*Engineering Optimization*
Vol. 41, No. 4, April 2009, 365–384

Taylor & Francis
Taylor & Francis Group

# Robust design of slow-light tapers in periodic waveguides

Almir Mutapcic[a]*, Stephen Boyd[a], Ardavan Farjadpour[b], Steven G. Johnson[b] and Yehuda Avniel[b]

[a] *Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA;* [b] *Research Laboratory of Electronics, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

This article concerns the design of tapers for coupling power between uniform and slow-light periodic waveguides. New optimization methods are described for designing *robust* tapers, which not only perform well under nominal conditions, but also over a given set of parameter variations. When the set of parameter variations models the inevitable variations typical in the manufacture or operation of the coupler, a robust design is one that will have a high yield, despite these parameter variations.

The ideas of successive refinement, and robust optimization based on multi-scenario optimization with iterative sampling of uncertain parameters, using a fast method for approximately evaluating the reflection coefficient, are introduced. Robust design results are compared to a linear taper, and to optimized tapers that do not take parameter variation into account. Finally, robust performance of the resulting designs is verified using an accurate, but much more expensive, method for evaluating the reflection coefficient.

**Keywords:** robust optimization; PDE-constrained optimization; shape optimization; coupling taper; periodic waveguide; slow-light; photonics

## 1. Introduction

In this article, a new approach for non-convex robust optimization, including a successive refinement technique to avoid local minima, is applied to the problem of designing waveguide taper transitions for the challenging case of 'slow-light' periodic structures. By robust optimization, it is meant that the resulting low-loss tapers still perform well even when manufacturing errors and other uncertainties are included; in contrast, the nominal optimum produced by straightforward optimization of this problem relies on delicate interference effects that are destroyed by any deviation from the design.

A standard component of optical and microwave devices is a waveguide taper, which couples light from one waveguide to another by means of a gradual transition. Although a sufficiently gradual taper approaches an 'adiabatic' limit of 100% transmission, in a practical setting the challenge is to design a taper as short as possible, or with as low a loss as possible for a given length. Perhaps the most challenging case is to design a short taper between an ordinary uniform waveguide and a periodic waveguide (Elachi 1976), a special case of a general class of periodic optical structures known as 'photonic crystals' (Joannopoulos *et al.* 1995, Johnson and Joannopoulos

*Corresponding author. Email: almirm@stanfordalumni.org

ISSN 0305-215X print/ISSN 1029-0273 online

DOI: 10.1080/03052150802576797
http://www.informaworld.com

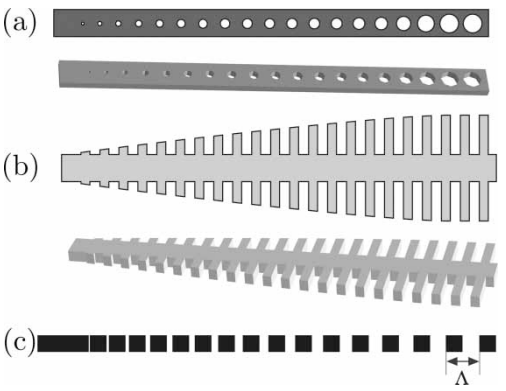


Figure 1. Various tapers between uniform and periodic dielectric waveguides. (a) Periodic sequence of holes, where taper varies the radius and period of the holes, in 2d or 3d. (b) Periodic set of flanges, where taper varies the width of the flange, in 2d or 3d. (c) Periodic sequence of dielectric blocks, where taper varies the period $\Lambda$ between the blocks. All three of these tapers, in 2d and 3d, can be efficiently optimized by the robust coupled-mode method, but this article focusses on (c) because it is also amenable to brute-force computation for verification purposes.

2001a). (Several such structures are depicted in Figure 1.) Periodic waveguides are both useful and challenging for the same reason: a periodic waveguide has a 'slow-light' band edge, for which the group velocity of light slows down as it approaches a certain frequency. Operating in this slow-light region is useful because it increases the interaction of the light with the material, enhancing nonlinearities (Xu *et al.* 2000b, Soljačić *et al.* 2002), tunability (Povinelli *et al.* 2005), gain (Yariv 1989), and other effects. However, as the group velocity decreases, the 'impedance mismatch' between the periodic and uniform waveguides increases, and a longer taper is generally required to achieve the same coupling loss (Povinelli *et al.* 2005). (If the waveguides are simply butt-coupled without a taper, the transmission goes to zero as the zero-velocity band-edge is approached (Sanchis *et al.* 2004).)

A variety of techniques have been employed to select a taper shape for coupling to periodic waveguides. Most of this work examines cases operating far from any band edge (so the group velocity is not small) and focuses on simple linear (constant-rate) tapers (Xu *et al.* 2000a, Mekis and Joannopoulos 2001, Palamaru and Lalanne 2001, Happ *et al.* 2001, Johnson *et al.* 2002, Talneau *et al.* 2002, Pottier *et al.* 2003, Bienstman *et al.* 2003, Chietera *et al.* 2004) or families of quadratic shapes (Khoo *et al.* 2005, Dossou *et al.* 2006, Zhang 2006). Genetic algorithms have also been employed to design couplers using arbitrarily placed scattering cylinders (Jiang *et al.* 2003, Hakansson, Sanchis, Sanchez-Dehesa and Marti 2005). (Non-taper-based couplers, from free space or parallel waveguides, have also been considered (Kuang *et al.* 2002, Prather *et al.* 2002, Barclay *et al.* 2004, Hakansson, Sanchez-Dehesa and Sanchis 2005)). Although this previous work did not explicitly account for uncertainties in the model parameters, the mostly small number of design parameters combined with the moderate group velocities help avoid non-manufacturable designs. As soon as the design involves optimization over a large number of free parameters, the nominal optimum tends to be a non-robust design that relies on delicate interference effects. (A similar result was observed as a strong frequency sensitivity in genetic optimization over many degrees of freedom (Hakansson, Sanchez-Dehesa and Sanchis 2005)). In an earlier article, some of the authors considered a slow-light periodic-waveguide coupler with higher-degree polynomial taper shapes, and used a simple regularization technique to avoid non-robust solutions (Povinelli *et al.* 2005).

The classic case of a tapered uniform waveguide has also been extensively studied, especially in the context of linear and quadratic polynomial shapes (Snyder 1970, Milton and Burns 1977, Constantinou and Jones 1992, Wei *et al.* 1997). (This is, of course, a special case of a periodic

waveguide: period zero.) More general optimization of uniform-waveguide couplers, allowing more arbitrary taper shapes (also called 'inverse design'), has also been considered using a variety of optimization techniques (Spühler *et al.* 1998, Felici and Engl 2001, Luyssaert *et al.* 2005). In this work, again, the initial problem was found to be ill-posed and to require a regularization to account for variations in the design variables (Felici and Engl 2001).

This article presents optimization methods for designing *robust* tapers, which not only perform well under nominal conditions, but also over a given set of parameter variations. The methods optimize over an arbitrary variable taper rate, described by hundreds (or thousands) of degrees of freedom, in order to find a design with performance orders of magnitude better than that of a simple linear (constant-rate) taper. Accurate techniques from coupled-mode theory (Johnson *et al.* 2002) are used to quickly explore different shapes; the results are validated against a direct numerical solution of Maxwell's equations (Bienstman 2001, 2006). Although the same coupled-mode and robust-optimization techniques are equally effective for arbitrary waveguide tapers, in order to directly validate the results against an explicit numerical solution of Maxwell's equations this article will focus on the structure of Figure 1(c), which can be efficiently solved by an eigenmode-expansion method (Bienstman 2001, 2006). Because the set of parameter variations models the inevitable variations typical in the manufacture or operation of the coupler, and is explicitly accounted for in the optimization, a robust design will have a high yield, despite these parameter variations.

There are several general models of parameter uncertainty, as well as general approaches for dealing with uncertain parameters. These can be broadly classified into three groups (which are closely related and connected): regularization, stochastic optimization, and worst-case robust optimization. In *regularization*, what the parameter variations are is a bit vague; the technique simply adds an extra cost term to the objective function that penalizes sensitive or non-robust designs.

In a *stochastic optimization* approach, there is a stochastic or probabilistic model for the parameter uncertainty; a taper design is chosen that, for example, minimizes an expected or average objective value, or directly maximizes the yield (*i.e.* the probability that a set of specifications will be met). Various methods can be used to approximately solve these stochastic design problems, such as sampling or scenario approaches (not unlike the method proposed in this article); see, *e.g.* (Calafiore and Dabbene 2006, Rockafellar and Wets 1991, Calafiore and Campi 2005, de Farias and Roy 2004, Prekopa 1995, Birge and Louveaux 1997). These methods are computationally expensive, and require knowledge of the probability distribution of the uncertain parameters, which typically is not known. Another approach that is based on a stochastic model of parameter uncertainty is Taguchi's method, where the goal is to find a design that achieves a target mean value of objective, while maintaining small variance. The method consists of a collection of heuristics and simple design procedures for achieving this goal, and is widely used in industry; see, *e.g.* (Taguchi *et al.* 2004, Beyer and Sendhoff 2007, Phadke 1995, Dehnad 1989, Tsui 1992, Yang and El-Haik 2003).

The third general approach, which is taken in this article, is *worst-case robust optimization* (or *minimax optimization*). Here the parameters are modelled as lying in some given set of possible values, but without any known distribution; a taper design is chosen, whose worst-case objective value, over the given set of possible uncertainties, is minimized. In this model, one does not rely on any knowledge of the distribution of uncertain parameters (which, indeed, need not be stochastic). While the worst-case approach is adopted in this article, the solution methods and ideas can be modified to handle at least some stochastic design formulations. There is no claim that worst-case robust optimization is superior to stochastic optimization or Taguchi's approach; but it is generally found that worst-case robust designs produced by the methods of this article also perform well when analysed under a stochastic model of parameter variation.

There have been several recent breakthroughs in worst-case robust optimization, for specific convex optimization problems and associated parameter uncertainty sets; see, *e.g.* (Ben-Tal and

Nemirovski 1998, 2002, Soyster 1973, El Ghaoui and Lebret 1997, El Ghaoui *et al.* 1998, Chandrasekaran *et al.* 1998, Bertsimas and Sim 2006, Goldfarb and Iyengar 2003). In these articles, specific robust optimization problems are re-formulated as other convex optimization problems, which can now be efficiently solved using, for example, interior-point methods (Boyd and Vandenberghe 2004, Ruszczyński 2006). Unfortunately, the taper design problem considered here is not convex, so none of these methods can be directly applied. Worst-case robust optimization has been applied in control (Zhou *et al.* 1996, Calafiore and Campi 2006), signal processing (Vorobyov *et al.* 2003, Lorenz and Boyd 2005), portfolio optimization (Rustem and Howe 2002, El Ghaoui *et al.* 2003), machine learning (Lanckriet *et al.* 2003), and other fields.

The outline of this article is as follows. In §2 the basic taper design problem is introduced, along with a piecewise-linear parametrization of the taper shape function that results in a finite dimensional problem. In §3 two methods for computing the magnitude of the reflected light due to miscoupling, as well as the gradient with respect to the taper shape variables, are described. The first method is approximate but very fast, and thus appropriate for use in an optimization routine, since it will be evaluated many times. The second method is a slower, brute-force method, which is presumably more accurate. In §4 the algorithm for robust taper design is described. This method is based on sequential linear programming, and a method for identifying the worst, or at least bad, values of the unknown parameters for a given taper shape. This basic method can and does get caught in poor local minima, but a successive refinement method described in §4.4 helps avoid this pitfall, and greatly improves the overall robustness of the taper designs. In §5 numerical results are presented, comparing the robust design to a linear taper, and also to optimized tapers that do not take parameter variations into account. Performance is verified using both the fast method for evaluating the reflection coefficient magnitude, as well as the slower, but more accurate, brute force method.

## 2. Nominal and robust taper design problems

### 2.1. *Taper shape and reflection magnitude*

Consider a taper with length $L$ that couples a uniform and a slow-light waveguide structure with period $\Lambda$. The taper is a quasi-periodic structure that is parametrized by the *taper shape function* $s : [0, 1] \to \mathbf{R}_+$. The argument of the taper shape function is the normalized length variable $u = z/L$, where $z$ is the physical coordinate along the taper. Each value of $s$ corresponds to an intermediate periodic structure between the taper endpoints, for example, in Figure 1, $s$ could correspond to the width of the flanges, the radius of the holes, or the separation of the blocks. The varying periodic structure described by $s(u)$ defines a taper as described in (Johnson *et al.* 2002); essentially, the taper matches the cross-section of the periodic structure $s(u)$ at $z = Lu$. The taper shape function is constrained at its starting and its final point, with $s(0) = 0$ denoting the starting uniform structure, and $s(1) = 1$ denoting the final periodic structure. Figure 2 illustrates a sample taper and its shape function, where in this case $s$ is simply proportional to the continuously varying width of the flanges.

Given a taper shape function, one can evaluate the magnitude of the reflection from an incoming light wave coupled from the uniform into the slow-light waveguide, for example by numerical simulation of the wave equation. The reflection magnitude is denoted $R$; it depends on the taper shape function $s$, as well as various parameters such as the refractive index (which might, indeed, vary spatially), the wavelength, and so on. These parameters are denoted by a vector $v \in \mathbf{R}^m$; to emphasize that $R$ is a function (or, since $s$ is a function, a functional) of the taper shape $s$ and the parameter vector $v$, it will sometimes be written as $R(s, v)$.

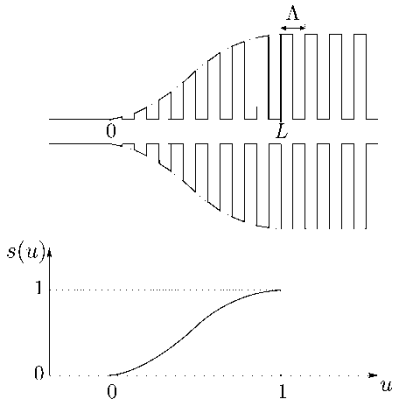


Figure 2. *Top.* A taper coupling uniform and slow-light waveguide structures. *Bottom.* Its taper shape function $s$.

Let $v_{\text{nom}}$ be the *nominal* value of $v$, *i.e.* a typical (or expected) value of the parameter vector. The *nominal reflection magnitude* is defined as

$$R_{\text{nom}}(s) = R(s, v_{\text{nom}}).$$

The nominal reflection magnitude is a functional of the taper shape function $s$, and gives the magnitude of the reflection when the parameter vector is equal to its nominal value.

### 2.2. *Parameter uncertainty and worst-case reflection magnitude*

Parameter uncertainty, which can be caused by manufacturing imperfections, wavelength variation, model parameter errors, *etc.*, is modelled by a set $\mathcal{V} \subseteq \mathbf{R}^m$. The set $\mathcal{V}$ can be thought of as the set of possible values of the parameter vector. It will be assumed that $v_{\text{nom}} \in \mathcal{V}$. As a simple (but important) example, $\mathcal{V}$ can be a finite set $\mathcal{V} = \{v_1, \ldots, v_K\}$. In this case the index $i$ is referred to as a *scenario*, with associated parameter vector $v_i$. As another common example, $\mathcal{V}$ can be a box in $\mathbf{R}^m$, for example, centred at the nominal parameter value,

$$\mathcal{V} = \{v \mid |v_i - v_{\text{nom},i}| \leq \xi_i,\ i = 1, \ldots, m\}, \tag{1}$$

where $\xi_i$ gives the radius or half-range of the variation in parameter $i$. (This type of parameter variation can be described as $v_i = v_{\text{nom},i} \pm \xi_i$.)

The performance of a taper design, in the presence of parameter uncertainty, is judged by the worst-case (largest possible) reflection magnitude over all possible $v \in \mathcal{V}$. The *worst-case reflection magnitude* is defined as

$$R_{\text{wc}}(s) = \sup_{v \in \mathcal{V}} R(s, v).$$

The worst-case magnitude reflection $R_{\text{wc}}$ is a functional of the taper shape function $s$. It is always the case that $R_{\text{wc}}(s) \geq R_{\text{nom}}(s)$ for any $s$; indeed, the ratio $R_{\text{wc}}(s)/R_{\text{nom}}(s)$ gives a measure of (worst-case) performance degradation of the taper, due to parameter variation.

For a scenario model of parameter uncertainty, *i.e.* when $\mathcal{V} = \{v_1, \ldots, v_K\}$, the worst-case reflection magnitude has the form

$$R_{\text{wc}}(s) = \max_{i=1,\ldots,K} R(s, v_i),$$

the maximum reflection magnitude over the $K$ scenarios. But in most cases, $R_{\text{wc}}(s)$ cannot be computed exactly, since this involves solving a non-convex optimization problem. It can be approximately computed, however, using several methods described in §3.4.

### 2.3. *Nominal and robust taper shape problems*

In the *nominal taper shape problem*, a taper shape function $s$ is found that minimizes the nominal reflection magnitude $R_{\text{nom}}$, subject to some constraints:

$$\begin{array}{ll} \text{minimize} & R_{\text{nom}}(s) \\ \text{subject to} & s(0) = 0, \quad s(1) = 1 \\ & 0 \le s(u) \le S^{\max}, \quad |s'(u)| \le D^{\max} \quad \text{for } 0 \le u \le 1. \end{array} \tag{2}$$

The optimization variable is the taper shape function $s : [0, 1] \to \mathbf{R}_+$. The problem parameters are the maximum allowed shape value $S^{\max}$, the maximum allowed taper slope $D^{\max}$, and, of course, the objective function $R_{\text{nom}}$. A solution of this problem is called a *nominal optimal taper*.

In the *robust taper shape problem*, the goal is to find a taper shape function $s$ that minimizes the worst-case reflection magnitude $R_{\text{wc}}$, subject to some constraints:

$$\begin{array}{ll} \text{minimize} & R_{\text{wc}}(s) \\ \text{subject to} & s(0) = 0, \quad s(1) = 1 \\ & 0 \le s(u) \le S^{\max}, \quad |s'(u)| \le D^{\max} \quad \text{for } 0 \le u \le 1. \end{array} \tag{3}$$

A solution of this problem is called a *robust optimal taper*. The main goal of this article is to present a tractable way to (approximately) solve the robust taper shape problem (3).

Both the nominal and robust taper shape problems (2) and (3) are infinite-dimensional optimization problems, since the optimization variable is a function (Anderson and Nash 1987), and they include semi-infinite constraints (Hettich and Kortanek 1993), *i.e.* an infinite set of constraints indexed by a continuous variable ($u$). Both of these issues will be (approximately) handled by searching over a finite-dimensional set of shape functions, for which the semi-infinite constraints can be expressed in a simple way. The complexity of the algorithm grows linearly with the dimension of the finite-dimensional parametrization, and easily scales to dimensions large enough (*e.g.* thousands) that errors due to the finite-dimensional parametrization are negligible.

A more fundamental issue is that the problems (2) and (3) are not convex (since the objectives are, in general, not convex), which makes it unlikely that the global solutions can be found efficiently. So one must settle for locally optimal solutions of the problems, which need not be globally optimal. In §4.4 a successive refinement approach is described, which appears to be quite resistant to getting trapped in poor local minima.

### 2.4. *Piecewise-linear taper shape parametrization*

It is assumed that the taper shape functions are piecewise-linear, parametrized by their values at $n$ fixed *control* or *break* or *grid* points $u_1, \ldots, u_n$, with $0 < u_1 < \cdots < u_n < 1$, and passing through the endpoints $s(0) = 0$ and $s(1) = 1$. This is illustrated for $n = 4$ grid points in Figure 3.

This can be expressed as

$$s(u) = x_1 f_1(u) + \cdots + x_n f_n(u) + f_{n+1}(u), \tag{4}$$

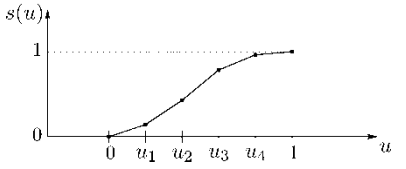


Figure 3. A piecewise-linear taper shape $s$ with $n = 4$, with grid-points $u_1, \ldots, u_4$. The taper shape satisfies $s(0) = 0$, $s(u_1) = x_1, \ldots, s(u_4) = x_4$, and $s(1) = 1$.

with

$$f_i(u) = \begin{cases} (u - u_{i-1})/(u_i - u_{i-1}) & u_{i-1} \leq u \leq u_i, \\ (u_{i+1} - u)/(u_{i+1} - u_i) & u_i \leq u \leq u_{i+1}, \\ 0 & \text{otherwise}, \end{cases}$$

where $u_0 = 0$ and $u_{n+1} = 1$, and

$$f_{n+1}(u) = \begin{cases} (u - 1)/(u_n - 1) & u_n \leq u \leq 1, \\ 0 & \text{otherwise}. \end{cases}$$

The vector $x \in \mathbf{R}^n$ is referred to as the *taper shape vector*. Evidently $s(u_i) = x_i$.

With this parametrization, the endpoint constraints $s(0) = 0$ and $s(1) = 1$ hold automatically, for any shape vector $x$. Moreover, the semi-infinite constraints

$$0 \leq s(u) \leq S^{\max}, \quad |s'(u)| \leq D^{\max} \quad \text{for } 0 \leq u \leq 1,$$

hold if and only if

$$\begin{aligned} &0 \leq x_i \leq S^{\max}, \quad i = 1, \ldots, n, \\ &|x_{i+1} - x_i| \leq D^{\max}(u_{i+1} - u_i), \quad i = 1, \ldots, n-1, \\ &|x_1| \leq D^{\max} u_1, \quad |1 - x_n| \leq D^{\max}(1 - u_n). \end{aligned} \tag{5}$$

These are a set of $4n$ linear inequalities on the shape vector $x$. The notation $x \in \mathcal{S}$ will be used to denote this, where $\mathcal{S}$ is the (polyhedral) set of $x$ for which (5) holds.

With some abuse of the notation, $R_{\text{nom}}(x)$ and $R_{\text{wc}}(x)$ will be used to denote the values of $R_{\text{nom}}(s)$ and $R_{\text{wc}}(s)$, for the shape function $s$ associated with the shape vector $x$. With piecewise-linear parametrization of taper shapes, the nominal taper design problem can be expressed as

$$\begin{array}{ll} \text{minimize} & R_{\text{nom}}(x) \\ \text{subject to} & x \in \mathcal{S}, \end{array} \tag{6}$$

and the robust taper design problem as

$$\begin{array}{ll} \text{minimize} & R_{\text{wc}}(x) \\ \text{subject to} & x \in \mathcal{S}. \end{array} \tag{7}$$

These are finite-dimensional optimization problems, with optimization variable $x \in \mathbf{R}^n$, and $4n$ linear inequality constraints.

## 3. Computation of reflection magnitude

To optimize the taper shape function $s$, one needs a rapid method to compute the reflected power fraction $R$ and its gradient for light incident on a particular taper structure. This article will employ two such methods, described below: a fast approximate method for the optimization (including the computation of the gradient), and a slower brute-force method for verification of the final design. In §3.4, some methods for estimating the worst-case reflection $R_{\text{wc}}$, given $s$ and the uncertainty parameter set $\mathcal{V}$, will be described.

### 3.1. *Coupled-mode theory*

In general, computing the reflection from an arbitrary structure could require an expensive solution of the complete Maxwell equations, evaluated to high accuracy in order to distinguish the tiny reflected field in a well-designed gradual taper. In the present case, however, the fact that the structure is *nearly* periodic (slowly varying) and the reflection is consequently small, can be exploited to utilize a fast semi-analytical method based on *coupled-mode theory*.

Coupled-mode theory, also known as *coupled-wave theory* or the *slowly-varying envelope approximation* (SVEA), involves an expansion of the electromagnetic field along the waveguide taper in terms of the eigenmodes (indexed by $k$) of a uniform periodic waveguide matching the cross-section at each point. The expansion coefficients $c_k$ in this basis are then determined by a set of ordinary differential equations for $dc_k/dz$ along the taper direction ($z$), where the different modes are coupled by terms proportional to the rate of change of the structure. Because the structure is slowly varying, the expansion coefficients approach an 'adiabatic' limit in which the $c_k$ are nearly constant. In this limit, the equations can be integrated approximately, to first-order in the taper rate, to yield a simple integral for the reflection coefficient. (Reflection dominates the loss in slow-light tapers.)

The most common form of coupled-mode theory was developed for nearly uniform waveguides (Marcuse 1991), but some of the authors have recently generalized this approach to strongly periodic waveguides of the type considered in this article (Johnson *et al.* 2002). The results of a simple first-order calculation were found to be nearly exact as long as the reflections were under 10%, making them ideal for the present case where the taper designs all have reflections well under 1%.

In particular, coupled-mode theory of a taper shape $s(u)$ with length $L$ leads to a first-order reflection amplitude $c_r$, where the fraction of reflected power is $R = |c_r|^2$, given by an integral of the form:

$$c_r\{s(u)\} = \int_0^1 du\, s'(u) \sum_k \frac{M_k[s(u)]}{\Delta\beta_k[s(u)]} e^{iL\int_0^u \Delta\beta_k[s(u')]\mathrm{d}u'}. \tag{8}$$

Here, $M_k$ and $\Delta\beta_k$ are given functions of the taper parametrization $s$. That is, each $s$ denotes a given intermediate periodic structure, $M_k(s)$ is a (complex-valued) coupling coefficient determined from the eigenfields of that structure, and $\Delta\beta_k(s)$ is a (real) phase-mismatch factor. The summation must in principle run over all integers $k$, but in practice only a handful of terms are required because the contributions decrease rapidly with $k$. (In particular, $\Delta\beta_k(s) = \Delta\beta(s) + 2\pi k/\Lambda(s)$, where $\Lambda(s)$ is the variable period along the taper.)

The derivation of these coupled-mode equations is rather complicated and will not be reproduced here.[1] The key point, however, is that the full Maxwell equations need only be solved once: a set of small calculations for the eigenmodes of the periodic structures at each $s$, by a spectral method (Johnson and Joannopoulos 2001*b*), yields the functions $M_k(s)$ and $\Delta\beta(s)$. One can then re-use these functions to compute the reflection for any taper shape $s(u)$ and any length $L$ by a single integral, which allows quick exploration and optimization over many different shapes.

The equations are the same regardless of the dimensionality of the problem, and have previously been used by some of the authors to compute taper reflections and perform simple optimizations in large three-dimensional structures where direct simulation was not possible (Povinelli *et al.* 2005).

### 3.2. *Coupled-mode reflection gradient*

To carry out taper shape optimization one will need to evaluate not only the reflection magnitude $R$ but also its functional derivative (gradient) $\partial R/\partial s$. In general, such gradients can be computed by an adjoint method (Cao *et al.* 2003), but in this case the problem is simple enough that one can derive the same thing without resorting to such cumbersome techniques.

In particular, since $R = |c_r|^2$ and $c_r$ is a summation over $k$, it suffices to compute the gradient of each $k$ term in the summation equation for $c_r$ above. Dropping the $k$ subscript for simplicity, each $k$ term corresponds to the functional:

$$c\{s(u)\} = \int_0^1 du\, s'(u) F[s(u)] e^{\int_0^u f[s(u')]du'}, \tag{9}$$

where $F(s) = M_k(s)/\Delta\beta_k(s)$ and $f(s) = iL\Delta\beta_k(s)$. The gradient $g(u)$ of this functional is defined by the first-order change of $c\{s(u)\}$ under a small change $\delta s(u)$ (where $\delta s(0) = \delta s(1) = 0$ to preserve the boundary conditions):

$$\delta c = c\{s + \delta s\} - c\{s\} = \int_0^1 g(u)\delta s(u) du. \tag{10}$$

The explicit gradient $g$ can be derived by substituting $s + \delta s$ into $c$, dropping terms higher than first-order in $\delta s$, and integrating by parts to eliminate the $\delta s'$ term. After some algebra, one obtains:

$$g(u) = -F[s(u)]f[s(u)]e^{\int_0^u f[s(u')]du'} + f'[s(u)] \int_u^1 d\tilde{u} s'(\tilde{u}) F[s(\tilde{u})] e^{\int_0^{\tilde{u}} f[s(u')]du'}, \tag{11}$$

which is a single integral in terms of $s(u)$ and the known functions $F$ and $f$ and their derivatives, which means that the gradient can be evaluated with roughly the same cost as evaluating $c\{s(u)\}$ (similar to what one would expect for adjoint methods).

In practice, of course, infinitely many degrees of freedom are not present in $s(u)$. As explained in §2.4, a piecewise-linear parametrization $s(u) = \sum_i x_i f_i(u)$, for 'tent' functions $f_i(u)$ and parameters $x_i$, is employed. One therefore needs only the finite-dimensional gradient with respect to the $x_i$:

$$\frac{\partial c}{\partial x_i} = \int_0^1 g(u) f_i(u)\, du. \tag{12}$$

The gradient of the reflection $R$ is then found by first summing $\partial c/\partial x_i$ over $k$ to obtain $\partial c_r/\partial x_i$, and then $\partial R/\partial x_i$ is the real part of $2c_r^* \partial c_r/\partial x_i$.

### 3.3. *Brute-force verification*

Because coupled-mode theory involves some approximations, it is also desirable to directly solve the Maxwell equations, with no assumptions, in order to verify the correctness of the solutions. Such a direct solution allows one to consider the effect of imperfections that violate the slow-taper assumption underlying coupled-mode theory; in particular, one can include rapid small variations in the structure corresponding to fabrication imperfections (*e.g.* surface roughness). The specific

computational method employed is an eigenmode-expansion, or transfer-matrix, method that is implemented in a free software package called CAMFR (Bienstman 2001, 2006).

CAMFR works by expanding the fields at every $z$ in terms of the eigenmodes of that cross-section, with perfectly-matched layer (PML) absorbing boundaries in the lateral directions (Bérenger 1994). In this sense, it is related to the classic coupled-mode method mentioned above (Marcuse 1991). Unlike the first-order integration above, however, CAMFR makes no assumption of small scattering or slow variation, and computes a complete transfer matrix at each point where the cross-section changes that couples all the modes according to the continuity conditions on the electromagnetic field. In this sense, it is a 'brute-force' method: it solves the complete Maxwell equations with no assumptions, to an arbitrary accuracy given enough computational time and memory (*i.e.* a large enough eigenmode basis). However, it is extremely efficient for structures like the one considered in Figure 1(c), in which the cross-section is piecewise uniform, since the uniform regions are handled analytically.

Moreover, CAMFR imposes the incident-wave boundary conditions (at $z = 0$ and $z = L$) analytically, thanks to its eigenmode basis, and hence can distinguish even a tiny reflection coefficient with high accuracy. It is most effective, however, when the two ends of the simulation are terminated by semi-infinite uniform waveguide, and so the CAMFR simulations are performed using a double taper, which tapers from uniform to periodic, then five periods in the periodic structure, and then tapers back from periodic to uniform.

### 3.4. *Worst-case reflection magnitude*

The problem of finding the worst, or at least a bad, value of the parameter $v \in \mathcal{V}$, for a given taper shape $s$, is called *pessimizing*, since the goal is to find the least favourable value of the parameter for the given shape. When $\mathcal{V}$ is finite, exact pessimizing can be carried out by evaluating $R$ under each scenario and taking the largest value found.

When $\mathcal{V}$ is infinite it is difficult to compute the exact value of the worst-case reflection magnitude

$$R_{\mathrm{wc}}(s) = \sup_{v \in \mathcal{V}} R(s, v),$$

along with a (worst-case) parameter $v^\star$ that achieves this supremum, since in general $R(s, v)$ is not concave in $v$ (and $\mathcal{V}$ need not be convex). Options for pessimizing include direct search methods (Nelder and Mead 1965, Wright 1995, Kolda *et al.* 2003), or any standard local optimization method such as sequential quadratic programming methods (Nocedal and Wright 1999, Gill *et al.* 2005). With any of these methods, the algorithm is run from a number of starting points in $\mathcal{V}$; the largest value of $R$ found is then an estimate of $R_{\mathrm{wc}}$.

When $\mathcal{V}$ is a box (1), one can easily guess a value of $v$ that often leads to large (if not largest) $R$. The gradient of $R$ with respect to $v$ is evaluated at $v_{\mathrm{nom}}$; the approximate pessimizer is then

$$v_i^\star \approx \begin{cases} v_{\mathrm{nom},i} + \xi_i, & \partial R/\partial v_i > 0 \\ v_{\mathrm{nom},i} - \xi_i, & \partial R/\partial v_i < 0. \end{cases}$$

(This is the maximizer of the first-order approximation of $R$ over $\mathcal{V}$.) This point can, of course, be used as the starting point for a local optimization method.

## 4. Taper shape optimization

In this section the methods for nominal, multi-scenario, and general robust taper shape optimization, for a given piecewise-linear taper shape parametrization, are described. The section finishes by explaining a method that applies to all three, in which the taper shape is successively refined.

### 4.1. *Nominal design*

A standard local linearization method, with a trust region constraint (see, *e.g.* Conn *et al.* (2000), Nocedal and Wright (1999, Chap. 4)) is used, but any other first-order nonlinear optimization method could have been used here. Let $x$ be the current shape vector parameter, which is assumed feasible (*i.e.* $x \in \mathcal{S}$). The reflection function $R_{\text{nom}}$ is replaced with the affine approximation

$$R_{\text{nom}}(x+\delta) \approx R_{\text{nom}}(x) + \nabla R_{\text{nom}}(x)^T \delta, \tag{13}$$

where $\delta \in \mathbf{R}^n$ is a proposed perturbation, and a trust region constraint

$$\|\delta\|_\infty = \max_i |\delta_i| \leq \rho$$

is added to ensure the approximation (13) is good enough for the algorithm to make adequate progress. The size of the trust region is scaled by $\rho > 0$, which is called the trust region radius. For example, for $\rho = 0.01 S^{\max}$, the trust region constraint limits the maximum perturbation to 1% of the maximum allowable shape function value. (Another trust region constraint that gives good results is the $l_2$-norm constraint on the perturbation and its difference, *i.e.* $\|\delta\|_2 \leq \rho_1$ and $\|\delta_{i+1} - \delta_i\|_2 \leq \rho_2$, where $\rho_1$ and $\rho_2$ are appropriate constants.)

To find $\delta$, the search direction subproblem

$$\begin{array}{ll} \text{minimize} & R_{\text{nom}}(x) + \nabla R_{\text{nom}}(x)^T \delta \\ \text{subject to} & x + \delta \in \mathcal{S}, \quad \|\delta\|_\infty \leq \rho, \end{array} \tag{14}$$

is solved, with optimization variable $\delta \in \mathbf{R}^n$. This gives the best possible update $\delta$, based on the linearized objective, and subject to the constraint that $x + \delta$ be feasible, and the trust region constraint.

The problem (14) can be expressed as a linear program (LP),

$$\begin{array}{ll} \text{minimize} & \nabla R_{\text{nom}}(x)^T \delta \\ \text{subject to} & 0 \leq x_i + \delta_i \leq S_{\max}, \quad i = 1, \ldots, n \\ & -D^{\max}(u_{i+1} - u_i) \leq x_{i+1} + \delta_{i+1} - x_i - \delta_i \leq D^{\max}(u_{i+1} - u_i), \quad i = 1, \ldots, n-1 \\ & -D^{\max} u_1 \leq x_1 + \delta_1 \leq D^{\max} u_1 \\ & -D^{\max}(1 - u_n) \leq 1 - x_n - \delta_n \leq D^{\max}(1 - u_n) \\ & -\rho \leq \delta_i \leq \rho, \quad i = 1, \ldots, n, \end{array} \tag{15}$$

with $n$ variables and $6n$ inequality constraints. This particular LP can be solved very efficiently, by exploiting its structure. Each constraint involves only a single variable $\delta_i$, or the difference between successive variables, $\delta_{i+1} - \delta_i$. As a result, the linear system that needs to be solved to determine the search direction in each step of an interior-point method is tridiagonal, and so can be solved extremely fast. Since only a few tens of iterations are needed to solve the LP, the entire cost is $O(n)$ flops. (See, *e.g.* Boyd and Vandenberghe (2004 Chap. 11, App. C).) Any general purpose LP solver that exploits sparsity will recognize and exploit this structure automatically (although not quite as efficiently as when the tridiagonal structure is recognized from the beginning).

Once the tentative update $\delta$ is computed, $R_{\text{nom}}(x+\delta)$ is evaluated. If this is not worse than the current point $x$, *i.e.* $R_{\text{nom}}(x+\delta) \leq R_{\text{nom}}(x)$, one accepts the update step and sets $x := x+\delta$, and increases the trust region radius $\rho$ (up to some maximum value $\rho^{\max}$). If the tentative point $x+\delta$ is worse than the current point, *i.e.* $R_{\text{nom}}(x+\delta) > R_{\text{nom}}(x)$, one reduces the trust region radius $\rho$ (down to some minimum value $\rho^{\min}$), and re-computes the step $\delta$. As a stopping criterion, a maximum number of iterations $N^{\max}$ can be specified, or the algorithm can terminate when, even with the minimum trust region radius, there is no improvement in $R_{\text{nom}}$. The overall algorithm is summarized below.

NOMINAL TAPER SHAPE (NTS) ALGORITHM.

**given** initial feasible $x$, initial $\rho > 0$, and parameters $\rho^{\min} > 0$, $\rho^{\max} > 0$, $\alpha^{\text{decr}} < 1$, $\alpha^{\text{incr}} > 1$, and $N^{\max}$.

**repeat**

1. Determine an update step $\delta$ by solving (15).
2. *Update.*

2a. **if** $R_{\text{nom}}(x+\delta) \leq R_{\text{nom}}(x)$

$x := x+\delta$; $\rho := \min\{\rho^{\max}, \alpha^{\text{incr}}\rho\}$

2b. **else**

$\rho := \max\{\rho^{\min}, \alpha^{\text{decr}}\rho\}$.

**until** stopping criterion is satisfied.

### 4.2. *Multi-scenario design*

The NTS algorithm is now extended to (approximately) solve the robust taper shape problem (7), when $\mathcal{V} = \{v_1, \ldots, v_K\}$. (This problem is referred to as the *multi-scenario taper shape problem*.) In this case, the objective in (7) is

$$R_{\text{wc}}(x) = \max_{i=1,\ldots,K} R(x, v_i). \tag{16}$$

The algorithm is similar to the NTS algorithm, but with two differences: a tentative update step based on the linearization of each term in the objective (16) is computed, and the objective (16) is evaluated when determining if whether to accept or reject the tentative update step. This new algorithm is called the multi-scenario taper shape (MSTS) algorithm. A similar algorithm for solving nonlinear discrete minimax (*i.e.* multi-scenario) problems was presented in Madsen and Schjaer-Jacobsen (1978), Jonasson and Madsen (1994), where the authors also give a proof of convergence to stationary points. An algorithm that solves a discrete minimax problem using penalty functions and trust-region methods can also be found in Erdmann and Santosa (2004).

In the update step calculation the affine approximation of $R$ is used for each scenario,

$$R(x, v_i) \approx R(x, v_i) + \nabla R(x, v_i)^T \delta,$$

and the objective (16) is approximated with the piecewise-linear function

$$R_{\text{wc}}(x) \approx \max_{i=1,\ldots,K} (R(x, v_i) + \nabla R(x, v_i)^T \delta).$$

At each iteration, the following problem is solved to compute a tentative update step $\delta$:

$$\begin{aligned} \text{minimize} \quad & \max_{i=1,\ldots,K} (R(x, v_i) + \nabla R(x, v_i)^T \delta) \\ \text{subject to} \quad & x+\delta \in \mathcal{S}, \quad \|\delta\|_\infty \leq \rho. \end{aligned}$$

This problem is easily solved, since it can be cast as the LP; see Boyd and Vandenberghe (2004, Chap. 4),

$$
\begin{array}{ll}
\text{minimize} & t \\
\text{subject to} & R(x, v_i) + \nabla R(x, v_i)^T \delta \leq t, \quad i = 1, \ldots, K \\
& 0 \leq x_i + \delta_i \leq S_{\max}, \quad i = 1, \ldots, n \\
& -D^{\max}(u_{i+1} - u_i) \leq x_{i+1} + \delta_{i+1} - x_i - \delta_i \leq D^{\max}(u_{i+1} - u_i), \quad i = 1, \ldots, n-1 \\
& -D^{\max} u_1 \leq x_1 + \delta_1 \leq D^{\max} u_1 \\
& -D^{\max}(1 - u_n) \leq 1 - x_n - \delta_n \leq D^{\max}(1 - u_n) \\
& -\rho \leq \delta_i \leq \rho, \quad i = 1, \ldots, n.
\end{array} \tag{17}
$$

Here the optimization variables are $\delta \in \mathbf{R}^n$ and $t \in \mathbf{R}$. This LP has $n + 1$ variables and $6n + K$ inequality constraints. Again, the associated LP coefficient matrix is sparse, except for the block of $K$ constraints generated by the first $K$ inequalities. The LP can be solved in $O(\min\{n, K\}^2 \max\{n, K\})$ flops. (See, *e.g.* Boyd and Vandenberghe (2004, Chap. 11, App. C).)

The MSTS algorithm outline is the same as the one for the NTS algorithm, except that the linear program (17) is solved in step 1, and $R_{\text{wc}}$ is evaluated instead of $R_{\text{nom}}$ in step 2.

### 4.3. *Robust design*

In this section, an algorithm to (approximately) solve the robust taper shape problem (7) with general parameter set $\mathcal{V}$ is presented. The algorithm is based on carrying out a sequence of multi-scenario robust designs, with an expanding set $\mathcal{V}^{\text{bad}}$ of scenarios that are found by approximate worst-case analysis.

ROBUST TAPER SHAPE (RTS) ALGORITHM.

**given** initial feasible $x$ and $\mathcal{V}^{\text{bad}} = \{v_{\text{nom}}\}$.

**repeat**

1. *Pessimization. (Approximate worst-case analysis.)*
   1a. Find approximate worst-case parameter value $v^\star$ for current $s$.
   1b. $\mathcal{V}^{\text{bad}} := \mathcal{V}^{\text{bad}} \cup \{v^\star\}$.
2. *Multi-scenario design.*
   Solve MSTS problem with $\mathcal{V}^{\text{bad}}$, starting from current shape.

**until** stopping criterion is satisfied.

As a stopping criterion, a maximum number of iterations $N^{\max}$ can be specified, or termination can occur when there is no improvement in $R_{\text{wc}}$. Several variations on the basic algorithm given above can be given. First, instead of producing just one bad parameter value, several bad parameter values can be found in the pessimization step, and these are appended to the set $\mathcal{V}^{\text{bad}}$. Second, one can add a step in which $\mathcal{V}^{\text{bad}}$ is pruned, *i.e.* parameter values are removed from the set.

### 4.4. *Successive refinement*

The taper design problems (6) and (7) are non-convex, and local methods, such as the local linearization method described above, can (and do) get stuck in poor locally optimal points. A common method to fix this problem is to run the algorithm multiple times, starting with different initial taper designs, picking the best design obtained among the runs of the algorithm. A method called *successive refinement*, however, seems to avoid the problem of getting caught in poor local minima, and eliminates the need for multiple runs from different starting points.

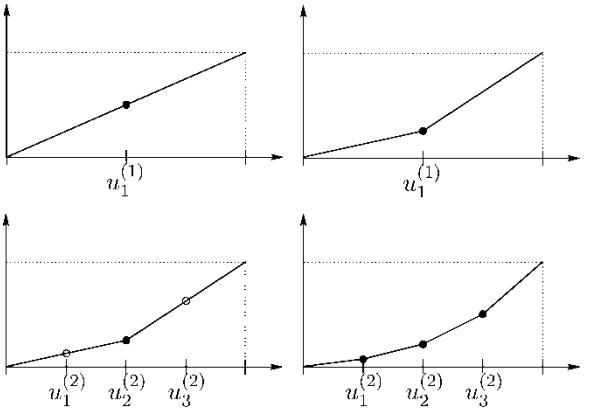


Figure 4. *Top left*. Linear taper with a single grid point. *Top right*. Full search performed to obtain a global optimum taper with a single grid point. *Bottom left*. Two new grid points added and taper values interpolated at $u_1^{(2)}$ and $u_3^{(2)}$. Optimization algorithm is run starting from this taper. *Bottom right*. New local optimum.

In successive refinement, a sequence of design problems with successively finer piecewise-linear taper shape functions is solved, in each case starting from the previous design. One starts with a single grid point, *i.e.* $n = 1$, and runs a global search of the optimal robust taper, which is tractable only for this single dimensional problem. One then adds two more grid points, in between 0 and the first grid point, and the first grid point and 1, so that $n = 3$, and runs the RTS algorithm described above, starting from the previous design. This is repeated until some maximum value of $n$ is reached. This is illustrated in Figure 4.

In numerical experiments the authors started with initial grid point at $1/2$, and in each successive refinement step, new grid points are added halfway in-between the old ones (and 0 and 1). At the $M$th refinement step there will be $n = 2^M - 1$ grid points, with values

$$u_i^{(M)} = i2^{-M}, \quad i = 1, \ldots, 2^M - 1.$$

This approach is related in spirit to the multigrid method (Briggs *et al.*. 2000), where the latter uses both successive refinements and coarsenings in order to speed up convergence of a linear solver rather than to avoid local minima. Successive refinement ideas have been successfully applied in circuit design (Chan *et al.* 2000), in motion estimation for video coding (Chun and Ra 1992), *etc*.

## 5. Numerical results

In this section some numerical results for a particular structure are presented.

### 5.1. *Taper geometry and uncertainty model*

The two-dimensional taper depicted in Figure 1(c), similar to the one considered in Johnson *et al.* (2002), will be optimized, in order to have a structure where the brute-force method is efficient (and thus can be used to validate the coupled-mode theory for a large number of values of the parameters). The periodic structure is a sequence of dielectric blocks with period $\Lambda_0$, size $0.4\Lambda_0 \times 0.4\Lambda_0$, and dielectric constant $\varepsilon = 12$. The blocks are surrounded by air ($\varepsilon = 1$). The electric field is polarized perpendicular to the 2d plane ('TM' polarization). As described in Johnson *et al.* (2002), this structure supports true localized guided modes by the mechanism

of index-guiding (Fan *et al.* 1995), and has a zero-group-velocity band edge at a frequency of $\omega\Lambda_0/2\pi = 0.2434$. The operating frequency is $\omega\Lambda_0/2\pi c = 0.23$,[2] which is slightly below the band edge, where the group velocity is under $c/4$ and the waveguide is single-mode at every point along the taper.

A uniform waveguide of width $0.4\Lambda_0$, which can be treated as a sequence of (touching) blocks with period $0.4\Lambda_0$, is tapered to the periodic structure by gradually spreading the blocks apart. That is, their period varies as $\Lambda(s) = \Lambda_0[s + 0.4(1 - s)]$, so that $s = 0$ corresponds to the uniform structure with pitch $0.4\Lambda_0$ and $s = 1$ corresponds to the periodic structure with pitch $\Lambda_0$. The problem is then to determine the function $s$ describing how fast the period (pitch) varies along the taper. The taper will be designed with length $L = 30\Lambda_0$, maximum shape value $S^{\max} = 1$, and maximum slope $D^{\max} = 5$.

On physical grounds, one expects the optimal taper to be more rapid at the $u = 0$ end corresponding to the uniform waveguide where the group velocity is larger, and to be more gradual at the $u = 1$ end corresponding to the periodic waveguide where the group velocity is low (and thus the structure is more sensitive to small changes (Povinelli *et al.* 2005)). This is precisely what is found, below, although the exact taper rate is difficult to predict *a priori*.

The following parameter uncertainty model is used. The operating frequency varies $\pm 1\%$ around its nominal value $\omega\Lambda_0/2\pi c = 0.23$; variation in the taper shape function is bounded at each grid point by $\pm 0.001$ around the current value, with the perturbed taper shape within the bounds 0 and 1. The shape variation is meant to model, for example, manufacturing variation.

### 5.2. *Pessimizing method*

The following method is used to carry out approximate worst-case analysis. At each of 20 values of $\omega$, uniformly spaced over the interval $[0.227, 0.233]$, the approximate worst-case shape perturbation at the current point $s$ is found using the derivative heuristic as described in §3.4, *i.e.*

$$s^\star(u) = \begin{cases} \min\{s(u) + 0.001, 1\} & \partial R/\partial s(u) > 0 \\ \max\{s(u) - 0.001, 0\} & \partial R/\partial s(u) < 0. \end{cases}$$

(The worst-case shape perturbation depends on $\omega$.) The reflection magnitude is evaluated for each $\omega$, with its associated approximate worst-case taper shape. The result is the approximate worst-case reflection magnitude over the shape uncertainty; maximizing over the 20 values of $\omega$ yields the approximation of $R_{\text{wc}}$.

The authors cannot claim that this pessimization heuristic gives the true worst-case value. However, it has been tested extensively, by attempting to find worse parameter values using other methods, such as derivative-free optimization, SQP, and simply sampling random parameter values in $\mathcal{V}$. In no case was a significantly worse value of the parameter found.

### 5.3. *Optimization*

Tapers were found using the NTS and RTS algorithms, with the following algorithm parameters: initial $\rho = 0.1S^{\max}$, $\rho^{\min} = 0.001S^{\max}$, $\rho^{\max} = 0.5S^{\max}$, $\alpha^{\text{decr}} = 0.75$, $\alpha^{\text{incr}} = 1.25$, and $N^{\max} = 150$, terminating also if no improvement is made. 10 iterations of successive refinement are used, with dyadic grid points, which results in a final taper design with $n = 1023$ grid points. Global optimization is carried out during the first step of the successive refinement, after which the obtained shape is used to construct initial points for the subsequent steps.

The NTS algorithm, and the MSTS algorithms carried out in each iteration of the RTS algorithm, usually terminate in 50–70 steps, due to no improvement in objective value. The RTS algorithm

converged after around 30–40 steps (each of which consisted of an approximate worst-case analysis and a multi-scenario optimization). For highest level of refinement, the RTS algorithm required a total of around 2000 basic iterations (each requiring an approximate worst-case analysis, the solution of an LP, etc.).

The algorithms were implemented in Matlab, solving the update step subproblems (15) and (17) using CVX (Grant *et al.* 2006), which calls the SeDuMi solver (Sturm 1999). The subproblem calculation for the NTS algorithm with $n = 1023$ variables (the last step in successive refinement) takes about a second, while the subproblem calculation for the MSTS algorithm with $n = 1023$ variables and $K = 50$ scenarios takes about ten seconds (on a personal computer). Solving the NTS problem required a total of around 40 seconds, and solving the RTS problem required a total of around 20 minutes. Had algorithms been implemented in C, using a custom LP solver for the particular structure that arises in these problems, these times would likely have been far smaller, by a factor exceeding 10.

### 5.4. *Results*

Figure 5 shows the nominal and robust taper designs obtained, together with the linear taper. The performance of these three taper designs is compared in Table 1, which gives the nominal reflection magnitude, and the approximate worst-case reflection magnitude computed using the fast coupled-mode solver, and also using the brute-force CAMFR solver. (In order to minimize numerical errors in the simulations the double taper setup as described in Johnson *et al.* (2002, Sec. 5) is used.) The calculations of approximate worst-case reflection magnitude, using the fast and brute-force methods, agree reasonably well. The robust optimal design gives a worst-case reflection that is around an order of magnitude better than the nominal design, and almost two orders of magnitude better than the simple linear taper design.

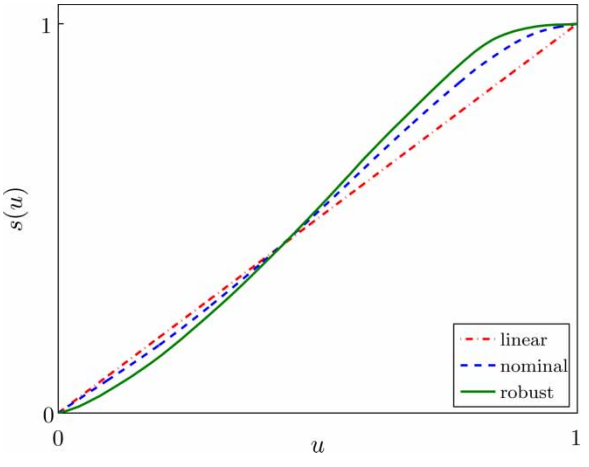


Figure 5. Linear, nominal optimal, and robust optimal designs.

Table 1. Summary of results for three tapers.

| | Linear taper | Nominal taper | Robust taper |
|---|---|---|---|
| $R_{\text{nom}}$ | $3.27 \times 10^{-3}$ | $1.75 \times 10^{-10}$ | $3.15 \times 10^{-8}$ |
| $R_{\text{wc}}$ (fast) | $8.60 \times 10^{-3}$ | $4.03 \times 10^{-4}$ | $2.09 \times 10^{-5}$ |
| $R_{\text{wc}}$ (brute) | $1.90 \times 10^{-2}$ | $9.77 \times 10^{-4}$ | $2.19 \times 10^{-5}$ |

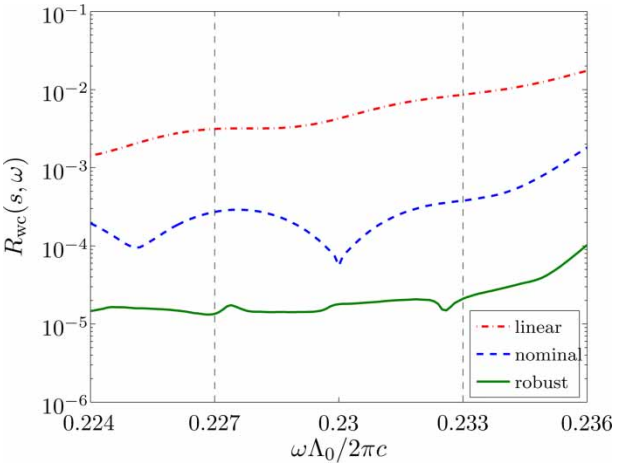


Figure 6. Partial worst-case values over the operating frequency generated by using the faster, but less accurate coupled-mode computations. The tapers were designed using the same coupled-mode model of the reflected power.

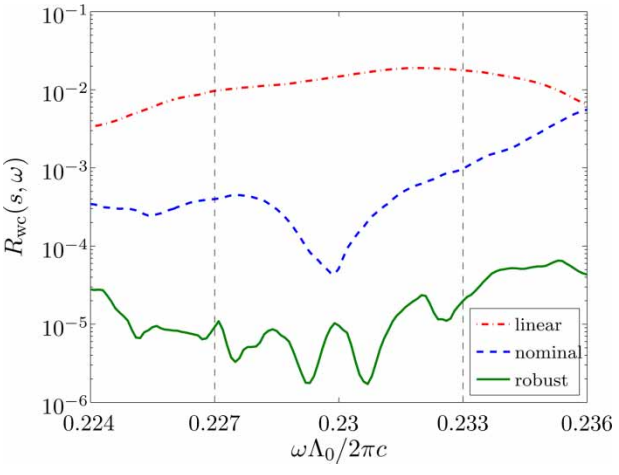


Figure 7. Partial worst-case values over the operating frequency generated by using the slower, but more accurate brute-force computations.

To give a little more insight into the performance of the three designs, the approximate worst-case reflection is plotted, over the shape parameter, as $\omega$ varies over the interval [0.224, 0.236], in Figure 6. Figure 7 shows the values computed using the brute-force method. This 'partial' worst-case values were generated by fixing a value of $\omega$ and approximately computing the worst reflection over all allowed taper shape perturbations. One can observe from Figure 6 that the partial worst-case value for the robust taper is approximately flat over the uncertain $\omega$ region, which is a desired characteristic of a robust (minimax) solution.

## 6. Conclusions

In this article, a new approach to non-convex robust optimization is presented, which is applied to the challenging problem of designing robust taper transitions to 'slow-light' periodic waveguides. The robust optimization algorithm is based on multi-scenario optimization with iterative sampling

of uncertain parameters, and uses fast and accurate coupled-mode computations in order to quickly explore different taper designs. The approach also uses the idea of successive refinement in order to avoid poor locally optimal points and to improve design robustness to taper shape uncertainty.

Experimental results verify that the obtained robust tapers perform well over a given set of parameter variations, while the optimized tapers that do not take parameter variation into account only perform well under nominal conditions. Robust performance of the designs is verified using an accurate, but much more expensive, method for evaluating the reflection coefficient.

The techniques presented in this article should be well suited for robust optimization of other non-convex PDE-based problems that lack most analytical guarantees. The authors are planning to apply robust optimization methods to the complex problems of microcavity design and superlensing. Other applications are also being considered.

## Acknowledgements

This work was supported in part by Dr. Dennis Healy of DARPA MTO, under award N00014-05-1-0700 administered by the Office of Naval Research, by the National Science Foundation under grant #0423905, by the Air Force Office of Scientific Research under grant #F49620-01-1-0365, and by MARCO Focus Center for Circuit & System Solutions contract #2003-CT-888.

## Notes

1. Our original derivation (Johnson *et al.* 2002) did not include an explicit shape function $s(u)$. However, it was noted that the coupling matrix elements were simply proportional to the taper rate, and this is what allows us to pull out the taper-rate dependence as an $s'(u)$ term in the integral.
2. It is convenient to use dimensionless frequency units of $2\pi c/\Lambda_0$, where $c$ is the speed of light in vacuum, because of the scale-invariance of Maxwell's equations (Joannopoulos *et al.* 1995).

## References

Anderson, E.J. and Nash, P., 1987. *Linear programming in infinite-dimensional spaces*. New York: John Wiley.

Barclay, P., *et al.*, 2004. Efficient input and output fiber coupling to a photonic crystal waveguide. *Optics Letters*, 29 (7), 697–699.

Ben-Tal, A. and Nemirovski, A., 1998. Robust convex optimization. *Mathematics of Operations Research*, 23 (4), 769–805.

Ben-Tal, A. and Nemirovski, A., 2002. Robust optimization: methodology and applications. *Mathematical Programming Series B*, 92 (3), 453–480.

Bérenger, J.-P., 1994. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114 (1), 185–200.

Bertsimas, D. and Sim, M., 2006. Tractable approximations to robust conic optimization problems. *Mathematical Programming*, 107 (1-2), 5–36.

Beyer, H.-G. and Sendhoff, B., 2007. Robust optimization - a comprehensive survey. *Computer Methods in Applied Mechanics and Engineering*, 196 (33–34), 3190–3218.

Bienstman, P., 2001. *Rigorous and efficient modelling of wavelength scale photonic components*. Thesis (PhD). Ghent University, Ghent, Belgium.

Bienstman, P., 2006. *CAMFR: cavity modelling framework software* [online]. Department of Information Technology (INTEC), Ghent University, Ghent, Belgium. Available from: camfr.sourceforge.net. [Accessed 1 October 2006].

Bienstman, P., *et al.*, 2003. Taper structures for coupling into photonic crystal slab waveguides. *Journal of the Optical Society of America – Part B, Optical Physics*, 20 (9), 1817–1821.

Birge, J.R. and Louveaux, F., 1997. *Introduction to stochastic programming*. New York: Springer.

Boyd, S. and Vandenberghe, L., 2004. *Convex optimization*. Cambridge, UK: Cambridge University Press.

Briggs, W., Henson, V., and McCormick, S., 2000. *A multigrid tutorial*. 2nd edn. Philadelphia, PA: SIAM.

Calafiore, G. and Campi, M., 2005. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102 (1), 25–46.

Calafiore, G. and Campi, M., 2006. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51 (5), 742–753.

Calafiore, G. and Dabbene, F., eds, 2006. *Probabilistic and randomized methods for design under uncertainty*. London, UK: Springer-Verlag.

Cao, Y., *et al.*, 2003. Adjoint sensitivity analysis for differential-algebraic equations: the adjoint DAE system and its numerical solution. *SIAM Journal on Scientific Computing*, 24 (3), 1076–1089.

Chan, T., *et al.*, 2000. Multilevel optimization for large-scale circuit placement. *In*: *Proceedings of the 2000 IEEE/ACM International Conference on Computer-Aided Design*. IEEE, 5-9 November, 2000 San Jose, CA, pp. 171–176.

Chandrasekaran, S., *et al.*, 1998. Parameter estimation in the presence of bounded data uncertainties. *SIAM Journal on Matrix Analysis and Applications*, 19 (1), 235–252.

Chietera, G., *et al.*, 2004. Numerical design for efficiently coupling conventional and photonic-crystal waveguides. *Microwave and Optical Technology Letters*, 42 (3), 196–199.

Chun, K. and Ra, J., 1992. Fast block matching algorithm by successive refinement of matching criterion. *Proceedings of the International Society for Optical Engineering*, 1818, 552–560.

Conn, A., Gould, N., and Toint, P., 2000. *Trust region methods*. Philadelphia, PA: SIAM.

Constantinou, C. and Jones, R., 1992. Path-integral analysis of an arbitrarily tapered, multimode, graded-index waveguide: the inverse-square-law and parabolic tapers. *Proceedings of the Institution of Electrical Engineers – Part B, Journal of Optoelectronics*, 139 (6), 365–375.

de Farias, D. and Roy, B.V., 2004. On constraint sampling in the linear programming approach to approximate dynamic programming. *Mathematics of Operations Research*, 29 (3), 462–478.

Dehnad, K., 1989. *Quality control, robust design, and the Taguchi method*. Pacific Grove, CA: Wadsworth Pub.

Dossou, K., *et al.*, 2006. Efficient couplers for photonic crystal waveguides. *Optics Communications*, 265 (1), 207–219.

El Ghaoui, L. and Lebret, H., 1997. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18 (4), 1035–1064.

El Ghaoui, L., Oks, M., and Oustry, F., 2003. Worst-case value-at-risk and robust portfolio optimization: a conic programming approach. *Operations Research*, 51 (4), 543–556.

El Ghaoui, L., Oustry, F., and Lebret, H., 1998. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9 (1), 33–52.

Elachi, C., 1976. Waves in active and passive periodic structures: a review. *Proceedings of the Institute of Electrical and Electronics Engineers*, 64 (12), 1666–1698.

Erdmann, G. and Santosa, F., 2004. Minimax design of optically transparent and reflective coatings. *Journal of the Optical Society of America – Part A, Optics, Image Science, and Vision*, 21 (9), 1730–1739.

Fan, S., *et al.*, 1995. Guided and defect modes in periodic dielectric waveguides. *Journal of the Optical Society of America – Part B, Optical Physics*, 12 (7), 1267–1272.

Felici, T. and Engl, H., 2001. On shape optimization of optical waveguides using inverse problem techniques. *Inverse Problems*, 17 (4), 1141–1162.

Gill, P., Murray, W., and Saunders, M., 2005. SNOPT: an SQP algorithm for large-scale constrained optimization. *SIAM Review*, 47 (1), 99–131.

Goldfarb, D. and Iyengar, G., 2003. Robust convex quadratically constrained programming. *Mathematical Programming*, 97 (3), 495–515.

Grant, M., Boyd, S., and Ye, Y., 2006. *CVX: matlab software for disciplined convex programming, version 1.0RC* [online]. Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA. Available from: www.stanford.edu/boyd/cvx/. [Accessed 10 December 2007].

Hakansson, A., Sanchez-Dehesa, J., and Sanchis, L., 2005. Inverse design of photonic crystal devices. *IEEE Journal On Selected Areas In Communications*, 23 (7), 1365–1371.

Hakansson, A., *et al.*, 2005. High-efficiency defect-based photonic-crystal tapers designed by a genetic algorithm. *Journal of Lightwave Technology*, 23 (11), 3881–3888.

Happ, T., Kamp, M., and Forchel, A., 2001. Photonic crystal tapers for ultracompact mode conversion. *Optics Letters*, 26 (14), 1102–1104.

Hettich, R. and Kortanek, K., 1993. Semi-infinite programming: theory, methods, and applications. *SIAM Review*, 35 (3), 380–429.

Jiang, J., *et al.*, 2003. Parallel microgenetic algorithm design for photonic crystal and waveguide structures. *Optics Letters*, 28 (23), 2381–2383.

Joannopoulos, J., Meade, R., and Winn, J., 1995. *Photonic Crystals: molding the flow of light*. Princeton, NJ: Princeton University Press.

Johnson, S.G., *et al.*, 2002. Adiabatic theorem and continuous coupled-mode theory for efficient taper transitions in photonic crystals. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 66 (2), 066608.1–066608.15.

Johnson, S.G. and Joannopoulos, J. (2001a). *Photonic crystals: the road from theory to practice*. New York: Kluwer Academic.

Johnson, S.G. and Joannopoulos, J.D. (2001b). Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis. *Optics Express*, 8 (3), 173–190.

Jonasson, K. and Madsen, K., 1994. Corrected sequential linear programming for sparse minimax optimization. *BIT Numerical Mathematics*, 34 (3), 372–387.

Khoo, E., Liu, A., and Wu, J., 2005. Nonuniform photonic crystal taper for high-efficiency mode coupling. *Optics Express*, 13 (20), 7748–7759.

Kolda, T., Lewis, R., and Torczon, V., 2003. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Review*, 45 (3), 385–482.

Kuang, W., *et al.*, 2002. Grating-assisted coupling of optical fibers and photonic crystal waveguides. *Optics Letters*, 27 (18), 1604–1606.

Lanckriet, G., *et al.*, 2003. A robust minimax approach to classiflcation. *Journal of Machine Learning Research*, 3 (3), 555–582.
Lorenz, R. and Boyd, S., 2005. Robust minimum variance beamforming. *IEEE Transactions on Signal Processing*, 53 (5), 1684–1696.
Luyssaert, B., *et al.*, 2005. Efficient nonadiabatic planar waveguide tapers. *Journal of Lightwave Technology*, 23 (8), 2462–2468.
Madsen, K. and Schjaer-Jacobsen, H., 1978. Linearly constrained minimax optimization. *Mathematical Programming*, 14, 208–223.
Marcuse, D., 1991. *Theory of dielectric optical waveguides*. 2nd edn. San Diego, CA: Academic Press.
Mekis, A. and Joannopoulos, J., 2001. Tapered couplers for efficient interfacing between dielectric and photonic crystal waveguides. *Journal of Lightwave Technology*, 19 (6), 861–865.
Milton, A.F. and Burns, W.K., 1977. Mode coupling in optical waveguide horns. *IEEE Journal of Quantum Electronics*, 13 (10), 828–835.
Nelder, J.A. and Mead, R., 1965. A simplex method for function minimization. *The Computer Journal*, 7 (4), 308–313.
Nocedal, J. and Wright, S., 1999. *Numerical optimization*. New York: Springer.
Palamaru, M. and Lalanne, P., 2001. Photonic crystal waveguides: out-of-plane losses and adiabatic modal conversion. *Applied Physics Letters*, 78 (11), 1466–1468.
Phadke, M., 1995. *Quality engineering using robust design*. Upper Saddle River, NJ: Prentice Hall.
Pottier, P., Ntakis, I., and La, R.D., 2003. Photonic crystal continuous taper for low-loss direct coupling into 2D photonic crystal channel waveguides and further device functionality. *Optics Communications*, 223 (4), 339–347.
Povinelli, M.L., Johnson, S.G., and Joannopoulos, J.D., 2005. Slow-light, band-edge waveguides for tunable time delays. *Optics Express*, 13, 7145–7159.
Prather, D., *et al.*, 2002. High-efficiency coupling structure for a single-line-defect photonic-crystal waveguide. *Optics Letters*, 27 (18), 1601–1603.
Prekopa, A., 1995. *Stochastic programming*. London, UK: Kluwer Academic Publishers.
Rockafellar, R.T. and Wets, R.J.-B., 1991. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16 (1), 119–147.
Rustem, B. and Howe, M., 2002. *Algorithms for worst-case design and applications to risk management*. Princeton, NJ: Princeton University Press.
Ruszczyński, A., 2006. *Nonlinear optimization*. Princeton, NJ: Princeton University Press.
Sanchis, P., *et al.*, 2004. Analysis of butt coupling in photonic crystals. *IEEE Journal of Quantum Electronics*, 40 (5), 541–550.
Snyder, A.W., 1970. Coupling of modes on a tapered dielectric cylinder. *IEEE Transactions on Microwave Theory and Technology*, 18 (7), 383–392.
Soljačić, M., *et al.*, 2002. Photonic-crystal slow-light enhancement of non-linear phase sensitivity. *Journal of the Optical Society of America – Part B, Optical Physics*, 19, 2052–2059.
Soyster, A., 1973. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21, 1154–1157.
Spühler, M.M., *et al.*, 1998. A very short planar silica spot-size converter using a nonperiodic segmented waveguide. *Journal of Lightwave Technology*, 16 (9), 1680–1685.
Sturm, J., 1999. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11, 625–653. Software available from: sedumi.mcmaster.ca.
Taguchi, G., Jugulum, R., and Taguchi, S., 2004. *Computer-based robust engineering: essentials for DFSS*. Milwaukee, WI: ASQ Quality Press.
Talneau, A., *et al.*, 2002. Low-reflection photonic-crystal taper for efficient coupling between guide sections of arbitrary widths. *Optics Letters*, 27 (17), 1522–1524.
Tsui, K.-L., 1992. An overview of Taguchi method and newly developed statistical methods for robust design. *IIE Transactions*, 24 (5), 44–57.
Vorobyov, S., Gershman, A., and Luo, Z.-Q., 2003. Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem. *IEEE Transactions on Signal Processing*, 51 (2), 313–324.
Wei, C., *et al.*, 1997. Integrated optical elliptic couplers: modeling, design, and applications. *Journal of Lightwave Technology*, 15 (5), 906–912.
Wright, M., 1995. Direct search methods: once scorned, now respectable. *Proceedings of the 1995 Dundee Biennial Conference in Numerical Analysis*, pp. 191–208.
Xu, Y., Lee, R.K., and Yariv, A. (2000a). Adiabatic coupling between conventional dielectric waveguides and waveguides with discrete translational symmetry. *Optics Letters*, 25 (10), 755–757.
Xu, Y., Lee, R.K., and Yariv, A., (2000b). Propagation and second-harmonic generation of electromagnetic waves in a coupled-resonator optical waveguide. *Journal of the Optical Society of America – Part B, Optical Physics*, 17 (3), 387–400.
Yang, K. and El-Haik, B., 2003. *Design for six sigma*. New York: McGraw-Hill Professional.
Yariv, A., 1989. *Quantum Electronics*. 3rd edn. New York: Wiley.
Zhang, B., 2006. Horn waveguide couplers for interfacing between two-dimensional photonic crystal and single mode planar dielectric waveguides. *Acta Physica Sinica*, 55 (4), 1857–1861.
Zhou, K., Doyle, J., and Glover, K., 1996. *Robust and optimal control*. Upper Saddle River, NJ: Prentice-Hall.