.

Improving Q factors of MCs has previously been studied and subjected to optimization in a 2D PhC slab with a triangular lattice pattern of air holes. These studies reveal that small mode volume and high Q factors are essential in the realization of high-performing active light-emitting devices such as zero-threshold lasers. Standard designs of PhC-MCs exploit simple circular defects. A general recipe relying on an analytically derived inverse

problem approach is developed in [18] restricted to geometries of circular shape. Alternatively, a more general inverse systematic approach based on a level-set approach for shape optimization is suggested in [19] that yields optimized designs of arbitrary convex shapes. Meanwhile, when considering the MC as side-coupled to a WG in filtering devices, it acts as a passive component. Consequently, the requirement for the mode volume becomes secondary, and the interaction between the PhC-MC and the PhC-WG is the primary target [20]. A detailed conceptual study of this interaction can be found in [21,22].

The purpose of this work is to utilize the free material distribution technique provided by topology optimization [23], and thereby not limiting the optimized design to any particular geometrical shape. The optimization method is formulated in the framework of the FETD method for TE-polarized modes, and it uses PMLs as ABCs (Section 2). We express the problem in a fashion that improves the Q factor by maximizing the stored magnetic energy in the exponentially decaying regime of the transient response of PhC-MC monopole and dipole modes. In the optimization process we consider a 2D PhC configuration with a triangular lattice of air holes in the dielectric GaAs (Section 3). The system is homogenous in the third dimension, whereby the total Q factor only depends on the in-plane Q factor. However, the method presented here is believed also to be valid for optimization of the Q factor for equivalent PhC slab devices, in which the out-of-plane Q factor limits the total Q factor. It can also immediately be adapted to TM modes. Finally (Section 4) by using temporal coupled-mode (CM) theory [24], the PhC filter with the optimized PhC-MC is analyzed to verify that the spectral performance of the filter is improved as desired.

2. FORMULATION OF THE TRANSIENT TOPOLOGY OPTIMIZATION METHOD

Throughout this paper we consider propagation of TE modes within 2D infinitely tall PhC structures. The medium inside the structures is assumed to be invariant in time and to occupy a composite of regions of a homogeneous dielectric material as a function of the plane (Cartesian) position vector $\mathbf{r}=(x,y)$ in the solution domain Ω_S. Due to the infinite extension in the third dimension, it is sufficient to solve the scalar wave equation for the transverse component of the magnetic field, $H_z(\mathbf{r},t)$. The numerical solution is sought by truncating Ω_S with a PML region Ω_{PML} as a means to minimize the nonphysical reflection from the boundary $\partial\Omega_S$ (see Fig. 1). Interpreting the material behavior in Ω_{PML} as anisotropic, dispersive, and lossy, the governing equation for $\mathbf{r}\in\Omega=\Omega_S\cup\Omega_{\mathrm{PML}}$ takes the following general form [17]:

$$\mu\mathcal{L}_1(t)H_z+\left[\frac{\partial}{\partial x}\left(\frac{\mathcal{L}_{2,x}(t)}{\varepsilon}\frac{\partial}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\mathcal{L}_{2,y}(t)}{\varepsilon}\frac{\partial}{\partial y}\right)\right]H_z=-\frac{\partial J_{B,z}}{\partial t}, \tag{1}$$

where $\varepsilon=\varepsilon(\mathbf{r})$ and $\mu=\mu(\mathbf{r})$ denote permittivity and permeability, respectively.

On the right-hand side of Eq. (1), $J_{B,z}(\mathbf{r},t)$ is the magnetic charge current, serving here as the excitation term

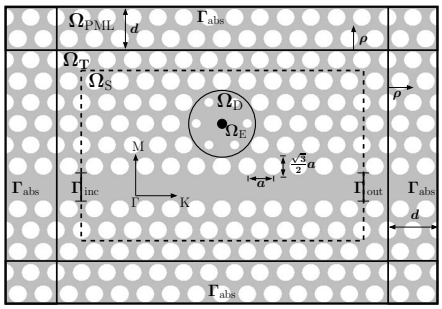


Fig. 1. Illustration of a $\Gamma-K$ directional PhC-WG structure as the computational domain. It consists of the solution domain Ω_S and the transition domain Ω_T that are truncated by the PML region Ω_{PML}. The circle encloses the scattering design region Ω_D. In Ω_E (black circular region) the energy is maximized. The PhC is built by blocks of size $a/2\times\sqrt{3}a/2$, where $\Omega_T\cup\Omega_S$ contains 30 $\times 14$ building blocks, and Ω_{PML} is extended with 24 and 4 blocks on both sides in directions $\Gamma-K$ and $\Gamma-M$, respectively.

in the case of a radiating source residing in Ω_S. The operator $\mathcal{L}_1(t)$ in Eq. (1) is given by

$$\mathcal{L}_1(t)=\frac{\partial^2}{\partial t^2}+\frac{\sigma_x+\sigma_y}{\varepsilon}\frac{\partial}{\partial t}+\frac{\sigma_x\sigma_y}{\varepsilon^2}, \tag{2}$$

and the other operator $\mathcal{L}_{2,p}(t)$ is given by

$$\mathcal{L}_{2,p}(t)=1-a_p\exp[-b_pt]\bar{u}(t)*,\quad p=x,y, \tag{3}$$

with $a_x=(\sigma_x-\sigma_y)/\varepsilon$, $a_y=(\sigma_y-\sigma_x)/\varepsilon$, and $b_p=\sigma_p/\varepsilon$. In Eq. (3), $\bar{u}(t)$ denotes the Heaviside step function, and $*$ stands for temporal convolution. The spatially dependent coordinate-wise conductivities σ_x and σ_y attenuate the field in Ω_{PML} and are thus only nonzero in Ω_{PML}. They are expressed in terms of the PML profile $\sigma(r)$ that is chosen here as [25]

$$\sigma(r)=\sigma_{\max}\left(\frac{\rho}{d}\right)^m,$$

$$\sigma_{\max}=-\frac{(m+1)\log_{10}(R_0)}{2dZ_0}. \tag{4}$$

In Eq. (4), ρ denotes the perpendicular distance from the PML interface, d is the width of the PML, m is the order of the PML profile, R_0 is the theoretical reflection coefficient at normal incidence (typically around the order of 10^{-5}), and Z_0 is the free-space material impedance. The impedance matching condition between Ω_S and Ω_{PML} is satisfied by requiring that the permittivity in Ω_{PML} [cf. Eqs. (2) and (3)] is determined by the value at the PML interface (i.e., when $\rho=0$) [25,26]. For convenience it is designated by $\varepsilon_i=\varepsilon(\rho=0)$.

If the PML is backed with the first-order Silver–Müller radiation boundary condition [27] on $\Gamma_{\mathrm{abs}}\subseteq\partial\Omega$,

$$\mathbf{n}\cdot(\varepsilon^{-1}\tilde{\nabla}H_z)-Y\frac{\partial H_z}{\partial t}=0\quad\text{on }\Gamma_{\mathrm{abs}}, \tag{5}$$

where $\mathbf{n}$ is the outward unit vector normal to the boundary, and $Y=\sqrt{\mu/\varepsilon}$ is the surface admittance, the nonphysical reflections will be further diminished [17]. Here, we have introduced $\tilde{\nabla}=(\mathcal{L}_{2,x}\partial/\partial x,\mathcal{L}_{2,y}\partial/\partial y)$ to abbreviate the

notation. Propagation of plane waves with direction $\boldsymbol{\rho}$ is generated in Ω_S by an incident field $H_z^{\mathrm{inc}}(\mathbf{r},t)$ that is introduced on $\Gamma_{\mathrm{inc}} \subset \partial\Omega_S$ as

$$\mathbf{n}\cdot(\varepsilon^{-1}\tilde{\nabla}H_z) - Y(\mathbf{n}\cdot\boldsymbol{\rho}-1)H_{0,z}^{\mathrm{inc}}\frac{\partial f}{\partial t} = 0 \quad \text{on } \Gamma_{\mathrm{inc}}, \tag{6}$$

where $H_{0,z}^{\mathrm{inc}}$ and $f(t)$, respectively, are the amplitude and the time evolution of the incident field. When the waves originate from a radiating point source in Ω_S, the magnetic charge current becomes $J_{B,z}(\mathbf{r},t)=\delta(\mathbf{r}-\mathbf{r}')g(t)$, where $\delta(\mathbf{r})$ denotes the Dirac delta function, $\mathbf{r}'$ is the location of the point source, and $g(t)$ is the temporal evolution of the source.

A. Finite-Element Implementation

To seek the FETD solution of Eq. (1), we employ Galerkin's method [17]. By introducing an appropriate testing function $T_z(\mathbf{r})$, the weak-form representation becomes

$$\int\int\int_V \left[\mu T_z \mathcal{L}_1(t)H_z(t) + \nabla T_z\cdot\tilde{\nabla}H_z(t) + T_z\frac{\partial J_{B,z}}{\partial t}\right]\mathrm{d}V + \int\int_S \left[Y_c T_z\frac{\partial H_z}{\partial t} + T_z U_z\right]\mathrm{d}S = 0, \tag{7}$$

where $U_z = Y_c(\mathbf{n}\cdot\boldsymbol{\rho}-1)H_{0,z}^{\mathrm{inc}}\partial f/\partial t$. Then, expanding the field as

$$H_z(\mathbf{r},t) = \sum_{i=1}^{N} N_i(\mathbf{r})u_i(t), \tag{8}$$

where N_i is the ith interpolating shape function, and assuming that μ, ε, σ_x, and σ_y are constant within each element yields the corresponding finite-element discretization as

$$\sum_{e=1}^{M}(\mathbf{T}^e\ddot{\mathbf{u}} + \mathbf{R}^e\dot{\mathbf{u}} + \mathbf{S}^e\mathbf{u} + \mathbf{g}^e - \mathbf{f}^e) = 0, \tag{9}$$

where N and M denote the numbers of nodes and elements, respectively; $\dot{(\,)}\equiv \mathrm{d}/\mathrm{d}t$; and $\ddot{(\,)}\equiv \mathrm{d}^2/\mathrm{d}t^2$. The square matrices $\mathbf{T}^e$, $\mathbf{R}^e$, and $\mathbf{S}^e$ are computed by

$$T_{ij}^e = \mu\langle N_i, N_j\rangle_{\Omega^e},$$

$$R_{ij}^e = \mu(\sigma_x+\sigma_y)\varepsilon_i^{-1}\langle N_i, N_j\rangle_{\Omega_{\mathrm{PML}}^e} + Y_c\langle N_i, N_j\rangle_{\Gamma_{\mathrm{ABS}}^e},$$

$$S_{p,ij}^e = \varepsilon^{-1}\langle \partial N_i/\partial p, \partial N_j/\partial p\rangle_{\Omega^e}, \quad p = x,y,$$

$$S_{ij}^e = \mu\sigma_x\sigma_y\varepsilon_i^{-2}\langle N_i, N_j\rangle_{\Omega_{\mathrm{PML}}^e} + S_{x,ij}^e + S_{y,ij}^e, \tag{10}$$

where $\langle\ ,\ \rangle_{\Omega^e}$ and $\langle\ ,\ \rangle_{\Gamma^e}$ mean integration over the volume and surface, respectively, of an element, and $\mathbf{u}=(u_1,\ldots,u_N)^T$. Since we model only the scalar field H_z it is adequate to consider N_i as the nodal based shape function of an element in order to satisfy the field continuity equations. The convolution and the excitation vectors, $\mathbf{g}^e$ and $\mathbf{f}^e$, respectively, are given by individual vector components:

$$g_i^e = \sum_j S_{x,ij}^e\psi_{x,j} + S_{y,ij}^e\psi_{y,j}, \quad e\in\Omega_{\mathrm{PML}},$$

$$f_i^e = -\langle N_i, \partial J_{B,z}/\partial t\rangle_{\Omega_S} + \langle N_i, U_z\rangle_{\Gamma_{\mathrm{inc}}^e}, \tag{11}$$

In Eq. (11) the elements of ψ_p are expressed by

$$\psi_{p,j} = a_p\,\exp[-b_p t]\bar{u}(t) * u_j(t), \quad p = x,y. \tag{12}$$

In Subsection 2.B it is described how the computationally cumbersome convolution term in Eq. (12) is resolved very efficiently to reduce the computational costs.

The PhC in Fig. 1 is reproduced by using building blocks of size $a/2\times\sqrt{3}a/2$, where a is the basic step length of the PhC, i.e., the lattice constant. These building blocks are discretized in 7×12 elements yielding 14 elements per inter-hole spacing in the computational mesh, which corresponds to 14 elements per wavelength in the dielectric material.

B. Time Integration

As a means to speed up the FETD solution in the iterative topology optimization process, we benefit from explicit integration schemes, because they do not require the solution of a matrix system within each time step. Additionally, they have a natural parallelizability. Here, we use a technique that renders the **T**-matrix diagonal when inversion is needed, and in the case of multiplication an averaged $\mathbf{T}_\delta$-matrix is used. The integration of the **T**- and **S**-matrices for a four-node rectangular bilinear element follows modified rules that can be found in [28], such that fourth-order dispersion error accuracy is achieved. A resolution of 14 elements per wavelength is found to be sufficient to obtain acceptable dispersion error. The semidiscrete version of Eq. (9) is given by

$$\mathbf{T}_d\dot{\mathbf{u}}^{n+1/2} = \mathbf{T}_\delta\mathbf{v}^{n-1/2},$$

$$\mathbf{T}_d\dot{\mathbf{v}}^n = -\mathbf{R}\dot{\mathbf{u}}^{n+1/2} - \mathbf{S}\mathbf{u}^n - \mathbf{g}^n + \mathbf{f}^n, \tag{13}$$

where $T_{d,ii}=\Sigma_j T_{ij}$, $T_{d,ij}=0$, and $\mathbf{T}_\delta=(1-\delta)\mathbf{T}_d+\delta\mathbf{T}$. The optimal combination factor is $\delta=(\tau^2-1)/2$ to obtain fourth-order accuracy, where $\tau=\Delta t_c c/\Delta x$, with Δt_c denoting the critical time step. It follows the Courant–Friedrichs–Lewy (CFL)-condition, i.e., $\Delta t_c \le 0.7071\Delta x/c$, where c is the vacuum speed of light. However, we choose $\Delta t = 0.9\Delta t_c$ to avoid or delay numerical instabilities, such as nonphysical oscillations and late-time instabilities in the PML. Half-step approximation is used for central time differences in this scheme, i.e., $\dot{\mathbf{u}}^{n+1/2}=(\mathbf{u}^n-\mathbf{u}^{n-1})/\Delta t$ and $\dot{\mathbf{v}}^n=(\mathbf{v}^{n+1/2}-\mathbf{v}^{n-1/2})/\Delta t$, and it is initiated by $\mathbf{u}^0=\dot{\mathbf{u}}^0=0$ and $\mathbf{v}^{1/2}=\Delta t\mathbf{T}_d^{-1}(\mathbf{f}^0-\mathbf{S}\mathbf{u}^0)/2$. The temporal derivatives (e.g., of the primal response **u** or of the analytical excitation) are evaluated by following the central second difference scheme approximations:

$$\ddot{\mathbf{u}}^n = \frac{\mathbf{u}^{n+1}-2\mathbf{u}^n+\mathbf{u}^{n-1}}{\Delta t^2},$$

$$\dot{\mathbf{u}}^n = \frac{\mathbf{u}^{n+1}-\mathbf{u}^{n-1}}{2\Delta t}. \tag{14}$$

The convolution term in Eq. (12) requires significant computation time and memory since the entire solution history is needed in the convolution integral. However, it can be recursively evaluated at time $n\Delta t$ instead as [17]

$$\psi_{p,j}^{n} = \exp[-b_p \Delta t]\psi_{p,j}^{n-1} + \frac{a_p \Delta t}{2}(u_j^n + \exp[-b_p \Delta t]u_j^{n-1}). \tag{15}$$

Since a_p and b_p are assumed to be constant within each element the recursive convolution relation in Eq. (15) is most easily implemented when the matrix-vector multiplication is carried out element-wise in the time-marching. In practice this means that we assign ψ_p^n separately to each element and update this according to Eq. (15).

C. Design Variables and Material Interpolation

In the design domain $\Omega_D \subseteq \Omega_S$ each finite element is associated with one (density) design variable x_e that varies continuously between $0 \leq x_e \leq 1$. All design variables are assembled into the global design vector $\mathbf{x} = (x_1, \ldots, x_M)^T$. By following the solid isotropic material with penalization (SIMP) scheme [23] the design variable is then used to interpolate between two candidate material phases, designated here as $(\,)^I$ for air and $(\,)^{II}$ for dielectric. Since relative magnetic permeability is very close to unity for dielectric materials, explicit design dependence is only restricted to the inverse of the relative permittivity. This is successfully adopted by a linear interpolation [6],

$$\varepsilon_r^{-1}(x_e) = (1 - x_e)(\varepsilon_r^I)^{-1} + x_e(\varepsilon_r^{II})^{-1}. \tag{16}$$

The continuous design parameterization above now facilitates the use of a gradient-based optimization algorithm (referred to as the *optimizer*) to find an optimized design. The shortcoming of the continuous material parameterization is the possible scenario of intermediate design variables appearing in the final design. In Subsection 2.E it is described how design variables close to discrete ($x_e = 0$ and $x_e = 1$) are efficiently obtained through penalization schemes.

D. Sensitivity Analysis

In our continuous optimization problem we need to compute the design sensitivities in order to use gradient-based optimization solvers. For large numbers of design variables the adjoint sensitivity approach offers a clever and computationally efficient alternative to direct sensitivity analysis [29]. As briefly described, the adjoint method makes the central processing unit time associated with the sensitivity analysis almost independent on the number of design variables by introducing an auxiliary problem (the adjoint problem) that needs to be solved backward in time.

The development of an explicit design sensitivity expression proceeds as follows [30]. Consider a design functional Φ that is defined here as a function of the design $\mathbf{x}$:

$$\Phi(\mathbf{x}) = \int_0^T F(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \mathbf{x}) \mathrm{d}t, \tag{17}$$

where T is the termination time of the transient simulation, and $\mathbf{u} \equiv \mathbf{u}(\mathbf{x}, t)$ where the design dependence is implicit. Now, we continue by expressing the sensitivities in terms of the residual of Eq. (9), given by

$$\mathbf{r}(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \mathbf{x}) = \mathbf{f} - (\mathbf{T}\ddot{\mathbf{u}} + \mathbf{R}\dot{\mathbf{u}} + \mathbf{S}\mathbf{u} + \mathbf{g}) = \mathbf{0}. \tag{18}$$

Combining the integrand in Eqs. (17) and (18) defines the augmented functional

$$\hat{F} = F(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{r}(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \mathbf{x}), \tag{19}$$

where the adjoint operator (Lagrange multiplier) $\boldsymbol{\lambda} \equiv \boldsymbol{\lambda}(\mathbf{x}, t)$ depends implicitly on the design and explicitly on time. The augmented design functional is identical to that in Eq. (17) with F replaced with $\hat{F}$, though, since $\mathbf{r} = \mathbf{0}$. By a clever choice of $\boldsymbol{\lambda}$, which can be chosen freely since $\mathbf{r} = \mathbf{0}$, the sensitivity analysis is significantly simplified as shown in the following.

Now, differentiating Eq. (17) with respect to each component $x_e \in \mathbf{x}$ by the chain rule yields the sensitivities

$$\frac{\partial \Phi}{\partial x_e} = \int_0^T \left(\frac{\partial F}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial x_e} + \frac{\partial F}{\partial \dot{\mathbf{u}}}\frac{\partial \dot{\mathbf{u}}}{\partial x_e} + \frac{\partial F}{\partial \ddot{\mathbf{u}}}\frac{\partial \ddot{\mathbf{u}}}{\partial x_e} + \frac{\partial F}{\partial x_e} + \frac{\partial \boldsymbol{\lambda}^T}{\partial x_e}\mathbf{r} + \boldsymbol{\lambda}^T \left[\frac{\partial \mathbf{r}}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial x_e} + \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}}\frac{\partial \dot{\mathbf{u}}}{\partial x_e} + \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}\frac{\partial \ddot{\mathbf{u}}}{\partial x_e} + \frac{\partial \mathbf{r}}{\partial x_e} \right] \right) \mathrm{d}t. \tag{20}$$

Utilizing integration by parts and that $\mathbf{r} = \mathbf{0}$, Eq. (20) is now rewritten as

$$\frac{\partial \Phi}{\partial x_e} = \left[\boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}}\frac{\partial \mathbf{u}}{x_e} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}\frac{\partial \dot{\mathbf{u}}}{x_e} - \dot{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}\frac{\partial \mathbf{u}}{\partial x_e} \right]_0^T + \int_0^T \left(\frac{\partial F}{\partial \dot{\mathbf{u}}}\frac{\partial \dot{\mathbf{u}}}{\partial x_e} + \frac{\partial F}{\partial \ddot{\mathbf{u}}}\frac{\partial \ddot{\mathbf{u}}}{\partial x_e} + \frac{\partial F}{\partial x_e} \right) \mathrm{d}t + \int_0^T \left(\frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}\ddot{\boldsymbol{\lambda}} - \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}}\dot{\boldsymbol{\lambda}} + \frac{\partial \mathbf{r}}{\partial \mathbf{u}}\boldsymbol{\lambda} - \frac{\partial F}{\partial \mathbf{u}} \right) \frac{\partial \mathbf{u}}{\partial x_e} \mathrm{d}t + \int_0^T \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial x_e} \mathrm{d}t. \tag{21}$$

The implicit system derivatives $\partial \mathbf{u}/\partial x_e$ are eliminated from the sensitivity expression by selecting the appropriate $\boldsymbol{\lambda}$. This process induces the adjoint problem

$$\frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}\ddot{\boldsymbol{\lambda}} - \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}}\dot{\boldsymbol{\lambda}} + \frac{\partial \mathbf{r}}{\partial \mathbf{u}}\boldsymbol{\lambda} = \frac{\partial F}{\partial \mathbf{u}}, \tag{22}$$

where $\partial F/\partial \mathbf{u}$ designates the adjoint load. If we introduce the time shift $t \equiv T - \tau$ for $\tau \in [0, T]$, the adjoint problem in terms of $\boldsymbol{\lambda}(\tau)$ becomes

$$\frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}\ddot{\bar{\boldsymbol{\lambda}}} + \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}}\dot{\bar{\boldsymbol{\lambda}}} + \frac{\partial \mathbf{r}}{\partial \mathbf{u}}\bar{\boldsymbol{\lambda}} = \frac{\partial F}{\partial \mathbf{u}}, \tag{23}$$

since $\partial/\partial t = -\partial/\partial \tau$. Then upon imposing $\mathbf{u}(0) = \dot{\mathbf{u}}(0) = \mathbf{0}$, solving Eq. (23) with the terminal condition $\bar{\boldsymbol{\lambda}}(0) = \dot{\bar{\boldsymbol{\lambda}}}(0) = \mathbf{0}$, and subsequently substituting $\boldsymbol{\lambda}(T - \tau) = \bar{\boldsymbol{\lambda}}(\tau)$, the sensitivity expression reduces to

$$\frac{\partial \Phi}{\partial x_e} = \int_0^T \left(\frac{\partial F}{\partial x_e} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{r}}{\partial x_e} \right) \mathrm{d}t. \tag{24}$$

In our case $\partial F/\partial \dot{\mathbf{u}}$ and $\partial F/\partial \ddot{\mathbf{u}}$ vanish since we only consider problems in which $F = F(\mathbf{u}, \mathbf{x})$. The result is an adjoint problem whose form is identical to that of the primal transient analysis in Eq. (9), albeit with another excita-

tion term. Hence, the same time integration scheme can be used to find the adjoint response.

The evaluation of the adjoint sensitivities proceeds as follows: After the transient primal analysis is concluded, the adjoint response $\boldsymbol{\lambda}$ is computed at the exact same time steps by reusing the exact same ABC setup from the primal analysis [30]. While integrating the adjoint response in time, the contribution to the integral in Eq. (24) is computed, which is plausible since $\boldsymbol{\lambda}$ and $\mathbf{u}$ are already known. Consequently, when integration of the adjoint response is concluded the sensitivities are obtained. This process only requires the storage of the primal response $\mathbf{u}(t)$ from which, when needed, $\ddot{\mathbf{u}}(t)$ and $\dot{\mathbf{u}}(t)$ are computed by Eq. (14).

More often, it is desirable to control the objective in a localized time interval $[T_1,T]\subseteq[0,T]$ (see [31]). The lower integration limit in Eq. (17) cannot simply be replaced with T_1, because then the brackets in Eq. (21) will not vanish. Alternatively, the time integral in Eq. (17) could be split into two integrals $\int_{T_1}^{T}=\int_0^T-\int_0^{T_1}$. However, this requires two adjoint analyses. In order to reduce the computational costs we suggest incorporating a localizing function in the design functional instead. The Heaviside step size function $\bar{u}(t-T_1)$ allows us to specify a time interval, such that the design functional becomes

$$\Phi(\mathbf{x})=\int_0^T F(\mathbf{u},\dot{\mathbf{u}},\ddot{\mathbf{u}},\mathbf{x})\bar{u}(t-T_1)\mathrm{d}t, \tag{25}$$

which modifies the adjoint load to become $[\partial F/\partial\mathbf{u}]\bar{u}(t-T_1)$.

E. Optimization Problem

In the present optimization formulation, the design functional $\Phi(\mathbf{x})$ is restricted to scalar measures. Here, we consider magnetic energy given by $\int_{\Omega_E}\frac{1}{2}\mu H_z^2\mathrm{d}\Omega$. Thus, F can in general be expressed as $\mathbf{u}^T\mathbf{Q}\mathbf{u}$, where $\mathbf{Q}$ is a matrix with the components $Q_{ij}^e=\frac{1}{2}\mu\langle N_i,N_j\rangle_{\Omega^e}$ for $e\in\Omega_E$. In PhC-MC design problems the Q factor can be enhanced by delaying the exponential energy decay [32]. To achieve this, we suggest including the localizing function in F to specify the elapsed time, after which the magnetic energy should be maximized. The optimization problem is now formulated as

$$\max_{\mathbf{x}\in\mathbf{R}^M}\Phi(\mathbf{x})=\log_{10}\left[\int_0^T\mathbf{u}^T\mathbf{Q}\mathbf{u}\bar{u}(t-T_1)\mathrm{d}t\right],$$

$$\text{s.t.:}\quad \text{Governing Eq. (9)}$$

$$\sum_e v_e x_e<Vf^*,\quad 0\le x_e\le 1,\quad e\in\Omega_D. \tag{26}$$

In the second constraint v_e is the element volume, and V is the total volume occupied by Ω_D; hence $V=\Sigma_e v_e$. The constraint serves as a restriction on the available amount of dielectric material, set by f^* herein, that is to be distributed in Ω_D. It is important to stress that it is not necessarily active in this problem. The logarithm in the objective is introduced to ensure better numerical scaling. The optimization process is initiated by a qualified design that is found via the trial and error approach or previously reported designs.

The solution of the optimization problem in Eq. (26) suffers from strong non-uniqueness leading to multiple local minima. Some of them are a result of the possible existence of degenerate modes [3]. Others stem from local resonance effects yielding a poor performance away from the target frequency [7]. Since the physics behind sharp resonance phenomena induces very sensitive behavior upon tiny design changes, we use a globally convergent optimizer. Thereby we stay in the vicinity of the initial design and avoid undesirable minima. Based on an initial design and the adjoint sensitivities, the design update is carried out by the globally convergent method of moving asymptotes that is suggested and implemented in Fortran 77 by Svanberg [33]. The Heaviside step function $\bar{u}(t-T_1)$ in Eq. (26) is regularized in a neighborhood ΔT centered at T_1 by

$$\bar{u}(t-T_1)\approx\frac{1}{2}\frac{\tanh[2\beta(t-T_1)/\delta T]}{\tanh[\beta]}+\frac{1}{2} \tag{27}$$

to avoid numerical problems. The parameter β dictating the curvature of the regularization and the size of δT are chosen such that the fast Fourier transform of $\mathbf{u}^T\mathbf{Q}\mathbf{u}\bar{u}(t-T_1)$ does not disclose any undesirable local resonance peaks away from the target frequency of the mode.

In order to avoid any non-manufacturable features in the structure we use density filtering techniques capable of controlling the minimum length scale of void and dielectric simultaneously [34,35]. However, this particular type of multiphase projection leads to intermediate design variables in the transition region between the material phases. We use the so-called *pamping* method [7] that introduces an artificial mass proportional damping contribution $\mathbf{R}^e=4qx_e(1-x_e)\omega_0\mathbf{T}^e$ and makes the existence of intermediate design variables expensive with respect to the objective. This penalization procedure is, however, only applicable in this form for maximization problems. Numerical experiments have proven $q=0.2(T-T_1)/T$ to be an adequate choice.

The state problem solver, providing the objective and sensitivity evaluation, is parallellized by using Fortran 90 MPI for all interprocessor communications to avoid computational bottlenecks in the storage of the primal response. A non-blocking communication strategy is implemented in the time integration to speed up the simulation time. The design update is carried out on a single processor, to which all necessary data for the optimizer are sent.

3. DESIGN OF A SIDE-COUPLED RESONANT CAVITY

The high index-contrast devices used in the present investigation are constructed by using a 2D PhC composed of a triangular lattice of air holes embedded in the dielectric material GaAs (see Fig. 1). By viewing the structures as infinite in the third dimension and using the radius of the air holes $r/a=0.35$, it supports a complete TE_z bandgap in the normalized frequency range $a/\lambda_0=0.21-0.33$, where

λ_0 denotes the wavelength in vacuum. The dielectric contrast between GaAs and air is 11.4 for wavelengths around 1.5 μm [3].

The side-coupled cavity in the 2D PhC with an infinite height has two distinct loss mechanisms. One concerns the leak of the cavity mode into the $\Gamma-K$ directional WG, and the other concerns that into the surrounding PhC. Hence, the total in-plane Q factor is given by

$$1/Q_{\text{in}} = 1/Q_e + 1/Q_0, \tag{28}$$

where Q_e is the quality factor of the cavity mode with respect to the WG, and Q_0 is the quality factor of the isolated cavity. From Eq. (28) it is deduced that Q_{in} increases when the cavity is far away from the WG, albeit at the expense of poor coupling. In fact, Q_{in} increases exponentially with the distance [20]. Since the ultimate goal is miniaturized integrated PhC circuit devices, the method of topology optimization is utilized here to improve Q_{in} for short distances while maintaining strong coupling. The first target in the design process is to improve Q_0 by considering the isolated cavity.

A. Isolated Cavity Design

Although an arbitrary initial condition could be employed, it is reasonable to search for a cavity geometry supporting a high Q mode by varying the radii of nearest neighboring holes to the cavity. The geometry in Fig. 2(a) that is surrounded by a two cell sizes thick PML pulls down a monopole H_z-mode from the air band at the frequency of $\omega_0 a/2\pi c = 0.3030$ with $Q_0 = 2.2\times10^5$. We calculate the Q factor by measuring the slope of the exponential decay of the energy of the cavity mode [32]

$$U(t) = U_0 \exp[-\omega_0 t/Q_0], \tag{29}$$

where U_0 represents the (initial) amount of energy stored in the cavity that is reached at time $T_{\max}$. The corresponding mode profile in Fig. 3(b) is extracted by the discrete Fourier transform of the response that is excited by a radiation point source in the center of the cavity with time evolution

$$g(t) = \exp[-(t-T_0)^2/\tau^2]\sin[2\pi\omega_0(t-T_0)], \tag{30}$$

where $\tau = 180c/a$ and $T_0 = 500a/c$.

The design region Ω_D for the coupled system depicted in Fig. 1 is also adopted for isolated MC optimization. In the center hereof the objective is evaluated in a circular region Ω_E with radius $3a/14$. We excite the system by reusing the time evolution in Eq. (30), and the simulation time is set to $T = 100,000\Delta t$. The choice of T_1 is not critical here since the isolated cavity has only a single loss mechanism. Additionally, we impose vertical and horizontal symmetry conditions through the center of the cavity and do not constrain the amount of the dielectric material, i.e., $f^* = 1$. The optimized design reached after 311 design iterations appears in Fig. 2(b). As expected it does not include any structural features that violate the minimum allowable length scale prescribed by the density filter with radius $2.5a/14$, and the length scale of the blending region between dielectric and air corresponds to the thickness of one finite element. Based on the normalized logarithmic energy decay in Fig. 3(a) the Q factor of the corresponding mode is computed to be $Q_0 = 4.8\times10^5$. Evidently, the initial design has been altered significantly to achieve this doubling of the Q factor, likewise the mode profile in Fig. 3(c).

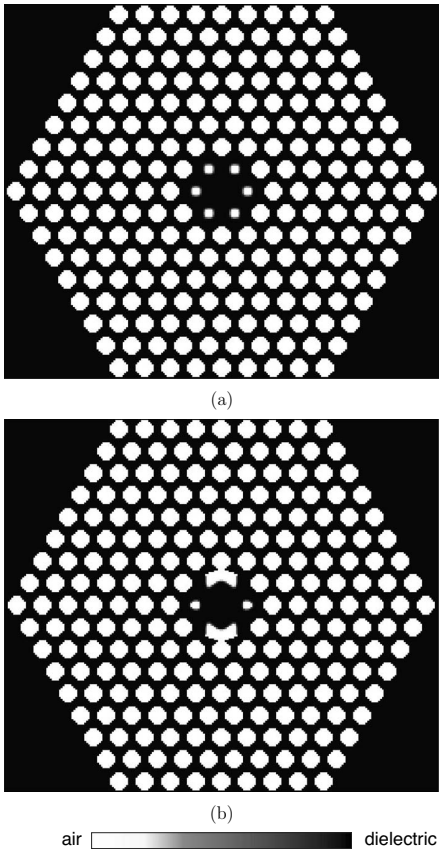


Fig. 2. Monopole mode. (a) Initial MC geometry. (b) Optimized MC geometry.

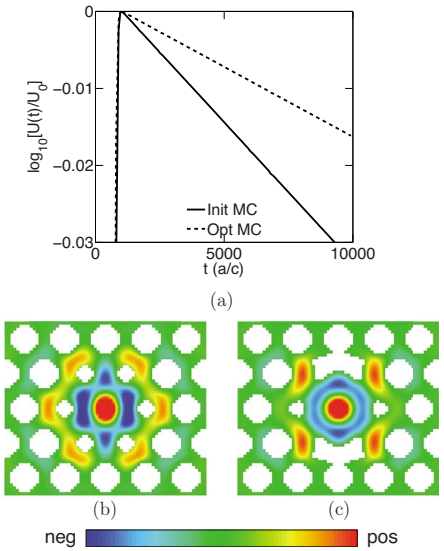


Fig. 3. (Color online) (a) Logarithmic envelope of normalized stored energy for the monopole mode. (b),(c) H_z-field distribution for the initial and optimized MC geometries, respectively. The material distribution is shown with $x_e^t = 0.6$ as threshold.

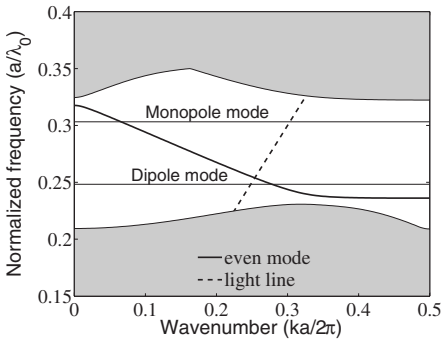


Fig. 4. Dispersion curve of the $\Gamma-K$ directional PhC-WG. The bandgap exists between the dielectric (lower) and air (higher) bands in the normalized frequency range $a/\lambda_0=0.21-0.33$. The horizontal lines represent the dipole mode frequency of 0.2480 and the monopole mode frequency of 0.3030.

In practical applications of the side-coupled MC the initial geometry and corresponding mode cannot be chosen uncritically. The influence of the light line in the dispersion diagram should then be taken into account. Unlike for the monopole mode, we identify a dipole mode at the frequency of $\omega_0 a/2\pi c=0.2480$ whose intersection point with the WG mode is below the light line; see the dispersion diagram in Fig. 4. The corresponding MC geometry appears in Fig. 5(a).

For the dipole mode Ω_D remains unchanged, and the radius of Ω_E is now doubled to $6a/14$. The system is excited by two point sources located in the eye of the dipole

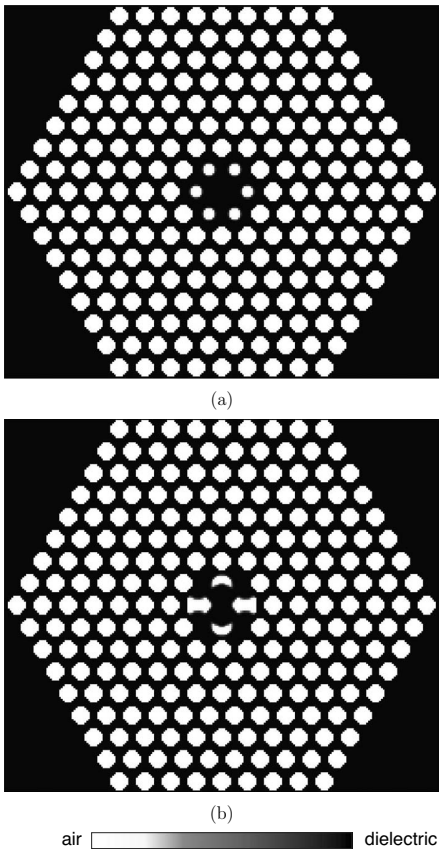


Fig. 5. Dipole mode. (a) Initial MC geometry. (b) Optimized MC geometry.

with opposite signs. Now, however, it has been necessary to constrain the amount of the dielectric material to $f^*=0.8$ to avoid ending up in local maxima bearing degenerate modes. The optimized geometry reached after 592 design iterations is displayed in Fig. 5(b). Compared to the mode in Fig. 6(b) of the initial design with $Q_0=9.9\times10^4$ the Q factor of the optimized design is slightly improved to $Q_0=11\times10^4$ [see Fig. 6(a)]. Similar to the monopole mode, the topological features of the optimized design prevent the cavity field in Fig. 6(c) from leaking into the surrounding PhC above and below the cavity. Instead, the leakage is only concentrated on the corners.

B. Coupled System Design

For the coupled system in Fig. 1 we consider, for the sake of completeness, both cavity modes despite the limited practical applicability of the monopole. Here, we impose only a vertical symmetry condition through the cavity center to increase the design freedom. The system is excited by a line source Γ_{inc} located inside the PhC-WG with an appropriate distance from the cavity and with temporal dependence given by Eq. (30). The entire structure is surrounded by a PML into which the PhC features are continued (see Fig. 1). The design process is initiated by reusing the optimized isolated cavity designs with similar Ω_D and Ω_E, and the simulation time is set to $T=120,000\Delta t$. Since the coupled system now supports multiple loss mechanism, the choice of T_1 becomes crucial in order to improve Q_{in}. Here, the three cases $T_1=0$, $=T_{\max}$, and $=4T_{\max}$ are considered. We compute Q_{in} by utilizing that ω_0/Q equals the full width at half-maximum (FWHM) of the resonant shape of the transmitted power ratio $|T|^2\equiv P_{\text{out}}/P_{\text{in}}$ for the coupled device. Similarly, however, by varying the separation between the PhC-WG and -MC determines Q_e. The power is computed by the Poynting vector at the flux plane Γ_{out}, given by

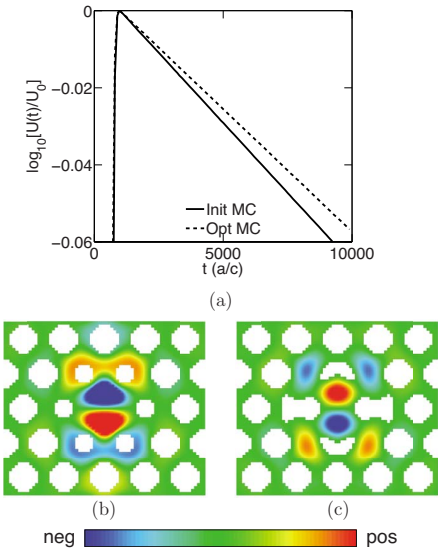


Fig. 6. (Color online) (a) Logarithmic envelope of normalized stored energy for the dipole mode. (b),(c) H_z-field distribution for the initial and optimized MC geometries, respectively. The material distribution is shown with $x_e^t=0.6$ as threshold.

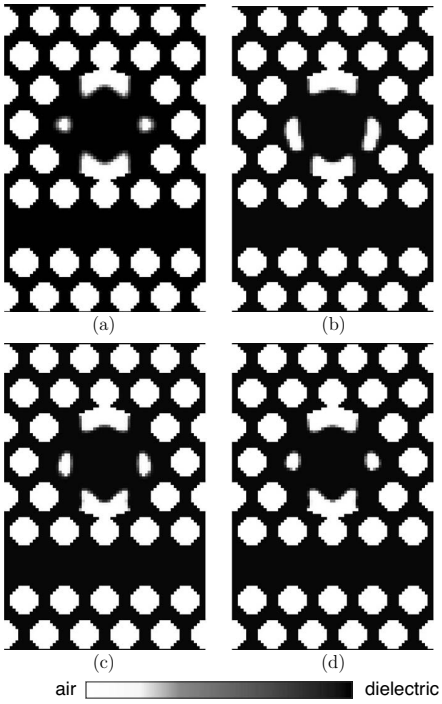


Fig. 7. Optimized designs for the monopole MC mode. (a) Isolated MC. Optimized coupled system geometries for (b) $T_1=0$, (c) $T_1=T_{\max}$, and (d) $T_1=4T_{\max}$.

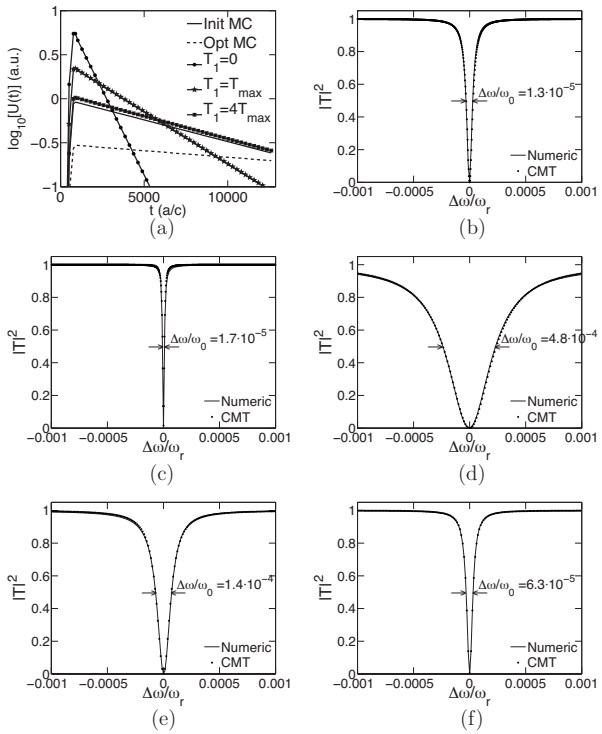


Fig. 8. Monopole mode. (a) Logarithmic envelope of stored energy $U(t)$. Transmission spectrum for (b) initial design, (c) isolated MC, coupled system (d) $T_1=0$, (e) $T_1=T_{\max}$, and (f) $T_1=4T_{\max}$.

$$P(\omega)=\frac{1}{2}\mathrm{Re}\left[\mathbf{n}\cdot\int_{\Gamma_{\mathrm{out}}}\mathbf{E}_\omega^*\times\mathbf{H}_\omega d\Gamma\right], \tag{31}$$

where $(\,)^*$ denotes complex conjugate, and $(\,)_\omega$ means discrete Fourier transform of the time dependent field. By removing the scattering cavity P_{in} can be determined at Γ_{out}. For a broadband analysis, the frequency span of the incoming Gaussian wave packet corresponds to that of the guided mode.

The optimized designs for the monopole and dipole modes follow from Figs. 7 and 9, respectively, and none of those contravene the similar minimum allowable length scale as above. For the monopole mode, the results in Figs. 7(b)–7(d) reveal that the optimization has caused minor redistribution of the dielectric material. According to Figs. 8(b)–8(f) the FWHM of the Lorentzian dip in transmission indicates that Q_{in} is deteriorated compared to the performance of the isolated cavity design. An overview of all relevant Q factors is presented in Table 1. The significance of choosing T_1 is illustrated by the envelope of the stored energy response inside the cavity in Fig. 8(a). When T_1 approaches zero the WG energy attempts to couple into the cavity immediately yielding a strong interaction and a subsequent rapid decay inside the cavity as a consequence of low Q_{in}. As T_1 grows beyond the $T_{\max}$-limit, the energy decay is postponed, which suggests that achieving high Q devices in principle counteracts strong coupling. Nevertheless, in all cases almost zero transmission is achieved independent of the difference in Q_{in}. This behavior can be understood by utilizing CM theory [24] to obtain an analytical expression for the transmission in the WG, given by

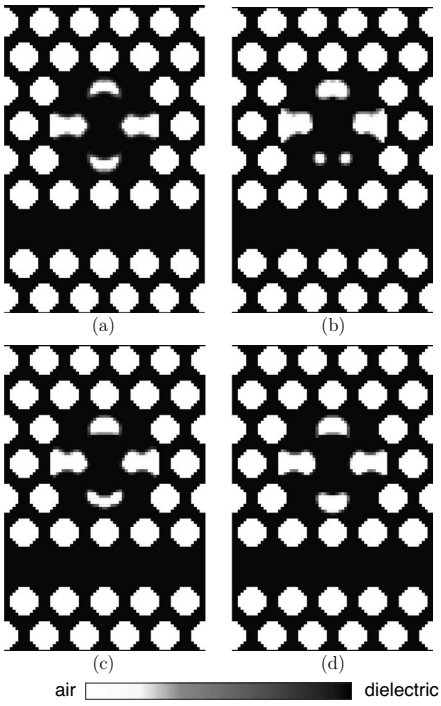


Fig. 9. Optimized designs for the dipole MC mode. (a) Isolated MC. Optimized coupled system geometries for (b) $T_1=0$, (c) $T_1=T_{\max}$, and (d) $T_1=4T_{\max}$.

Table 1. $Q[10^3]$ Factors for Coupled System Configurations

	Monopole			Dipole		
	Q_{in}	Q_e	Q_0	Q_{in}	Q_e	Q_0
Init. MC	16	17.3	220	2.1	2.1	99
Opt. MC	60	68.3	474	4.1	4.2	110
$T_1=0$	2.1	2.1	139	2.0	2.0	90
$T_1=T_{\max}$	7	7.3	168	3.3	3.4	100
$T_1=4T_{\max}$	16	17.2	239	6.0	6.2	140

$$|T(\omega)|^2 = 1 - \frac{\frac{1}{4Q_e^2}\left(1+2\frac{Q_e}{Q_0}\right)}{\left(\frac{\omega-\omega_0}{\omega_0}\right)^2 + \frac{1}{4Q_e}\left(1+\frac{Q_e}{Q_0}\right)^2}. \quad (32)$$

At resonance, i.e., $\omega=\omega_0$, it simplifies to

$$|T(\omega_0)|^2 = \frac{Q_{\text{in}}^2}{Q_0^2} \approx \frac{Q_e^2}{Q_0^2} + O([Q_e/Q_0]^3), \quad (33)$$

where the last approximation is valid if $Q_0 \gg Q_e$. Table 1 indicates that it holds true for the present system, and additionally that Q_e and Q_{in} have the same order of magnitude. Hence at resonance according to Eq. (33), strong coupling is retained even upon considerable improvement of Q_{in} (and Q_e). Consequently, the optimization formulation makes it possible to control the counteracting relation between high Q and strong coupling. One should also expect that Q_0 is largest for the isolated cavity optimization. Table 1 supports this fact for the monopole. However, the dipole mode favors the largest Q_0 for unsymmetrical cavity design in the vertical direction. This is only obtainable for the coupled system ($T_1=4T_{\max}$) as a result of the symmetry conditions in both directions for the isolated cavity optimization.

For the dipole mode the original cavity geometry in Fig. 5(a) and the optimized ones in Fig. 9 exhibit practically zero transmission at resonance [see Figs. 10(b)–10(f)]. The isolated cavity optimization only improves Q_{in} by a factor of 2. In contrast to the monopole, optimizing for $T_1=4T_{\max}$ further increases Q_{in} by 50% as a result of minor design changes [compare Figs. 9(a) and 9(d)], yielding a total improvement of 185%. The design result for $T_1=0$ in Fig. 5(a) shows substantial material redistribution in the interaction region in order to enhance coupling. However, the envelopes of the energy decay in Fig. 10(a) only display marginal changes among the various designs in the coupling performance. As expected the H_z-field plot in Fig. 11 of the best dipole candidate ($T_1=4T_{\max}$) shows nearly zero transmission at resonance.

The use of a 2D PhC of infinite height in this paper omits the influence of the out-of-plane energy losses, which constitutes a limiting factor with regards to improving the total Q factor in PhC slabs. Thus, improving the in-plane Q does not necessarily imply the same trend for the out-of-plane Q [22]. Furthermore, we cannot take the third dimension into account by the effective index method since it holds limited accuracy for high index-contrast structures or near the cutoff. However, our experience from previous studies (cf. [9–12]) is that 2D

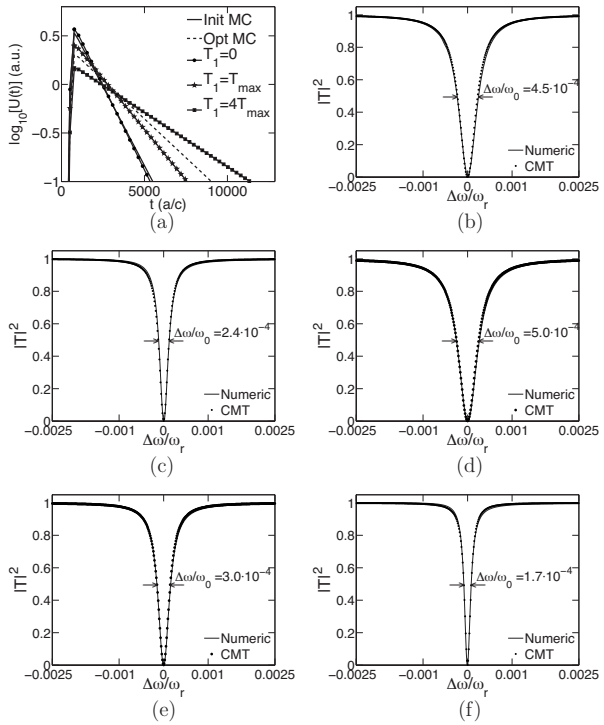


Fig. 10. Dipole mode. (a) Logarithmic envelope of stored energy $U(t)$. Transmission spectrum for (b) initial design, (c) isolated MC, coupled system (d) $T_1=0$, (e) $T_1=T_{\max}$, and (f) $T_1=4T_{\max}$.

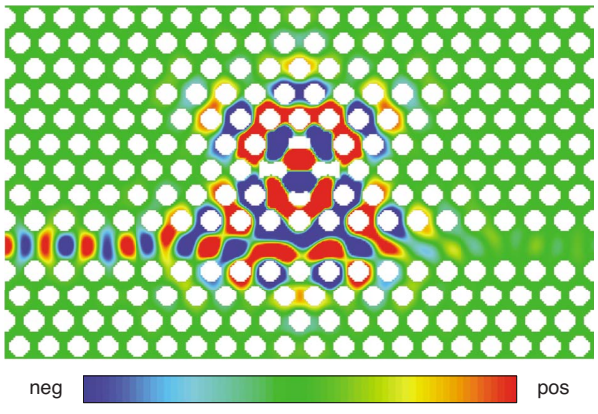


Fig. 11. (Color online) H_z-field distribution for the best optimized WG-side-coupled MC candidate for the dipole mode when $T_1=4T_{\max}$. The material distribution is shown with $x_e^t=0.6$ as threshold.

optimized designs in general yield good behavior in three dimensions as well. In future studies we will verify the obtained designs by three-dimensional simulations and possibly extend the optimization to three dimensions as well.

4. CONCLUSION

In this paper, we have developed a design method based on topology optimization for transient response, and we have used it to design a miniaturized PhC-WG-side-coupled PhC-MC device with an improved Q factor while maintaining strong coupling. Frankly we have shown that the transient optimization formulation makes us capable of controlling the counteracting relation between high Q factor and strong coupling.

The optimization algorithm relies on a 2D FETD model for TE-polarized waves that uses PMLs as absorbing boundary conditions (ABCs), which has not been reported before. To manage the material distribution of air and dielectric we associate a continuously varying design variable to each element in the design domain enclosing the MC. We suggest that the in-plane Q factor can be improved by maximizing the stored cavity energy in the decaying regime of the transient response. Manufacturable designs are achieved by filtering techniques that control the minimum length scale of air and dielectric simultaneously.

The design process is threefold. First, by trial and error we alternate the radii of nearest neighboring holes to the cavity to find a well-performing geometry. This serves as an initial guess in the optimization of the isolated cavity, and if the performance is improved, this is subsequently used to start the optimization of the coupled device. We have optimized the coupling to a monopole and a dipole MC mode, and in both cases we managed to improve the in-plane Q factors by 275% and 185%, respectively, compared to the original design. To study the spectral performance of the optimized designs we have derived an analytical expression for the transmission based on coupled-mode (CM) theory. These agree very well with numerical simulations.

ACKNOWLEDGMENTS

This work is supported by the Eurohorcs/ESF European Young Investigator Award (EURYI) [36] through the grant "Synthesis and topology optimization of optomechanical systems." The Danish Center for Scientific Computing is gratefully acknowledged. The authors are grateful to members of the TopOpt-group [37] for fruitful discussions related to the presented work.

REFERENCES

1. S. John, "Strong localization of photons in certain disordered dielectric superlattices," Phys. Rev. Lett. **58**, 2486–2489 (1987).
2. E. Yablonovitch, "Inhibited spontaneous emission in solid-state physics and electronics," Phys. Rev. Lett. **58**, 2059–2062 (1987).
3. J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light* (Princeton U. Press, 2008).
4. M. Burger, S. J. Osher, and E. Yablonovitch, "Inverse problem techniques for the design of photonic crystals," IEICE Trans. Electron. **E87C**, 258–265 (2004).
5. J. S. Jensen and O. Sigmund have prepared a manuscript to be called "Topology optimization for nano-photonics—a review".
6. J. S. Jensen and O. Sigmund, "Systematic design of photonic crystal structures using topology optimization: Low-loss waveguide bends," Appl. Phys. Lett. **84**, 2022–2024 (2004).
7. J. S. Jensen and O. Sigmund, "Topology optimization of photonic crystal structures: a high-bandwidth low-loss T-junction waveguide," J. Opt. Soc. Am. B **22**, 1191–1198 (2005).
8. W. R. Frei, D. A. Tortorelli, and H. T. Johnson, "Topology optimization of a photonic crystal waveguide termination to maximize directional emission," Appl. Phys. Lett. **86**, 111114 (2005).
9. P. I. Borel, A. Harpoth, L. H. Frandsen, M. Kristensen, P. Shi, J. S. Jensen, and O. Sigmund, "Topology optimization and fabrication of photonic crystal structures," Opt. Express **12**, 1996–2001 (2004).
10. L. H. Frandsen, A. Harpoth, P. I. Borel, M. Kristensen, J. S. Jensen, and O. Sigmund, "Broadband photonic crystal waveguide 60 degrees bend obtained utilizing topology optimization," Opt. Express **12**, 5916–5921 (2004).
11. P. I. Borel, L. H. Frandsen, A. Harpoth, M. Kristensen, J. S. Jensen, and O. Sigmund, "Topology optimised broadband photonic crystal Y-splitter," Electron. Lett. **41**, 69–71 (2005).
12. J. S. Jensen, O. Sigmund, L. H. Frandsen, P. I. Borel, A. Harpoth, and M. Kristensen, "Topology design and fabrication of an efficient double 90(circle) photonic crystal waveguide bend," IEEE Photon. Technol. Lett. **17**, 1202–1204 (2005).
13. 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," Int. J. Numer. Methods Eng. **71**, 1261–1296 (2007).
14. J. Dahl, J. S. Jensen, and O. Sigmund, "Topology optimization for transient wave propagation problems in one dimension design of filters and pulse modulators," Struct. Multidiscip. Optim. **36**, 585–595 (2008).
15. L. R. Yang, A. V. Lavrinenko, J. M. Hvam, and O. Sigmund, "Design of one-dimensional optical pulse-shaping filters by time-domain topology optimization," Appl. Phys. Lett. **95**, 261101 (2009).
16. J. S. Jensen, "Space-time topology optimization for one-dimensional wave propagation," Int. J. Numer. Methods Eng. **198**, 705–715 (2009).
17. J. Jin and D. J. Riley, *Finite Element Analysis of Antennas and Arrays* (Wiley, 2007).
18. D. Englund, I. Fushman, and J. Vuckovic, "General recipe for designing photonic crystal cavities," Opt. Express **13**, 5961–5975 (2005).
19. W. R. Frei, H. T. Johnson, and K. D. Choquette, "Optimization of a single defect photonic crystal laser cavity," J. Appl. Phys. **103**, 033102 (2008).
20. Y. Akahane, M. Mochizuki, T. Asano, Y. Tanaka, and S. Noda, "Design of a channel drop filter by using a donor-type cavity with high-quality factor in a two-dimensional photonic crystal slab," Appl. Phys. Lett. **82**, 1341–1343 (2003).
21. C. J. M. Smith, R. M. De la Rue, M. Rattier, S. Olivier, H. Benisty, C. Weisbuch, T. F. Krauss, R. Houdre, and U. Oesterle, "Coupled guide and cavity in a two-dimensional photonic crystal," Appl. Phys. Lett. **78**, 1487–1489 (2001).
22. G. H. Kim, Y. H. Lee, A. Shinya, and M. Notomi, "Coupling of small, low-loss hexapole mode with photonic crystal slab waveguide mode," Opt. Express **12**, 6624–6631 (2004).
23. M. P. Bendsøe and O. Sigmund, *Topology Optimization: Theory, Methods, and Applications*, 2nd ed. (Springer Verlag, 2004).
24. H. A. Haus, *Waves and Fields in Optoelectronics* (Prentice-Hall, 1984).

25. A. F. Oskooi, L. Zhang, Y. Avniel, and S. G. Johnson, "The failure of perfectly matched layers, and towards their redemption by adiabatic absorbers," Opt. Express **16**, 11376–11392 (2008).
26. M. Koshiba, Y. Tsuji, and S. Sasaki, "High-performance absorbing boundary conditions for photonic crystal waveguide simulations," IEEE Microw. Wirel. Compon. Lett. **11**, 152–154 (2001).
27. S. Silver, *Microwave Antenna Theory and Design* (McGraw-Hill, 1949).
28. B. Yue and M. N. Guddati, "Dispersion-reducing finite elements for transient acoustics," J. Acoust. Soc. Am. **118**, 2132–2141 (2005).
29. D. A. Tortorelli and P. Michaleris, "Design sensitivity analysis: overview and review," Inverse Probl. Eng. **1**, 71–105 (1994).
30. Y. S. Chung, C. Cheon, I. H. Park, and S. Y. Hahn, "Optimal shape design of microwave device using FDTD and design sensitivity analysis," IEEE Trans. Microwave Theory Tech. **48**, 2289–2296 (2000).
31. D. A. Tortorelli and R. B. Haber, "First-order design sensitivities for transient conduction problems by an adjoint method," Int. J. Numer. Methods Eng. **28**, 733–752 (1989).
32. J. D. Jackson, *Classical Electrodynamics* (Wiley, 1999).
33. K. Svanberg, "A class of globally convergent optimization methods based on conservative convex separable approximations," SIAM J. Optim. **12**, 555–573 (2002).
34. O. Sigmund, "Morphology-based black and white filters for topology optimization," Struct. Multidiscip. Optim. **33**, 401–424 (2007).
35. J. K. Guest, "Topology optimization with multiple phase projection," Comput. Methods Appl. Mech. Eng. **199**, 123–135 (2009).
36. http://www.esf.org/euryi.
37. http://www.topopt.dtu.dk.

[P3]

Struct Multidisc Optim
DOI 10.1007/s00158-011-0642-y

RESEARCH PAPER

Topology optimization of pulse shaping filters using the Hilbert transform envelope extraction

Boyan Stefanov Lazarov · René Matzen · Yuriy Elesin

Received: 20 December 2010 / Accepted: 20 February 2011

Abstract Time domain topology optimization is applied to design pulse shaping filters. The objective function depends on the pulse envelope, which is extracted by utilizing the Hilbert transform. The gradients with respect to the topology optimization variables are derived, and the optimization methodology is demonstrated for pulse delaying and pulse splitting. The formulation is applicable for nonlinear structures and signals consisting of broad range of frequencies.

Keywords Topology optimization · Hilbert transform · Envelope extraction

1 Introduction

The focus in this article is on the topology optimization of waveguides where the objective function depends on the envelope of the output signal. Such optimized designs as well as the methodology for obtaining them, can be utilized in a wide range of physical areas—in control of electromagnetic waves in optical waveguides and design of logical elements for photonic processors, in coupled opto-mechanical and acoustic sensors, for sound and vibration isolation, inverse problems in seismology, biomedical imaging and structural health monitoring, etc. The structures are obtained by using alternating layers of different materials in 1D or by distributing several materials in space for multidimensional problems. Periodic structures possess filtering properties and have been investigated by many authors using analytical, numerical and experimental methods (e.g. Brillouin 1953; Mead 1975; Kushwaha et al. 1993). Waves attenuate along these structures within specific bands of frequencies called stop bands or band gaps. The periodic structures act as a filter for frequencies inside the band gap.

Simple filter designs (e.g. band-gaps) can be easily obtained by utilizing analytical techniques, however the design of more complicated filters, like pulse shaping or pulse delaying devices require utilization of numerical techniques such as topology optimization. Topology optimization is an iterative optimization procedure, which produces optimal material distribution. The design domain is split into multiple cells and a design variable is assigned to each of them. The design variables take value 0 if the cell is empty, or 1 if the cell is filled with material. In order to make the optimization problem solvable by using gradient based optimization techniques, the design variables are allowed to vary continuously between 0 and 1. In each optimization iteration the design variables are updated based on fulfillment of prescribed constraints and minimization of a given objective. Initially topology optimization has been developed for mechanical systems (Bendsøe and Sigmund 2004) and later its field of applicability has been extended to include applications in electromagnetics—antenna design and photonics (Sigmund and Jensen 2003; Jensen and Sigmund 2004, 2011; Nomura et al. 2007).

Topology optimization has been applied to wave propagation problems modeled in the frequency domain (Sigmund and Jensen 2003; Jensen and Sigmund 2005)

B. S. Lazarov (✉) · R. Matzen · Y. Elesin
Department of Mechanical Engineering, Solid Mechanics, Technical University of Denmark, Nils Koppels Alle, B. 404, 2800 Lyngby, Denmark
e-mail: bsl@mek.dtu.dk

R. Matzen
e-mail: rmat@mek.dtu.dk

Y. Elesin
e-mail: yuel@mek.dtu.dk

Published online: 01 April 2011

and in the time domain (Nomura et al. 2007; Dahl et al. 2008; Matzen et al. 2010). Time domain topology optimization has been introduced relatively recently and the main advantage compared to the frequency domain topology optimization is lower computational cost for wave problems consisting of broad range of frequencies. For non-linear material behavior where the energy can be exchanged between different frequency components, models in the frequency domain can represent the system behavior approximately for weak interaction and stationary response (Jensen 2011). In contrast, time domain methods can model strongly non-linear or active (time-controllable) materials without any limitations.

Practical applications of the optimized wave guides and filters require the utilization of signal generators and detectors. The generation or the detection of short signals is an unreliable process, e.g. see the discussion in Yang et al. (2009), and in practical applications a signal is composed of one or several modulated high frequency waves. In digital signal processing, the actual information is carried by the signal envelope. A way to extract the envelope information in topology optimization for single frequency modulated wave is reported in Yang et al. (2009). Here a more robust alternative based on the Hilbert transform (e.g. Cohen 1995) is introduced and demonstrated for pulse shaping and delaying devices. The Hilbert transform envelope extraction can be utilized for a broader range of problems: waves with multiple frequencies (i.e. pulses) and non-linear wave guides and filters.

2 Time domain topology optimization

The considered physical problem is wave propagation in a wave guide. The model setup is shown in Fig. 1. Pulses are propagating from left to right and the aim is to find material distribution in the middle region, i.e. the design domain, which minimizes given objective. The objective depends on the pulse envelope. The envelope of the signal is obtained by using the Hilbert transform as

$$s(t) = \sqrt{u(t)^2 + \hat{u}(t)^2} \tag{1}$$

where $u(t)$ is the signal, $\hat{u}(t)$ is the Hilbert transform and $s(t)$ is the signal envelope. The Hilbert transform is defined in the time domain as convolution between the Hilbert transformer $1/(\pi t)$ and a function $u(t)$. Examples of signals, their Hilbert transforms and the envelopes are shown in Fig. 2. The transformed signal is obtained from the original one by using $\pi/2$ phase shift. For both examples the envelope is smooth and slowly varying function. The envelope extraction is applicable for signal centered around a single frequency Fig. 2a, as well as for signal consisting of a broad range of frequencies Fig. 2b. More details about Hilbert transform can be found in most textbooks on signal processing (e.g. Cohen 1995) and a short introduction can be found in Appendix A.

2.1 Wave equation

The physical behavior of the system can be described by the following scalar wave equation

$$\frac{\partial}{\partial x}\left(A\frac{\partial \Psi}{\partial x}\right) - B\frac{\partial^2 \Psi}{\partial t^2} = 0, \quad x \in \Omega \tag{2}$$

where Ψ is a scalar quantity which can represent:

1. $u(x,t)$—displacements for elastic wave propagation
2. $p(x,t)$—pressure for acoustic problems
3. $e(x,t)$—electric field in electromagnetics

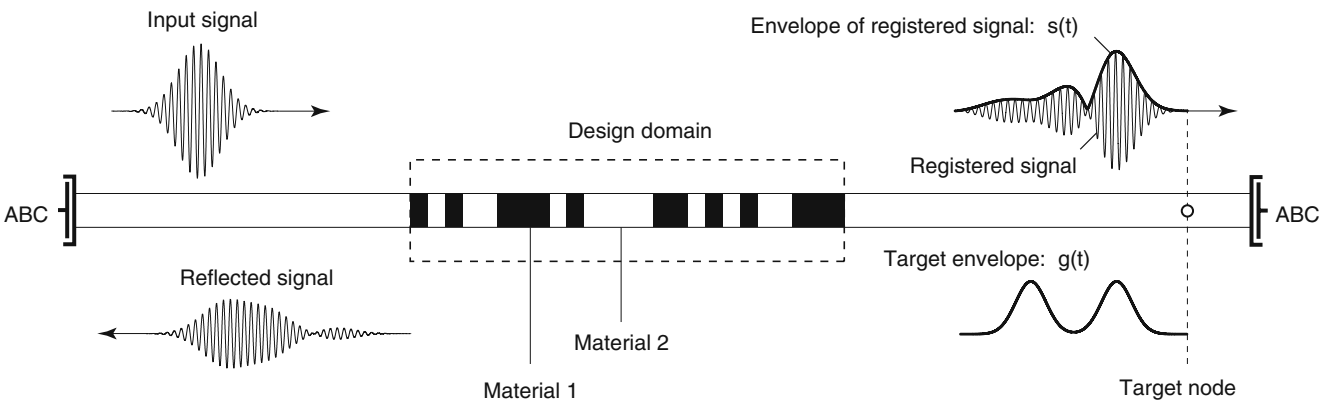


Fig. 1 Numerical model setup for the optimization experiments. Absorbing boundary condition (ABC) is applied in both ends of the modeling domain. Incident signal is generated in the *left* boundary using the incident-scattered field formulation. The signal and its envelope are extracted in the target node on the *right* side of the design domain

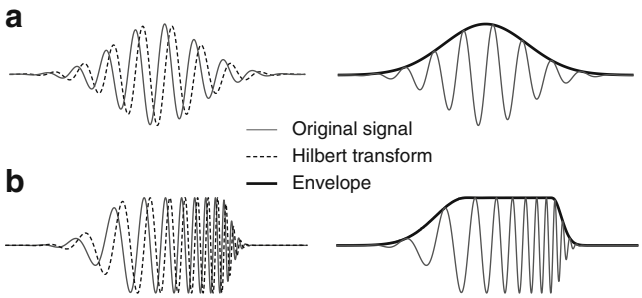


Fig. 2 Pulse envelope obtained by using Hilbert transform

$\Psi(x,t)$ is a function of both the time t and the spatial coordinate $x \in \Omega$. The material parameters A and B are equal to:

1. $A = E(x)$ the Young's modulus and $B = \rho(x)$ the mass density for linear elasticity
2. $A = 1/\rho(x)$ the inverse of the mass density and $B = 1/\kappa(x)$ the inverse of the bulk modulus for linear acoustics
3. $A = 1/\mu(x)$ the inverse of the material permeability and $B = \epsilon(x)$ the permittivity in electromagnetics

In order to model an infinite wave guide, first order absorbing boundary conditions are applied at both ends of the modeling domain. The wave field can be split in two parts—scattered and incident field.

$$\Psi = \Psi^{\text{sc}} + \Psi^{\text{inc}} \tag{3}$$

The boundary condition at the right end is given as

$$\frac{\partial \Psi^{\text{sc}}}{\partial x} + c\frac{\partial \Psi^{\text{sc}}}{\partial t} = 0 \tag{4}$$

and at the left end can be written as

$$\frac{\partial \Psi}{\partial x} + c\frac{\partial \Psi}{\partial t} = \frac{\partial \Psi^{\text{inc}}}{\partial x} + c\frac{\partial \Psi^{\text{inc}}}{\partial t} \tag{5}$$

The value of Ψ^{inc} at the left end of the modeling domain is known a priori. The coefficient $c = \sqrt{AB}$ is the so-called surface admittance. For elastic wave propagation the two boundary conditions (4) and (5) can be modeled as two dash-pots with damping coefficient $c = \sqrt{E\rho}$. It should be pointed out that for high dimensional problems in 2D and 3D the above boundary condition cannot absorb the outgoing waves. In these cases an alternative such as perfectly matched layers (PML) (e.g. Jin and Riley 2007; Matzen et al. 2010) can be utilized for avoid the boundary reflections.

2.2 Discretization

The computational domain is discretized into finite elements. A design variable $\rho_i \in [0,1]$ is associated with each element located in the design domain (see Fig. 1). By applying the standard Galerkin discretization procedure (Zienkiewicz et al. 2005), the discrete form of equations (2), (4) and (5) becomes

$$\mathbf{M}(\boldsymbol{\rho})\,\ddot{\mathbf{u}} + \mathbf{C}(\boldsymbol{\rho})\,\dot{\mathbf{u}} + \mathbf{K}(\boldsymbol{\rho})\,\mathbf{u} = \mathbf{f}(t), \quad t \in [0,T] \tag{6}$$

where $\mathbf{u}$ is the nodal displacement vector, $\dot{(\cdot)}$ and $\ddot{(\cdot)}$ denote first and second derivative with respect to time, $\mathbf{f}(t)$ is a vector with the system input. In linear elasticity $\mathbf{M}(\boldsymbol{\rho})$, $\mathbf{C}(\boldsymbol{\rho})$ and $\mathbf{K}(\boldsymbol{\rho})$ are the standard mass, damping and stiffness matrices, respectively. The damping matrix is populated only with entries coming from the boundary conditions, since the wave guide is considered to be lossless. The vector $\boldsymbol{\rho}$ consists of all design variables associated with the elements in the design domain.

The material properties for each design element are obtained by using the following linear interpolation[1]

$$\begin{aligned} A_e &= (1-\rho_e)\,A^I + \rho_e A^{II} \\ B_e &= (1-\rho_e)\,B^I + \rho_e B^{II} \end{aligned} \tag{7}$$

where (A^I, B^I) and (A^{II}, B^{II}) represent the material properties for material I and II, respectively, and ρ_e is the design variable associated with element e. The above interpolation scheme makes it possible to apply a gradient based optimization algorithm. The main drawback is the appearance of intermediate (gray) values in the final design (between 0 and 1), which are difficult to interpret. Several techniques have been developed to ensure black and white designs (Bendsøe and Sigmund 2004; Yang et al. 2009; Wang et al. 2011). Obtaining length scale and crisp black and white designs for wave problems is a challenging problem deserving a more detailed study, which is outside of the scope of this article.

2.3 Optimization problem

The focus in this section is to define a pulse shaping and pulse delaying strategy based on a signal envelope extraction. The envelope $s(t)$ in the time domain is defined by

[1] The interpolation scheme corresponds to the Solid Isotropic Material with Penalization (SIMP) scheme, where the penalization parameter is set to 1. SIMP with $p \geq 1$ is effective for suppressing intermediate(gray) values in the design, in the cases when an active volume constraint is introduced in the optimization problem.

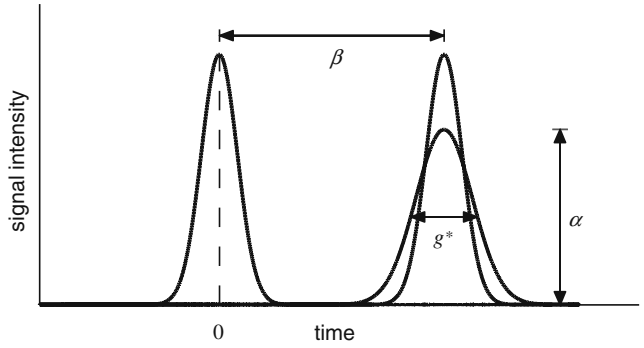


Fig. 3 Pulses with a group delay 0 and β. The parameter β is introduced as an design variable. The g^*-parameter controls the temporal spreading of the output pulse envelope. The α-parameter mainly specify the transmitted energy, i.e. the amplitude decrease

(1). The aim of the optimization is to obtain a signal at the registration point (Fig. 1), with an envelope following a prescribed function $g\,(t)$. The objective is defined as

$$f = \frac{\int_0^T \left[s\,(t)^2 - g\,(t)^2\right]^2 \mathrm{d}t}{\int_0^T g\,(t)^4\,\mathrm{d}t} \tag{8}$$

The term $s\,(t)$ is squared in the above equation in order to avoid special handling of the case when $u\,(t) = 0$ and $\hat{u}\,(t) = 0$. The objective is normalized with respect to the pulse area and obtained results are reported in percentages. The pulse shaping optimization problem can be written in the following form

$$\begin{aligned} &\min_{\rho} : f\,(\boldsymbol{\rho}) \\ &s.t. : \mathbf{M}\,(\boldsymbol{\rho})\,\ddot{\mathbf{u}} + \mathbf{C}\,(\boldsymbol{\rho})\,\dot{\mathbf{u}} + \mathbf{K}\,(\boldsymbol{\rho})\,\mathbf{u} = \mathbf{f}\,(t)\,, \quad t \in [0, T] \\ &\quad 0 \le \boldsymbol{\rho} \le 1 \end{aligned} \tag{9}$$

where the objective is given by (8). Due to the dispersive properties of the optimized structure, the different frequency components of a pulse traveling through the device arrive at the registration point shifted with respect to each other. Therefore, an initially Gaussian shaped pulse changes its shape and amplitude through the optimization domain (Fig. 3). The distortion and the amplitude decrease become significant when the objective is to delay the pulse. In order to account for these changes the pulse delay optimization problem is defined as

$$\begin{aligned} &\min_{\rho} : -\beta \\ &s.t. : \mathbf{M}\,(\boldsymbol{\rho})\,\ddot{\mathbf{u}} + \mathbf{C}\,(\boldsymbol{\rho})\,\dot{\mathbf{u}} + \mathbf{K}\,(\boldsymbol{\rho})\,\mathbf{u} = \mathbf{f}\,(t)\,, \quad t \in [0, T] \\ &\quad g_1\,(\boldsymbol{\rho}, \beta) \le 0 \\ &\quad 0 \le \boldsymbol{\rho} \le 1 \end{aligned} \tag{10}$$

where β is the pulse delay and the constraint function $g_1\,(\rho, \beta)$ is given as follows

$$g_1\,(\rho, \beta) = \frac{\int_0^T \left[s\,(t)^2 - \alpha g\,(t-\beta)^2\right]^2 \mathrm{d}t}{g^* \int_0^T \alpha^2 g\,(t)^4\,\mathrm{d}t} - 1 \tag{11}$$

The parameter α controls the amplitude scaling and the parameter g^* controls mainly the envelope distortion (spreading). The meaning of the parameters is shown in Fig. 3.

The first formulation can be used to find an optimized structure which delays the input pulse with a prescribed fixed value by minimizing the difference between the output signal and prescribed delayed signal. The second formulation tries to find the maximum delay so that the constraint (11) is fulfilled.

2.4 Sensitivity analysis

Sensitivities of the objective and the constraints with respect to all design variables are needed in order to solve the optimization problems. For large numbers of design variables the adjoint sensitivity approach offers a computationally effective alternative to direct sensitivity analysis (Tortorelli and Michaleris 1994; Kang et al. 2006). The objective or the constraint function is augmented with product of Lagrangian multiplier vector $\boldsymbol{\lambda}\,(t)$ and the residual of the discretized state problem

$$\psi = \int_0^T G\left(\hat{\mathbf{u}}, \mathbf{u}, \boldsymbol{\rho}, t\right) + \boldsymbol{\lambda}^{\mathsf{T}}\,(t)\,\mathbf{r}\,(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \boldsymbol{\rho}, t)\,\mathrm{d}t \tag{12}$$

where

$$\mathbf{r}\,(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}, \boldsymbol{\rho}) = \mathbf{f} - (\mathbf{M}\,(\boldsymbol{\rho})\,\ddot{\mathbf{u}} + \mathbf{C}\,(\boldsymbol{\rho})\,\dot{\mathbf{u}} + \mathbf{K}\,(\boldsymbol{\rho})\,\mathbf{u}) \tag{13}$$

The function $G\left(\hat{\mathbf{u}}, \mathbf{u}, \rho, t\right)$ in (12) represents the integrand in the original objective or constraint function. For zero residual the augmented function coincides with the original one. Derivative of the objective function is obtained by differentiating (12) with respect to the design variables

$$\begin{aligned} \frac{\partial \psi}{\partial \rho_e} = \int_0^T &\frac{\partial G}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \rho_e} + \frac{\partial G}{\partial \hat{\mathbf{u}}}\frac{\partial \hat{\mathbf{u}}}{\partial \rho_e} + \frac{\partial G}{\partial \rho_e} + \frac{\partial \boldsymbol{\lambda}^{\mathsf{T}}}{\partial \rho_e}\mathbf{r} \\ &+ \boldsymbol{\lambda}^{\mathsf{T}}\left[\frac{\partial \mathbf{r}}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \rho_e} + \frac{\partial \mathbf{r}}{\partial \dot{\mathbf{u}}}\frac{\partial \dot{\mathbf{u}}}{\partial \rho_e} + \frac{\partial \mathbf{r}}{\partial \ddot{\mathbf{u}}}\frac{\partial \ddot{\mathbf{u}}}{\partial \rho_e} + \frac{\partial \mathbf{r}}{\partial \rho_e}\right]\mathrm{d}t \end{aligned} \tag{14}$$

Integrating by parts and rearranging the above equation yields

$$\begin{aligned} \frac{\partial \psi}{\partial \rho_e} = &\left[-\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{C}\frac{\partial \mathbf{u}}{\partial \rho_e} - \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{M}\frac{\partial \dot{\mathbf{u}}}{\partial \rho_e} + \dot{\boldsymbol{\lambda}}^{\mathsf{T}}\mathbf{M}\frac{\partial \mathbf{u}}{\partial \rho_e}\right]_0^T \\ &- \int_0^T \left[\ddot{\boldsymbol{\lambda}}^{\mathsf{T}}\mathbf{M}\frac{\partial u}{\partial \rho_e} - \dot{\boldsymbol{\lambda}}^{\mathsf{T}}\mathbf{C}\frac{\partial \mathbf{u}}{\partial \rho_e} + \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{K}\frac{\partial u}{\partial \rho_e}\right]\mathrm{d}t \end{aligned}$$

$$+\int_0^T \left[\frac{\partial G}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \rho_e}+\frac{\partial G}{\partial \hat{\mathbf{u}}}\frac{\partial \hat{\mathbf{u}}}{\partial \rho_e}\right]\mathrm{d}t$$

$$+\int_0^T \frac{\partial G}{\partial \rho_e}+\boldsymbol{\lambda}^{\mathsf{T}}\left[-\frac{\partial \mathbf{M}}{\partial \rho_e}\ddot{\mathbf{u}}-\frac{\partial \mathbf{C}}{\partial \rho_e}\dot{\mathbf{u}}-\frac{\partial \mathbf{K}}{\partial \rho_e}\mathbf{u}+\frac{\partial f}{\partial \rho_e}\right]\mathrm{d}t \tag{15}$$

The above equation holds for arbitrary $\boldsymbol{\lambda}$, therefore the Lagrangian multipliers can be chosen to eliminate the second and the third integral in (15). Using the identity

$$\int_0^T \frac{\partial G}{\partial \hat{\mathbf{u}}}\frac{\partial \hat{\mathbf{u}}}{\partial \rho_e}\mathrm{d}t=\int_0^T -\frac{\widehat{\partial G}}{\partial \hat{\mathbf{u}}}\frac{\partial \mathbf{u}}{\partial \rho_e}\mathrm{d}t \tag{16}$$

and collecting all terms with $\frac{\partial \mathbf{u}}{\partial \rho_e}$, the following adjoint equation is obtained from the above requirement

$$\ddot{\boldsymbol{\lambda}}^{\mathsf{T}}\mathbf{M}-\dot{\boldsymbol{\lambda}}^{\mathsf{T}}\mathbf{C}+\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{K}=\frac{\partial G}{\partial \mathbf{u}}-\frac{\widehat{\partial G}}{\partial \hat{\mathbf{u}}} \tag{17}$$

The term $\frac{\widehat{\partial G}}{\partial \hat{\mathbf{u}}}$ denotes the Hilbert transform of $\frac{\partial G}{\partial \hat{\mathbf{u}}}$. The evaluation of the right hand side of (17) is performed in several steps as follows:

1. the Hilbert transform $\hat{\mathbf{u}}$ is obtained from the response $\mathbf{u}$
2. the envelope $s\,(t)$ is computed using (1)
3. the derivatives $\frac{\partial G}{\partial \hat{\mathbf{u}}}$ and $\frac{\partial G}{\partial \mathbf{u}}$ are computed as functions of time
4. the Hilbert transform $\frac{\widehat{\partial G}}{\partial \hat{\mathbf{u}}}$ is computed using $\frac{\partial G}{\partial \hat{\mathbf{u}}}$ evaluated in the previous step

The proof of the identity given by (16) is presented in Appendix B.

The first term in (15) is evaluated at time $t=0$ and $t=T$. As the initial velocities and displacements do not depend on the material distribution, the term vanishes at $t=0$. To eliminate the remaining part of the term at time $t=T$, the terminal conditions for the adjoint equation (17) are set to be

$$\boldsymbol{\lambda}\,(T)=0,\quad \dot{\boldsymbol{\lambda}}\,(T)=0 \tag{18}$$

Utilizing the solution of (17) and (18), the sensitivity expression reduces to

$$\frac{\partial \psi}{\partial \rho_e}=\int_0^T \frac{\partial G}{\partial \rho_e}+\boldsymbol{\lambda}^{\mathsf{T}}\left[-\frac{\partial \mathbf{M}}{\partial \rho_e}\ddot{\mathbf{u}}-\frac{\partial \mathbf{C}}{\partial \rho_e}\dot{\mathbf{u}}-\frac{\partial \mathbf{K}}{\partial \rho_e}\mathbf{u}+\frac{\partial f}{\partial \rho_e}\right]\mathrm{d}t \tag{19}$$

which consists of only explicit derivatives with respect to the design variables. They are computed based on the numerical discretization and the material interpolation. Usually the first term in the integral $\frac{\partial G}{\partial \rho_e}$ is zero.

The terminal value problem (17) can be converted to initial value problem by substituting t with $t=T-\tau$, where $\tau\in[0,T]$. Setting $\bar{\boldsymbol{\lambda}}\,(\tau)=\boldsymbol{\lambda}\,(T-\tau)$, (17) becomes

$$\mathbf{M}\ddot{\bar{\boldsymbol{\lambda}}}+\mathbf{C}\dot{\bar{\boldsymbol{\lambda}}}+\mathbf{K}\bar{\boldsymbol{\lambda}}=\frac{\partial G}{\partial \mathbf{u}}-\frac{\widehat{\partial G}}{\partial \hat{\mathbf{u}}} \tag{20}$$

with initial conditions $\bar{\boldsymbol{\lambda}}\,(0)=0$ and $\dot{\bar{\boldsymbol{\lambda}}}\,(0)=0$.

The adjoint problem in this form is identical to the state problem. The only difference is the input to the differential equation. Therefore the same solver used for obtaining the state problem solution can be utilized for obtaining the solution of the adjoint equation.

The sensitivity of $g_1\,(\boldsymbol{\rho},\beta)$ with respect to β is computed as

$$\frac{\partial g_1}{\partial \beta}=\frac{\int_0^T -4\left[s\,(t)^2-\alpha g\,(t-\beta)^2\right]\alpha g\,(t-\beta)\,\frac{\partial g(t-\beta)}{\partial \beta}\mathrm{d}t}{g^*\int_0^T \alpha^2 g\,(t)^4\,\mathrm{d}t} \tag{21}$$

2.5 Time integration

As no analytical solution to (6) and (20) exists in the general case, the solution is obtained numerically at discrete time points. The choice of time integration scheme is extremely important for topology optimization, as the integration has to be performed twice after each update of the design variables. Explicit schemes with lumped mass matrices are preferred over implicit ones, due to their lower computational cost. In each topology optimization iteration step, first the state response is computed and after that the adjoint sensitivities are computed utilizing the state solution for evaluating the adjoint input. The adjoint response is computed at the same time points as for the state solution. During the integration process the value of the integrand in (19) is computed and its contribution is added to the sensitivities. The evaluation of the integrand in (19) is based on the solution for $\boldsymbol{\lambda}$ and $\mathbf{u}$ at the current time step. In order to implement the above scheme, it is necessary to store the state response $\mathbf{u}\,(t)$.

The time derivatives $\dot{\mathbf{u}}\,(t)$ and $\ddot{\mathbf{u}}\,(t)$ are computed based on finite difference scheme. At the n^{th} time step they are given as

$$\dot{\mathbf{u}}_n=\frac{\mathbf{u}_{n+1}-\mathbf{u}_{n-1}}{2\Delta t} \tag{22}$$

$$\ddot{\mathbf{u}}_n=\frac{\mathbf{u}_{n+1}-2\mathbf{u}_n+\mathbf{u}_{n-1}}{\Delta t^2} \tag{23}$$

where $\Delta t = t_{n+1} - t_n$. Inserting the above expressions in (6) and rearranging the term gives

$$\left(\frac{1}{\Delta t^2}\mathbf{M} + \frac{1}{2\Delta t}\mathbf{C}\right)\mathbf{u}_{n+1} = \mathbf{f}_n - \left(\frac{2}{\Delta t^2}\mathbf{M} + \mathbf{K}\right)\mathbf{u}_n - \left(\frac{1}{\Delta t^2}\mathbf{M} - \frac{1}{2\Delta t}\mathbf{C}\right)\mathbf{u}_{n-1} \quad (24)$$

where $\mathbf{f}_n$ is the system input at time t_n. Equation (24) provides the solution at time t_{n+1}, based on the system response at time steps t_n and t_{n-1}. The integration procedure starts with initial conditions $u_0 = 0$ and $u_{-1} = 0$. For the sensitivity analysis the second time derivative of the response is computed as

$$\ddot{\mathbf{u}}_0 = \mathbf{M}^{-1} f_0 \quad (25)$$

The time step is chosen according to Courant-Friedrichs-Lewy (CFL) condition to ensure stable integration, i.e. $\Delta t \leq \Delta t_c = \Delta x / c$ where Δx is the spatial step and c is the wave speed.

In (24) it is desirable to have diagonal mass $\mathbf{M}$ and damping $\mathbf{C}$ matrices in order to avoid the solution of a linear system, which will add significant computational cost to the integration process. Various lumping techniques have been presented in the literature. Among them spectral finite element discretization leads to fully diagonal mass matrix for nodal based elements (Cohen 2002; Peng et al. 2009). For acoustic problems the employment of modified central scheme (Yue and Guddati 2005) results in diagonal mass matrix and fourth order dispersion error. In electromagnetic problems, application of the trapezoidal integration rule to vector-based elements results in diagonal mass matrix without compromising the accuracy (Jin and Riley 2007).

2.6 Implementation details

The time domain optimization process consists of the following steps

1. Integration of the state response using (24).
2. Evaluating the objective and the constraint functions using the trapezoidal numerical integration
3. Integration of the adjoint response using (24) with the forcing term evaluated using the state response obtained in step 1. At each time step the integrand in (19) is evaluated and the contribution to the sensitivities is added by using the trapezoidal integration rule.
4. Update the design variables $\boldsymbol{\rho}$ using optimizer (MMA)

The Method of Moving Asymptotes (MMA) (Svanberg 1987) is utilized for solving the optimization problem. The optimization iteration is stopped when the infinite norm of the design vector increment is smaller than 10^{-3}.

A major challenge for the proposed method is the storage of the solution for the state problem, especially for larger structures and long simulation times. In order to attack large problems the state solution is stored in selected time points. For computing the adjoint solution the state problem solution is recomputed between each of the selected time points. Using such scheme decreases the memory requirement by several orders of magnitude. This method, however increases the computational cost by adding one additional simulation of the state equation.

3 Numerical experiments

Two numerical experiments, which correspond to the two optimization problems defined in Section 2.3, are presented in this section. The first numerical experiment is pulse delay, where the aim is to delay a pulse as much as possible and at the same time restricting the allowable envelope distortion. The objective for the first numerical experiment is given by (10). A well known example of such an experiment is the slow light-matter interaction technology which is achieved by utilizing pulse delaying structures. It might find high applicability in the future optical networks and information processing systems, e.g. to achieve low power consumption and efficient optical buffer devices (Baba 2008). The method presented here works equivalently for other physical problems as well. Among other potential applications are ultrasonic (Arenberg 1948) and surface acoustic wave (Reindl et al. 2001) (SAW) delay lines.

The second numerical experiment is pulse shaping example where the aim is the pulse recorded in the registration point to follow predefined envelope. Optimized designs providing such features can find utilization in data decoding devices and optical networks. In addition the methodology can be utilized for solving inverse problems in seismology or medical applications.

3.1 Pulse delay

Slow light can be realized in structures with composite dielectric media (Joannopoulos et al. 2008). Such composite structures are realized with permeability μ equal to the vacuum permeability, i.e. $\mu_r = 1$, and the only material parameter left for interpolation is the permittivity. The interpolation is done between air and dielectric material, i.e. $A^I = A^{II} = \mu_r$ and $B^I = \epsilon_0$ and $B^{II} = \epsilon_r \epsilon_0$.

The structure is excited by an incident plane-wave Gaussian pulse centered around t_0

$$\Psi^{\text{inc}}(x,t) = \Psi_0^{\text{inc}} \sin\left(kx - \omega_c (t - t_0)\right) \mathrm{e}^{-\frac{(t-t_0-x/v_p)^2}{T_0^2}} \quad (26)$$

where Ψ_0^{inc} is the amplitude of the pulse, $k = \omega_c / v_p$ is the wave number, ω_c is denoting the center angular frequency, $v_p = c/n$ is the phase velocity, $n = \sqrt{\mu_r \epsilon_r}$ is the refractive index, and c is the speed of light in vacuum. T_0 controls the width of the pulse.

The presented examples in Figs. 4 and 5 show the influence of the pulse bandwidth, the allowable envelope dispersion and the pulse amplitude decrease on the slow-down effect (Lenz et al. 2001; Kashyap 2010). It should be emphasized that the results serve to illustrate that the optimization formulation is able to represents the expected physical effects, rather than providing new results with respect to slow light generation. The physical parameters used in the presented study are shown in Table 1. The core material is silicon dioxide (SiO_2) with effective refractive index $n = 1.446$.

The results from the numerical experiments are shown in Figs. 4 and 5. The initial material layout consists of homogeneous core material. The envelope of the transmitted pulse is measured at location $0.9L_s$. The spatial discretization provides 20 uniformly spaced elements per wave length in the dielectric material. To investigate the significance of the allowable envelope distortion to the slow down effect, the numerical experiment start with value of g^* equal to zero, which is increased gradually to $g^* = 0.045$. Comparison of the results between Fig. 4a and b, and the results in Fig. 5a and b, indicates a slow-down effect enhancement by allowing larger envelope spreading. By increasing $g^* = 0.015$ to $g^* = 0.045$ the light is slowed down from 0.05ps to 0.11ps and 0.28 $\rightarrow$ 0.39ps. Broadening of the delayed pulse is inevitable in achieving slow light for linear material (Lenz et al. 2001). Comparison of Fig. 4a and c shows how the increase of the parameter α influence the delay. Increasing α from 0.7 to 0.8 has minor effect on the delay (0.05 $\rightarrow$ 0.06ps and 0.27 $\rightarrow$ 0.28ps for Fig. 5). The influence of the pulse bandwidth and the delay of the transmitted pulse can be seen by comparing Figs. 4 and 5.

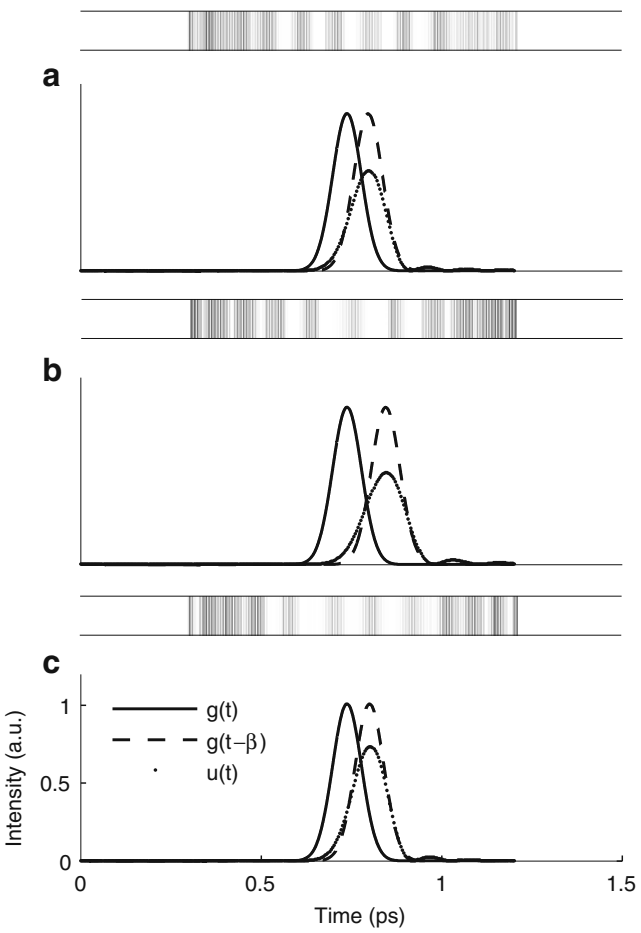


Fig. 4 Envelope of the signal intensity (*bottom*) obtained using the optimized structure (*top*) for $T_0 = 0.08$ps. The settings for the allowable temporal spreading and the pulse amplitude are as follows: **a**: g_{I}^*, and α_{I} with a group delay $\beta = 0.05$ps. **b**: g_{II}^*, and α_{I} with $\beta = 0.11$ps. **c**: g_{I}^*, and α_{II} with $\beta = 0.06$ps

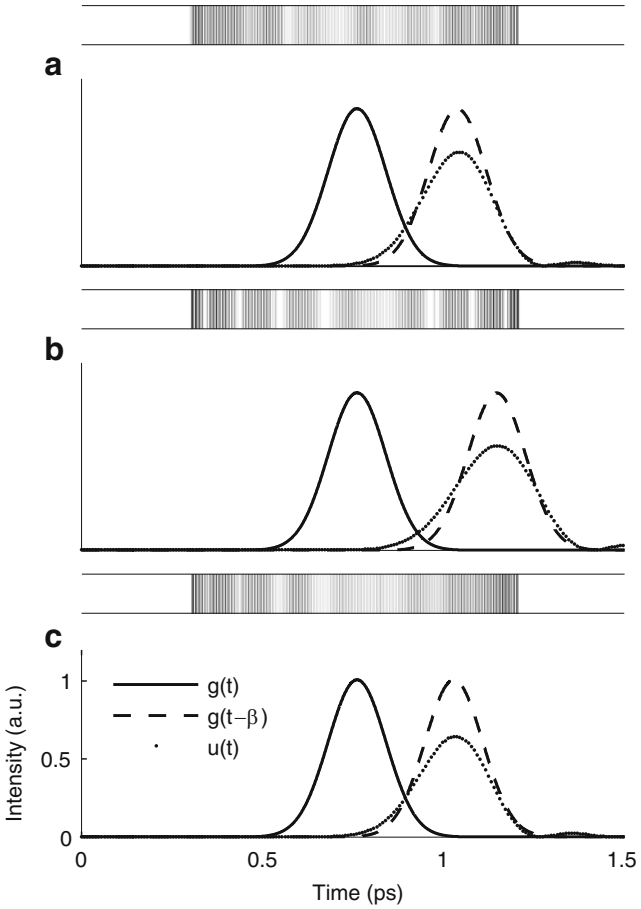


Fig. 5 Envelope of the signal intensity (*bottom*) obtained using the optimized structure (*top*) with $T_0 = 0.16$ps and the simulation time $T = 2.5$ps. The settings for the allowable temporal spreading and the pulse amplitude are as follows: **a**: g_{I}^*, and α_{I} with a group delay $\beta = 0.28$ps. **b**: g_{II}^*, and α_{I} with $\beta = 0.39$ps. **c**: g_{I}^*, and α_{II} with $\beta = 0.27$ps

Table 1 List of physical and discretization parameters used and their values

Parameters	Values	
Structure length, L_s	0.1mm	
Design domain length, L_{FBG}	$0.6L_s$	
Center frequency, f_c	193.55 THz	
Relative permittivity, ε_r	2.09	
Simulation time, T	1.2 ps	
Time step, Δt	$\Delta t = \Delta t_c$	
Number of elements, M	2,000	
Pulse delay optimization:	①	⑪
Pulse width, T_0	0.08 ps	0.16 ps
Amplitude decrease, α	0.7	0.8
Temporal spreading, g^*	0.015	0.045

Increasing $T_0 = 0.08$ps to $T_0 = 0.16$ps leads to slow-down ($0.11 \rightarrow 0.39$ps).

3.2 Pulse shaping

The objective in the numerical experiments presented in this section is to convert an incident Gaussian pulse into two subsequent Gaussian pulses after passing though the design domain in Fig. 1. The length of the entire modeling area (L_s) is 25 μm and the length of the design domain is $0.6L_s$. The modeling area is discretized with 3,000 uniformly spaced elements. The incident pulse prescribed by (26) has the following parameters: $t_0 = 0$, $\omega_c = 400\pi$ THz, $T_0 = 0.02$ ps and $\Psi_0^{\text{inc}} = 1$. The target envelope is prescribed by two Gaussian envelopes with $T_0 = 0.02$ ps and $\Psi_0 = 0.5$ shifted 0.09 ps from each other.

Two materials with different electric permittivities are used in the design domain: air ($\epsilon_r = 1$) and dielectric with arbitrary (but prescribed initially) permittivity varying from $\epsilon_r = 1.5$ to $\epsilon_r = 16$. Initially, the design domain includes randomly distributed air stripes, so that the probability of each element to be made of air is 1%. The delay of the entire target envelope is defined for each series of experiments with different dielectric materials separately. It coincides with the time when the maximum of the envelope of the registered pulse passes the target node for the non-optimized (initial) distribution of the material inside the modeling domain (e.g. the delay $t_0 = 0.092$ ps for the experiments with dielectric material having permittivity $\epsilon_r = 1.5$ and $t_0 = 0.301$ ps for the experiments with dielectric material having permittivity $\epsilon_r = 16$).

Twelve series of experiments with different permittivities of the dielectric material are performed. Sixteen numerical simulations are performed in each of these series with random initial distribution of air inside the design domain (192 numerical simulations in total). Results of the experiments are shown in Fig. 6. The objective function (and therefore the conformance of the envelope of the signal to the target envelope) quickly improves for dielectric materials having higher permittivities. This behavior can be associated with two factors:

– The effective length of the design domain is higher for materials with high permittivity as the wavelength of the signal passing through these materials is lower. This makes it possible for the design domain to introduce more substantial modifications into the incident signal.
– Higher refractive index of the dielectric materials makes the design domain more flexible in terms of altering the

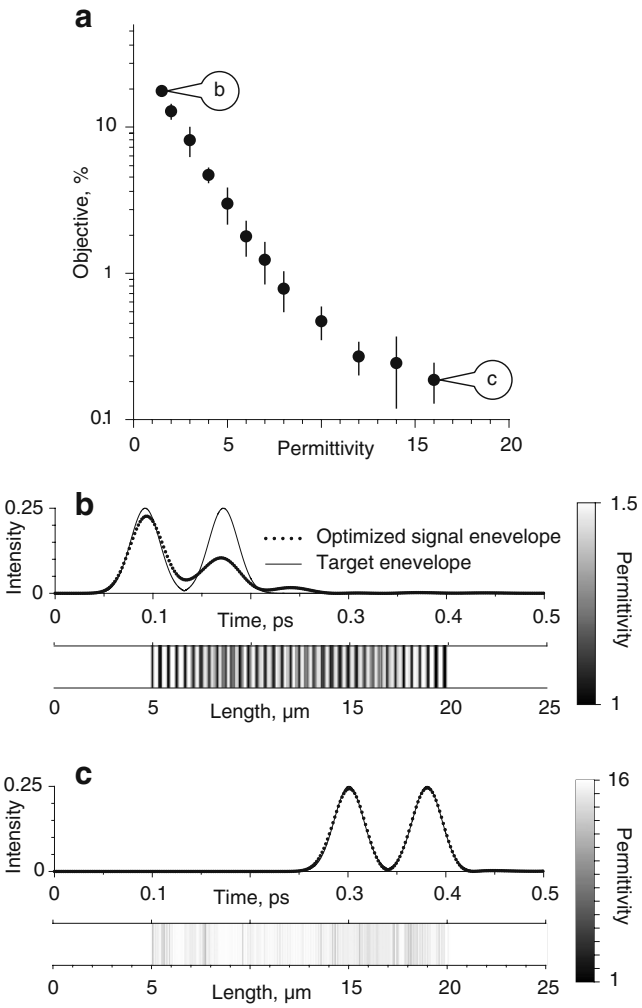


Fig. 6 **a**. Normalized objective vs. permittivity of the dielectric material used in the entire modeling domain for the pulse shaping experiment. *Error bars* indicate the standard deviation resulting from 16 independent numerical simulations with random initial distribution of air inside the design domain. **b**, **c**. Envelope of the signal intensity (*top*) obtained using the optimized structure with the lowest objective (*bottom*) for the dielectric materials with $\epsilon_r = 1.5$ (**b**) and $\epsilon_r = 16$ (**c**)

original signal due to the wider spectrum of material properties.

In all presented experiments the optimization procedure results in intermediate materials inside the design domain which can cause problems during manufacturing process. Various techniques exist for enforcing black and white design, including penalization of the objective function (Dahl et al. 2008; Yang et al. 2009), implicit morphological filters imposed on the design variables (Sigmund 2007) and robust design (Sigmund 2009; Wang et al. 2011). However, by forcing the structure to have either air or dielectric but not a linear combination of both, limits the set of possible solution and in this case the improvement of the objective for the dielectrics with higher permittivities might only be expected from longer design domain.

A problem which can be clearly seen on the presented structures is the lack of length scale in the designs. The optimized result consists of oscillating design with oscillations comparable with the elements dimensions. Such solutions are difficult for realization in practice and yield questionable numerical accuracy. In order to avoid such fast oscillations in the optimized design, the optimization problem has to be regularized. The regularization here is achieved by applying a density filter (Bruns and Tortorelli 2001). The design field in the state equations, the objective and the constraint equations is replaced with the filtered design field. The filtered density values $\tilde{\boldsymbol{\rho}}$ for each element are computed as follows

$$\tilde{\rho}_e = \frac{\sum_{i\in\mathbb{B}_e} w_i \rho_i}{\sum_{i\in\mathbb{B}_e} w_i} \tag{27}$$

where $\tilde{\rho}_e$ is the filtered density for element e, w_i is a weighting factor which depends on the distance between the centers of element e and element i and is computed as

$$w_i = R - |x_e - x_i| \tag{28}$$

for $|x_e - x_i| \leq R$. For all other elements the weight is zero. $\mathbb{B}_e$ is the index set of all elements located in the filter support domain, i.e. all elements for which $w_i \neq 0$.

The objective, the constraint functions and the state equations depend on the filtered density. The sensitivities are computed using the procedure in Section 2.4 and by applying the chain rule

$$\frac{\partial f}{\partial \rho_e} = \frac{\partial f}{\partial \tilde{\rho}_e}\frac{\partial \tilde{\rho}_e}{\partial \rho_e} \tag{29}$$

Applying density filtering restricts the solution space, and the performance of the optimized structure decreases for large filter radius. This can be observed in Fig. 7a. For small filter diameter the objective is on the same order as for the unfiltered design. However, for filter diameter larger

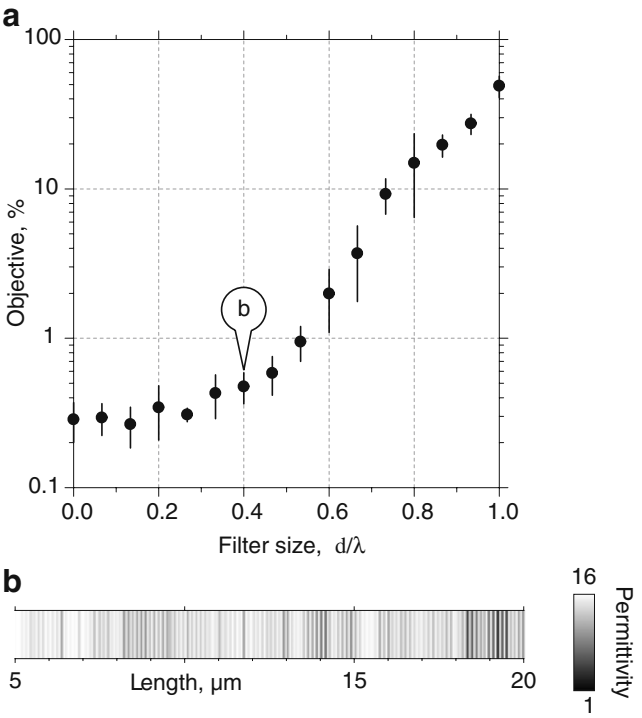


Fig. 7 **a**. Normalized objective vs. filter size (normalized by the central wavelength of the signal) for the pulse shaping experiment. *Error bars* indicate the standard deviation resulting from 12 independent numerical simulations with random initial distribution of air inside the design domain. **b**. Design domain obtained with the filter size 40% of the wavelength of the signal

than $0.4 - 0.5d/\lambda$, where λ is the wave length, the objective quickly deteriorates. The design for $d/\lambda = 0.4$ is presented at Fig. 7b. It consists of gray areas, which create also difficulties for practical realizations of the optimized structure. As mentioned earlier, obtaining crisp black and white design is an important subject which will be discussed in details in following articles.

4 Conclusions

The application of time domain topology optimization for design of pulse shaping devices is demonstrated. The information transferred through the optimized wave guide is considered to be represented by the pulse envelope. Such representation improves the robustness of the device with respect to signal generation and registration. The pulse envelope is extracted by using the Hilbert transform. The derivation of the sensitivities when the objective depends on the pulse envelope is demonstrated, and the optimization process has been demonstrated with several designs for pulse delaying and pulse splitting devices. The presented approach is applicable to signals consisting of broad

range of frequencies. Therefore the optimization methodology opens the possibility for applying time domain topology optimization in non-linear devices such as photonic diodes, transistors and logical elements. The presented algorithm is extendable to 2D and 3D space and it is applicable in a broad range of engineering areas, including acoustics and linear elasticity.

Acknowledgments This work was financially supported by Villum Fonden (via the NATEC Centre of Excellence), a Center of Advanced User Support (CAUS) grant from the Danish Center of Scientific Computing (DCSC), by the Elite Research Prize from the Danish Minister of Research, by Eurohorcs/ESF European Young Investigator Award (EURYI, http://www.esf.org/euryi) through the grant "Synthesis and topology optimization of optomechanical systems", and M.B. Richters foundation. The authors wish to thank professor Ole Sigmund and professor Jakob Søndergaard Jensen for their suggestions and valuable discussions.

Appendix A: Hilbert transform

Hilbert transform can be found in many textbooks on signal processing (e.g. Cohen 1995). It plays an important role in this article and, therefore, a short introduction and discussion on the numerical implementation are included for completeness. The Hilbert transform $\hat{f}(t)$ for a function $f(t)$ is defined as

$$\hat{f}(t) = \frac{1}{\pi} \mathrm{P} \int_{-\infty}^{\infty} \frac{f(t)}{t-\tau} \mathrm{d}\tau \tag{30}$$

when the integral exists. Usually, it is not possible to compute the integral as an ordinary improper integral due to the pole at $\tau = t$. In such cases the integral is computed as the Cauchy principal value. P in (30) stands for Cauchy principal value. The Hilbert transform can be easily computed numerically by using Fourier transform. If $F(\omega)$ is the Fourier transform of $f(t)$, the Fourier transform of $\hat{f}(t)$ is computed as

$$\hat{F}(\omega) = -i \operatorname{sgn}(\omega) F(\omega) \tag{31}$$

where $i = \sqrt{-1}$, and sgn is the signum function defined as follows

$$\operatorname{sgn}(\omega) = \begin{cases} 1 & \text{for } \omega \geq 0 \\ 0 & \text{for } \omega = 0 \\ -1 & \text{for } \omega \leq 0 \end{cases} \tag{32}$$

The Hilbert transform $\hat{f}(t)$ can be computed by taking the inverse Fourier transform of (31). In practical application the Fourier transform and the inverse Fourier transform are replaced with the Fast Fourier Transform and its inverse.

Appendix B: Proof of the identity (16)

The aim of this section is to demonstrate the derivation of the identity given by (16). In order to simplify the notations the derivative $\partial G / \partial \hat{u}$ is substituted with $b(t)$

$$\frac{\partial G}{\partial \hat{u}} \frac{\partial \hat{u}}{\partial \rho_e} = b(t) \frac{\partial \hat{u}}{\partial \rho_e} \tag{33}$$

where $b(t)$ can be computed from the time history of the response $u(t)$ and its Hilbert transform $\hat{u}(t)$. Assuming that the system response and the system input are zero outside the time interval $[0, T]$, the following integrals are equivalent

$$\int_0^T b(t) \frac{\partial \hat{u}}{\partial \rho_e} \mathrm{d}t = \int_{-\infty}^{\infty} b(t) \frac{\partial \hat{u}}{\partial \rho_e} \mathrm{d}t \tag{34}$$

The Hilbert transform $\partial \hat{u} / \partial \rho_e$ of $\partial u / \partial \rho_e$ is given as

$$\frac{\partial \hat{u}}{\partial \rho_e} = \frac{1}{\pi} \mathrm{P} \int_{-\infty}^{\infty} \frac{1}{t-\tau} \frac{\partial u}{\partial \rho_e} \mathrm{d}\tau \tag{35}$$

Introducing (35) into (34) and after some algebra the following equality is obtained

$$\begin{aligned} \int_{-\infty}^{\infty} b(t) \frac{\partial \hat{u}}{\partial \rho_e} \mathrm{d}t &= \int_{-\infty}^{\infty} b(t) \frac{1}{\pi} \mathrm{P} \int_{-\infty}^{\infty} \frac{1}{t-\tau} \frac{\partial u}{\partial \rho_e} \mathrm{d}\tau \mathrm{d}t \\ &= \int_{-\infty}^{\infty} \frac{\partial u}{\partial \rho_e} \left(-\frac{1}{\pi} \mathrm{P} \int_{-\infty}^{\infty} \frac{b(t)}{\tau - t} \mathrm{d}t \right) \mathrm{d}\tau \quad (36) \\ &= \int_{-\infty}^{\infty} -\hat{b}(\tau) \frac{\partial u}{\partial \rho_e} \mathrm{d}\tau = \int_0^T -\hat{b}(t) \frac{\partial u}{\partial \rho_e} \mathrm{d}t \end{aligned} \tag{37}$$

which concludes the derivation of the identity

$$\int_0^T \frac{\partial G}{\partial \hat{\mathbf{u}}} \frac{\partial \hat{\mathbf{u}}}{\partial \rho_r} \mathrm{d}t = \int_0^T -\widehat{\frac{\partial G}{\partial \hat{\mathbf{u}}}} \frac{\partial \mathbf{u}}{\partial \rho_r} \mathrm{d}t \tag{38}$$

References

Arenberg DL (1948) Ultrasonic solid delay lines. J Acoust Soc Am 20(1):1–26

Baba T (2008) Slow light in photonic crystals. Nature Photonics 2(8):465–473

Bendsøe MP, Sigmund O (2004) Topology optimization—theory, methods and applications. Springer Verlag, Berlin

Brillouin L (1953) Wave propagation in periodic structures. Dover

Bruns T, Tortorelli D (2001) Topology optimization of non-linear elastic structures and compliant mechanisms. Comput Methods Appl Mech Eng 190(26–27):3443–3459

Cohen GC (2002) Higher order numerical methods for transient wave equations. Springer-Verlag, Berlin

Cohen L (1995) Time-frequency analysis. Prentice-Hall, New Jersey

Dahl J, Jensen JS, Sigmund O (2008) Topology optimization for transient wave propagation problems in one dimension design of filters and pulse modulators. Struct Multidisc Optim 36(6):585–595

Jensen J, Sigmund O (2004) Systematic design of photonic crystal structures using topology optimization: low-loss waveguide bends. Appl Phys Lett 84(12):2022–2024

Jensen J, Sigmund O (2011) Topology optimization for nano-photonics. Laser Photonics Rev 5(2):308–321

Jensen JS (2011) Topology optimization of nonlinear optical devices. Struct Multidisc Optim. doi:10.1007/s00158-011-0640-0

Jensen JS, Sigmund O (2005) Topology optimization of photonic crystal structures: a high-bandwidth low-loss t-junction waveguide. J Opt Soc Am B Opt Phys 22(6):1191–1198

Jin J, Riley DJ (2007) Finite element analysis of antennas and arrays. John Wiley & Sons Inc

Joannopoulos J, Johnson S, Winn J, Meade R (2008) Photonic crystals: molding the flow of light. Princeton University Press, Princeton

Kang BS, Park GJ, Arora JS (2006) A review of optimization of structures subjected to transient loads. Struct Multidisc Optim 31(2):81–95

Kashyap R (2010) Fiber bragg gratings. In: Fiber Bragg gratings, 2nd edn. Academic Press, Boston

Kushwaha MS, Halevi P, Dobrzynski L, Djafari-Rouhani B (1993) Acoustic band structure of periodic elastic composites. Phys Rev Lett 71(13):2022–2025

Lenz G, Eggleton BJ, Madsen CK, Slusher RE (2001) Optical delay lines based on optical filters. IEEE J Quantum Electron 37(4):525–532

Matzen R, Jensen JS, Sigmund O (2010) Topology optimization for transient response of photonic crystal structures. J Opt Soc Am B 27(10):2040–2050

Mead D (1975) Wave propagation and natural modes in periodic systems: I. mono-coupled systems. J Sound Vib 40(1):1–18

Nomura T, Sato K, Taguchi K, Kashiwa T, Nishiwaki S (2007) Structural topology optimization for the design of broadband dielectric resonator antennas using the finite difference time domain technique. Int J Numer Methods Eng 71(11):1261–1296

Peng HK, Meng G, Li FC (2009) Modeling of wave propagation in plate structures using three-dimensional spectral element method for damage detection. J Sound Vib 320(4–5):942–954

Reindl L, Ruppel CCW, Berek S, Knauer U, Vossiek M, Heide P, Oreans L (2001) Design, fabrication, and application of precise saw delay lines used in an fmcw radar system. IEEE Trans Microwave Theor Tech 49(4):787–794

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

Sigmund O (2009) Manufacturing tolerant topology optimization. Acta Mech Sin 25(2):227–239

Sigmund O, Jensen J (2003) Systematic design of phononic band-gap materials and structures by topology optimization. Philos Trans Royal Soc Lond Ser A Math Phys Eng Sci 361:1001–1019

Svanberg K (1987) The method of moving asymptotes—a new method for structural optimization. Int J Numer Methods Eng 24:359–373

Tortorelli DA, Michaleris P (1994) Design sensitivity analysis: overview and review. Inverse Probl Eng 1(1):71–105

Wang F, Lazarov B, Sigmund O (2011) On projection methods, convergence and robust formulations in topology optimization. Struct Multidisc Optim. doi:10.1007/s00158-010-0602-y

Yang LR, Lavrinenko AV, Hvam JM, Sigmund O (2009) Design of one-dimensional optical pulse-shaping filters by time-domain topology optimization. Appl Phys Lett 95(26):261,101

Yue B, Guddati MN (2005) Dispersion-reducing finite elements for transient acoustics. J Acoust Soc Am 118(4):2132–2141

Zienkiewicz OC, Taylor RL, Zhu J (2005) The finite element method. Elsevier

[P4]

INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng (2011)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nme.3205

An efficient finite element time-domain formulation for the elastic second-order wave equation: A non-split complex frequency shifted convolutional PML

René Matzen*, †

Department of Mechanical Engineering, Solid Mechanics, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

SUMMARY

The perfectly matched layer (PML) technique has demonstrated very high efficiency as absorbing boundary condition for the elastic wave equation recast as a first-order system in velocity and stress in attenuating non-grazing bulk and surface waves. This paper develops a novel convolutional PML formulation based on the second-order wave equation with displacements as the only unknowns to annihilate spurious reflections from near-grazing waves. The derived variational form allows for the use of e.g. finite element and the spectral element methods as spatial discretization schemes. A recursive convolution update scheme of second-order accuracy is employed such that highly stable, effective time integration with the Newmark-beta (implicit and explicit with mass lumping) method is achieved. The implementation requires minor modifications of existing displacement-based finite element software, and the stability and efficiency of the proposed formulation is verified by relevant two-dimensional benchmarks that accommodate bulk and surface waves. Copyright © 2011 John Wiley & Sons, Ltd.

Received 22 November 2010; Revised 3 March 2011; Accepted 9 March 2011

KEY WORDS: perfectly matched layers; absorbing boundary conditions; elastic wave equation; finite element time-domain discretization; implicit/explicit time integration

1. INTRODUCTION

Solving open-region elastodynamic problems implies that the infinite region exterior to the radiating/scattering object has to be truncated with some type of artificial boundary such that the computational domain becomes limited. Hence, a boundary condition that absorbs elastic waves leaving the computational domain independent of direction and frequency needs to be introduced. Such a condition serves, as far as possible, as a transparent boundary that should yield perfect transmission of the scattered/radiated field and thereby minimize the non-physical reflections from the boundary. During the last 35 years, several absorbing boundary condition (ABC) techniques have been developed in the endeavor to complete this task [1–9]. Formulations based on non-local conditions (i.e. an exact representation of the infinite medium) are difficult to implement and cannot manage arbitrary material heterogeneity, whereas their local peers only exhibit good performance under tailored circumstances (frequency dependent, specific angles of incidence).

*Correspondence to: René Matzen, Department of Mechanical Engineering, Solid Mechanics, Technical University of Denmark, DK-2800 Lyngby, Denmark.
†E-mail: rmat@mek.dtu.dk

In 1994, Berenger introduced the concept of transparent perfectly matched interface to electromagnetic (EM) waves [10]. It separates two half-spaces, one of which is dispersive, i.e. lossy, and is known as the perfectly matched layer (PML). By construction, the PML is a local ABC and renders the interface reflection-less in the continuous case while outwardly propagating waves are attenuated exponentially independent of frequency, angle of incidence and polarization. The formulation was based on a field-splitting procedure recasting the governing equation in the PML into a *non-Maxwellian* form as means to avoid convolution operations in the time-domain. Despite the introduction of spurious reflections upon a finite-difference time-domain (FDTD) discretization, it reveals remarkable absorptive and stable late-time capabilities redeeming the additional costs associated with doubling the number of unknowns as a result of the field-splitting.

After its introduction, the PML concept has elegantly been generalized to become a superior ABC technique by interpreting the absorbing layer as *complex coordinate stretched* [11] or *artificially anisotropic* [12]. The latter maintains the original *Maxwellian* form of the governing equation, whereby the PML can be readily extended to problems involving irregular geometries, as e.g. encountered by finite/spectral element (FE/SE) methods [13]. Nevertheless, both interpretations have facilitated a straightforward adoption of the PML concept to other physical wave propagation problems [14, 15].

The first real applications of PMLs to elastic waves were based on the split-field, velocity–stress formulation [16–19], later extended to cases involving anisotropic media [20]. To obtain a unified formulation instead, a non-splitting velocity/displacement–stress approach has been suggested, in which the expensive temporal convolution operations, consequently emerging, are resolved by an effective recursive update strategy [21]. This is also known as the convolutional PML (C-PML) and includes the option for *complex frequency shifted* (CFS) [22] coordinate stretching that improves the behavior of the discrete PML for near-grazing waves [23–29]. In the aforementioned studies the spatial operators are approximated by finite differences (FDs). Similarly, the general trend in an FE/SE time-domain framework involves split-field [30–32] and non-split [33] PML formulations typically ensuing in a mixed scheme with velocity/displacement and stresses as unknowns to avoid temporal convolution operations. However, when using implicit time marching to circumvent the Courant–Friedrichs–Lewy (CFL) constraint between the time step and the smallest element sizes in explicit schemes [34] for unstructured meshes, the mixed formulation based on an FE discretization could lead to an increase in computational costs. This is due to system matrices with extra non-zero entries (and broken symmetries) implying a growth in storage and the number of flops to be performed at each time step, since, apart from displacements, stresses additionally become unknowns (see e.g. [33]). Furthermore, since existing software for FE modeling of solid mechanics and elasticity problems predominantly solves for displacements only there is a need for a PML formulation for the elastic second-order wave equation. The proposed method supports a straightforward switch to explicit time integration with mass lumping which is of high relevance in very large three-dimensional (3D) elastodynamic problems.

To the author's knowledge five attempts to solve the above challenge have been reported. In [35] Komatitisch and Tromp manage to eliminate the stress terms; although at the expense of splitting the displacement field into four components. The resulting discrete equations are either third-order, or second-order coupled with a first-order equation. An FE discretized second-order wave equation for the displacement field is almost established in [36, 37]. However, the formulation requires computation of the strains that are obtained by tricky coordinate transformations of the displacement gradients. Even though the rather involved implicit time integration in [36] has been replaced by an explicit scheme in [37] to enhance computational efficiency the complexity persists. Recent work by Martin et al. [38] demonstrates a highly stable hybrid technique rendering SE discretized governing equations for the interior non-split in displacement-only, that couples to a mixed velocity–stress form for the PML previously developed in [39]. Another recent work in [40] presents a C-PML model for the second-order wave equation that contains auxiliary memory variables to avoid the convolution operators. However in reality, the scheme is mixed since each memory variable is governed by an additional equation of first-order in which the stresses are needed. The common outcome of these formulations is rather complex implementations that necessitate fundamental re-structuring of existing FE codes. It appears that a non-split variational

DOI: 10.1002/nme

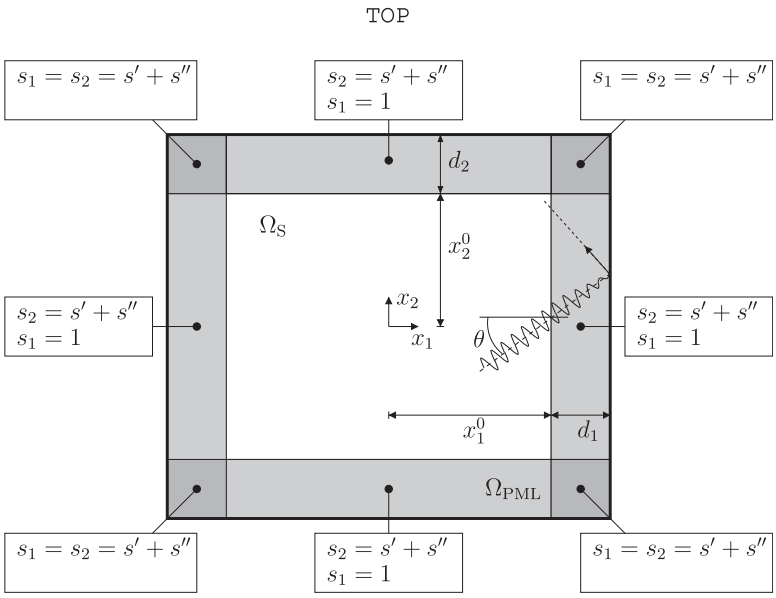


Figure 1. Computational domain truncated by PML layer. A penetrating wave with angle θ is depicted. For the complex frequency shifted convolutional PML $s' = \kappa \geqslant 1$ and $s'' = \beta/(\alpha + \mathrm{j}\omega)$ is a imaginary number, where $\beta \geqslant 0$ and $\alpha \geqslant 0$.

formulation of the C-PML for the elastic second-order wave equation with displacement as the single unknown has yet to be suggested.

1.1. Complex frequency shifted coordinate stretching

The key concept of the PML interpretation that relies on complex frequency coordinate stretching is a coordinate transformation in which the spatial variables are mapped independently onto the complex space by a complex stretching function. In particular, this mapping replaces propagating waves with exponentially decaying waves as soon as the waves passes the PML interface located at $x_1 = x_1^0$ along the spatial direction x_1, see Figure 1. The attenuation occurs over the PML layer width d_1, and continues after the remainder of the wave has been reflected from the PML boundary at $x_1 = x_1^0 + d_1$.

In the two-dimensional (2D) PML region Ω_{PML}, the original spatial coordinate variables x_i $(i = 1, 2)$ are replaced in each distinct spatial direction by a stretched coordinate, that is defined as [11, 41]

$$x_i' = \int_0^{x_i} s_i(\tilde{x}_i, \omega) \mathrm{d}\tilde{x}_i, \quad i = 1, 2 \tag{1}$$

The apostrophe $(\,)'$ indicates the stretched version of the subtended function, ω is the angular frequency, and s_i $(i = 1, 2)$ are the CFS stretched-coordinate metrics proposed by Kuzuoglu and Mittra [22]

$$s_i(x_i, \omega) = \kappa_i(x_i) + \frac{\beta_i(x_i)}{\alpha_i(x_i) + \mathrm{j}\omega}, \quad i = 1, 2 \tag{2}$$

where 'j' denotes the imaginary unit. In Equation (2) $\beta_i \geqslant 0$ is a coordinate-wise real function that controls the attenuation of the propagating waves. The coordinate-wise real functions $\kappa_i \geqslant 1$ and $\alpha_i \geqslant 0$ serve to enhance attenuation of evanescent and near-grazing waves. In the stretched coordinate metrics the spatial variation of the planar wave amplitude that propagate

DOI: 10.1002/nme

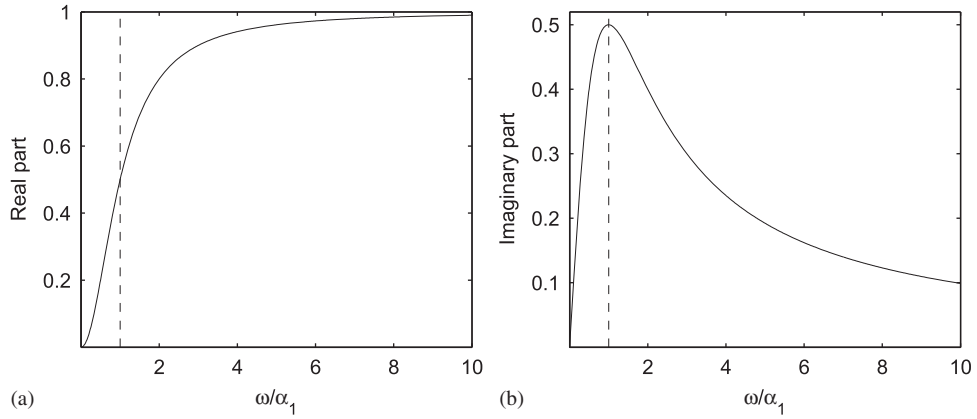


Figure 2. Real and imaginary part of the decay factor $(\omega^2-\mathrm{j}\alpha_1\omega)/(\omega^2+\alpha_1^2)$ as a function of ω/α_1: (a) the real part resembles the behavior of a Butterworth filter, where high frequency propagating wave components are subject to attenuation, while at low frequencies the propagating waves remain unchanged. Conversely, (b) the imaginary part indicates evanescent wave attenuation at low frequencies.

along x_1, i.e. $\mathrm{e}^{-\mathrm{j}k_1x_1}$, modifies to [39, 42, 43]

$$\mathrm{e}^{-\mathrm{j}k_1x_1'}=\mathrm{e}^{-\mathrm{j}k_1\int\kappa_1\mathrm{d}\tilde{x}_1}e^{-\frac{k_1}{\omega}\frac{\omega^2-\mathrm{j}\alpha_1\omega}{\omega^2+\alpha_1^2}\int\beta_1\mathrm{d}\tilde{x}_1} \tag{3}$$

where we, for simplicity, have assumed that α_1 is constant. Since propagating waves at grazing incidence become evanescent in the direction perpendicular to the boundary layer, strong reflections will inevitably occur from the standard PML ($\kappa_1=1$ and $\alpha_1=0$), as this is only designed to effectively absorb outward propagating waves. However, according to Equation (3) κ_1 serves as a scaling function in attenuating evanescent waves. In addition to that, if we plot the real and imaginary part of $(\omega^2-\mathrm{j}\alpha_1\omega)/(\omega^2+\alpha_1^2)$ as a function of ω/α_1 we realize (see Figure 2(b)) that shifting the pole of stretching onto the imaginary axis by α_1 causes attenuation of evanescent waves at low frequencies. Simultaneously (see Figure 2(a)), propagating waves will be attenuated for high frequencies and, according to Figure 2(b), subject to maximum phase shift at $\omega=\alpha_1$. Essentially, the complex frequency shift implements a Butterworth filter [39, 44], and it has been successfully implemented using mixed non-split FD and [23, 25–29] and split-field FE/SE formulations [38, 39].

The aim of this paper is to combine the CFS-PML model in Equation (2) and, for simplicity, 2D wave propagation in a linear elastic medium occupying an open domain $\Omega=\Omega_\mathrm{S}\cup\Omega_\mathrm{PML}\subset\mathbb{R}^2$ in a time interval $\mathbb{I}=[0,T]$ governed by

$$\text{(Eq. of motion)}\quad \nabla\cdot\boldsymbol{\sigma}+\rho\boldsymbol{p}=\rho\ddot{\boldsymbol{u}} \tag{4a}$$

$$\text{(Constitutive law)}\quad \boldsymbol{\sigma}=\boldsymbol{C}:\boldsymbol{\varepsilon} \tag{4b}$$

$$\text{(Strain–displ.)}\quad \boldsymbol{\varepsilon}=\tfrac{1}{2}[\nabla\boldsymbol{u}+(\nabla\boldsymbol{u})^\mathrm{T}] \tag{4c}$$

such that we obtain an FE discretized second-order wave equation with displacements as the only unknowns. In Equation (4), $\boldsymbol{u}=(u_1,u_2)^\mathrm{T}$ and $\boldsymbol{p}=(p_1,p_2)^\mathrm{T}$ are the displacement and body force vectors, respectively, ρ is the mass density, $\boldsymbol{\sigma}$, $\boldsymbol{\varepsilon}$, and $\boldsymbol{C}$ are stress, strain, and constitutive tensors, respectively. $\nabla=(\partial/\partial x_1,\partial/\partial x_2)^\mathrm{T}$ is the divergence operator, and $\ddot{(\,)}=\partial^2/\partial t^2$. The formulation presented in this paper implies minor modifications to the original second-order wave equation, in which the convolution terms act as (internal) forces with the displacements from the previous time step appearing as the driving source. The proposed method facilitates a simple switch between explicit and implicit time integration, and the use of either FE or SE in the integration of the element matrices is free of choice. The extension to 3D follows the same approach as in 2D.

The remainder of the paper is organized as follows: In Section 2 the continuous frequency domain equations are derived (Section 2.1), followed by the time domain counterparts (Section 2.2). Next,

DOI: 10.1002/nme

the variational form is established from which the FE system matrices are derived (Section 2.3), and the section ends with a presentation of the time integration method (Section 2.4). In Section 3 the formulation is verified by various numerical experiments, and the section is completed by a numerical stability analysis of the proposed scheme (Section 3.4), followed by conclusions in Section 4.

2. PML EQUATIONS FOR PLANE MOTION

2.1. Frequency domain equations

The basic concept of the PML can be introduced by considering the frequency-domain counterparts of the equations of motion, constitutive law, and the kinematic conditions (4) in the stretched coordinate metrics

$$\text{(Eq. of motion)} \quad \nabla_s \cdot \widehat{\boldsymbol{\sigma}} + \rho \widehat{\boldsymbol{p}} = -\omega^2 \rho \widehat{\boldsymbol{u}} \tag{5a}$$

$$\text{(Constitutive law)} \quad \widehat{\boldsymbol{\sigma}} = \boldsymbol{C} : \widehat{\boldsymbol{\varepsilon}} \tag{5b}$$

$$\text{(Strain–displ.)} \quad \widehat{\boldsymbol{\varepsilon}} = \tfrac{1}{2}[\nabla_s \widehat{\boldsymbol{u}} + (\nabla_s \widehat{\boldsymbol{u}})^{\mathrm{T}}] \tag{5c}$$

The caret $\widehat{(\,)}$ designates the frequency-domain equivalent of the subtended function, which is obtained by the Fourier transform [45] by assuming vanishing initial conditions for the displacement field. For a 2D isotropic elastic medium, the constitutive tensor entries are given by

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{6}$$

with μ and λ denoting the Lamé coefficients, and where δ_{ij} is Kronecker's delta. In Equation (5) the ∇_s operator is defined in 2D by $\nabla_s = (\partial/\partial x_1', \partial/\partial x_2')^{\mathrm{T}}$. Now, by use of the chain-rule the ∇_s operator becomes

$$\nabla_s = \left(\frac{1}{s_1}\frac{\partial}{\partial x_1}, \frac{1}{s_2}\frac{\partial}{\partial x_2} \right)^{\mathrm{T}} \tag{7}$$

Thus, in a Cartesian coordinate system ∇_s can be interpreted in terms of the standard ∇, although with x_1 and x_2-axes scaled by s_1 and s_2, respectively. By restoring non-stretched divergence ∇ in Equation (5) it is possible to obtain a uniform formulation that governs wave propagation in the entire computational domain, including the PML region. To accomplish this, we need first to modify the equation of motion and introduce a stretched complex-valued stress tensor $\widehat{\boldsymbol{\sigma}}'$ herein [16]

$$\nabla \cdot \widehat{\boldsymbol{\sigma}}' + \rho \widehat{\boldsymbol{p}} = -\omega^2 \rho' \widehat{\boldsymbol{u}} \tag{8}$$

where $\rho' = \rho s_1 s_2$. The stretched tensor in Eq. (7) can be expressed in terms of the original stress tensor $\widehat{\boldsymbol{\sigma}}$ by

$$\widehat{\boldsymbol{\sigma}}' = s_1 s_2 \begin{bmatrix} \dfrac{1}{s_1} & \\ & \dfrac{1}{s_2} \end{bmatrix} \cdot \widehat{\boldsymbol{\sigma}} = s_1 s_2 \Lambda \cdot \widehat{\boldsymbol{\sigma}} \tag{9}$$

Second, in order to render the kinematic condition non-stretched the constitutive tensor also needs to be modified as follows:

$$\widehat{C}'_{ijkl} = C_{ijkl} \frac{s_1 s_2}{s_i s_k}, \quad i, j, k, l = 1, 2 \tag{10}$$

where no Einstein summation over indices i and k applies. This leads to an artificial anisotropic material, since $\widehat{C}'_{1111} \neq \widehat{C}'_{2222}$, with minor, however not major, symmetry properties conserved. Writing out the constitutive law for the anisotropic medium, one will also find that the symmetry

DOI: 10.1002/nme

of the stress tensor is destroyed, i.e. $\widehat{\sigma}'_{ij} \neq \widehat{\sigma}'_{ji}$ when $i \neq j$. Notice that the transformations in Equations (9) and (10) alternate the PML interpretation from being based upon complex coordinate stretching to an anisotropic material model.

For a general FE implementation it is convenient to adopt a formulation based on Voigt notation [46]. To reestablish the symmetry of the constitutive matrix (despite the non-symmetric tensor counterpart), the constitutive law, based on displacements, now reads as

$$\begin{Bmatrix} \widehat{\sigma}'_{11} \\ \widehat{\sigma}'_{22} \\ \frac{1}{2}(\widehat{\sigma}'_{21}+\widehat{\sigma}'_{12}) \\ \frac{1}{2}(\widehat{\sigma}'_{21}-\widehat{\sigma}'_{12}) \end{Bmatrix} = \begin{bmatrix} \widehat{C}'_{11} & C_{12} & & \\ C_{12} & \widehat{C}'_{22} & & \\ & & \widehat{C}'_{66} & \widehat{C}''_{66} \\ & & \widehat{C}''_{66} & \widehat{C}'''_{66} \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x_1} & 0 \\ 0 & \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} & -\frac{\partial}{\partial x_1} \end{bmatrix} \begin{Bmatrix} \widehat{u}_1 \\ \widehat{u}_2 \end{Bmatrix} \tag{11}$$

where $\widehat{C}'_{11} = s_2 C_{11}/s_1$, $\widehat{C}'_{22} = s_1 C_{22}/s_2$, and the shear-related entries in the constitutive matrix are given by

$$\widehat{C}'_{66} = \frac{C_{66}}{4}\left(\frac{s_1}{s_2}+\frac{s_2}{s_1}+2\right) \tag{12a}$$

$$\widehat{C}''_{66} = \frac{C_{66}}{4}\left(\frac{s_1}{s_2}-\frac{s_2}{s_1}\right) \tag{12b}$$

$$\widehat{C}'''_{66} = \frac{C_{66}}{4}\left(\frac{s_1}{s_2}+\frac{s_2}{s_1}-2\right) \tag{12c}$$

For a compact version of Equation (11) we write $\widehat{\boldsymbol{\sigma}}'_{\mathrm{v}} = \widehat{\mathbf{C}}' \partial \widehat{\boldsymbol{u}}$, from which we identify the kinematic (strain–displacement) operator as

$$\partial = \begin{bmatrix} \frac{\partial}{\partial x_1} & 0 \\ 0 & \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} & -\frac{\partial}{\partial x_1} \end{bmatrix} \tag{13}$$

the subscript $(\,)_{\mathrm{v}}$ indicates that Voigt notation is applied to second-order tensors. As preparation for the construction of the time-domain formulation, we split the constitutive matrix into a non-stretched part and two stretched parts associated with s_1/s_2 and s_2/s_1, respectively, i.e.

$$\widehat{\mathbf{C}}' = \underbrace{\mathbf{D}_0 \otimes \mathbf{C}}_{=\mathbf{C}_0} + \frac{s_1}{s_2}\underbrace{\mathbf{D}_1 \otimes \mathbf{C}}_{=\mathbf{C}_1} + \frac{s_2}{s_1}\underbrace{\mathbf{D}_2 \otimes \mathbf{C}}_{=\mathbf{C}_2} \tag{14}$$

where $\mathbf{C}$ is the non-stretched constitutive matrix given by

$$\mathbf{C} = \begin{bmatrix} \lambda+2\mu & \lambda & & \\ \lambda & \lambda+2\mu & & \\ & & \mu & \mu \\ & & \mu & \mu \end{bmatrix} \tag{15}$$

DOI: 10.1002/nme

and $\otimes$ denotes the matrix operator that scales the entry in the first matrix by the element residing at the same location in the second matrix. From Equation (11) we deduce the matrix operators $\mathbf{D}_0$, $\mathbf{D}_1$, and $\mathbf{D}_2$ as

$$\mathbf{D}_0 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}, \quad \mathbf{D}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} \end{bmatrix}, \quad \mathbf{D}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \tag{16}$$

For the three-dimensional (3D) case a similar split needs to be employed in order to separate the non-stretched and stretched terms with different frequency dependencies from each other. The constitutive matrix for 3D generalization can be found in [47]. The anisotropic interpretation of the PML is designed to ease the implementation in an FE framework compared to the difficulties occurring when the FE formulation derives from complex coordinate stretching interpretation. The latter idea is pursued in [36, 37], and the approach renders a semi-discrete form that is almost second-order in time. However, it contains an internal force term, whose computation requires storage of strains, that in addition need to be integrated in time at every time step. The strains are obtained from rather complicated coordinate transformations of the displacement gradients.

2.2. Time domain equations

Next, we establish the time-domain counterparts of the stretched frequency-dependent motion Equation (8) and the displacement-based constitutive law in Equation (11). This conversion is aided by the inverse Fourier transform [45]. The stretched stress tensor and the displacement vector transform directly to $\boldsymbol{\sigma}'_{\mathrm{v}}$ and $\boldsymbol{u}$. Meanwhile, special treatment to the constitutive matrix in Equation (11) and the right-hand side in Equation (8) is needed, since they contain s_i $(i=1,2)$. Applying the inverse Fourier transform (e.g. provided in Maple 14 [48]) yields

$$\nabla \cdot \boldsymbol{\sigma}' + \rho \boldsymbol{p} = \mathscr{L}_0(t)\boldsymbol{u} \tag{17a}$$

$$\boldsymbol{\sigma}'_{\mathrm{v}} = \mathbf{C}' \partial \boldsymbol{u} \tag{17b}$$

with

$$\mathbf{C}' = \mathbf{C}_0 + \mathscr{L}_1(t)\mathbf{C}_2 + \mathscr{L}_2(t)\mathbf{C}_2 \tag{17c}$$

where the operator $\mathscr{L}_0(t)$ is the inverse Fourier transform of $-\omega^2 s_1 s_2$, and it is, based on the definition in Equation (2), given by

$$\mathscr{L}_0(t) = P_0 \frac{\partial^2}{\partial t^2} + P_1 \frac{\partial}{\partial t} + P_2 + \alpha_1^2 P_{3,1} \mathrm{e}^{-\alpha_1 t} \bar{u}(t)^* + \alpha_2^2 P_{3,2} \mathrm{e}^{-\alpha_2 t} \bar{u}(t)^* \tag{18}$$

In Equations (17b) and (17c) $\mathscr{L}_1(t)$ and $\mathscr{L}_2(t)$ are the inverse Fourier transform of s_1/s_2 and s_2/s_1, respectively, and are given by

$$\mathscr{L}_1(t) = \frac{\kappa_1}{\kappa_2} + \beta_1^2 P_{3,1}^{-1} \mathrm{e}^{-\alpha_1 t} \bar{u}(t)^* - P_{4,2} \mathrm{e}^{-(\beta_2 + \kappa_2 \alpha_2)t/\kappa_2} \bar{u}(t)^* \tag{19a}$$

$$\mathscr{L}_2(t) = \frac{\kappa_2}{\kappa_1} + \beta_2^2 P_{3,2}^{-1} \mathrm{e}^{-\alpha_2 t} \bar{u}(t)^* - P_{4,1} \mathrm{e}^{-(\beta_1 + \kappa_1 \alpha_1)t/\kappa_1} \bar{u}(t)^* \tag{19b}$$

where $\bar{u}(t)$ is the Heaviside step function, and $*$ denotes the temporal convolution. In Equations (18) and (19), the spatially varying functions P_0, P_1, P_2, $P_{3,i}$, and $P_{4,i}$ $(i=1,2)$ are introduced as

DOI: 10.1002/nme

follows:

$$P_0 = \kappa_1 \kappa_2 \tag{20a}$$

$$P_1 = \kappa_2 \beta_1 + \kappa_1 \beta_2 \tag{20b}$$

$$P_2 = \beta_1 \beta_2 - \beta_1 \alpha_1 \kappa_2 - \beta_2 \alpha_2 \kappa_1 \tag{20c}$$

$$P_{3,1} = \frac{\beta_1[\kappa_2(\alpha_1 - \alpha_2) - \beta_2]}{\alpha_1 - \alpha_2}, \quad P_{3,2} = \frac{\beta_2[\kappa_1(\alpha_2 - \alpha_1) - \beta_1]}{\alpha_2 - \alpha_1} \tag{20d}$$

$$P_{4,1} = \frac{\beta_1[\kappa_1\kappa_2(\alpha_1 - \alpha_2) + \kappa_2\beta_1 - \kappa_1\beta_2]}{\kappa_1^2[\kappa_1(\alpha_1 - \alpha_2) + \beta_1]}, \quad P_{4,2} = \frac{\beta_2[\kappa_1\kappa_2(\alpha_2 - \alpha_1) + \kappa_1\beta_2 - \kappa_2\beta_1]}{\kappa_2^2[\kappa_2(\alpha_2 - \alpha_1) + \beta_2]} \tag{20e}$$

Note that $P_0 = 1$ and the remaining functions in Equations (20) vanish in the computational domain Ω_S, since $\kappa_i = 1$ and $\beta_i = \alpha_i = 0$ $(i = 1, 2)$. The distribution of κ_i, β_i, and α_i $(i = 1, 2)$ in the PML region appears in Figure 1. Special considerations are needed in the corner regions of the PML, in particular when $\alpha_1 = \alpha_2$. In this case $P_{3,1}$ and $P_{3,2}$ simplify to $P_{3,1} = \beta_1 \kappa_2$ and $P_{3,2} = \beta_2 \kappa_1$, respectively. Additionally, we have that $\beta_1 = \beta_2$ and $\kappa_1 = \kappa_2$, from which the operators in Equations (19) simplify to $\mathscr{L}_1(t) = \mathscr{L}_2(t) = 1$, while $\mathscr{L}_0(t)$ remains unchanged. Note, by choosing the spatial variation of κ_i, α_i, and β_i $(i = 1, 2)$ thoughtfully one avoids the singularities of $P_{4,i}$ $(i = 1, 2)$ when $\kappa_2(\alpha_1 - \alpha_2) + \beta_2$ and $\kappa_1(\alpha_2 - \alpha_1) + \beta_1$ vanish.

2.3. Finite element implementation

For a displacement-based FE solution, the weak form of Equation (17) is required. This solution is sought by using the principle of virtual work over the domain Ω with boundary Γ. We introduce, as in [39], the space of admissible kinematic displacement field defined as $\mathcal{K}^t = \{\boldsymbol{u}(\boldsymbol{x}, t) : \Omega \times \mathbb{I} \to \mathbb{R}^{d+1} | \boldsymbol{u} \in H^1(\Omega) \ \forall t \in \mathbb{I}\}$, where $H_0^1(\Omega)$ is the Hilbert space of vector functions and their weak partial derivatives that are square integrable on Ω and vanish at Γ. Then, by associating the space of admissible displacement variations at a given time t, $\delta\mathcal{K} = \{\delta\boldsymbol{u}(\boldsymbol{x}) : \Omega \to \mathbb{R}^d | \delta\boldsymbol{u} \in H_0^1(\Omega)\}$, the variational problem consists of finding $\boldsymbol{u} \in \mathcal{K}^t$, such that $\forall \delta\boldsymbol{u} \in \delta\mathcal{K}$ and $\forall t \in \mathbb{I}$

$$\int_\Omega (\delta\boldsymbol{u})^{\mathrm{T}} \rho \mathscr{L}_0 \boldsymbol{u} \mathrm{d}\Omega + \int_\Omega (\partial\delta\boldsymbol{u})^{\mathrm{T}} \boldsymbol{\sigma}'_{\mathrm{v}} \mathrm{d}\Omega - \int_\Omega (\delta\boldsymbol{u})^{\mathrm{T}} \boldsymbol{p} \mathrm{d}\Omega - \int_\Gamma (\delta\boldsymbol{u})^{\mathrm{T}} \boldsymbol{T} \mathrm{d}\Gamma = 0 \tag{21}$$

where $\boldsymbol{\sigma}'_{\mathrm{v}}$ follows from Equation (17b). The contour integral in Equation (21) vanishes, since we, for the sake of simplicity, assume a free-surface boundary condition on Γ, i.e. the traction is $\boldsymbol{T} = \boldsymbol{0}$. The system is loaded through the body force $\boldsymbol{p}$.

Now, we expand the displacement field at any point within the element as follows [49, 50]:

$$\boldsymbol{u} \approx \sum_{k=1}^{m_e} N_k \mathbf{I} \mathbf{d}_k^e = [N_1\mathbf{I}, \ldots, N_{m_e}\mathbf{I}] \begin{Bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_{m_e} \end{Bmatrix}^e = \mathbf{N}\mathbf{d}^e \tag{22}$$

where $\mathbf{d}_k^e = (d_1, d_2)^{\mathrm{T}}$ contains the discrete displacement values at node k inside element e with m_e nodes, $\mathbf{I}$ denotes the 2×2 identity matrix, and N_k is the nodal-based shape function. If we then separate the convolution terms in the operators $\mathscr{L}_0$, $\mathscr{L}_1$, and $\mathscr{L}_2$ we obtain the ordinary differential equation

$$\mathbf{M}\ddot{\mathbf{d}} + \mathbf{Z}\dot{\mathbf{d}} + \mathbf{K}\mathbf{d} + \mathbf{h} + \mathbf{g} = \mathbf{f} \tag{23}$$

DOI: 10.1002/nme

where the system matrices $\mathbf{M}$, $\mathbf{Z}$, and $\mathbf{K}$ are assembled from their respective element-level constituent square matrices $\mathbf{M}^e$, $\mathbf{Z}^e$, and $\mathbf{K}^e$, whose sub-matrix entries are given by

$$\mathbf{M}_{ij}^e = \int_{\Omega^e} \rho P_0 \mathbf{N}_i^{\mathrm{T}} \mathbf{N}_j \mathrm{d}\Omega \tag{24a}$$

$$\mathbf{Z}_{ij}^e = \int_{\Omega^e} \rho P_1 \mathbf{N}_i^{\mathrm{T}} \mathbf{N}_j \mathrm{d}\Omega \tag{24b}$$

$$\mathbf{K}_{ij}^e = \int_{\Omega^e} \mathbf{B}_i^{\mathrm{T}} \left(\mathbf{C}_0 + \frac{\kappa_1}{\kappa_2}\mathbf{C}_1 + \frac{\kappa_2}{\kappa_1}\mathbf{C}_2 \right) \mathbf{B}_j \mathrm{d}\Omega + \int_{\Omega} \rho P_2 \mathbf{N}_i^{\mathrm{T}} \mathbf{N}_j \mathrm{d}\Omega \tag{24c}$$

Here, $\mathbf{B}_i = \partial \mathbf{N}_i$ denotes the strain–displacement matrix, and $\mathbf{d}$ is the unknown global vector with the discrete displacement values, i.e. $\mathbf{d} = (\mathbf{d}_1, \ldots, \mathbf{d}_M)^{\mathrm{T}}$, where M is the total number of nodes. The force vector $\mathbf{f}$ in Equation (23) is determined by

$$\mathbf{f}_i^e = \int_{\Omega^e} \mathbf{N}_i^{\mathrm{T}} \boldsymbol{p} \mathrm{d}\Omega \tag{25}$$

Further, in the convolution vectors $\mathbf{h} = \sum_e \mathbf{h}^e$ and $\mathbf{g} = \sum_e \mathbf{g}^e$ in Equation (23) the element contributions are, respectively, given by individual sub-vectors

$$\mathbf{h}_i^e = \int_{\Omega^e} \rho \mathbf{N}_i^{\mathrm{T}} \sum_j [\mathbf{N}_j (\alpha_1^2 P_{3,1} \boldsymbol{u}_{1,j}^+ + \alpha_2^2 P_{3,2} \boldsymbol{u}_{2,j}^+)] \mathrm{d}\Omega \tag{26a}$$

$$\mathbf{g}_i^e = \int_{\Omega^e} \mathbf{B}_i^{\mathrm{T}} \sum_j \left[\mathbf{C}_1 \mathbf{B}_j \left(\frac{\beta_1^2}{P_{3,1}} \mathbf{u}_{1,j}^+ - P_{4,2} \mathbf{u}_{2,j}^{++} \right) \right. \tag{26b}$$

$$\left. + \mathbf{C}_2 \mathbf{B}_j \left(\frac{\beta_2^2}{P_{3,2}} \mathbf{u}_{2,j}^+ - P_{4,1} \mathbf{u}_{1,j}^{++} \right) \right] \mathrm{d}\Omega \tag{26c}$$

where $\mathbf{u}_{p,j}^+ = \mathrm{e}^{-\alpha_p t} \bar{u}(t) * \mathbf{d}_j$ and $\mathbf{u}_{p,j}^{++} = \mathrm{e}^{-(\beta_p + \kappa_p \alpha_p)t/\kappa_p} \bar{u}(t) * \mathbf{d}_j$. The computational disadvantage of the occurrence of the convolution terms in Equation (26) is the need for their successive evaluation at each time steps. However, since all spatial element information can be computed prior to time stepping and subsequently stored it is only necessary to update the global auxiliary memory vectors $\mathbf{u}_p^+$ and $\mathbf{u}_p^{++}$ $(p = 1, 2)$ in the integration points of the element. The representation in Equation (26) has to be followed when large element sizes appear in the mesh, because then the assumption of constant variation of the PML parameters β_p, κ_p, and α_p $(p = 1, 2)$ will not suffice anymore [51].

Meanwhile, introducing the requirement of constant PML parameters inside each element, e.g. for bilinear shape functions, simplifies the convolution vectors in the resultant ordinary differential equation that now reads as

$$\mathbf{M}\ddot{\mathbf{d}} + \mathbf{Z}\dot{\mathbf{d}} + \mathbf{K}\mathbf{d} + \widetilde{\mathbf{h}} + \widetilde{\mathbf{g}} = \mathbf{f} \tag{27}$$

In the computation of the modified convolution vectors $\widetilde{\mathbf{h}}$ and $\widetilde{\mathbf{g}}$ there is no need for spatial integration each time step anymore. As a result, their expressions reduce to

$$\widetilde{\mathbf{h}}_i^e = \sum_j \left(\mathbf{M}_{1,ij}^e + \frac{\beta_1^2}{P_{3,1}} \mathbf{K}_{1,ij}^e \right) \mathbf{u}_{1,j}^+ + \left(\mathbf{M}_{2,ij}^e + \frac{\beta_2^2}{P_{3,2}} \mathbf{K}_{2,ij}^e \right) \mathbf{u}_{2,j}^+ \tag{28a}$$

$$\widetilde{\mathbf{g}}_i^e = \sum_j \mathbf{K}_{2,ij}^e P_{4,1} \mathbf{u}_{1,j}^{++} + \mathbf{K}_{1,ij}^e P_{4,2} \mathbf{u}_{2,j}^{++} \tag{28b}$$

DOI: 10.1002/nme

where the element matrices $\mathbf{M}_p$ and $\mathbf{K}_p$ $(p=1,2)$ are given by

$$\mathbf{M}^e_{p,ij} = \int_\Omega \rho\alpha_p^2 P_{3,p}\mathbf{N}_i^{\mathrm{T}}\mathbf{N}_j \mathrm{d}\Omega \tag{29a}$$

$$\mathbf{K}^e_{p,ij} = \int_\Omega -\mathbf{B}_i^{\mathrm{T}}\mathbf{C}_p\mathbf{B}_j \mathrm{d}\Omega \tag{29b}$$

Here, it is crucial to stress that $\mathbf{u}^{+}_{p,j}$ and $\mathbf{u}^{++}_{p,j}$ are local meaning that they are defined on each element of the mesh and are discontinuous from one element to another, since the PML parameters are now element-wise constant. In practice this implies that e.g. for Q nodal quadrilateral elements the integration requires the update and storage of $8Q$ memory variables. The following section will present a recursive update technique that effectively avoids integration of the entire solution history at each time step in the evaluation of the convolution terms when progressing Equations (23) and (27) in time. Finally, on the exterior of the computational domain the homogeneous Dirichlet boundary condition is imposed, i.e. $\mathbf{d}=\mathbf{0}$, as this seems to stabilize the time integration for long times of resolution [52].

2.4. Time integration

In the time marching of the FE discretized wave equations (23) or (27) we approximate the spatial derivatives and non-derivatives by the Newmark-beta time integration method [53]. As a result, the elastic displacement field is progressed in time as follows:

$$\begin{aligned}
&\left(\frac{1}{\Delta t^2}\mathbf{M}+\frac{1}{2\Delta t}\mathbf{Z}+\beta\mathbf{K}\right)\mathbf{d}^{n+1}\\
&\quad=\left(\frac{2}{\Delta t^2}\mathbf{M}-(1-2\beta)\mathbf{K}\right)\mathbf{d}^{n}-\left(\frac{1}{\Delta t^2}\mathbf{M}-\frac{1}{2\Delta t}\mathbf{Z}+\beta\mathbf{K}\right)\mathbf{d}^{n-1}-\beta\mathbf{g}^{n+1}-(1-2\beta)\mathbf{g}^{n}-\beta\mathbf{g}^{n-1}\\
&\qquad-\beta\mathbf{h}^{n+1}-(1-2\beta)\mathbf{h}^{n}-\beta\mathbf{h}^{n-1}+\beta\mathbf{f}^{n+1}+(1-2\beta)\mathbf{f}^{n}+\beta\mathbf{f}^{n-1}
\end{aligned} \tag{30}$$

where Δt is the time step. When $\beta=0$, the formula reduces to explicit time integration. However, it can be shown that the integration scheme becomes unconditionally stable for non-active, and non-dispersive media when $\beta\geqslant 1/4$ with second-order accuracy preserved [54]. Thus, unlike explicit methods in which the choice of Δt is bounded by the CFL condition the implicit method has no limiting time step. The price for this convenient assurance in the time marching is the necessity of solving a linear system at each time step. The important advantage is that the choice of Δt is not limited by the smallest FE in highly irregular mesh geometries through the CFL condition. Instead, it depends solely on the accuracy with which the temporal variation of the field is reproduced. In practical cases the elastodynamic problems often involve large 3D domains, in which cases matrix inversion is not possible. If implicit time integration is still required then a direct solution of the linear matrix system needs to be substituted by iterative methods that are rather involved to parallelize. Alternatively, explicit methods with mass lumping could be used to avoid matrix inversion, and they are also relatively straightforward to parallelize based on a non-blocking communication strategy, because matrix–vector multiplication can be performed locally. The scalability is believed to be linear. With the proposed method explicit time integration with mass lumping is achieved for $\beta=0$ and by computing the mass matrix as $\mathbf{M}^e_{ii}=\sum_j \mathbf{M}^e_{ij}$ and $\mathbf{M}^e_{ij}=\boldsymbol{0}$ for $i\neq j$ [49]. To be consistent all mass-like terms in Equations (24) and (29) should be evaluated similarly. Notice that for non-dispersive media the choice of time step is bounded by the CFL conditions for explicit time integration.

We have learned from the derivation in Section 2.2, that, regardless of the PML technique, the presence of the convolution term in the time-domain equations (17) is inevitable as soon as

DOI: 10.1002/nme

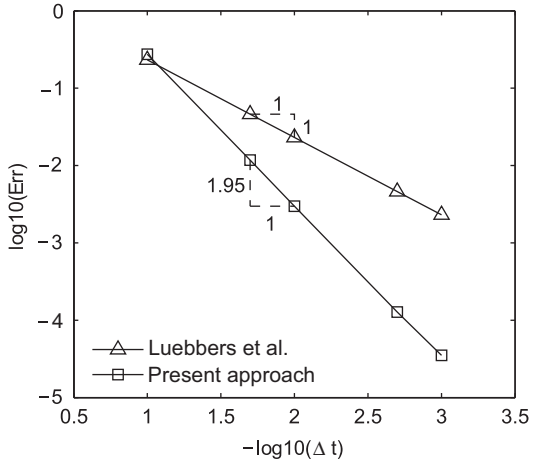


Figure 3. Small example that illustrates the convergence of the recursive convolution update scheme in Equation (36) compared to the scheme of Luebbers *et al.* when $K(t)=\mathrm{e}^{-b\Delta t}\bar{u}(t)$ and $u(t)=\sin(2\pi t)$. The analytical convolution is $\psi_a(t)=\bar{u}(t)[2\pi(\mathrm{e}^{-b\Delta t}-\cos(2\pi t))+b\sin(2\pi t)]/(b^2+4\pi^2)$. We define the error measure as $\mathrm{Err}=||\psi_a(t)-\psi_r(t)||/||\psi_a(t)||$, where $||\cdot||$ is the Euclidean norm.

the governing equation in the PML region needs to be recast in a format that only contains the displacement as unknown. The general appearance of the convolution term is

$$\begin{aligned}\psi(t) &= [a\mathrm{e}^{-bt}\bar{u} * u](t)\\ &= \int_0^t a\mathrm{e}^{-b\tau}\bar{u}(\tau)u(t-\tau)\mathrm{d}\tau\end{aligned} \tag{31}$$

where a, b are real constants, and u is the state field variable of analysis. Time convolution is by definition a computational burden since it requires storage of the entire solution history in the time marching. Fortunately, however, it has been shown by Luebbers and Hunsberger [21] that with a kernel function on a simple exponential form $K(t)=a\mathrm{e}^{-bt}\bar{u}(t)$ it can be recursively updated, and thus only requires information from the previous time step. They assume that the field of analysis is constant over $n\Delta t$ through the point $(n\Delta t, u(n\Delta t))$ yielding first-order accuracy. We have found that if $\psi(t)$ is still discretized through the same point, but one half time step earlier than $u(t)$ in the time marching, i.e. $u(t)$ is now constant in $[(n-1/2)\Delta t, (n+1/2)\Delta t]$ in Equation (31), the accuracy of the recursive update improves to second-order, see Figure 3. Thus, it aligns with the intrinsic second-order accuracy of the Newmark-beta method. For this case, the convolution integral is discretely evaluated as [55]

$$\begin{aligned}\psi^n &= \sum_{k=0}^{n} u^{n-k}\int_{(k-1/2)\Delta t}^{(k+1/2)\Delta t} K(\tau)\mathrm{d}\tau\\ &= u^n\int_0^{\Delta t/2} K(\tau)\mathrm{d}\tau + \sum_{k=0}^{n-1} u^{n-k-1}\int_{(k+1/2)\Delta t}^{(k+3/2)\Delta t} K(\tau)\mathrm{d}\tau\end{aligned} \tag{32}$$

Here, the integrals for $b\neq 0$ are given by

$$\zeta^0 = \int_0^{\Delta t/2} K(\tau)\mathrm{d}\tau = \frac{a}{b}[1-\mathrm{e}^{-b\Delta t/2}] \tag{33a}$$

$$\zeta^{k+1/2} = \int_{(k+1/2)\Delta t}^{(k+3/2)\Delta t} K(\tau)\mathrm{d}\tau = \frac{a}{b}[1-\mathrm{e}^{-b\Delta t}]\mathrm{e}^{-b(k+1/2)\Delta t} \tag{33b}$$

DOI: 10.1002/nme

and for $b=0$ they reduce to $\zeta^0=a\Delta t/2$ and $\zeta^{k+1/2}=a\Delta t$, respectively. For convenience, we introduce Ψ^n for the last term in Equation (32) that is now rewritten to

$$\Psi^n=\sum_{k=0}^{n-1}u^{n-k-1}\zeta^{k+1/2}$$

$$=\zeta^{1/2}u^{n-1}+\sum_{k=0}^{n-2}u^{n-k-2}\zeta^{k+3/2} \tag{34}$$

and since $\zeta^{k+3/2}=\mathrm{e}^{-b\Delta t}\zeta^{k+1/2}$ we realize that

$$\sum_{k=0}^{n-2}u^{n-k-2}\zeta^{k+3/2}=\mathrm{e}^{-b\Delta t}\sum_{k=0}^{n-2}u^{n-k-2}\zeta^{k+1/2} \tag{35}$$

which leads to the recurrence relation

$$\Psi^n=\zeta^{1/2}u^{n-1}+\mathrm{e}^{-b\Delta t}\Psi^{n-1} \tag{36}$$

Summing up the above algebraic manipulations, the evaluation of ψ^n boils down to the following two-step process:

$$\Psi^n=\zeta^{1/2}u^{n-1}+\mathrm{e}^{-b\Delta t}\Psi^{n-1} \tag{37a}$$

$$\psi^n=\zeta^0u^n+\Psi^n \tag{37b}$$

where ζ^0 and $\zeta^{1/2}$ are given by Equation (33). The recurrence relation in Equation (36) saves substantial computation time and memory. Evidently, we only need information about u^n and values from the previous time step to compute ψ^n. Meanwhile, the cornerstone of the unconditionally stable Newmark-beta method relies on the weighted average approximation of the non-differentiated term, i.e. $\psi(t)=\beta\psi^{n+1}+(1-2\beta)\psi^n+\beta\psi^{n-1}$. Since ψ^{n+1} requires u^{n+1} there exists information of the unknown displacement field towards which we progress in time on the right-hand side of Equation (31). Hence, unconditionally stable time marching is only achieved when all terms containing u^{n+1} are moved to the left-hand side. Since the unconditional stability is only strictly valid for non-dispersive media, it cannot be guaranteed to hold in the PML region. However, numerous tests have shown that the suggested scheme is highly stable.

Compared to some FE-based non-split mixed formulations, see e.g. [33], the proposed displacement-based FE scheme reduces the number of unknowns from $5M$ (u_1, u_2, σ_{11}, σ_{22}, and σ_{12}) to $2M$ (u_1, and u_2) in 2D; in 3D $9M$ unknowns are reduced to $3M$. The algorithmic features of the proposed scheme are symmetric system matrices that need to be stored in conjunction with the solution $\mathbf{d}$ and the auxiliary memory variables $\mathbf{u}_p^+$ and $\mathbf{u}_p^{++}$ ($p=1,2$ in 2D) in the convolution terms, which essentially act as forcing terms that can be recursively updated. Furthermore, it supports an easy switch between implicit and explicit time integration with mass lumping.

3. NUMERICAL EXPERIMENTS

In this section, we conduct numerical experiments that demonstrate the excellent absorbing properties of the suggested formulation. We limit the study to homogeneous material hosts, such that the medium only contains pure pressure, shear, and surface waves [26]. More importantly, the convenience of using homogeneous media is the existence of an analytical solution, against which we validate the numerical solution from the proposed formulation. It should be emphasized that the primary scope of this paper is merely to demonstrate the performance of C-PML when it is employed directly to the wave equation recast as a second-order system with displacement as the single unknown field. This without the need of introducing a mixed formulation approach although it is claimed inconceivable in [33]. The first example, which is the simplest, involves an

DOI: 10.1002/nme

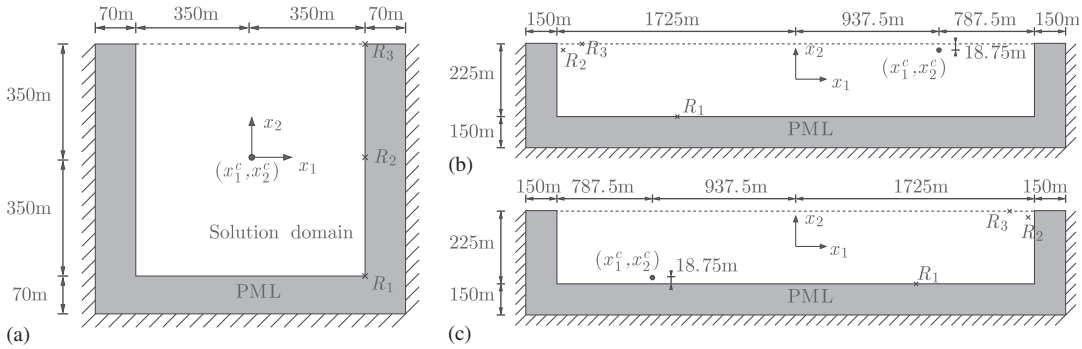


Figure 4. PML-truncated semi-infinite domains in two dimensions subjected to a vertically downward point-source located at (x_1^c, x_2^c) driven by an explosive Ricker wavelet. R_1, R_2, and R_3 indicate the location of the receivers at which the wave response is recorded: (a) experiment 1; (b) experiment 2; and (c) experiment 3. The domains are not illustrated to scale.

explosive point-source residing inside a isotropic semi-infinite domain, see Figure 4(a). Examples 2 and 3 in Figure 4(b) and (c), respectively, address the issue that near-grazing waves often exhibit degrading PML-performance [19, 25–27, 56] by using elongated domains with wave sources near the free surface and the bottom-PML interface. The FORTRAN code EX2DDIR of Berg *et al.* [57] has been used to compute the exact analytical solution of the responses that originate from a vertical directional point-source in a two-dimensional isotropic elastic half-space with a free surface.

The material properties of the elastic isotropic half-plane considered in the three cases are density $\rho = 2000\,\mathrm{kg/m^3}$, longitudinal wave velocity $c_p \approx 774.6\,\mathrm{m/s}$, and shear wave velocity $c_s = c_p/2$ in a state of plane strain. The explosive point-source is defined as

$$\boldsymbol{p}(\boldsymbol{x}, t) = T(t)\delta(\boldsymbol{x} - \boldsymbol{x}^c)\boldsymbol{d}(\theta) \tag{38}$$

where δ denotes the Dirac delta function, and $\boldsymbol{x}^c$ is the location of the point-source with temporal evolution $T(t)$. The unit vector $\boldsymbol{d} = (\cos\theta, \sin\theta)^{\mathrm{T}}$ prescribes the direction of the source, where the angle θ is measured counterclockwise from the x_1 axis. For comparison with the analytical solution we choose the angle $\theta = -\pi/2$. As temporal variation of the source a modified Ricker wavelet is employed, which is given by

$$T(t) = (0.5 + a_1(t - t_d)^2)\mathrm{e}^{a_1(t-t_d)^2} \tag{39}$$

where t_d is the source delay time, $a_1 = -(\pi f_c)^2$, and $f_c = 1.0\,\mathrm{Hz}$ is the dominant frequency of the source.

Based on the stretched-coordinate metrics, the spatial dependence of the CFS PML parameters is chosen in the x_i direction as follows [33, 35, 39, 40, 58]:

$$\kappa_i = 1 + \kappa_{\max,i}\left(\frac{x_i - x_i^0}{d_i}\right)^{n_1} \tag{40a}$$

$$\beta_i = \beta_{\max,i}\left(\frac{x_i - x_i^0}{d_i}\right)^{n_1+n_2} \tag{40b}$$

$$\alpha_i = \alpha_{\max,i}\left(\frac{d_i - x_i + x_i^0}{d_i}\right)^{n_3} \tag{40c}$$

DOI: 10.1002/nme

with $\alpha_{\max,i}=2\pi f_\alpha$ and

$$\beta_{\max,i}=-\frac{(1+n_1+n_2)c_p\log_{10}(R_0)}{2d_i} \tag{40d}$$

$$\kappa_{\max,i}=-\frac{(1+n_1)b_i\log_{10}(R_0)}{2d_i} \tag{40e}$$

for $i=1,2$. Here, R_0 is the theoretical reflection coefficient at normal incidence, b_i is a characteristic length of the domain (e.g. integer multiples of the element size), x_i^0 and d_i $(i=1,2)$ are the starting location and the thickness of the PML layer, respectively. Further, f_α is a characteristic cut-off frequency that designs the absorbing behavior of the C-PML as described in Section 1.1. In all experiments, we have, unless otherwise stated, chosen the following values for the profile parameters: $f_\alpha=0.5$ [39, 42], $b=10\Delta x_i$, $R_0=10^{-8}$ [33], $n_1=3$, $n_2=0$, and $n_3=1$. This set of variables does not induce any singularities for $P_{4,i}$ in Equation (20e), and for $1/P_{3,i}$ $(i=1,2)$ in Equations (26) and (28).

For the discretization of the PML and the solution domain quadrilateral bilinear FEs are employed. It follows naturally from this choice of element types that the PML parameters in Equations (40) are assumed constant inside each element which leads to the use of the second-order ordinary differential equation in Equation (27). In the time integration we use the implicit version of the Newmark-beta method, i.e. $\beta=1/4$ to have unconditionally stable time marching. Although the time step is merely free of choice with regards to stability considerations it is essential to choose a time step that reproduces the temporal variation of the source with an accuracy such that the numerical and exact solution will be comparable. We have found that a suitable choice is based on the CFL condition

$$\Delta t_c=\min\left(\frac{1}{c_p},\frac{1}{c_s}\right)\left(\frac{1}{\Delta x_1^2}+\frac{1}{\Delta x_2^2}\right)^{-1/2} \tag{41}$$

For equisized elements the criterion reduces to $\Delta t_c=\Delta x_1\min(c_p^{-1},c_s^{-1})/\sqrt{2}$. The FE scheme is implemented in the Matlab programming environment. However, for the recursive update scheme in Equation (37), which needs to be carried out separately for the nodal displacements of each element, we make Mex function calls to avoid the latency from which Matlab inherently suffers when executing FE assembly procedures in for-loops.

3.1. Experiment 1

The numerical analysis is conducted by reducing a semi-finite elastic medium to a $350\,\text{m}\times 350\,\text{m}$ computational domain surrounded by 70 m thick PML layer, see Figure 4(a). The element edge size is $\Delta x_1=\Delta x_2=3.6458\,\text{m}$ which leads to a time step $\Delta t_c=0.0033\,\text{s}$, and the simulation runs for 10.0 s requiring 3030 time steps for completion. The excitation is driven by the Ricker wavelet source in Equation (39) that is located at the center of the solution domain and with a delay $t_d=1.0\,\text{s}$. We record the horizontal and vertical transient responses in three receiving positions R_1, R_2, and R_3 located at $(350\,\text{m},-350\,\text{m})$, $(350\,\text{m},0\,\text{m})$, and $(350\,\text{m},350\,\text{m})$, respectively. The monitored time evolutions appear in Figure 5, and they behave perfect in accordance with the analytical solution. In all plots, the responses based on the standard PML implementation are also included, i.e. for $\alpha_i=0$, and $\kappa_i=1$ $(i=1,2)$ in Equations (20). Clearly, the standard PML and C-PML strategies perform almost equally. This congruence is a result of the point-source settings that do not generate strong evanescent waves in the PML layer, in which case the C-PML outperforms the classical PML approach.

Figure 6 shows the snapshots of the displacement field magnitude taken at four different time steps. After approximately 1.0 s it becomes clear, that two primary wave trains (the fastest associated to the P-carrier followed by a stronger S) have been created by the directional point-source. Both

DOI: 10.1002/nme

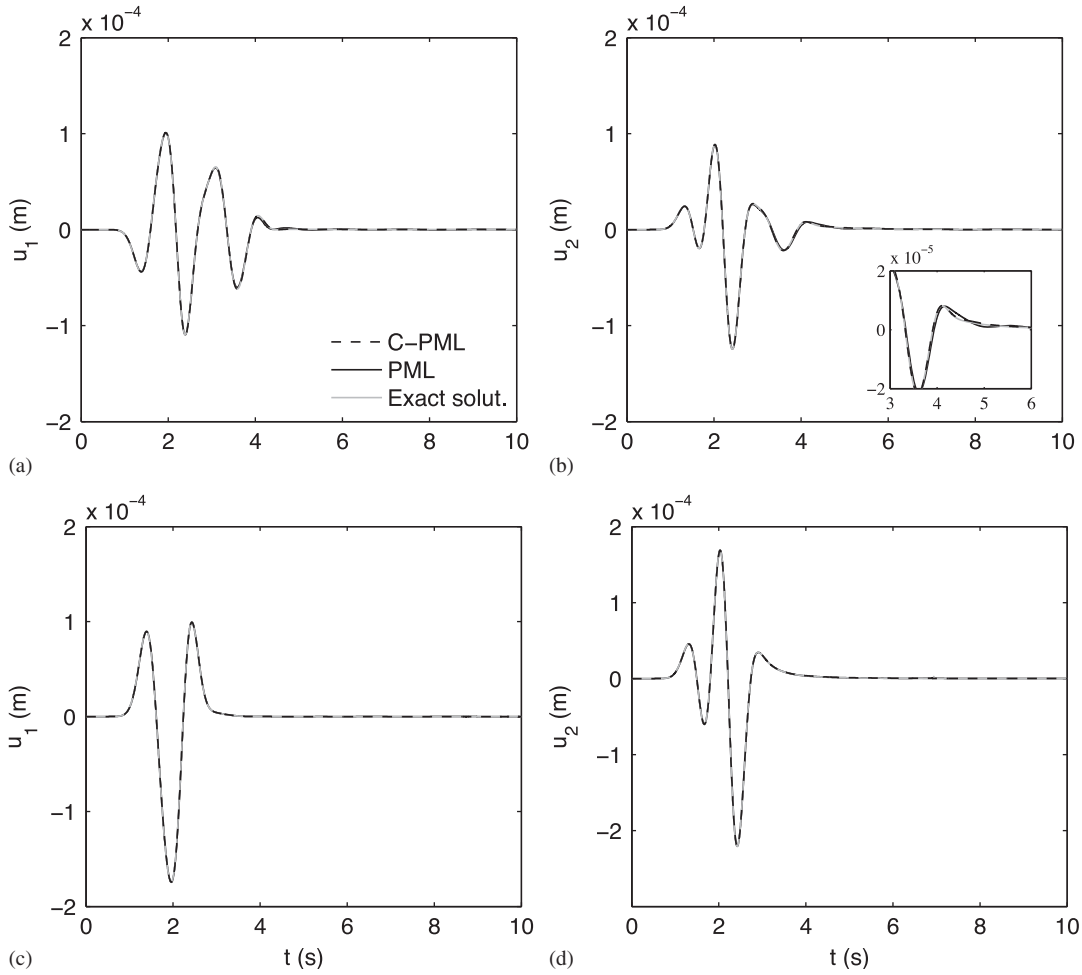


Figure 5. Experiment 1: C-PML, standard PML, and analytical transient response comparison of the horizontal u_1 (left) and the vertical u_2 (right) at three receiving positions (a)–(b) R_1, and (c)–(d) R_3 located at (350 m, −350 m), and (350 m, 350 m), respectively.

waves propagate to the sides and towards the bottom of the truncated domain and penetrate the PML at normal incidence. However, due to momentum flux conservation, an upward directional field appears which will split into reflected bulk and Rayleigh waves at the free surface, see Figure 6(b) and (c). Frankly, the u_2 response at R_3 in Figure 5(f) indicates a weak S wave around 1.5 s prior to a much stronger Rayleigh wave around 2.0 s. According to Figure 6(d), only bulk waves remain in the domain after 2.8 s. They travel at grazing incidence towards the bottom as P- and S-waves (compare Figure 5(a) and (b)), and create a weak evanescent field in the PML side regions. This phenomenon explains the small deviations in Figure 5(b) of the standard PML response at the last receiving position R_1 before the waves leave the computational domain. The standard PML is not capable of attenuating the evanescent field, and while traveling towards the bottom this dispersion error accumulates, which is eventually recorded at R_1.

3.2. Experiment 2

The purpose of this experiment is to address the issue that the standard C-PML suffers from poor performance when an evanescent wave field is generated in the boundary layer as a result of near-grazing incidence. It originates from Rayleigh waves penetrating into the domain. To accommodate this scenario, we have reduced a semi-finite domain with a free surface to a 3450 m × 225 m elongated domain truncated by a 150 m thick PML layer, as illustrated in Figure 4(b). Both regions are

DOI: 10.1002/nme

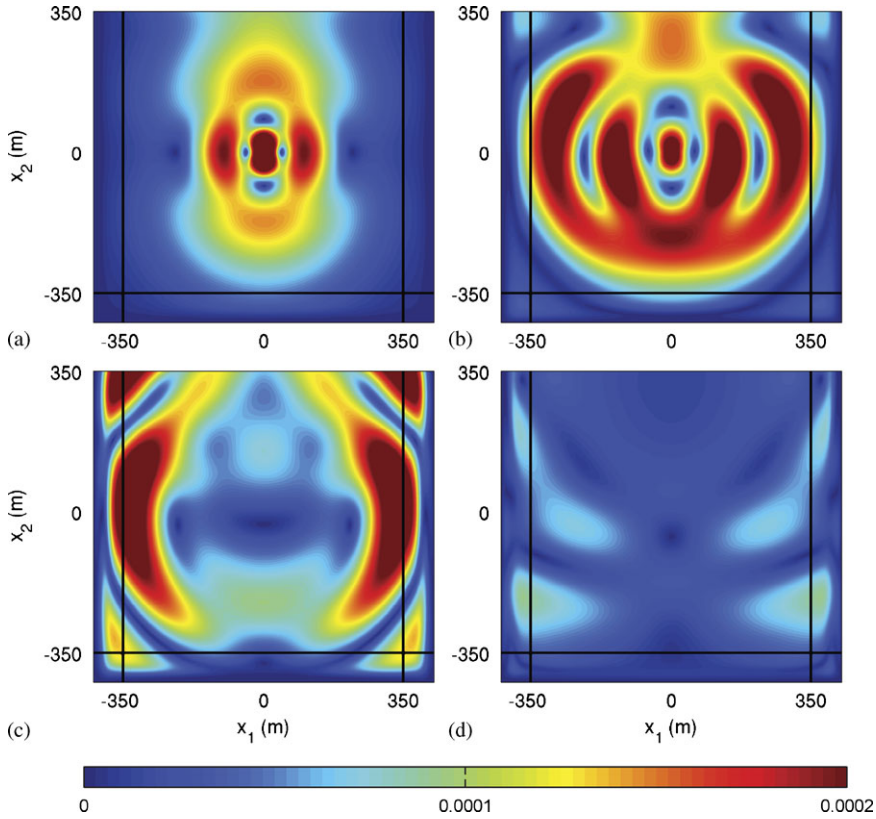


Figure 6. Experiment 1: Snapshots of the displacement field magnitude defined as the norm $||\boldsymbol{u}|| = \sqrt{u_1^2 + u_2^2}$ at time steps: (a) 0.5 s, (b) 1.0 s, (c) 2.0 s, and (d) 2.8 s. The waves are generation by a Ricker wavelet source located at $(x_1^c, x_2^c) = (0\,\text{m}, 0\,\text{m})$.

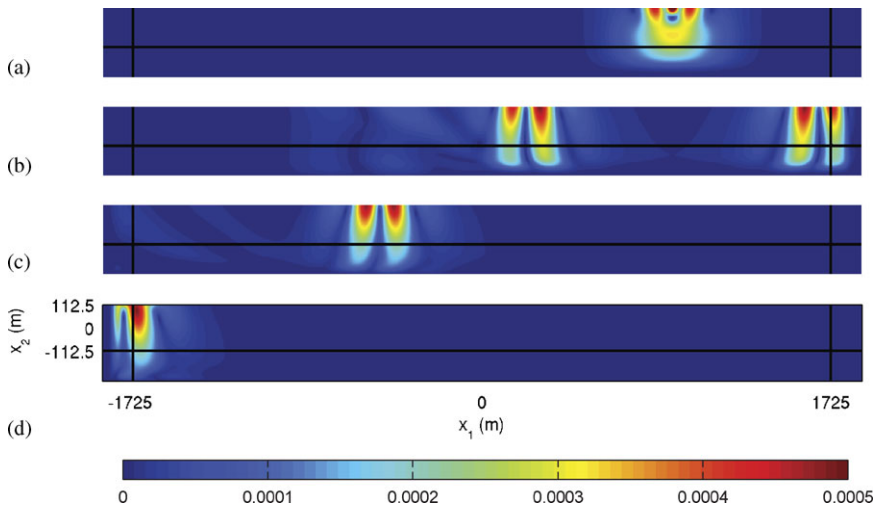


Figure 7. Experiment 2: Snapshots of the displacement field magnitude defined as the norm $||\boldsymbol{u}|| = \sqrt{u_1^2 + u_2^2}$ at time steps: (a) 1.0 s, (b) 3.0 s, (c) 5.0 s, and (d) 8.5 s. The waves are driven by a Ricker wavelet source located at $(x_1^c, x_2^c) = (937.50\,\text{m},\ 93.75\,\text{m})$.

DOI: 10.1002/nme

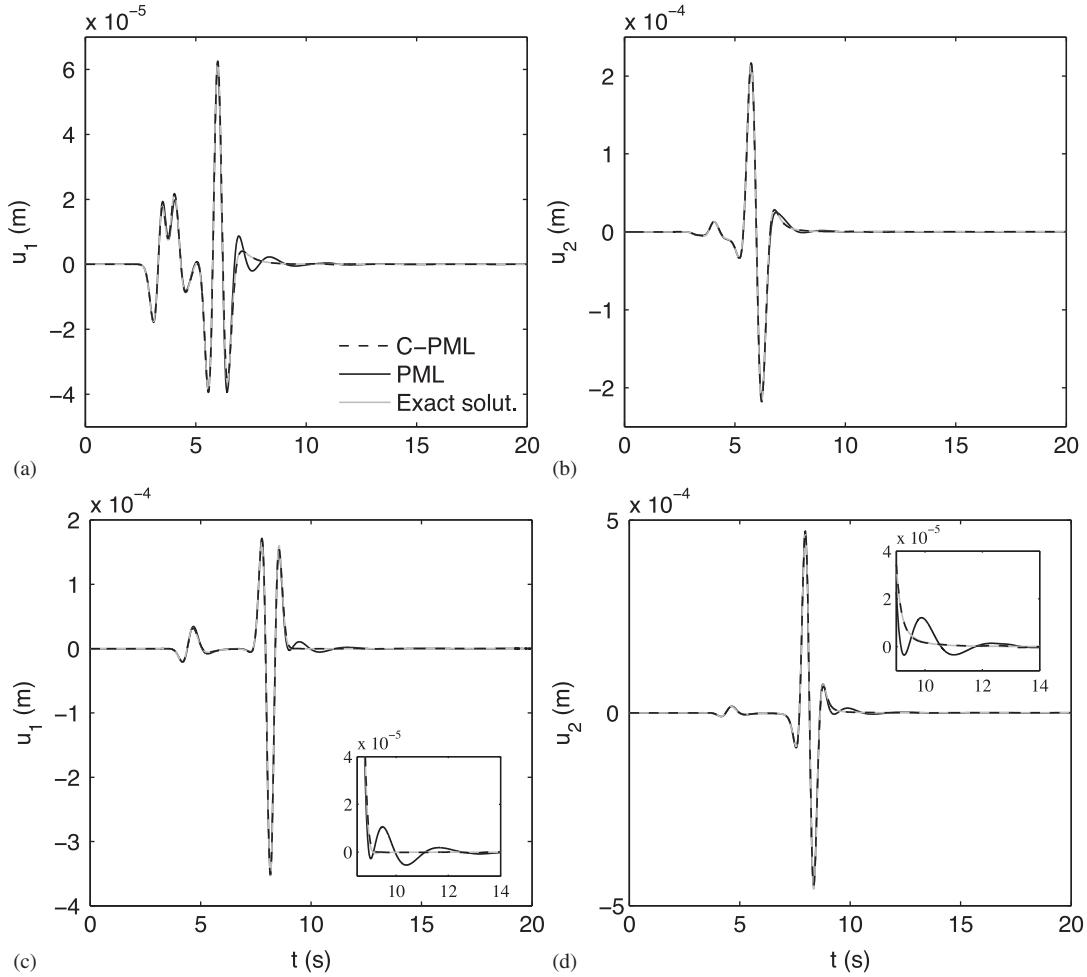


Figure 8. Experiment 2: C-PML, standard PML, and analytical transient response comparison of the horizontal u_1 (left) and the vertical u_2 (right) at three receiving positions (a)–(b) R_1, and (c)–(d) R_3 located at $(-862.5\,\text{m}, -112.5\,\text{m})$, and $(-1650.0\,\text{m}, 112.5\,\text{m})$, respectively.

discretized by equisized quadrilateral elements with edge size $\Delta x_1 = \Delta x_2 = 4.6875\,\text{m}$. We run the simulation for 20.0 s that yields 4674 time steps with a time step $\Delta t = 0.0043\,\text{s}$. The Ricker wavelet source from experiment 1 is reconstituted and is now located at $(937.50\,\text{m}, 93.75\,\text{m})$. We have allocated three receivers at $(-862.5\,\text{m}, -112.5\,\text{m})$, $(-1706.25\,\text{m}, 93.75\,\text{m})$, and $(-1650.0\,\text{m}, 112.5\,\text{m})$ to track u_1 and u_2 responses. In Figure 8 the response records clearly display the poor performance of the classical C-PML compared to the C-PML implementation which are in perfect agreement with the exact solution, even with an inactive evanescent scaling factor, i.e. $\kappa_i = 1$ $(i = 1, 2)$. The degrading functionality of the standard C-PML is magnified in the inset of Figure 8(c) and (d), and it is primarily caused by a strong evanescent field that for geometrical reasons does not penetrate the PML region very deep [27] while the surface wave propagates at grazing incidence to the left, see Figure 7.

As already mentioned the surface waves carry a large amount of the energy identified as the large amplitude u_2 responses in Figure 8(b) and (d) occurring approximately around 5.5 and 7.5, respectively. In addition, before the arrival of the P- (see Figure 8(c)) and Rayleigh waves very weak S-waves are recorded around 4.5 s. It is well known from analytical studies of this Lamb's problem that the energy of the Rayleigh waves eclipses far away from the driving source [59]. Figure 8(a) also indicates that in the vicinity of the bottom PML layer a P-wave appears.

DOI: 10.1002/nme

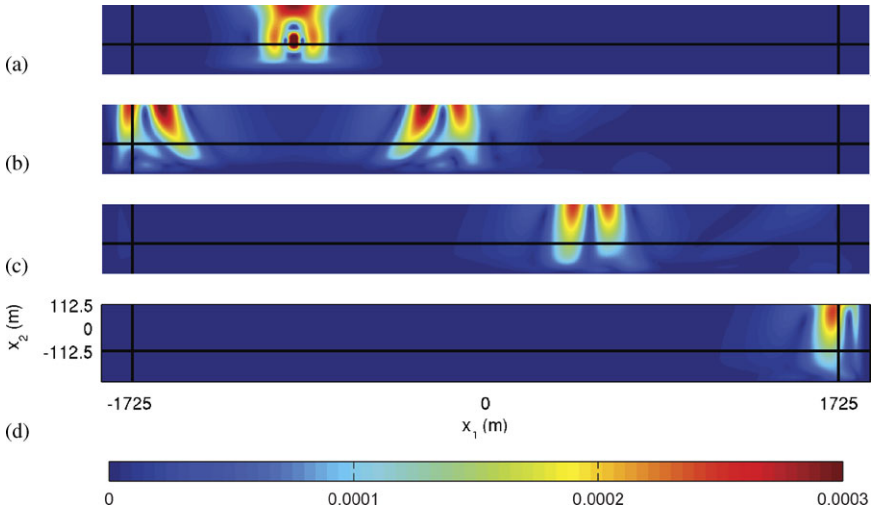


Figure 9. Experiment 3: Snapshots of the displacement field magnitude defined as the norm $\|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2}$ at time steps: (a) 1.0 s, (b) 3.0 s, (c) 5.0 s, and (d) 8.5 s. The waves are driven by a Ricker wavelet source located at $(x_1^c, x_2^c) = (-937.50\,\text{m}, -93.75\,\text{m})$.

3.3. *Experiment 3*

In this experiment explicit time integration with mass lumping is employed. The domain and discretization settings in conjunction with the temporal variation of the directional point-source follow the same choices as for the numerical experiment in Section 3.2. To avoid or delay numerical instabilities, such as nonphysical oscillations and late-time instabilities in the PML we choose the time step to $\Delta t = 0.95\Delta t_c$ yielding 4921 time steps. The location of the source is altered to be near the bottom-PML instead, i.e. $(x_1^c, x_2^c) = (-937.50\,\text{m}, -93.75\,\text{m})$. Consequently, surface and pressure waves are not formed instantly as in experiments 2, but rather after the P- and S-wave have interfered/interacted with the surface. If we consider the snapshots of the displacement magnitude at different time steps in Figure 9, and simultaneously compare Figure 10(a), (c) to (b), (d), respectively, we notice that Rayleigh waves still dominate. The near bottom-PML location of the source gives rise to an inclined wave vector away from the surface at 3.0 s as illustrated in Figure 9(b). The wave vector eventually aligns with the direction parallel to the bottom layer, see Figure 9(c) and (d). Conversely, the wave motion of experiment 2, shown in Figure 7, remains almost perpendicular to a vertical cross section at any time.

In this experiment it is crucial to enforce the scaling of evanescent attenuation by $\kappa_i > 1$ $(i = 1, 2)$. Otherwise, the P-wave response in receiving position R_1 at the bottom PML-interface will be polluted by energy reentering the domain constituted as spurious reflected evanescent waves, as depicted in the inset of Figure 10(a).

3.4. *Late-time stability and energy conservation*

The illustration of the late-time behavior of the proposed PML implementation is based on long-time simulation of numerical experiments 2 and 3 from Sections 3.2 and 3.3, respectively. In particular, we record the total (elastic and kinetic) energy decay inside the physical solution domain Ω_S given by

$$E = \tfrac{1}{2}\dot{\mathbf{d}}^{\mathrm{T}}\mathbf{M}\dot{\mathbf{d}} + \tfrac{1}{2}\mathbf{d}^{\mathrm{T}}\mathbf{K}\mathbf{d} \tag{42}$$

and increase the simulation time from 20 to 320 s. The mathematical proof of unconditional stability is a rather involved task, for which reason we rely entirely on the numerical results. It is a well-established fact that many PML models (e.g in electromagnetics) suffer from strong instabilities at

DOI: 10.1002/nme

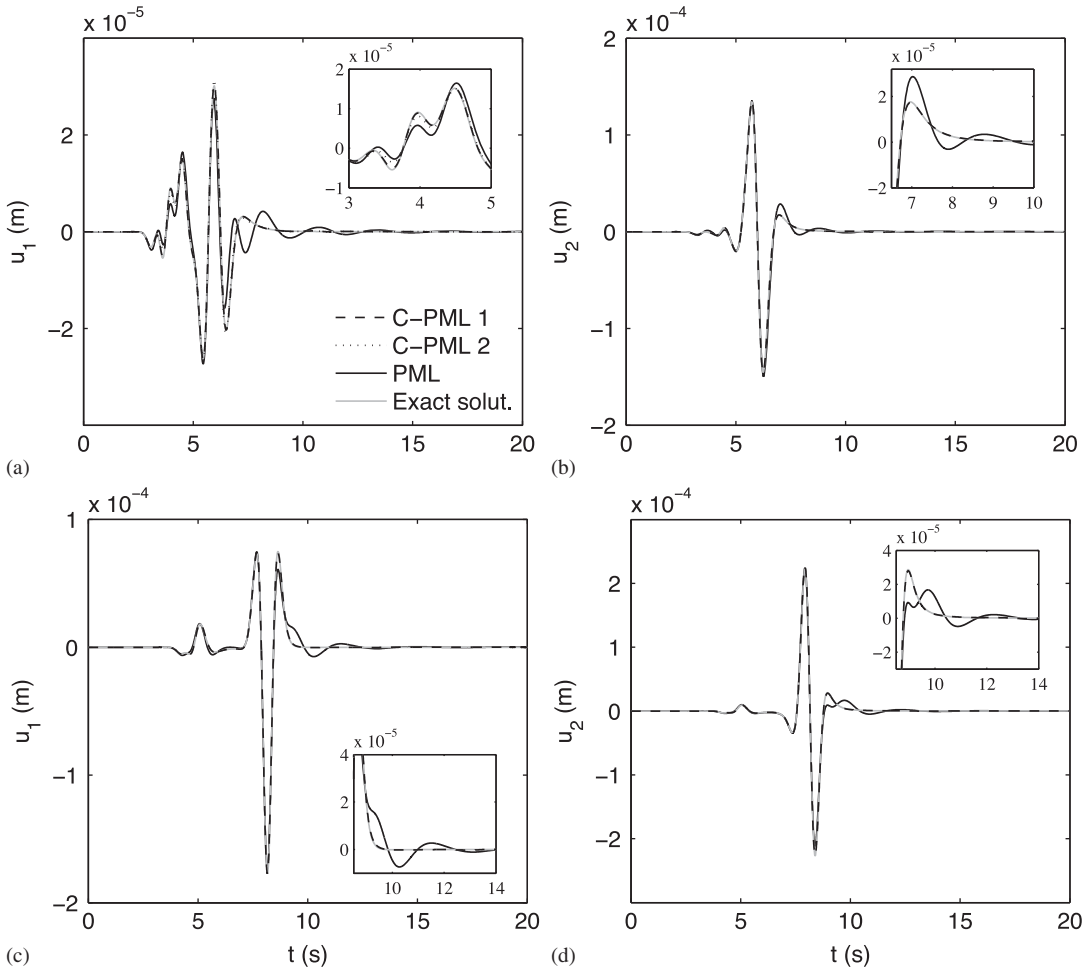


Figure 10. Experiment 3: C-PML, standard PML, and analytical transient response comparison of the horizontal u_1 (left) and the vertical u_2 (right) at three receiving positions (a)–(b) R_1, and (c)–(d) R_3 located at (862.5 m, −112.5 m), and (1650.0 m, 112.5 m), respectively. Note, that the response based on two C-PML models are recorded at R_1: one with (C-PML 1) evanescent scaling active, i.e. $\kappa_i > 1$, and one without (C-PML 2).

late times [60–62]. However, it has also been shown that the inclusion of the CFS filter improves the stability significantly [63, 64].

As previously emphasized, the unconditionally stable Newmark-beta method only applies to non-active, non-dissipative media when $\beta \geqslant 1/4$ [54]. One should therefore be careful when generalizing this criterion to the proposed implementation, because attenuation in the PML region is driven by dissipation. We verify unconditional stability by reusing the above energy recording approach for increasing time steps $\Delta t = 2\Delta t_c$, $4\Delta t_c$, $8\Delta t_c$, where Δt_c stems from Equation (41). The results appear in Figure 11(b) and (d). Clearly, the proposed scheme exhibits highly stable behavior. As confirmation, we have also conducted experiments for $\Delta t = 16\Delta t_c$ and $32\Delta t_c$ with a very long simulation of 128 000 s requiring 930 000 time steps. They are not included here because the energy is significantly overestimated due to an inaccurate representation of the temporal variation of the source. The long simulation has also been performed for mass lumped explicit time integration with $\Delta t = 0.95\Delta t_c$. Both results reveal stable time-marching. As noticed, the stability properties of the proposed scheme for anisotropic media have not been analyzed. However, as various numerical investigations indicate [20, 27, 33, 40, 60], we believe that this scheme will also suffer from intrinsic instabilities in the case of anisotropic media. Hence, the instability growth in the PML is not a consequence of the discretization but depends on the physical properties of the anisotropic medium

DOI: 10.1002/nme

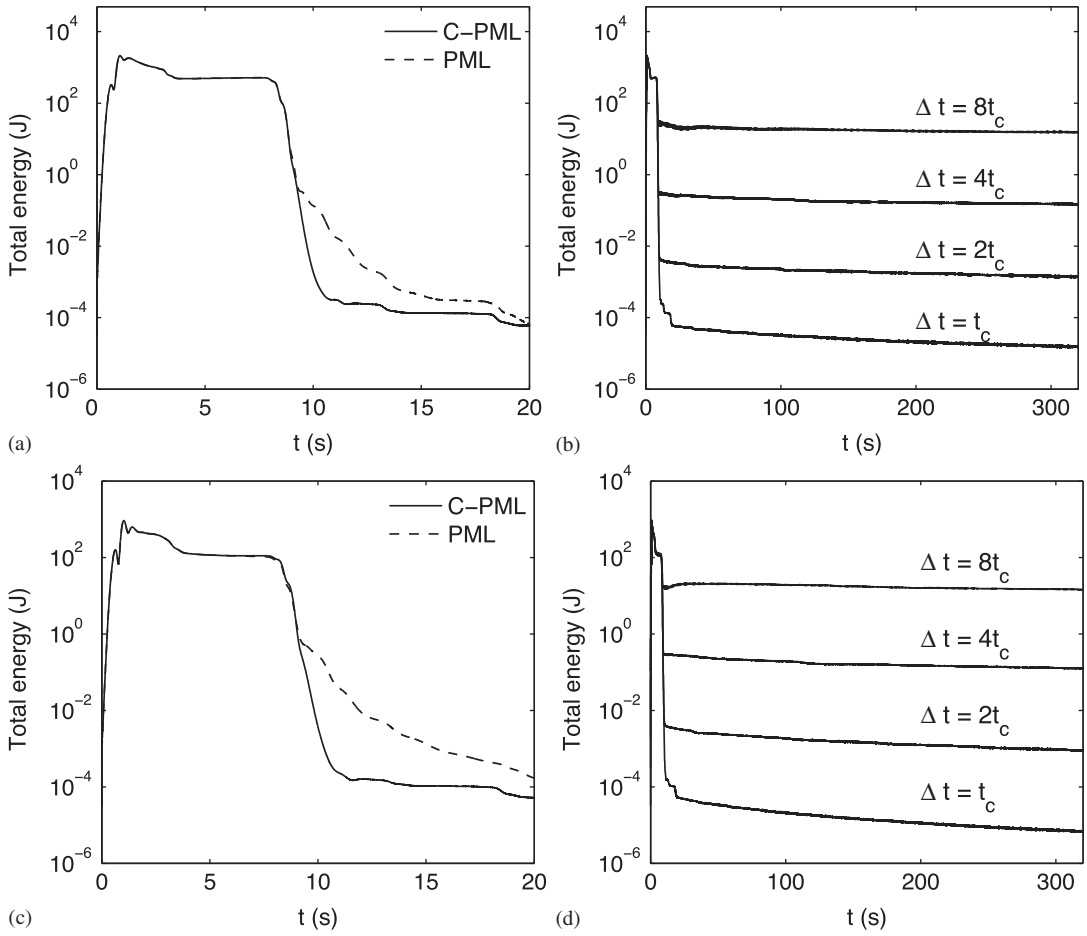


Figure 11. Energy decay in the computational domain for experiment 2 (top) and experiment 3 (bottom) with implicit time integration: (a), (c) performance comparison of the C-PML and the standard PML implementation and (b), (d) unconditional stability verification for time steps $\Delta t = \Delta t_c$, $2\Delta t_c$, $4\Delta t_c$, and $8\Delta t_c$.

[27, 60]. Meanwhile, the recent development of the multi-axial PML (M-PML) has shown to delay the instabilities significantly [32].

Apart from addressing stability properties the total energy decay is also an accurate measure for the efficiency of the discrete C-PML model. Figure 11(a) and (c) shows that between approximately 0 and 1.5 s energy is introduced in the medium by the source. Initially, this energy is carried by P-, S-, and Rayleigh waves, and between roughly 8.0 and 10.5 s the waves leave the physical domain starting with the P-wave. The Rayleigh wave is the last one to be absorbed in the PML around 10s, indicated by the steep decay. Hereafter, all remaining energy is purely spurious. Evidently, the standard PML exhibits degrading performance since the energy residue after 11.5 s e.g. in experiment 3 is 0.1 J compared to the C-PML implementation that yields a remaining energy of 0.0004 J.

4. CONCLUSIONS AND FUTURE WORK

In this paper, we suggest an efficient FE time-domain implementation of the C-PML for the elastic second-order wave equation in which the displacement is the single unknown field.

To obtain a unified FE formulation, we utilize that the material behavior in the PML can be interpreted as anisotropic. Effectively, this requires modification of the kinematic relation between

DOI: 10.1002/nme

strains and displacements as well as the constitutive matrix to render it symmetric, such that the constitutive law is directly applicable in the variational form of the FE implementation. We realize that the appearance of the four time convolution integrals in 2D acting as auxiliary memory variables (or internal forces) is inevitable when a pure displacement-based formulation is desired. However, the condition of the convolutional integrals facilitates the use of a recursive update procedure that avoids the computational burden inherently associated with convolution operations. In the time marching we use the general Newmark-beta method which is highly recommendable for irregular geometries in order to avoid the inherent limitation of the CFL condition. Thus, to align with the accuracy of the Newmark-beta method we use a recursive update scheme of second-order accuracy which results in extremely stable time integration. However, it has also been demonstrated that the proposed scheme supports an easy switch to explicit time integration with mass lumping. The efficiency of the suggested formulation is tested on relevant numerical examples comprised by homogeneous, isotropic half-spaces subjected to a directional point-source to produce near-grazing waves. The results show perfect agreement with analytical solutions.

Although we do not complete the stability analysis of the formulation by including arbitrary anisotropic material behavior in the physical domain, we believe, similar to the previously reported findings, that the scheme will suffer from late-time instabilities. Hence, future work includes implementation of stabilization techniques in terms of the M-PML [32], which, naturally, needs to be verified for anisotropic materials. Additionally, the 3D model will be implemented in the future. The primary purpose of this paper is to show that an efficient C-PML FE implementation is indeed possible when e.g. existing displacement-based FE codes need to handle open-region elastic wave propagation problems.

ACKNOWLEDGEMENTS

This work is supported by the Eurohorcs/ESF European Young Investigator Award (EURYI, http://www.esf.org/euryi) through the grant ‘Synthesis and topology optimization of optomechanical systems’, by M.B. Richters foundation, and by the Danish Center for Scientific Computing (www.dcsc.dk). The author is grateful to the partners from the TopOpt-group (http://www.topopt.dtu.dk) for fruitful discussions related to the presented work. In particular, the help and support from my supervisors Jakob S. Jensen and Ole Sigmund are gratefully acknowledged.

REFERENCES

1. Lindman EL. Free-space boundary-conditions for the time-dependent wave-equation. *Journal of Computational Physics* 1975; **18**(1):66–78. DOI: 10.1016/0021-9991(75)90102-3.
2. Clayton R, Engquist B. Absorbing boundary conditions for acoustic and elastic wave equations. *Bulletin of the Seismological Society of America* 1977; **67**(6):1529–1540.
3. Cerjan C, Kosloff D, Kosloff R, Reshef M. A nonreflecting boundary-condition for discrete acoustic and elastic wave-equations. *Geophysics* 1985; **50**(4):705–708. DOI: 10.1190/1.1441945.
4. Higdon RL. Numerical absorbing boundary-conditions for the wave-equation. *Mathematics of Computations* 1987; **49**(179):65–90.
5. Givoli D. Nonreflecting boundary-conditions. *Journal of Computational Physics* 1991; **94**(1):1–29. DOI: 10.1016/0021-9991(91)90135-8.
6. Higdon RL. Absorbing boundary-conditions for elastic-waves. *Geophysics* 1991; **56**(2):231–241. DOI: 10.1016/0021-9991(92)90016-R.
7. Peng CB, Toksoz MN. An optimal absorbing boundary-condition for elastic-wave modeling. *Geophysics* 1995; **60**(1):296–301. DOI: 10.1190/1.1443758.
8. Hagstrom T, Hariharan SI. A formulation of asymptotic and exact boundary conditions using local operators. *Applied Numerical Mathematics* 1998; **27**(4):403–416.
9. Guddati MN, Lim KW. Continued fraction absorbing boundary conditions for convex polygonal domains. *International Journal for Numerical Methods in Engineering* 2006; **66**(6):949–977. DOI: 10.1002/nme.1574.
10. Berenger JP. A perfectly matched layer for the absorption of electromagnetic-waves. *Journal of Computational Physics* 1994; **114**(2):185–200. DOI: 10.1006/jcph.1994.1159.
11. Chew WC, Weedon WH. A 3D perfectly matched medium from modified Maxwells equations with stretched coordinates. *Microwave and Optical Technology Letters* 1994; **7**(13):599–604. DOI: 10.1002/mop.4650071304.
12. Sacks ZS, Kingsland DM, Lee R, Lee JF. A perfectly matched anisotropic absorber for use as an absorbing boundary condition. *IEEE Transactions on Antennas and Propagation* 1995; **43**(12):1460–1463. DOI: 10.1109/8.477075.

13. Gedney SD. An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices. *IEEE Transactions on Antennas and Propagation* 1996; **44**(12):1630–1639. DOI: 10.1109/8.546249.
14. Teixeira FL, Chew WC. Differential forms, metrics, and the reflectionless absorption of electromagnetic waves. *Journal of Electromagnetic Waves and Applications* 1999; **13**(5):665–686. DOI: 10.1163/156939399X01104.
15. Teixeira FL, Chew WC. Unified analysis of perfectly matched layers using differential forms. *Microwave and Optical Technology Letters* 1999; **20**(2):124–126. DOI: 10.1002/(SICI)1098-2760(19990120)20:2<124::AID-MOP12>3.0.CO;2-N.
16. Chew WC, Liu QH. Perfectly matched layers for elastodynamics: a new absorbing boundary condition. *Journal of Computational Acoustic* 1996; **4**(4):341–359.
17. Liu QH, Tao JP. The perfectly matched layer for acoustic waves in absorptive media. *Journal of the Acoustical Society of America* 1997; **102**(4):2072–2082. DOI: 10.1121/1.419657.
18. Liu QH. Perfectly matched layers for elastic waves in cylindrical and spherical coordinates. *Journal of the Acoustical Society of America* 1999; **105**(4):2075–2084. DOI: 10.1121/1.426812.
19. Collino F, Monk P. The perfectly matched layer in curvilinear coordinates. *SIAM Journal on Scientific Computing* 1998; **19**(6):2061–2090. DOI: 10.1137/S1064827596301406.
20. Collino F, Tsogka C. Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media. *Geophysics* 2001; **66**(1):294–307. DOI: 10.1190/1.1444908.
21. Luebbers RJ, Hunsberger F. Fdtd for *N*th-order dispersive media. *IEEE Transactions on Antennas and Propagation* 1992; **40**(11):1297–1301. DOI: 10.1109/8.202707.
22. Kuzuoglu M, Mittra R. Frequency dependence of the constitutive parameters of causal perfectly matched anisotropic absorbers. *IEEE Microwave and Guided Wave Letters* 1996; **6**(12):447–449. DOI: 10.1109/75.544545.
23. Roden JA, Gedney SD. Convolution PML (CPML): an efficient FDTD implementation of the CFS-PML for arbitrary media. *Microwave and Optical Technology Letters* 2000; **27**(5):334–339. DOI: 10.1002/1098-2760(20001205)27:5⟨334::AID-MOP14⟩3.0.CO;2-A.
24. Wang TL, Tang XM. Finite-difference modeling of elastic wave propagation: a nonsplitting perfectly matched layer approach. *Geophysics* 2003; **68**(5):1749–1755. DOI: 10.1190/1.1620648.
25. Drossaert FH, Giannopoulos A. A nonsplit complex frequency-shifted PML based on recursive integration for FDTD modeling of elastic waves. *Geophysics* 2007; **72**(2):T9–T17. DOI: 10.1190/1.2424888.
26. Drossaert FH, Giannopoulos A. Complex frequency shifted convolution PML for FDTD modelling of elastic waves. *Wave Motion* 2007; **44**(7–8):593–604. DOI: 10.1016/j.wavemoti.2007.03.003.
27. Komatitsch D, Martin R. An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. *Geophysics* 2007; **72**(5):SM155–SM167. DOI: 10.1190/1.2757586.
28. Martin R, Komatitsch D, Ezziani A. An unsplit convolutional perfectly matched layer improved at grazing incidence for seismic wave propagation in poroelastic media. *Geophysics* 2008; **73**(4):T51–T61. DOI: 10.1190/1.2939484.
29. Martin R, Komatitsch D. An unsplit convolutional perfectly matched layer technique improved at grazing incidence for the viscoelastic wave equation. *Geophysical Journal International* 2009; **179**(1):333–344. DOI: 10.1111/j.1365-246X.2009.04278.x.
30. Becache E, Fauqueux S, Joly P. Fictitious domains, mixed finite elements and perfectly matched layers for 2-D elastic wave propagation. *Journal of Computational Acoustic* 2003; **9**(3):1175–1201.
31. Cohen G, Fauqueux S. Mixed spectral finite elements for the linear elasticity system in unbounded domains. *SIAM Journal on Scientific Computing* 2005; **26**(3):864–884. DOI: 10.1137/S1064827502407457.
32. Meza-Fajardo KC, Papageorgiou AS. A nonconvolutional, split-field, perfectly matched layer for wave propagation in isotropic and anisotropic elastic media: Stability analysis. *Bulletin of the Seismological Society of America* 2008; **98**(4):1811–1836. DOI: 10.1785/0120070223.
33. Kucukcoban S, Kallivokas LF. Mixed perfectly-matched-layers for direct transient analysis in 2D elastic heterogeneous media. *Computational Methods in Applied Mechanics and Engineering* 2010; DOI: 10.1016/j.cma.2010.07.013.
34. Courant R, Friedrichs K, Lewy H. Über die partiellen differenzengleichungen der mathematischen physik. *Mathematische Annalen* 1928; **100**:32–74.
35. Komatitsch D, Tromp J. A perfectly matched layer absorbing boundary condition for the second-order seismic wave equation. *Geophysical Journal International* 2003; **154**(1):146–153. DOI: 10.1046/j.1365-246X.2003.01950.x.
36. Basu U, Chopra AK. Perfectly matched layers for transient elastodynamics of unbounded domains. *International Journal for Numerical Methods in Engineering* 2004; **59**(8):1039–1074. DOI: 10.1002/nme.896.
37. Basu U. Explicit finite element perfectly matched layer for transient three-dimensional elastic waves. *International Journal for Numerical Methods in Engineering* 2004; **77**(8):151–176. DOI: 10.1002/nme.2397.
38. Martin R, Komatitsch D, Gedney SD. A variational formulation of a stabilized unsplit convolutional perfectly matched layer for The isotropic or anisotropic seismic wave equation. *CMES—Computer Modeling in Engineering and Sciences* 2008; **37**(3):274–304.
39. Festa G, Vilotte JP. The Newmark scheme as velocity–stress time-staggering: an efficient PML implementation for spectral element simulations of elastodynamics. *Geophysical Journal International* 2005; **161**(3):789–812. DOI: 10.1111/j.1365-246X.2005.02601.x.
40. Li YF, Matar OB. Convolutional perfectly matched layer for elastic second-order wave equation. *Journal of the Acoustical Society of America* 2010; **127**(3):1318–1327. DOI: 10.1121/1.3290999.

41. Chew WC, Jin JM, Michielssen E. Complex coordinate stretching as a generalized absorbing boundary condition. *Microwave and Optical Technology Letters* 1997; **15**(6):363–369. DOI: 10.1002/(SICI)1098-2760(19970820)15:6⟨363::AID-MOP8⟩3.0.CO;2-C.
42. Berenger JP. Application of the CFS PML to the absorption of evanescent waves in waveguides. *IEEE Microwave and Wireless Components Letters* 2002; **12**(6):218–220. DOI: 10.1109/LMWC.2002.1010000.
43. Berenger JP. Numerical reflection from FDTD-PMLs: A comparison of the split PML with the unsplit and CFS PMLs. *IEEE Transactions on Antennas and Propagation* 2002; **50**(3):258–265. DOI: 10.1109/8.999615.
44. Butterworth S. On the theory of filter amplifiers. *Experimental Wireless and the Wireless Engineer* 1930; **7**:536–541.
45. Bracewell R. *The Fourier Transform and its Applications* (3rd edn). McGraw-Hill Science/Engineering/Math: New York, June 1999.
46. Nye F. *Physical Properties of Crystals*. Oxford University Press: Clarendon, 1985.
47. Harari I, Albocher U. Studies of FE/PML for exterior problems of time-harmonic elastic waves. *Computer Methods in Applied Mechanics and Engineering* 2006; **195**(29–32):3854–3879. DOI: 10.1016/j.cma.2005.01.024.
48. Maple 14. Reference Manual for Maple 14 Maplesoft, Canada. Available from: www.maplesoft.com.
49. Zienkiewicz OC, Taylor RL, Zhu JZ. *The Finite Element Method*: *its Basis and Fundamentals* (6th edn). Elsevier Butterworth-Heinemann: London, 2005.
50. Hughes TJR. *The Finite Element Method*: *Linear Static and Dynamic Finite Element Analysis*. Dover: New York, NY, 2000.
51. Jiao D, Jin JM, Michielssen E, Riley DJ. Time-domain finite-element simulation of three-dimensional scattering and radiation problems using perfectly matched layers. *IEEE Transactions on Antennas and Propagation* 2003; **51**(2):296–305.
52. Cohen GC. *Higher-order Numerical Methods for Transient Wave Equations*. Springer: Berlin, 2002.
53. Newmark NM. A method of computation for structural dynamics. *ASCE Journal of the Engineering Mechanics Division* 1959; **85**(3):67–94.
54. Chilton RA, Lee R. The discrete origin of FETD-Newmark late time instability, and a correction scheme. *Journal of Computational Physics* 2007; **224**(2):1293–1306. DOI: 10.1016/j.jcp.2006.11.021.
55. Abenius E, Edelvik F, Johansson C. Waveguide truncation using UPML in the finite element time-domain method. *Technical Report*, 2005; 026. DOI: 10.1.1.142.1839.
56. Winton SC, Rappaport CM. Specifying PML conductivities by considering numerical reflection dependencies. *IEEE Transactions on Antennas and Propagation* 2006; **48**(7):1055–1063. DOI: 10.1109/8.876324.
57. Berg P, If F, Nielsen P, Skovgaard O. Analytical reference solutions. In *Modeling the Earth for Oil Exploration*, Helbig K (ed.). Pergamon: Brussels, Belgium, 1994; 421–427.
58. Collino F, Monk P. Optimizing the perfectly matched layer. *Computer Methods in Applied Mechanics and Engineering* 1998; **164**(1–2):157–171. DOI: 10.1016/S0045-7825(98)00052-8.
59. Lamb H. On the propagation of tremors over the surface of an elastic solid. *Philosophical Transactions of the Royal Society of London Series A—Containing Papers of a Mathematical or Physical Character* 1904; **203**:1–42.
60. Becache E, Fauqueux S, Joly P. Stability of perfectly matched layers, group velocities and anisotropic waves. *Journal of Computational Physics* 2003; **188**(2):399–433. DOI: 10.1016/S0021-9991(03)00184-0.
61. Wang S, Teixeira FL. Some remarks on the stability of time-domain electromagnetic simulations. *IEEE Transactions on Antennas and Propagation* 2004; **52**(3):895–898. DOI: 10.1109/TAP.2004.825486.
62. Donderici B, Teixeira FL. Conformal perfectly matched layer for the mixed finite element time-domain method. *IEEE Transactions on Antennas and Propagation* 2008; **56**(4):1017–1026. DOI: 10.1109/TAP.2008.919215.
63. Becache E, Petropoulos PG, Gedney SD. On the long-time behavior of unsplit perfectly matched layers. *IEEE Transactions on Antennas and Propagation* 2004; **52**(5):1335–1342. DOI: 10.1109/TAP.2004.827253.
64. Abarbanel S, Gottlieb D, Hesthaven JS. Long time behaviour of the perfectly matched layer equations in computational electromagnetics. *Journal of Scientific Computing* 2002; **17**(1–4):405–422. DOI: 10.1023/A:1015141823608.

[P5]

Topology optimization of slow light structures in the time domain

René Matzen,* Jakob S. Jensen, and Ole Sigmund

Department of Mechanical Engineering, Technical University of Denmark
Nils Koppels Allé, Building 404, 2800 Kgs. Lyngby, Denmark

**Corresponding author: rmat@mek.dtu.dk*

A pulse delaying optimization scheme based on topology optimization for transient response of photonic crystal structures (PhCs) is formulated to obtain slow light devices. The optimization process is started from a qualified W1 PhC waveguide design with group index $n_g \approx 40$ obtained from a simple Edisonian parameter search. Based on this, the proposed pulse delaying and subsequent pulse restoring strategies yield a design that increases the group index by 75% to $n_g \approx 70 \pm 10\%$ for an operational full width at half maximum bandwidth $B_{\mathrm{FWHM}} = 6\mathrm{nm}$, and simultaneously minimizes interface penalty losses between the access ridge and W1 PhC waveguide. To retain periodicity and symmetry the active design set is limited to the in-/outlet region and a distributed supercell, and manufacturability is further enhanced by density filtering techniques combined with material phase projections.

OCIS codes: 000.4430, 130.5296, 230.7400.

1. Introduction

It is well-known that the speed of light is an upper bound for waves conveying matter, energy or information [1]. However, after the discovery of the slow-light phenomenon as a result of material dispersion, it seems that no lower limit for the group velocity at which light can travel exists [2], [3]. Since then, this promising technology has attracted much attention, as the application prospects are numerous, e.g. in future optical networks and information processing systems. To mention only a few remarkable properties, slow-light offers the opportunity for accurate time-domain processing of optical signals, low power consumption in optical switching devices [4], spatial pulse compression [5], optical buffering [6], and enhancement of weak linear as well as nonlinear light-matter interaction processes [7–10]. In

the realization of on-chip integration of slow-light devices equipped with these features, photonic crystals (PhCs) show very promising candidacy. This is due to their intrinsic photonic band gap (PBG) property that inhibits the existence of optical modes at certain frequency bands [11–13]. The PBG facilitates the structural engineering of PhC waveguides (-WG) [14] and micro-cavities (-MC) [15], that separately or in combinations exhibit unique tunable (group-velocity) dispersion characteristics. In contrast to conventional semiconductors PhCs exploit all-optical processes, thus leaving it as an obvious choice of platform for designing efficient slow-light devices [16–20]. At very low group velocities, however, this advantage is counteracted by losses that scale as the inverse square of the group velocity due to increased density of states [21].

Recently, inverse problem techniques have surfaced as competitive design tools to previous Edisonian approaches in the engineering of PhCs [22]. In the present paper we will employ the methodology behind topology optimization [23] as a means to structurally design devices that slow down the speed of light. The design process is based on a time-domain analysis, and the setup consists of a PhC-WG of finite length with in- and outlet connections to ridge waveguides. The entire structure is surrounded by a perfectly matched layer (PML) as absorbing boundary conditions (ABCs).

Topology optimization has previously exploited the tunability with which the PhC-WGs and -MCs are naturally born to optimize various PhC devices with large enhancement in optical properties. This comprises frequency-domain optimization of a 90-degree bend [24], a T-junction [25], and a termination [26]. Other efficient PhC-WG components have also been devised and experimentally verified in [27–30]. Even though the above examples rely on frequency-domain analyses, time-domain optimization is continuously maturing as an attractive approach. The reason hereto is its ability to handle broadband signals by short pulse excitation with a single analysis cycle, and it allows for local permutations at frequency level. Hence, time-domain optimization facilitates the treatment of active media and non-linearities that give rise to frequency modulation. Additionally, it can manage time-domain processing of optical signals, such as pulse shaping [31, 32] and pulse delay [33], as well as optimization of PhC notch filters [34]. A comprehensive review on topology optimization of nano-photonic devices is provided in [35].

In comparison to slow light devices based on material resonant enhancement e.g. caused by electromagnetically induced transparency (EIT) [36], the structural engineering counterpart is highly preferred because it accommodates a simpler control of light. The reason hereto is that signals very often in practice propagate in transparent media, i.e. at the operation frequency, far away from any material resonances [37]. However, both approaches are subject to the intrinsic detrimental effect that large time delays are only possible near large changes in the amplitude response (near band edges of filter pass and/or stop bands) and are therefore

accompanied by severe amplitude distortion. Very recently, topology optimization based on eigenvalue analyses has been used to enhance the slow light performance of PhC-WGs by engineering/tailoring their dispersion properties [38, 39]. Both approaches include manufacturing constraints, and in particular the latter successfully incorporates robustness to further enhance manufacturability [40, 41].

The optimization process is completed in two consecutive steps: First, a *pulse delay* step where the objective is the time delay of the pulse while its shape is preserved through a relaxed signal shaping constraint. Next, a *pulse restoring step* is carried out in which the pulse shape is tailored to match the original reference pulse in the exit ridge waveguide, but now with the maximized delay obtained in first step fixed in a pulse conforming objective function.

2. Formulation

The method presented in this paper supports propagation of transverse electric (TE) modes within two-dimensional (2D) optical waveguide structures. The medium residing inside these structures occupies a composite of regions of homogenous dielectric material in the xy-plane and is invariant in the z direction.

With these assumptions, the behavior is governed by the following scalar-wave equation

$$\begin{aligned} &\frac{\partial}{\partial x}\left(A\frac{\partial \Psi}{\partial x}\right)+\frac{\partial \Psi}{\partial y}\left(A\frac{\partial \Psi}{\partial y}\right)-\frac{B}{c^2}\frac{\partial^2 \Psi}{\partial t^2}=0 \\ &\Psi = H_z(\mathbf{r},t), \quad A=1/\varepsilon_r(\mathbf{r}), \quad B=1 \end{aligned} \tag{1}$$

where H_z is the transverse components of the magnetic field, as a function of position $\mathbf{r} = (x, y)$ and time t, c is vacuum speed of light, and $\varepsilon_r(\mathbf{r})$ is the relative permittivity that only varies spatially. The structures are surrounded by PML regions to minimize nonphysical reflection from the boundaries. The formulation in Eq. (1) then needs to be modified to handle the anisotropic, dispersive PML behavior, which can be found in [34]. The formulation can equally be employed for transverse magnetic (TM) modes by letting $\Psi = E_z(\mathbf{r}, t)$, $A = 1$, and $B = \varepsilon_r(\mathbf{r})$, where E_z denotes the out-of-plane electric field component.

Since we consider macroscopic systems containing dielectric material we utilize that no fundamental *length* scale exists by introducing $(x', y') = (x, y)/a$, where a is the lattice constant or the *pitch* of the PhC [43]. Based on Eq. (1) this leads to scaled time $t' = tc/a$ and scaled angular frequency $\omega' = \omega a/c$. We use the scaled parameters in the remainder of this paper unless otherwise stated, and it is straight forward to switch between these and the non-scaled parameters as soon as a is known a priori.

To realize the slow light device we exploit the unique tunable dispersion properties of a W1 PhC-WG given in Fig. 1, where a single row of holes has been removed. As illustrated,

the waveguide has in- and outlet connections to straight ridge waveguides, and the crystal is formed by perforated air holes with radius $r/a = 0.30$ in the dielectric material GaAs. This yields a band gap in the normalized frequency $a/\lambda = 0.21 - 0.30$ where λ denotes the wavelength. We will assume that the pitch of the lattice is $a = 370$nm.

To seek a numerical solution to Eq. (1) the computational domain including PML regions is divided into bilinear square elements. Then, applying the standard finite-element (FE) technique [44] yields the system of ordinary differential equations

$$\sum_{e=1}^{M}(\mathbf{T}^e\ddot{\mathbf{u}} + \mathbf{R}^e\dot{\mathbf{u}} + \mathbf{S}^e\mathbf{u} + \mathbf{g}^e - \mathbf{f}^e) = 0 \tag{2}$$

where $\dot{(\,)} = \mathrm{d}/\mathrm{d}t$, $\ddot{(\,)} = \mathrm{d}^2/\mathrm{d}t^2$, and M denotes the number of elements. We integrate Eq. (2) by a dispersion reducing scheme that can be found in [34]. Herein, the expressions for the element-level constituent system matrices $\mathbf{T}^e$, $\mathbf{R}^e$, $\mathbf{S}^e$, $\mathbf{g}^e$ and $\mathbf{f}^e$ are also derived.

3. Characterizing and quantifying slow light behavior

There are different ways to determine if a (PhC) structure for a given geometry or material distribution exhibits slow light behavior. A common method is to study the dispersion relation $\omega(k)$ from which the group velocity can be computed by

$$v_g \equiv \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{c}{n_g} \tag{3}$$

where k is the wave number. In the rightmost expression the dimensionless *group index* n_g is introduced as an alternative measure for the slow light behavior, and we can obtain that by numerical differentiation, once $\omega(k)$ is known.

We have performed an eigenvalue analysis to obtain the dispersion characteristics for the supercell structure in Fig. 2a. In conjunction with the dispersion diagram in the left part of Fig. 2b, the group index versus normalized frequency plot in Fig. 2c reveals that inside the band gap $a/\lambda = 0.21-0.30$, the waveguide mode exhibits (semi)-slow light behavior with very little group velocity dispersion at group velocity $v_g \approx c/(40 \pm 10\%)$. The useful bandwidth of the structure (i.e. the range over which the group index remains constant within $\pm 10\%$) is $B_{\mathrm{FWHM}} \approx 12$nm centered around the normalized frequency $a/\lambda = 0.2163$. Consequently, the given supercell structure serves as a suitable starting guess the for optimization process. The geometry has been found via a parameter search in which the radius of the three nearest neighboring holes to the waveguide has been varied [20, 45].

For time domain simulations it is not that straightforward to retrieve the dispersion characteristics as they appear in the band diagram [46]. Another reliable way to detect slow light behavior is to analyze the transmission spectrum, which we construct by exciting the inlet

ridge waveguide with a short pulse. We obtain the transmission spectrum by integrating the frequency dependent Poynting flux over a vertical wall located after the PhC-WG. Since the PhC-WG sample has a PhC mirror mechanism: the PhC-WG and the non-optimized in- and outlet, it functions as a so-called *Fabry-Pérot* (F-P) cavity. This leads to F-P fringes in the transmitted energy, and the group velocity can be determined by [16].

$$v_g = c\frac{2L_c\Delta\lambda}{\lambda^2} \tag{4}$$

where $\Delta\lambda$ is the free spectral range between adjacent F-P peaks, and L_c is the length of the (closed) cavity. The transmission spectrum in the right part of Fig. 2b reveals F-P oscillations in the normalized frequency $a/\lambda = 0.21 - 0.25$, thus insinuating slow light behavior. Furthermore, the spectrum shows an unambiguous correlation between the eigenvalue analysis and our FETD data. Apart from the energy transmittance Fig. 3 shows the energy reflectance recorded at a vertical wall in front of the PhC-WG, from which the energy balance is computed. We clearly see that for the wavelength range corresponding to the linear part of the guided dispersion curve all input energy is transmitted through the PhC-WG, automatically yielding almost zero reflection. Hence, very little coupling losses between the ridge and the PhC waveguide occur. The energy balance is almost conserved here due to high mode confinement in the PhC-WG. Around the band gap edges the energy balance is disrupted partly due to poor coupling at the PhC-WG entrance resulting in reflected and surface modes not captured by the vertical wall in front of the PhC-WG. The poor mode confinement in the slow light region can give rise to a lateral mode flow which is not included in the computed energy flow and therefore further explains why the energy is not preserved.

Where the above techniques require frequency information, the time-of-flight (ToF) method, see e.g. [47, 48], relies on time observations of the pulse delay η as it traverses the PhC-WG to compute the group velocity by

$$v_g = \frac{\eta}{L} \tag{5}$$

where L is the length of the PhC-WG. For the device in Fig. 1 the time delay η can be determined by recording the probe pulse at a given point in the exit ridge waveguide with and without PhC-WG. For a precise measure we extract the envelope of the probe pulses and define the delay as the temporal distance between the envelope peaks. It is important to be sufficiently downstream in the ridge waveguide to avoid potential near-field disturbances from surface modes living along the in-/outlet interface.

The definition in Eq. 5 gives us a quantitative measure, namely the delay η, which can be adopted in the formulation of a time domain optimization problem.

4. Optimization problem

The optimization problem with the setup given in Fig. 1 is based on pulse shaping of the envelope $r(t)$ of the probe signal $u(t)$ registered at Ω_{E} in the ridge waveguide. The aim is to distribute air and dielectric material in the design domain Ω_{D}, such that envelope signal obtained at multiple registration points is delayed as much as possible while following a prescribed envelope, obtained from the same input signals propagating in a straight ridge waveguide. The pulse conforming measure is defined as [33]

$$f = \frac{\displaystyle\int_0^T \sum_{i\in\Omega_{\mathrm{E}}} [r_i(t) - \alpha p_i(t-\eta)]^2 \mathrm{d}t}{\displaystyle\int_0^T \sum_{i\in\Omega_{\mathrm{E}}} \alpha^2 p_i(t)^2 \mathrm{d}t} \tag{6}$$

where $r_i(t)$ denotes the envelope of the probe signal $u_i(t)$ registered at the ith point in Ω_{E}, $p_i(t)$ is the prescribed envelope function delayed η and scaled by α, and T is the termination time of the transient simulation. To render a physically interpretable measure, we have normalized with respect to the prescribed pulse envelope area(s). The envelope of the probe signal $u_i(t)$ is retrieved by utilizing the Hilbert transform [33].

For the design parametrization each element in the design domain Ω_{D} is linked to one (density) variable ρ_e that varies continuously between $0 \leq \rho_e \leq 1$. By adopting a similar concept to that of the solid isotropic material with penalization (SIMP) scheme, the design variable serves to interpolate the inverse permittivity linearly between two candidate materials [23, 49]

$$\varepsilon_r^{-1}(\rho_e) = (1-\rho_e)(\varepsilon_r^{\mathrm{I}})^{-1} + \rho_e(\varepsilon_r^{\mathrm{II}})^{-1} \tag{7}$$

where $(\cdot)^{\mathrm{I}}$ and $(\cdot)^{\mathrm{II}}$ designate air and dielectric material, respectively. All design variables are assembled into the global design vector $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_M)^T$. The continuous design parametrization in Eq. (7) enables the use of gradient-based optimization algorithm of the globally convergent method of moving asymptotes (GCMMA) [50] to find optimized designs.

To formulate a pulse delaying strategy, we introduce $\eta' = \eta/s$ as a design variable with s denoting a scaling factor. The shape of the pulse is controlled by treating the design response measure in Eq. (6) as a constraint. The objective is to maximize the delay, which we formulate as

$$\begin{aligned} \min_{\boldsymbol{\rho}\in\mathbf{R}^M, \eta'\in\mathbf{R}} \quad & -\eta' \\ \text{s.t.:} \quad & \text{Governing equation (2)} \\ & g(\boldsymbol{\rho}, \eta') = f(\boldsymbol{\rho}, \eta')/g^* - 1 \leq 0 \\ & 0 \leq \rho_e \leq 1, \qquad e \in \Omega_{\mathrm{D}} \\ & 0 \leq \eta' \leq \eta'_{\max} \end{aligned} \tag{8}$$

where $0 < g^* \leq 1$ has been introduced to allow for broadening and distortion of the delayed pulse in Fig. (4) after traversing the PhC-WG. The maximum allowable distortion is obtained for $g^* = 1$ (i.e. the initial pulse shape has been completely destroyed) and is gradually diminished when $g^* \to 0$ (i.e. the output pulse shape is perfect). The gradients of g with respect to the structurally related design variables are found through the adjoint sensitivity analysis method [33, 34, 51]. The sensitivity of g with respect to the time related design variable η' is given by

$$\frac{\partial g}{\partial \eta'} = \frac{\displaystyle\int_0^T \sum_{i \in \Omega_{\mathrm{E}}} -2\left[r_i(t) - \alpha p_i(t - s\eta')\right] \alpha \frac{\partial p_i(t - s\eta')}{\partial \eta'} \mathrm{d}t}{\displaystyle g^* \int_0^T \sum_{i \in \Omega_{\mathrm{E}}} \alpha^2 p_i(t)^2 \mathrm{d}t} \tag{9}$$

where the change of variable $\Pi = t - s\eta'$ yields $\partial p_i / \partial \eta' = -s \partial p_i / \partial \Pi$ which we approximate by finite differences. It is important to stress that the delay variable $\eta = s\eta'$ is several orders of magnitudes larger compared to the structurally related design variables that vary between 0 and 1. Hence, the sensitivity expression in Eq. (9) has to be scaled adequately through s in order to gain a well-posed optimization problem. A strategy for choosing the scaling parameter is $0 \leq \eta / s \leq 3$ yielding $\eta'_{\max} = 3$.

Once a material distribution is obtained for a maximum delay and a given allowable distortion g^* and α, we will start from the resulting design and use the pulse shaping response function in Eq. (6) as the objective, now with a fixed η, to minimize the pulse distortion. The optimization problem is now formulated as

$$\begin{aligned} \min_{\boldsymbol{\rho} \in \mathbf{R}^M} \quad & f(\boldsymbol{\rho}, \eta') \\ \text{s.t.:} \quad & \text{Governing equation (2)} \\ & 0 \leq \rho_e \leq 1, \qquad e \in \Omega_{\mathrm{D}} \end{aligned} \tag{10}$$

Hence, the optimization is completed in two consecutive steps: ① a *pulse delay* step cf. Eq. (8), followed by ② a *pulse restoring* step cf. Eq. (10). The pulse shaping constraint in the delay formulation is applied in a single point in the ridge waveguide center sufficiently downstream to allow for larger pulse delays. To achieve perfect transmission for the output signal in the subsequent restoring step, we apply the pulse shaping objective in multiple points along a vertical line, such that the ridge waveguide mode profile will be inherited automatically in the registered output signal.

5. Maximizing the delay

We excite the ridge waveguide by an incident plane-wave Gaussian pulse with a $1/e$-intensity half width T_0 centered around t_0

$$\Psi^{\text{inc}}(\mathbf{x}, t) = \Psi(y) \sin[kx - \omega_c(t - t_0)] e^{-\frac{(t-t_0-x/v_p)^2}{T_0^2}} \tag{11}$$

where $\Psi(y)$ is the amplitude profile analytically known for ridge waveguides, $k = \omega_c/v_p$ is the wave number with ω_c denoting the center angular frequency. Further, $v_p = c/n$ is the phase velocity with $n = \sqrt{\mu_r \varepsilon_r}$ designating the refractive index. By utilizing the inverse Fourier transform, the conversion from FWHM bandwidth B_{FWHM} to Gaussian $1/e$ width is obtained by

$$T_0 = \frac{\sqrt{4 \ln 2}}{\pi B_{\text{FWHM}}} \tag{12}$$

For a specified bandwidth B_{FWHM} the optimization formulation in Eq. (8) can successfully be deployed to obtain a device exhibiting slow light behavior by exciting the structure with the incident wave in Eq. (11) in which T_0 is determined by Eq. (12). To give the optimization problem design freedom and opportunity for a significant slow down effect we choose an operational FWHM bandwidth $B_{\text{FWHM}} = 6\text{nm}$, and the driving frequency is $\omega_c a/2\pi c = 0.2163$.

To retain a certain periodicity and symmetry of the final design, we only allow for design freedom in the x-mirrored part of the in-/outlet region Ω_{IO} as well as in the x-y mirrored part of the supercell Ω_{SC} distributed periodically downstream over the lattice distance $17a$ in the design domain Ω_{D} (see the symmetry details in Fig. 5). Even though the actual design update is carried out only in a small part of the design domain, we still need to compute the sensitivities in the entire design domain, and subsequently map and sum those properly for the 'active' design variables. In Fig. 1 the active design set $\Omega_{\text{IO}} \cup \Omega_{\text{SC}}$ is highlighted in relation to the entire design domain Ω_{D}. It should be combined with Fig. 5 to clarify the symmetry conditions.

The geometrical restrictions will reduce the optimization design space whereby potentially well-performing designs are disregarded because nothing dictates that a symmetric and periodic structure should be the best performing. However, by limiting the optimization to the in-/outlet region and the supercell, we can directly determine the dispersion characteristics of the resulting structure by a simple eigenvalue analysis of the supercell. The periodicity also enhances manufacturability, that is further boosted by standard density filtering of the design variables with filter radius $R = 2.5\Delta x$ combined with material phase projection [52]. Even though the applied projection function does not preserve the minimum structural length scale for some settings [39], it results in well-defined designs, see Fig. 6b, -c, -d and 8c. It

should be emphasized that whenever an optimized design is presented we only illustrate the in-/outlet region in conjunction with the supercell structure.

The first part of the optimization has been dedicated to investigate the design as well as the time delay response to three different values for the relaxation parameter $g^* = 0.001$, 0.005, and 0.05 and $\alpha = 1$ of the prescribed pulse envelope area in the pulse shaping constraint. The designs and time delays appear in Fig. 6. For the lowest relaxation factor $g^* = 0.001$ the pulse is delayed 1.5ps which doubles to 3ps for $g^* = 0.05$. However, the price for the largest delay is a considerable broadening of the pulse and amplitude distortion which corresponds to an unfavorable reduction of the useful bandwidth of the device. In practice heavy distortion implies that some of the information carried by the pulse has been lost, and for signal processing devices it is important to preserve the pulse shape to a certain degree. Further, the peak intensity reduction of a pulse makes it less effective for driving nonlinear effects.

Based on an eigenvalue analysis of the optimized supercell structures the group index versus normalized frequency plot in Fig. 7 confirms that the maximized delays give rise to increased group indices. However, the bandwidth in which the optimized devices can be considered useful for slow light purposes is simultaneously decreased. The enhanced time delays induce very strong wavelength dependence of the group index. In particular, we see that the 'flat-band' region for the initial guess evolves into an s-shaped kink whose center is moving towards northwest, causing bandwidth shrinkage. For the largest relaxation value the average group index is $n_g \approx 160$, for which, however, the 10% criterion is not satisfied. This points to the fact, that a bandwidth extension is only obtained at the cost of the group index. Frankly, this detrimental effect is inherently associated to slow light systems, leaving the designer of such systems with a difficult task.

All the optimized designs have reached convergence[1] within 300 optimization iterations. In addition to changing the supercell layout, the geometry of the in-/outlet region has been altered to lower the interface penalty losses, when coupling the access waveguide mode into the PhC-WG. For real applications it is important to terminate the PhC properly, since coupling losses, together with backscattering losses, as a result of fabrication disorders, will degrade the functionality of the slow light device. Intriguingly, a comparison of the optimized supercell structures in Fig. 6a, -b, -c and the initial supercell layout in Fig. 2a reveals that the optimization has reduced the size of the air inclusions. We consider this a logical action in order to reduce backscattering losses as they originate from the overlap of the optical modes and the hole surfaces [45].

[1]For the convergence criteria we use that the relative change of the objective function should be less 10^{-4}.

6. Restoring the ridge waveguide pulse shape

In Fig. 7b it is seen that even for the lowest relaxation factor $g^* = 0.001$ we obtain a supercell design with a group index variation around $n_g \approx 70$ that is not confined to the allowable $\pm 10\%$ band for the specified 6nm bandwidth. To tweak this unsatisfied behavior we have executed the *pulse restoring* step with the design in Fig. 6a as the initial design and the maximized delay as the fixed time delay in the pulse shaping objective. Whereas the delay formulation typically reaches convergence within 300 optimization iterations (it naturally depends on the relaxation parameters), the pulse shape formulation has required 1710 iterations for converging to the design in Fig. 8b. We believe the reason for this rather high number of iterations lies in the fact that the optimization has two concerns; while modification of the supercell structure will result in a change of the waveguide mode, the in-/outlet design needs to be simultaneously adjusted to eliminate coupling losses. To meet and balance both regards, the optimization will then progress in small increments.

As it appears in Fig. 7b, the optimized design yields a group index variation that satisfies the $\pm 10\%$ criterion. It is obvious to pose the question why the pulse shaping optimization strategy does not result in a completely flat-band region. One explanation could be that the optimization has ended up in a local minimum or that a flat-band region is not physically obtainable. We should also remember that as designers we are competing against the intrinsic dispersion property of the PhC-WG, that the group index can only be increased at the cost of the operational bandwidth. It is also noted in Fig. 7b that the group index versus normalized frequency curve is shifted slightly towards lower frequencies compared to the start design. Based on a comparison of the supercell designs in Fig. 2a and 8c showing that the amount of dielectric material has increased moderately, the horizontal shift of group index curve is expected, since the frequency of the guided mode scales by $1/\sqrt{\varepsilon}$ in a medium of dielectric constant ε [39].

The transmission, reflection and energy balance spectrum for the finite waveguide structure assembled from the supercell design in Fig. 8b without the optimized in-/outlet region is presented in Fig. 9. It shows clearly defined Fabry-Pérot (F-P) fringes with decreased spacing towards the slow light wavelength region. In the slow light region near the band edge the wavelength spacing is measured to $\Delta\lambda = 5$nm that for an effective F-P cavity length $L_c = 13a$ yields $v_g \approx c/65$. The spectrum also contains a dip at $\lambda = 1375$nm with almost zero transmittance. To explain this we need to consider the band diagram for the optimized supercell in Fig. 8b, showing that a guided mode (red curve) has been pulled down from the air band. This mode has an odd mode profile and, apart from living in the waveguide, it can also live in the upper/lower part of the design domain between the waveguide and the surrounding PhC. A comparison of transmittance against the band diagram in Fig. 8b shows that anti-crossing of the even (black curve) and the odd (red curve) modes at $\lambda \approx 1375$ gives

rise to a mode gap inside the band gap, which explains the dip in transmission and in the energy balance [17] .

Finally, Fig. 10 shows that almost perfect transmission, i.e. $|T|^2 \approx 0.98$, and almost zero reflection are achieved in the wavelength range $B_{\mathrm{FWHM}} = 1710 \pm 3\mathrm{nm}$ specified by the source, if we include the optimized in-/outlet design. This indicates that no penalty losses are present at the interface due to a mode mismatch, and the associated H_z field pattern in Fig. 11a for the dominant source wavelength also confirms perfect transmission compared to the poorly performing start design with a field pattern given in Fig. 11b.

7. Conclusion

In this study a topology optimization methodology has been presented to design slow light structures subjected to transient responses. The slow light devices are obtained in two consecutive steps: A pulse delaying step, followed by a pulse restoring step. The first step employs an optimization formulation that delays the probe pulse as much as possible while controlling the allowable pulse distortion through a pulse-shaping constraint. In the second step a pulse-shaping objective is used for the maximum delay obtained from the previous step to restore the analytically known mode profile of the ridge waveguide, and thereby reduce the amount of GVD such that the group index variation satisfies the 10% rule [45] for the operational bandwidth.

The optimization process starts from a suitable W1 PhC-WG with in- and outlet connections to ridge waveguides used to feed the system with an incoming pulse-shaped signal and, after traversing the PhC-WG, to monitor the probe signal, respectively. The initial PhC-WG, which has been found through a simple parameter study by varying the radii of the nearest neighboring holes, has group index $n_g \approx 40$ with very low GVD.

To retain periodicity and symmetry in the final designs the active design set is limited to the in-/outlet region as well as to the supercell structure inside the entire design domain. The manufacturability is further enhanced by standard density filtering techniques combined with material phase projection. With these settings, the pulse delaying optimization step yields a structure with group index variation around $n_g \approx 70$ (i.e. a 75% increase) that does not meet the 10% criterion. However, the subsequent pulse restoring step fixes this issue, and the resulting pulse shape is very close to the equivalent one monitored in a straight ridge waveguide.

The energy flow inside the structures of finite length and an eigenvalue analysis of the supercell confirm that the initial as well as the optimized structures hold the expected slow light properties.

Acknowledgments

This work is supported from the Eurohorcs/ESF European Young Investigator Award [53] through the grant "Synthesis and topology optimization of optomechanical systems", as well as M.B. Richters foundation and the Danish Center for Scientific Computing are gratefully acknowledged. The authors are grateful to members of the TopOpt-group [54] for fruitful discussions related to the presented work.

References

1. G. Diener, "Superluminal group velocities and information transfer," Phys. Lett. A **223**(5):327–331 (1996).
2. L. V. Hau, S. E. Harris, Z. Dutton, and C. H. Behroozi, "Light speed reduction to 17 metres per second in an ultracold atomic gas," Nat. **397**(6720), 594–598 (1999).
3. M. M. Kash, V. A. Sautenkov, A. S. Zibrov, L. Hollberg, G. R. Welch, M. D. Lukin, Y. Rostovtsev, E. S. Fry, and M. O. Scully, "Ultraslow group velocity and enhanced nonlinear optical effects in a coherently driven hot atomic gas," Phys. Rev. Lett. **82**(26), 5229–5232 (1999).
4. T. Baba, "Slow light in photonic crystals," Nat. Phot. **2**(8), 465–473 (2008).
5. T. J. Karle, Y. J. Chai, C. N. Morgan, I. H. White, and T. F. Krauss, "Observation of pulse compression in photonic crystal coupled cavity waveguides," Journal Of Lightwave Technology **22**(2), 514–519 (2004).
6. R. S. Tucker, P. C. Ku, and C. J. Chang-Hasnain, "Slow-light optical buffers: Capabilities and fundamental limitations," Journal Of Lightwave Technology **23**(12), 4046–4066 (2005).
7. B. J. Eggleton, C. M. de Sterke, and R. E. Slusher, "Bragg solitons in the nonlinear schrodinger limit: experiment and theory," J. Opt. Soc. Am. B **16**(4), 587–599 (1999).
8. M. Soljacic, S. G. Johnson, S. H. Fan, M. Ibanescu, E. Ippen, and J. D. Joannopoulos, " Photonic-crystal slow-light enhancement of nonlinear phase sensitivity," J. Opt. Soc. Am. B **19**(9), 2052–2059 (2002).
9. M. Soljacic and J. D. Joannopoulos, "Enhancement of nonlinear effects using photonic crystals," Nat. Mat. **3**(4), 211–219 (2004).
10. T. F. Krauss, "Slow light in photonic crystal waveguides," J. Phys. D: Appl. Phys. **40**(9), 2666–2670 (2007).
11. K. Ohtaka, "Energy band of photons and low-energy photon diffraction," Phys. Rev. B **19**(10), 5057–5067 (1979).
12. E. Yablonovitch, "Inhibited Spontaneous Emission In Solid-state Physics and Electronics," Phys. Rev. Lett. **58**(20), 2059–2062 (1987).
13. S. John, "Strong localization of photons in certain disordered dielectric superlattices," Phys. Rev. Lett. **58** (23), 2486–2489 (1987).
14. T. F. Krauss, R. M. DeLaRue, and S. Brand, "Two-dimensional photonic-bandgap structures operating at near infrared wavelengths," Nat. **383**(6602), 699–702 (1996).
15. Y. Akahane, T. Asano, B. S. Song, and S. Noda, "High-q photonic nanocavity in a two-dimensional photonic crystal," Nat. **425**(6961), 944–947 (2003).
16. X. Letartre, C. Seassal, C. Grillet, P. Rojo-Romeo, P. Viktorovitch, M. L. dYerville, D.

Cassagne, and C. Jouanin, "Group velocity and propagation losses measurement in a single-line photonic-crystal waveguide on Inp membranes," Appl. Phys. Lett. **79**(15), 2312–2314 (2001).

17. M. Notomi, K. Yamada, A. Shinya, J. Takahashi, C. Takahashi, and I. Yokohama, "Extremely large group-velocity dispersion of line-defect waveguides in photonic crystal slabs," Phys. Rev. Lett. **87**(25), 253902 (2001).
18. H. Gersen, T. J. Karle, R. J. P. Engelen, W. Bogaerts, J. P. Korterik, N. F. van Hulst, T. F. Krauss, and L. Kuipers, "Real-space observation of ultraslow light in photonic crystal waveguides," Phys. Rev. Lett. **94**(7), 073903 (2005).
19. Y. A. Vlasov, M. OBoyle, H. F. Hamann, and S. J. McNab, "Active control of slow light on a chip with photonic crystal waveguides," Nat. **438**(7064), 65–69 (2005).
20. L. H. Frandsen, A. V. Lavrinenko, J. Fage-Pedersen, and P. I. Borel, "Photonic crys- tal waveguides with semi-slow light and tailored dispersion properties," œ**14**(20), 9444–9450 (2006).
21. L. O'Faolain, T. P. White, D. O'Brien, X.Yuan, M. D. Settle, and T. F. Krauss, "Dependence of extrinsic loss on group velocity in photonic crystal waveguides," Opt. Express **15**, 13129–13138 (2007).
22. M. Burger, S. J. Osher, and E. Yablonovitch, "Inverse problem techniques for the design of photonic crystals," IEICE Trans Electron **E87C** (3), 258–265 (2004).
23. M. P. Bendsøe and O. Sigmund, *Topology Optimization; Theory, Methods, and Applications* (Springer Verlag Berlin Heiddelberg New York, 2nd edition, 2004).
24. J. S. Jensen and O. Sigmund, "Systematic design of photonic crystal structures using topology optimization: Low-loss waveguide bends," Appl. Phys. Lett. **84** (12), 2022–2024 (2004).
25. J. S. Jensen and O. Sigmund, "Topology optimization of photonic crystal structures: a high-bandwidth low-loss T-junction waveguide," J. Opt. Soc. Am. B **22** (6), 1191–1198 (2005).
26. W. R. Frei, D. A. Tortorelli, and H. T. Johnson, "Topology optimization of a photonic crystal waveguide termination to maximize directional emission," Appl. Phys. Lett. **86** (11), 111114 (2005).
27. P. I. Borel, A. Harpoth, L. H. Frandsen, M. Kristensen, P. Shi, J. S. Jensen, and O. Sigmund, "Topology optimization and fabrication of photonic crystal structures," Opt. Express **12** (9), 1996–2001 (2004).
28. L. H. Frandsen, A. Harpoth, P. I. Borel, M. Kristensen, J. S. Jensen, and O. Sigmund, "Broadband photonic crystal waveguide 60 degrees bend obtained utilizing topology optimization," Opt. Express **12** (24), 5916–5921 (2004).
29. P. I. Borel, L. H. Frandsen, A. Harpoth, M. Kristensen, J. S. Jensen, and O. Sigmund,

"Topology optimised broadband photonic crystal Y-splitter," Elec Lett **41** (2), 69–71 (2005).

30. J. S. Jensen, O. Sigmund, L. H. Frandsen, P. I. Borel, A. Harpoth, and M. Kristensen, "Topology design and fabrication of an efficient double 90(circle) photonic crystal waveguide bend," IEEE Photonics Tech. Lett. **17** (6), 1202–1204 (2005).
31. J. Dahl, J. S. Jensen, and O. Sigmund, "Topology optimization for transient wave propagation problems in one dimension Design of filters and pulse modulators," Struct multidisc optim **36** (6), 585–595 (2008).
32. L. R. Yang, A. V. Lavrinenko, J. M. Hvam, and O. Sigmund, "Design of one-dimensional optical pulse-shaping filters by time-domain topology optimization," Appl. Phys. Lett. **95**(26), 261101 (2009).
33. B. Lazarov, R. Matzen, and Y. Elesin, "Topology optimization of pulse shaping filters using the Hilbert transform envelope extraction," Struct. Multidiscip. Optim. (2011). DOI: 10.1007/s00158-011-0642-y.
34. R. Matzen, J. S. Jensen, and O. Sigmund, "Topology optimization for transient response of photonic crystal structures," J. Opt. Soc. Am. B **27**(10), 2040–2050 (2010).
35. J. S. Jensen, and O. Sigmund, "Topology optimization for nano-photonics," Laser & Phot. Rev. **5**(2), 308–321 (2011).
36. M. D. Lukin and A. Imamoglu, "Controlling photons using electromagnetically induced transparency," Nat. **413**(6853), 273–276 (2001).
37. G. Lenz, B. J. Eggleton, C. K. Madsen, and R. E. Slusher, "Optical delay lines based on optical filters," IEEE Jour. Quant. Elect. **37**(4), 525–532 (2001).
38. R. Stainko and O. Sigmund, "Tailoring dispersion properties of photonic crystal waveguides by topology optimization," Wav. Rand. Compl. Med. **17**(4), 477–489 (2007).
39. F. Wang, J. S. Jensen, and O. Sigmund, "Robust topology optimization of photonic crystal waveguides with tailored dispersion properties," J. Opt. Soc. Am. B **28**(3), 387–397 (2011).
40. O. Sigmund, "Manufacturing tolerant topology optimization," Acta Mech. Sinica **25**(2), 227–239 (2009).
41. F. Wang, B. Lazarov, and O. Sigmund, "On projection methods, convergence and robust formulations in topology optimization," Struct. Multidiscip. Optim. (2010). DOI: 10.1007/s00158-010-0602-y.
42. J. S. Jensen, "Topology optimization of nonlinear optical devices," Struct. Multidiscip. Optim. (2010). DOI: 10.1007/s00158-011-0640-0.
43. J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic crystals : molding the flow of light* (Princeton University Press, 2008).
44. J. Jin and D. J. Riley, *Finite Element Analysis of Antennas and Arrays* (John Wiley &

Sons Inc, 2007).

45. S. A. Schulz, L. O'Faolain, D. M. Beggs, T. P. White, A. Melloni, and T. F. Krauss, "Dispersion engineered slow light in photonic crystals: a comparison," J Opt. **12**(10), 104004 (2010).
46. A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. D. Joannopoulos, and S. G. Johnson,"MEEP: A flexible free-software package for electromagnetic simulations by the FDTD method, Comput. Phys. Commun. **181**, 687702 (2010).
47. K. Inoue, N. Kawai, Y. Sugimoto, N. Carlsson, N. Ikeda, and K. Asakawa, "Observation of small group velocity in two-dimensional algaas-based photonic crystal slabs," Phys. Rev. B **65**(12): 121308 (2002).
48. R. S. Jacobsen, A. V. Lavrinenko, L. H. Frandsen, C. Peucheret, B. Zsigri, G. Moulin, J. Fage-Pedersen, and P. I. Borel, "Direct experimental and numerical determination of extremely high group indices in photonic crystal waveguides," Opt. Express **13**, 7861–7871 (2005).
49. O. Sigmund and J. S. Jensen, "Systematic design of phononic band-gap materials and structures by topology optimization," Phil. Trans. Roy. Soc. London Ser. A-math. Phys. Eng. Scienc. **361**(1806), 1001–1019 (2003).
50. K. Svanberg, "A class of globally convergent optimization methods based on conservative convex separable approximations," SIAM J Opt **12** (2), 555–573 (2001).
51. D. A. Tortorelli and P. Michaleris, "Design sensitivity analysis: Overview and review," Inv. Probl. Eng. **1**(1), 71–105 (1994).
52. O. Sigmund, "Morphology-based black and white filters for topology optimization," Struct multidisc optim **33** (4-5), 401–424 (2007).
53. http://www.esf.org/euriy.
54. http://www.topopt.dtu.dk.

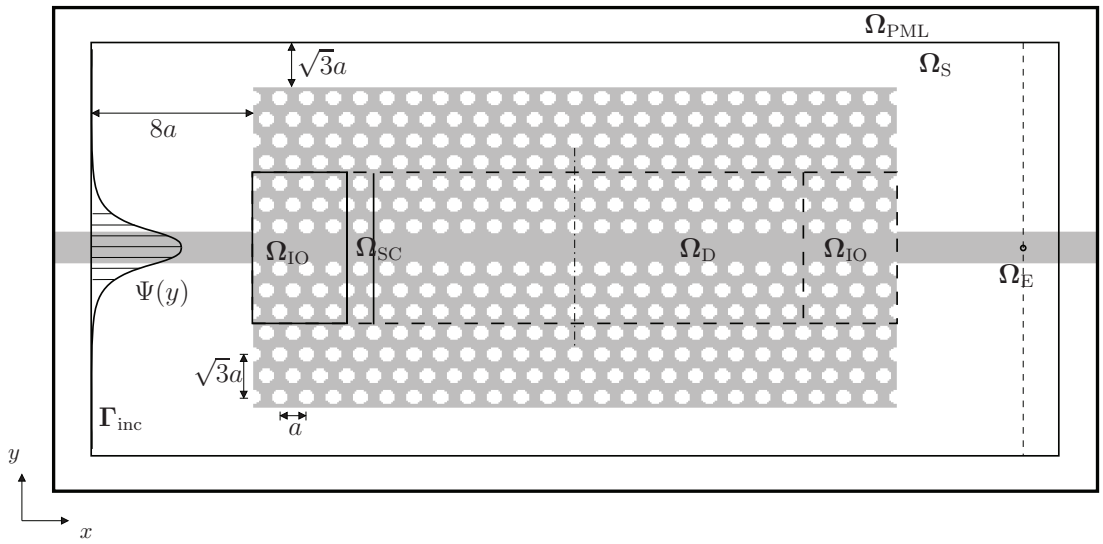
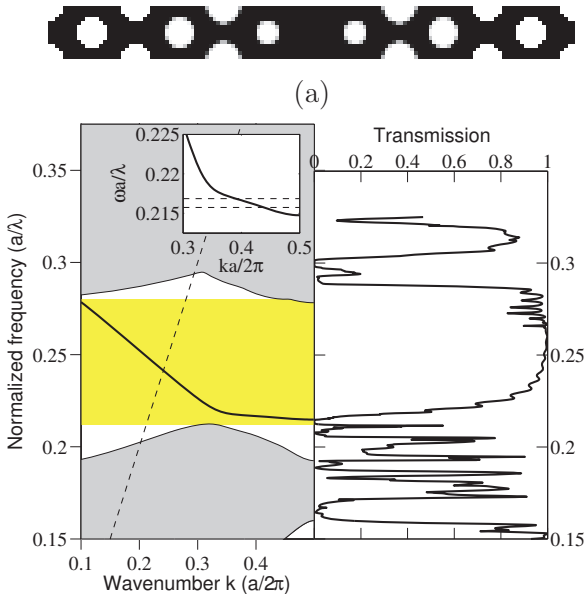
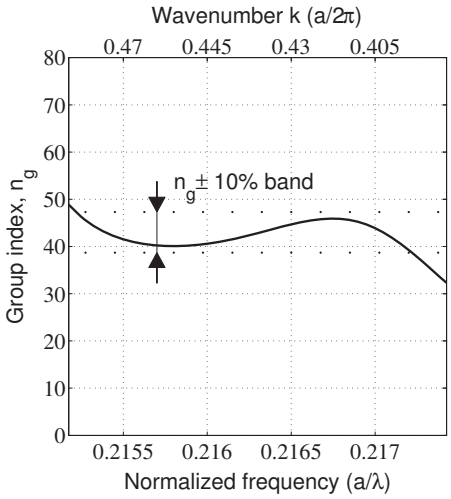


Fig. 1. Ridge wave guide optimization setup. The computational domain contains a solution region Ω_{S}, PML region Ω_{PML}, and design region Ω_{D} encapsulated by the dashed boundary. The 'active' design set consists of y mirrored in-/outlet region Ω_{IO} and a supercell Ω_{SC} that is a subset of Ω_{D}. Whenever an optimized design is presented we only show the material distribution in Ω_{IO} and Ω_{SC}. The waveguide mode is excited at Γ_{inc} with an analytically given amplitude profile $\Psi(y)$ [42]. The objective is measured at multiple point(s) on Ω_{E}.



(a)

(b)



(c)

Fig. 2. **(a)** The rotated supercell used in the eigenvalue computation to obtain the **(b)** band diagram. This shows the normalized frequencies versus normalized wave vectors for an even (solid) and odd (omitted here) PhC-WG mode in the band gap, and the dashed line indicates the light line above which leaky modes live. The inset of the graph magnifies the dispersion properties of the guided mode where the dotted lines confine a region with low dispersion. **(c)** Group index n_g versus normalized frequency.

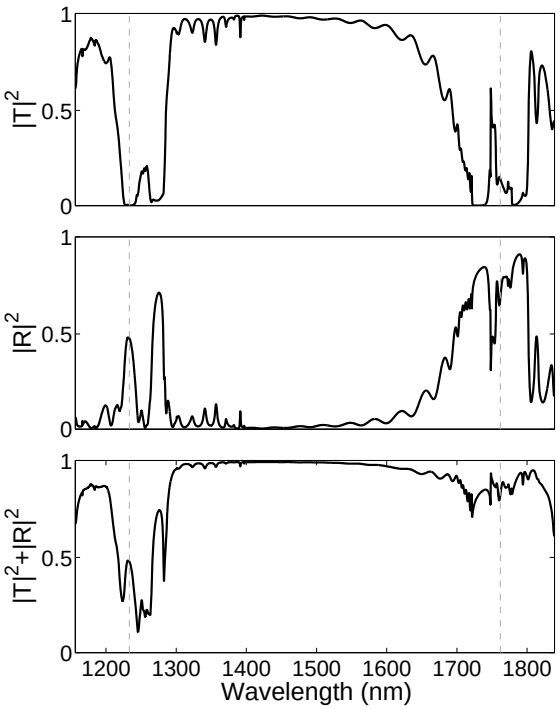


Fig. 3. Transmittance $|T|^2$, reflectance $|R|^2$, and energy balance $|T|^2 + |R|^2$ through in- and outlets as a function of the wavelength for the initial solution (and thus the sum of the reflection and transmission does not equal one). The interval between the dashed lines indicates the band gap region. The crystal lattice is assumed to be $a = 370$ nm.

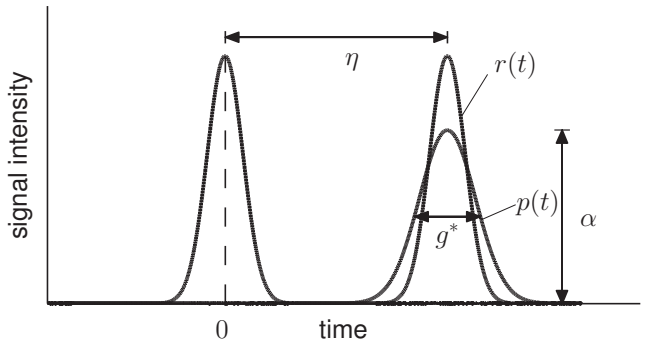


Fig. 4. Pulse delay strategy for pulses with a group delay 0 and η. The parameter η is introduced as a design variable. The g^*-parameter controls the temporal spreading of the output pulse envelope. The α-parameter mainly specifies the transmitted energy, i.e. the amplitude decrease.

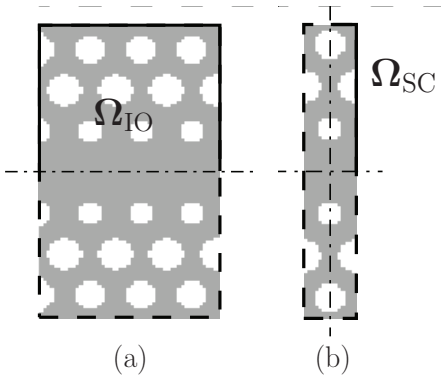


Fig. 5. Symmetry conditions for the 'active' design set. **(a)** x symmetry is imposed for Ω_{IO}. **(b)** x-y symmetry is imposed for Ω_{SC}. This reduces the number of active design variables further.

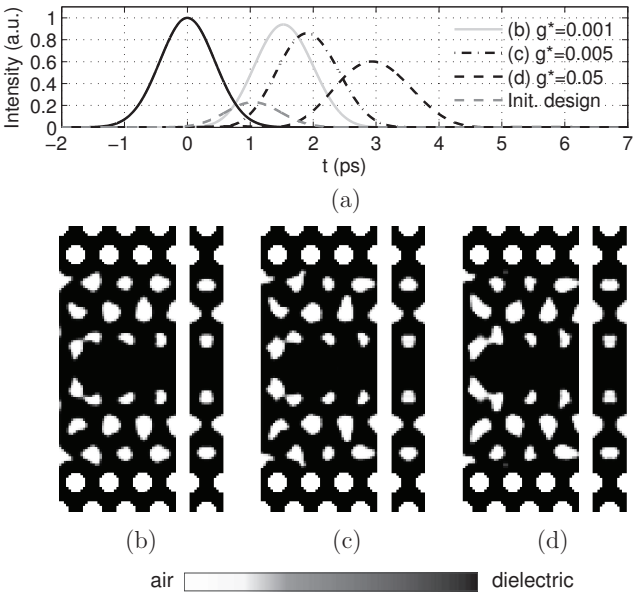
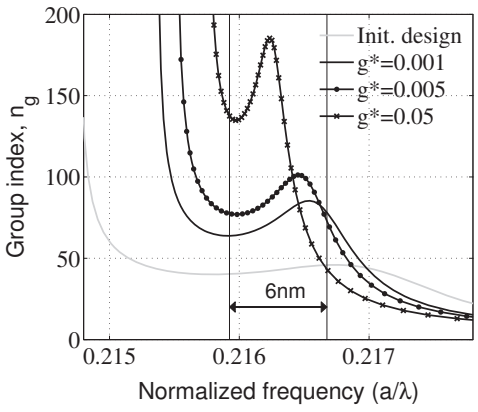
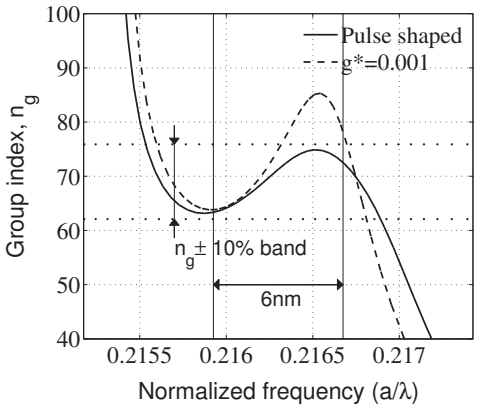


Fig. 6. Pulse delayed slow light designs. **(a)** Transient intensity response of the ridge waveguide pulse peaking at $t = 0$ps, and time delayed pulse for $g^* = 0.001$, 0.005, and 0.05 peaking at $t = 1.5$ps, 2.0ps, and 3.0ps, respectively. In/-outlet and supercell design for **(b)** $g^* = 0.001$, **(c)** $g^* = 0.005$, and **(d)** $g^* = 0.05$.



(a)



(b)

Fig. 7. Group index versus normalized frequency. **(a)** the optimized design obtained from the delay formulation in Eq. (8) for three different values for the pulse relaxation: $g^* = 0.001$, 0.005, and 0.05. **(b)** A comparison between the pulse delayed design for $g^* = 0.001$ and the pulse delayed-shaped design.

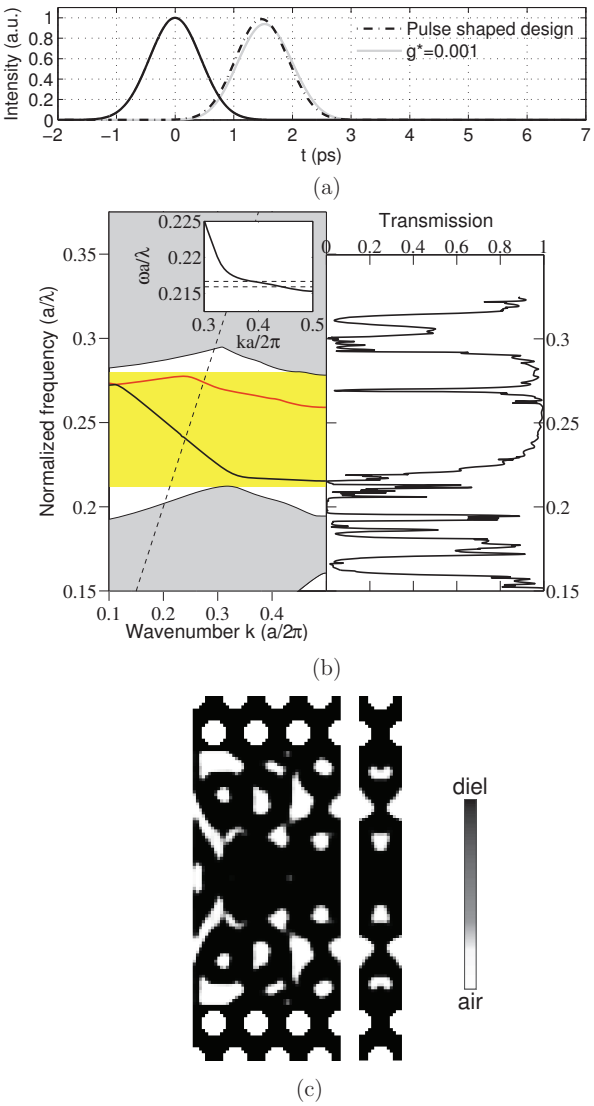


Fig. 8. Pulse delayed-shaped slow light design. **(a)** Transient intensity response of the initial envelope peaking at $t = 0$ps, and the optimized envelope peaking at $t = 1.5$ps with and without pulse shaping. **(b)** Band diagram for the optimized supercell structure and the transmission spectrum for the structure of finite length without optimized in-/outlets **(c)** .

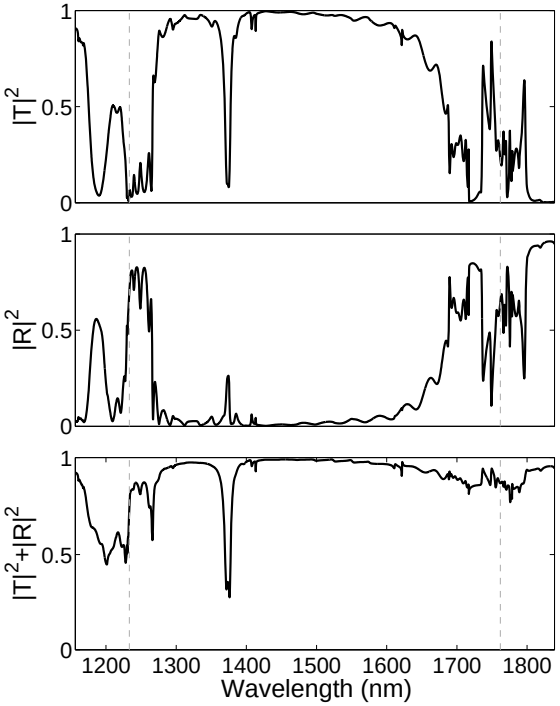


Fig. 9. Optimized supercell design without in-/outlet design. Transmittance $|T|^2$, reflectance $|R|^2$, and energy balance $|T|^2 + |R|^2$ as a function of the wavelength. The interval between the dashed lines indicates the band gap region.

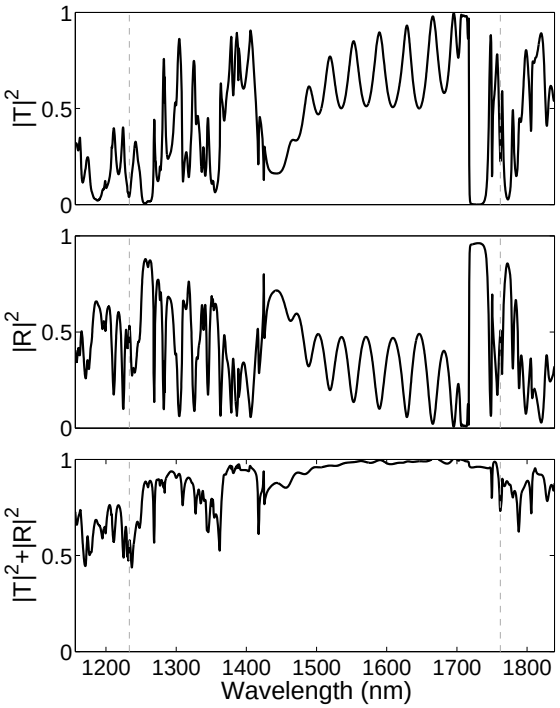


Fig. 10. In-/outlet and supercell design in Fig. 11b. Transmittance $|T|^2$, reflectance $|R|^2$, and energy balance $|T|^2 + |R|^2$ as a function of the wavelength. The interval between the dashed lines indicates the band gap region.

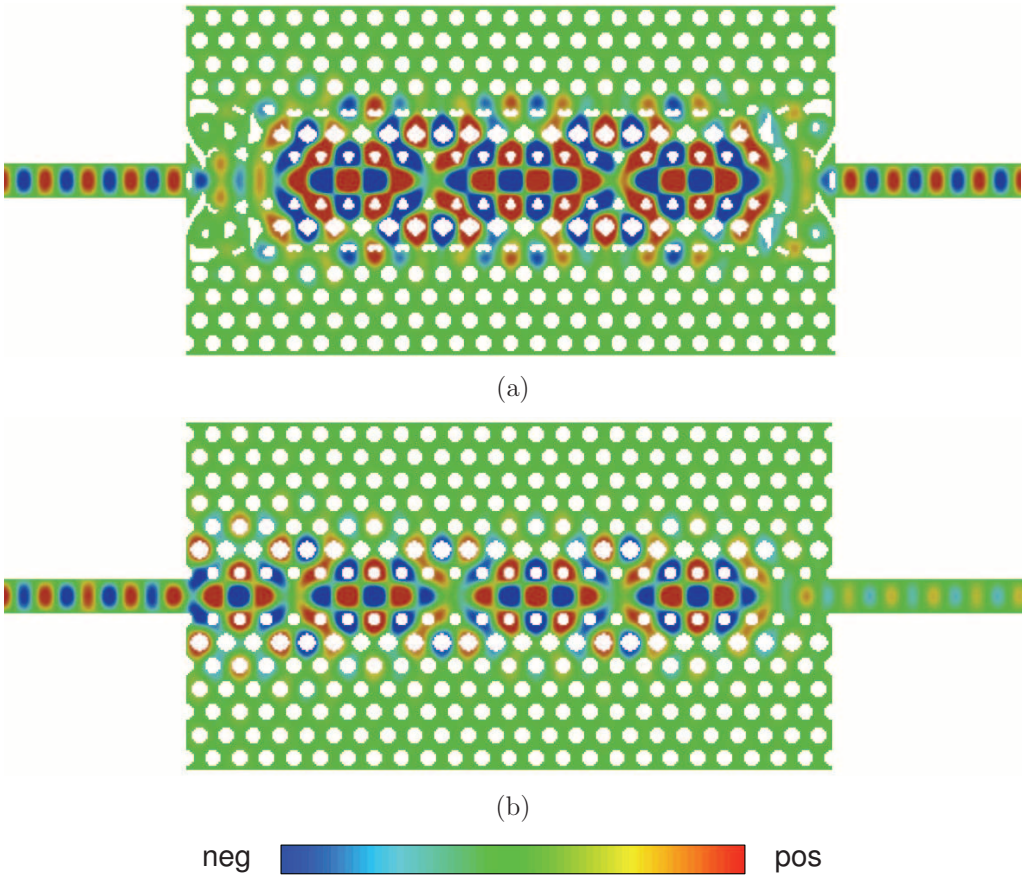


Fig. 11. H_3 field pattern for slow light device. **(a)** The optimized structure first by delaying and subsequently shaping the pulse. **(b)** The initial geometry. The material distribution is shown with 0.6 threshold.

DTU Mechanical Engineering
Section of Solid Mechanics
Technical University of Denmark

Nils Koppels Allé, Bld. 404
DK- 2800 Kgs. Lyngby
Denmark
Phone (+45) 45 25 42 50
Fax (+45) 45 93 14 75
www.mek.dtu.dk
ISBN: 978-87-90416-57-7

DCAMM
Danish Center for Applied Mathematics and Mechanics

Nils Koppels Allé, Bld. 404
DK-2800 Kgs. Lyngby
Denmark
Phone (+45) 4525 4250
Fax (+45) 4593 1475
www.dcamm.dk
ISSN: 0903-1685