

# Numerical Optimization of a Waveguide Transition Using Finite Element Beam Propagation

W. Dörfler

S. Findeisen

Preprint 13/01

INSTITUT FÜR WISSENSCHAFTLICHES RECHNEN  
UND MATHEMATISCHE MODELLBILDUNG



**Anschriften der Verfasser:**

Prof. Dr. Willy Dörfler  
Institut für Angewandte und Numerische Mathematik  
Karlsruher Institut für Technologie (KIT)  
D-76128 Karlsruhe

Dipl.-Math. techn. Stefan Findeisen  
Institut für Angewandte und Numerische Mathematik  
Karlsruher Institut für Technologie (KIT)  
D-76128 Karlsruhe

# NUMERICAL OPTIMIZATION OF A WAVEGUIDE TRANSITION USING FINITE ELEMENT BEAM PROPAGATION

W. DÖRFLER AND S. FINDEISEN

*Karlsruhe Institute of Technology, Department of Mathematics, D-76128 Karlsruhe, Germany*

ABSTRACT. We consider a waveguide with principal guiding direction for which the beam propagation method is applicable. A simulation method based on a finite element method for Maxwell's equation with 3D Nedelec elements is developed. The power loss of the waveguide is minimised by varying a finite number of shape parameters. We validate the method by comparing our findings to some published results.

## 1. INTRODUCTION

The beam propagation method (BPM) is a tool for the simulation of waveguides with varying cross section in the principal guiding direction. Typical examples are Y-junctions, S-shaped and tapered waveguides. The main interest is to compute the power loss of a given wave which occurs in a waveguide transition. At first beam propagation was used with fast Fourier transform (FFT-BPM) [Lif03, Ch. 5.2] and finite differences (FD-BPM) [Lif03, Ch. 5.3] to solve the (scalar) Helmholtz equation. Nowadays BPM is used with finite elements (FE-BPM) to solve the three-dimensional Maxwell problem [SGBV98]. The great success of BPM lies in its easy implementation and speed and it has been shown that under appropriate assumptions the beam propagation method leads to reasonable good results. Difficulties arise in fast varying structures according to the principal guiding direction and since BPM propagates only in one direction, neglecting back reflections [RDL81]. To weaken these assumptions is still a topic of research (e.g. [Had92], [CL04], [SA04], [LB09]). In this article we developed a FE-BPM method for the three-dimensional Maxwell problem and used an optimization method in order to minimize the losses of an electromagnetic wave passing through a waveguide transition. The paper is organized as follows: First we derive a model using the time-harmonic Maxwell equation and the fundamental assumption (1.4) in Section 2.1, together with initial and boundary conditions (Section 2.2). In Section 3 we describe an optimization algorithm based on the Gauß-Newton method. In order to validate our method, results are presented showing a linear taper and a sinusoidal shaped waveguide transition in Sections 4.1 and 4.2. Finally we used the Gauß-Newton algorithm to optimize power transmission in a S-shaped waveguide in Section 4.3.

To fix ideas let the cross section of the waveguide be  $\mathbb{R}^2$  with variables  $x, y$ , and the guiding direction be the  $z$ -coordinate. We consider the waveguide between  $z = 0$  and  $z = L > 0$ . At a fixed coordinate  $z$  it consists of a bounded simply connected core region and a cladding as the remainder of the cross-sectional plane, see Figure 1 (with radially symmetric, in  $x, y$ , material regions). The material properties are described by the (relative) permittivity function  $(x, y, z) \mapsto \varepsilon_r(x, y, z)$ . For  $z$  outside  $[0, L]$  we assume

---

Received by the editors January 29, 2013.

*Key words and phrases:* numerical optimization, waveguide transition

*E-mail addresses:* [willy.doerfler@kit.edu](mailto:willy.doerfler@kit.edu) (W. Dörfler), [stefan.findeisen@kit.edu](mailto:stefan.findeisen@kit.edu) (S. Findeisen)

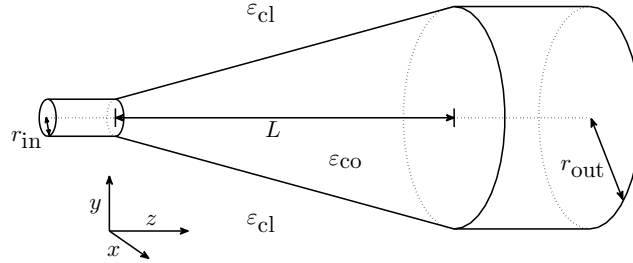


FIGURE 1. Structure of a tapered waveguide.

that  $\varepsilon_r$  is independent of  $z$  and has a constant value  $\varepsilon_{co}$  in the core and  $\varepsilon_{cl}$  in the cladding, respectively. We consider *weakly guided modes* of a frequency  $\omega$ , where

$$k_{co} = \frac{\omega}{c} \sqrt{\varepsilon_{co}} \gtrsim k_{cl} = \frac{\omega}{c} \sqrt{\varepsilon_{cl}}. \quad (1.1)$$

In this case the electric field can be set to

$$\mathbf{E}(x, y, z) = \mathbf{u}(x, y, z) e^{i\beta z}, \quad (1.2)$$

with (given) propagation constant  $\beta \in \mathbb{R}$ . Here  $e^{i\beta z}$  can be understood as a fast oscillating wave moving in  $z$ -direction, whereas  $\mathbf{u}(x, y, z)$  is a field amplitude of slow variation in  $z$ -direction. For guided modes  $\beta$  has to be real and fulfills

$$k_{co} > \beta \geq k_{cl}. \quad (1.3)$$

Furthermore it is assumed that the variation in direction of  $z$  is small in the sense of

$$\partial_z^2 \mathbf{u}(x, y, z) \approx \mathbf{0}, \quad (1.4)$$

see [Lif03, Ch. 5.1]. This allows us the possibility to reduce the problem so that it can be solved recursively as it is shown in the next section. Notice that the beam propagation approach is only valid if the assumptions (1.2) and (1.4) are fulfilled. Hence BPM is mainly used for long transitions with small bends.

## 2. DERIVATION OF THE MODEL

**2.1. The beam propagation equation.** We consider the time-harmonic Maxwell equations for the electric field  $\mathbf{E}$

$$\nabla \times \nabla \times \mathbf{E} = k^2 \mathbf{E} \quad (2.1)$$

( $\nabla = [\partial_x, \partial_y, \partial_z]^\dagger$ ) for the wave number

$$k = \frac{\omega}{c} \sqrt{\varepsilon_r}$$

together with the divergence condition  $\nabla \cdot (k^2 \mathbf{E}) = 0$ . This equation will be considered on a bounded domain  $\Omega := Q_R \times [0, L] \subset \mathbb{R}^3$ ,  $Q_R := [-R, R]^2 \subset \mathbb{R}^2$  with  $R > 0$  sufficiently large. The issue of boundary conditions is postponed to the next section.

Inserting (1.2) for  $\mathbf{E}$  and using (1.4) and  $\mathbf{e}_z = [0, 0, 1]^\dagger$ ,  $\nabla_\perp = [\partial_x, \partial_y, 0]^\dagger$ , we obtain for the function  $(x, y, z) \mapsto \mathbf{u}(x, y, z) \in \mathbf{H}(\text{curl}; \Omega) := \{\mathbf{v} : \Omega \rightarrow \mathbb{R}^3 : \mathbf{v}, \nabla \times \mathbf{v} \text{ square integrable over } \Omega\}$  the equations

$$\begin{aligned} & -\beta^2 \mathbf{e}_z \times \mathbf{e}_z \times \mathbf{u} + i\beta (\nabla_\perp \times \mathbf{e}_z \times \mathbf{u} + \mathbf{e}_z \times \nabla_\perp \times \mathbf{u}) \\ & \quad + \nabla_\perp \times \nabla_\perp \times \mathbf{u} + 2i\beta \mathbf{e}_z \times \mathbf{e}_z \times \partial_z \mathbf{u} + \nabla_\perp \times \mathbf{e}_z \times \partial_z \mathbf{u} + \mathbf{e}_z \times \nabla_\perp \times \partial_z \mathbf{u} \\ & = k^2 \mathbf{u} \end{aligned} \quad (2.2)$$

(note  $a \times b \times c \equiv a \times (b \times c)$ ). Now we multiply (2.2) with  $\bar{\mathbf{v}}$  for arbitrary  $\mathbf{v} \in \mathbf{V}(\Omega) := \{\mathbf{v} \in \mathbf{H}(\text{curl}; \Omega) : \mathbf{n} \times \mathbf{v} = \mathbf{0} \text{ on } \partial Q_R \times [0, L]\}$ , integrate by parts over  $\Omega$  and thereby get the *weak equation*

$$\begin{aligned} & \beta^2 \int_\Omega (\mathbf{e}_z \times \mathbf{u}) \cdot \overline{\mathbf{e}_z \times \mathbf{v}} + i\beta \int_\Omega (\mathbf{e}_z \times \mathbf{u} \cdot \overline{\nabla_\perp \times \mathbf{v}} - \nabla_\perp \times \mathbf{u} \cdot \overline{\mathbf{e}_z \times \mathbf{v}}) \\ & \quad + \int_\Omega (\nabla_\perp \times \mathbf{u} \cdot \overline{\nabla_\perp \times \mathbf{v}} - k^2 \mathbf{u} \cdot \bar{\mathbf{v}}) \\ & = 2i\beta \int_\Omega \mathbf{e}_z \times \partial_z \mathbf{u} \cdot \overline{\mathbf{e}_z \times \mathbf{v}} - \int_\Omega (\mathbf{e}_z \times \partial_z \mathbf{u} \cdot \overline{\nabla_\perp \times \mathbf{v}} - \nabla_\perp \times \partial_z \mathbf{u} \cdot \overline{\mathbf{e}_z \times \mathbf{v}}). \end{aligned}$$

Introducing the sesqui-linear forms

$$\begin{aligned} k_z(\mathbf{u}, \mathbf{v}) &= \int_\Omega (\nabla_\perp \times \mathbf{u} \cdot \overline{\nabla_\perp \times \mathbf{v}} - k^2 \mathbf{u} \cdot \bar{\mathbf{v}}), \\ g(\mathbf{u}, \mathbf{v}) &= \int_\Omega (\mathbf{e}_z \times \mathbf{u} \cdot \overline{\nabla_\perp \times \mathbf{v}} - \nabla_\perp \times \mathbf{u} \cdot \overline{\mathbf{e}_z \times \mathbf{v}}), \\ m(\mathbf{u}, \mathbf{v}) &= \int_\Omega (\mathbf{e}_z \times \mathbf{u} \cdot \overline{\mathbf{e}_z \times \mathbf{v}}), \end{aligned}$$

equation (2.2) is expressed in variational form by

$$\beta^2 m(\mathbf{u}, \mathbf{v}) + i\beta g(\mathbf{u}, \mathbf{v}) + k_z(\mathbf{u}, \mathbf{v}) = 2i\beta m(\partial_z \mathbf{u}, \mathbf{v}) - g(\partial_z \mathbf{u}, \mathbf{v}) \quad (2.3)$$

for all  $\mathbf{v} \in \mathbf{V}(\Omega)$ .

For discretisation we use a curl-conforming finite element method [Mon03, Ch. 6.3] which ensures that the divergence condition is satisfied in a weak sense: We decompose our three dimensional finite computational domain  $\Omega$  into rectangular hexahedra. Our approximation space  $\mathbf{V}_h$  is a subset of the piecewise trilinear functions that are globally in  $\mathbf{H}(\text{curl}; \Omega)$  (*Nedélec edge elements*). It contains the gradients of the piecewise trilinear functions, i. e. the functions  $\mathbf{v}_h \in \mathbf{V}_h$  that can be expressed as  $\mathbf{v}_h = \nabla q_h$  with  $q_h \in Q_h \subset H^1(\Omega) = \{q : \Omega \rightarrow \mathbb{R} : q, \nabla q \text{ square integrable over } \Omega\}$  [Mon03, Thm. 6.12]. By using this in the weak formulation of (2.1) one receives that

$$\int_\Omega k^2 \mathbf{E} \cdot \overline{\nabla q_h} = 0$$

for all  $q_h \in Q_h$ , which is the weak formulation of the divergence condition.

The basis of this approximation space  $\mathbf{V}_h$  is denoted by  $\{\varphi_\nu : \nu = 1, \dots, N\}$ . The nodes of the basis are located in the centers of the edges. Our aim is to find a vector  $U = [\alpha_\nu]_{\nu=1, N} \in \mathbb{C}^N$  such that  $\mathbf{u} = \sum_{\nu=1}^N \alpha_\nu \varphi_\nu$  solves (2.3) for all  $\mathbf{v} \in \mathbf{V}_h$ .

We choose  $N_z \in \mathbb{N}$  and define the  $N_z + 2$  points  $z_i := i\Delta z$ , for  $i = -1, \dots, N_z$ , with interval length  $\Delta z := L/N_z$  (the first interval will be needed for initial conditions). Then we decompose the set  $\Omega$  into  $N_z + 1$  slices  $\Omega_i := [-R, R]^2 \times [z_{i-1}, z_i]$  for  $i = 0, \dots, N_z$ . The nodes of the chosen edge element

discretisation are either located at  $z = z_i$  or  $z = z_{i-1/2} := (z_i + z_{i-1})/2$ , for  $i = 0, \dots, N_z$ , as mentioned above.

We number the basis elements of this finite element space such that degrees of freedom with nodes located at the same coordinate  $z_i$  are grouped together, i. e.  $\varphi_{i,j}$  is the  $j$ -th basis function with node in  $Q_R \times \{z_i\}$  for  $j = 1, \dots, M$ , while degrees of freedom located at nodes in  $Q_R \times \{z_{i-1/2}\}$  are denoted as  $\varphi_{i-\frac{1}{2},j'}$  for  $j' = 1, \dots, M'$ . Hence our approximation space is spanned by the set  $\{\varphi_{i,j}, \varphi_{i-\frac{1}{2},j'} : i = 0, \dots, N_z, j = 1, \dots, M, j' = 1, \dots, M'\}$  and we receive corresponding coefficient vectors  $U_i = [\alpha_{i,j}]_{j=1,\dots,M} \in \mathbb{C}^M$  and  $U_{i-\frac{1}{2}} = [\alpha_{i-\frac{1}{2},j'}]_{j'=1,\dots,M'} \in \mathbb{C}^{M'}$ , respectively. Finally, we approximate the first order derivatives in  $z$ -direction by the forward and backward difference quotients  $\partial_z \varphi_{i-1,j}, \partial_z \varphi_{i,j} \approx (\varphi_{i,j} - \varphi_{i-1,j})/\Delta z$  and  $\partial_z \varphi_{i-\frac{3}{2},j}, \partial_z \varphi_{i-\frac{1}{2},j} \approx (\varphi_{i-\frac{1}{2},j} - \varphi_{i-\frac{3}{2},j})/\Delta z$ , respectively, and neglect the second order derivatives according to (1.4). By inserting this approach into our variational form (2.3) and testing with  $\varphi_{i,j}, \varphi_{i-1/2,j'}$  for  $i, j, j'$  as above, we end up with a system of equations

$$\begin{bmatrix} D_0 & E_1 & 0 & \dots & \dots & \dots \\ 0 & D_1 & E_2 & \ddots & \dots & \dots \\ \dots & \ddots & \ddots & \ddots & \ddots & \dots \\ \dots & \dots & \ddots & D_{N_z-2} & E_{N_z-1} & 0 \\ \dots & \dots & \dots & 0 & D_{N_z-1} & E_{N_z} \end{bmatrix} \begin{bmatrix} \tilde{U}_0 \\ \tilde{U}_1 \\ \vdots \\ \tilde{U}_{N_z-1} \\ \tilde{U}_{N_z} \end{bmatrix} = 0, \quad (2.4)$$

with  $D_i := \Delta z A_i + 2B$ ,  $E_i := \Delta z A_i - 2B$ , where the matrices  $A_i, B \in \mathbb{C}^{M+M', M+M'}$  and vectors  $\tilde{U}_i, \tilde{U}_{i+1} \in \mathbb{C}^{M+M'}$  are defined by

$$A_i := \begin{bmatrix} [a_{i;l,j}]_{jl} & [a_{i;l',j}]_{j'l'} \\ [a_{i-1/2;l,j'}]_{j'l'} & [a_{i-1/2;l',j'}]_{j'l'} \end{bmatrix}, \quad B := \begin{bmatrix} [b_{l,j}]_{jl} & [b_{l',j}]_{j'l'} \\ [b_{l,j'}]_{j'l'} & [b_{l',j'}]_{j'l'} \end{bmatrix},$$

$$\tilde{U}_i := \begin{bmatrix} U_i \\ U_{i-\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} [\alpha_{i,j}]_j \\ [\alpha_{i-\frac{1}{2},j'}]_{j'} \end{bmatrix},$$

and the matrix entries are given by

$$a_{i;j,l} := \beta^2 m(\varphi_{i,l}, \varphi_{i,j}) + i\beta g(\varphi_{i,l}, \varphi_{i,j}) + k_i(\varphi_{i,l}, \varphi_{i,j}), \quad b_{j,l} := 2i\beta m(\varphi_{i,l}, \varphi_{i,j}) - g(\varphi_{i,l}, \varphi_{i,j}).$$

The notation  $a_{i;..}$  expresses the fact that the bilinear form  $a$  depends on  $i$  via  $k$ . The bilinear form  $b$ , however, does not depend on  $i$ .

Since the unknown sub-vector  $\tilde{U}_{i+1}$  only depends on  $\tilde{U}_i$ , (2.4) can be solved recursively for given  $\tilde{U}_0$  (provided in Section 2.2). This recursion yields the *beam propagation equation*

$$\left(B - \frac{1}{2}\Delta z A_i\right)\tilde{U}_i = \left(B + \frac{1}{2}\Delta z A_{i-1}\right)\tilde{U}_{i-1} \quad \text{for all } i = 1, \dots, N_z. \quad (2.5)$$

Thus we have to solve a system of equations in each of the  $N_z$  steps. For the chosen finite element method, the resulting  $(M + M') \times (M + M')$ -systems are sparse, that means the number of non-zero entries is only proportional to  $M + M'$ .

For computation we used the deal.II-library (Differential Equations Analysis Library [BHK07]). The resulting systems were solved by using the Generalized Minimal Residual Method (GMRES) and incomplete LU-decomposition as a preconditioner (ILU). The method was chosen from the PETSc-library (Portable, Extensible Toolkit for Scientific Computation [BBB<sup>+</sup>12]). For the termination of the algorithm we have chosen a tolerance of  $10^{-10}$  for the residual in the euclidian norm. The number of iterations was usually below 20, so that we did not use a restart.

We observed an experimental second order convergence in the  $L^2(Q_R)^3$ -norm, which is also used to compute the transmitted power, see Section 2.3. Since there is no exact solution available, we used an extrapolation technique to estimate the error.

**2.2. Initial and boundary conditions.** The input waveguide is the part of the structure for  $z \leq 0$ , while the output waveguide is the part for  $z \geq L$ . In both parts  $\varepsilon_r$ , and thus  $k$ , is assumed to be independent on  $z$ . The input waveguide is excited by one of it's eigenmodes with corresponding propagating constant  $\beta_{\text{in}}$ . It can be obtained by inserting the  $z$ -invariant ansatz

$$\mathbf{E}(x, y, z) = \boldsymbol{\psi}_{\text{in}}(x, y)e^{i\beta_{\text{in}}z}$$

into the time-harmonic Maxwell equation (2.1). This leads to the equation

$$\nabla_{\beta_{\text{in}}} \times \nabla_{\beta_{\text{in}}} \times \boldsymbol{\psi}_{\text{in}}(x, y) = k^2(x, y)\boldsymbol{\psi}_{\text{in}}(x, y) \quad (2.6)$$

for  $\boldsymbol{\psi}_{\text{in}} : \mathbb{R}^2 \rightarrow \mathbb{C}^3$ , where  $\nabla_{\beta_{\text{in}}} = [\partial_x, \partial_y, i\beta_{\text{in}}]^\dagger$  is a modified gradient operator. The aim is to find values for  $\beta_{\text{in}} \in \mathbb{R}$  that admit existence of nontrivial solutions  $\boldsymbol{\psi}_{\text{in}}$  of (2.6), that is we have to solve an eigenvalue problem where  $\beta_{\text{in}}$  appears quadratically.

Using the same technique as in Section 2.1, equation (2.6) can be rewritten in variational form as

$$\beta_{\text{in}}^2 m(\boldsymbol{\psi}_{\text{in}}, \mathbf{v}) + i\beta_{\text{in}} g(\boldsymbol{\psi}_{\text{in}}, \mathbf{v}) + k(\boldsymbol{\psi}_{\text{in}}, \mathbf{v}) = 0 \quad \text{for all } \mathbf{v} \in \mathbf{V}(Q_R).$$

This equation is discretised as follows: We pose the problem on  $Q_R \times [z_{-1}, z_0]$  with the technique of the previous section (on the same mesh in the  $x, y$ -plane) and with periodic boundary conditions in  $z$ -direction. This leads to a system of equations

$$\beta_{\text{in}}^2 MU^{\text{in}} + i\beta_{\text{in}} GU^{\text{in}} + KU^{\text{in}} = 0, \quad (2.7)$$

for the vector  $U^{\text{in}}$  containing the expansion coefficients in the finite element basis and where the matrices  $M, G, K \in \mathbb{C}^{N, N}$  are defined by

$$M := [m(\boldsymbol{\varphi}_{\nu'}, \boldsymbol{\varphi}_{\nu})]_{\nu, \nu'}, \quad G := [g(\boldsymbol{\varphi}_{\nu'}, \boldsymbol{\varphi}_{\nu})]_{\nu, \nu'}, \quad K := [k(\boldsymbol{\varphi}_{\nu'}, \boldsymbol{\varphi}_{\nu})]_{\nu, \nu'}.$$

To improve the accuracy and to avoid unphysical reflections from the boundary  $\partial Q_R$  we used a perfectly matched layer (PML) [Ged96] near  $\partial Q_R$ . This technique is also used in setting up the matrices  $A_i$  and  $B$  in the previous section.

The quadratic eigenvalue problem (2.7) was solved by a linearisation technique as described in [BDD<sup>+</sup>00, Ch. 9]. To solve the linear eigenvalue problem we used an iterative Krylov–Schur solver, chosen from the SLEPc-library [CRR12]. In this way a discrete set of modes  $\boldsymbol{\psi}_{\text{in}}$  and corresponding propagating constants  $\beta_{\text{in}}$  of the input waveguide are computed, where a numbering is introduced that is based on the ordering  $\beta_{\text{in}}^{(1)} \geq \beta_{\text{in}}^{(2)} \geq \dots$ . Since modes are unique up to scaling and multiplicity they have to be organized to form an orthonormal set in  $L^2(Q_R)^3$ . The same computations are done to obtain the modes  $\boldsymbol{\psi}_{\text{out}}$  and corresponding propagating constants  $\beta_{\text{out}}$  of the output waveguide (that is, we do the analogous computation on  $Q_R \times [z_{Nz}, z_{Nz} + \Delta z]$ ). Note that (1.3) has to be fulfilled for guided modes. Since our waveguides are weakly guiding,  $\beta^{(n)} := (\beta_{\text{in}}^{(n)} + \beta_{\text{out}}^{(n)})/2$  is taken as an approximation. In our case we set  $U^{\text{in}} := \boldsymbol{\psi}_{\text{in}}^{(1)}$  and  $\beta := \beta^{(1)} = (\beta_{\text{in}}^{(1)} + \beta_{\text{out}}^{(1)})/2$ , since we are interested in the behaviour of the first guided mode.

Applying the same numbering as before, we can write  $\tilde{U}_0 = [U_0^{\text{in}}, U_{-1/2}^{\text{in}}]$  and this provides us with the initial conditions for the iteration (2.5).

**2.3. Power losses.** The modes of the input and the output waveguide computed in the previous section can be used to approximate the power loss during a waveguide transition, see [BL82]. Let  $\psi_{\text{in}}^{(n)}$  and  $\psi_{\text{out}}^{(n)}$  be the  $n$ -th mode of the input and output waveguide, respectively. Then the output field  $\mathbf{u}_{\text{out}} = \mathbf{u}(\cdot, \cdot, L)$ , which is the computed solution  $\mathbf{u}$  at  $z = L$ , can be written as a linear combination of the output modes

$$\mathbf{u}_{\text{out}} = \sum_{\ell} a_{\ell} \psi_{\text{out}}^{(\ell)} \quad \text{with} \quad a_{\ell} := \int_{Q_R} \mathbf{u}_{\text{out}} \cdot \overline{\psi_{\text{out}}^{(\ell)}} \in \mathbb{C}. \quad (2.8)$$

Hence the power transmitted from the first into the  $n$ -th mode is given by

$$P_{1,n} := \left| \int_{Q_R} \psi_{\text{in}}^{(1)} \cdot \overline{a_n \psi_{\text{out}}^{(n)}} \right|^2 = |a_n|^2 \left| \int_{Q_R} \psi_{\text{in}}^{(1)} \cdot \overline{\psi_{\text{out}}^{(n)}} \right|^2. \quad (2.9)$$

In the following we specify the transmitted Power by  $100P_{1,n}$  in %.

### 3. OPTIMIZATION APPROACH

**3.1. Optimization problem.** In this section we describe the optimization of a transition between two waveguides with radial symmetry. Let the radius of the core be  $r_{\text{in}}$  at the input ( $z = 0$ ) and  $r_{\text{out}}$  at the output ( $z = L$ ). The radius of the waveguide is given by a function  $r : [0, L] \rightarrow \mathbb{R}_{>0}$  with  $r(0) = r_{\text{in}}$  and  $r(L) = r_{\text{out}}$ .

**Optimization model.** *The optimization model is given by*

- (i) a nodal grid  $\mathbf{z} = \{\xi_0 = 0, \xi_1, \xi_2, \dots, \xi_M, \xi_{M+1} = L\}$ ,
- (ii) model values  $\mathbf{m} \in \mathcal{M} \subset \mathbb{R}^M$ , here  $m_l = r(\xi_l)$  for  $l = 1, \dots, M$ ,  
with extensions  $m_0 := r(0)$ ,  $m_{M+1} := r(L)$ ,
- (iii) an interpolated profile  $z \mapsto \Phi(z)$  with  $\Phi(\xi_l) = m_l$  for  $l = 0, \dots, M + 1$ .

$m(\xi_l)$  gives the location of the material boundary at a point  $\xi_l$ . Notice that optimization algorithms without regularisation become unstable if the wavelength  $\lambda$  is larger than the discretisation of the nodal grid [EF01]. In the set  $\mathcal{M}$  we will also formulate restrictions on reasonable connections. For example we will include boundary conditions  $m_0$ ,  $m_{M+1}$  and a condition of the type  $0 < m_{\min} \leq m_l \leq m_{\max} < \text{diam}(\Omega)/2$ .  $\Phi$  is used to get interpolated values  $\Phi(z)$  for  $z \in [0, L]$  in the boundary value problems of Section 2.

**Minimisation problem.** *For given  $\mathbf{z}$ ,  $\mathcal{M}$ ,  $\Phi$  and  $n$  find suitable model values  $\mathbf{m}_*$ , such that the function*

$$\mathbf{m} \mapsto \mathcal{J}(\mathbf{m}) := \frac{1}{2} |1 - P_{1,n}(\mathbf{m})|^2$$

*is minimised, i. e.,*

$$\mathcal{J}(\mathbf{m}_*) = \min_{\mathbf{m} \in \mathcal{M}} \mathcal{J}(\mathbf{m}). \quad (3.1)$$

**3.2. Gauß–Newton method.** Starting from  $\mathbf{m}^{(0)} \in \mathcal{M}$ , we computed a sequence of approximations  $\mathbf{m}^{(k)}$  for  $k > 0$  by the Gauß–Newton method [DS96, Kap. 10.2] [NW99, Kap. 10.3]. It is derived by replacing, in step  $k$ , the functional  $\mathcal{J}$  in the minimization problem (3.1) by the linear part of it's Taylor series at  $\mathbf{m}^{(k)}$ , which results in the following definition of the next iterate  $\mathbf{m}^{(k+1)}$

$$\mathcal{J}(\mathbf{m}^{(k+1)}) = \min_{\mathbf{m} \in \mathcal{M}} \frac{1}{2} |1 - P_{1,n}(\mathbf{m}^{(k)}) - J_{k,n}(\mathbf{m} - \mathbf{m}^{(k)})|^2, \quad (3.2)$$



with  $J_{k,n} := \nabla P_{1,n}(\mathbf{m}^{(k)}) \in \mathbb{R}^{1,M}$ . Assuming that the first derivative vanishes in the minimum  $\mathbf{m}^{(k+1)}$ , we set the gradient of the function on the right in (3.2) to zero and obtain the relation

$$J_{k,n}(\mathbf{m}^{(k+1)} - \mathbf{m}^{(k)}) = 1 - P_{1,n}(\mathbf{m}^{(k)}) \quad (3.3)$$

for  $k \geq 0$ . If we let  $\mathbf{m}^{(k+1)} - \mathbf{m}^{(k)} = \mathbf{s}^{(k)} \sim J_{k,n}^T \in \mathbb{R}^{M,1}$ , we easily find

$$\mathbf{s}^{(k)} = \frac{1 - P_{1,n}(\mathbf{m}^{(k)})}{|J_{k,n}|^2} J_{k,n}^T.$$

This minimization step is refined by introducing a suitable damping parameter  $\alpha^{(k)}$  and setting

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha^{(k)} \mathbf{s}^{(k)}.$$

This parameter  $\alpha^{(k)}$  results from a step size control which is chosen to ensure that the so called *Armijo–Goldstein condition* [DS96, Kap. 6.3] [NW99, Kap. 3.1]

$$\mathcal{J}(\mathbf{m}^{(k+1)}) \leq \mathcal{J}(\mathbf{m}^{(k)}) - \mu \alpha^{(k)} |\nabla \mathcal{J}(\mathbf{m}^{(k)}) (\mathbf{m}^{(k+1)} - \mathbf{m}^{(k)})| \quad (3.4)$$

holds for  $\mathbf{m}^{(k+1)}$  for some given  $\mu \in (0, 1)$  (e. g.  $\mu = 0.5$ ). To get a suitable value for  $\alpha^{(k)}$  we used the *Armijo step size control*: Starting with  $\alpha^{(k,0)} = 2\alpha^{(k-1)}$  (to allow coarsening) we check  $\mathbf{m}^{(k,l)} = \mathbf{m}^{(k)} + \alpha^{(k,l)} \mathbf{s}^{(k)} \in \mathcal{M}$  and whether (3.4) holds for  $\mathbf{m}^{(k,l)}$  to stop the iteration on  $l$  with  $\mathbf{m}^{(k+1)} = \mathbf{m}^{(k,l)}$  and  $\alpha^{(k+1)} = \alpha^{(k,l)}$ , or continue with  $\alpha^{(k,l+1)} = \alpha^{(k,l)}/2$  and  $l \rightarrow l + 1$ .

To obtain the gradients in this algorithm, actually  $J_{k,n}$ , we used difference quotients  $\partial_l P_{1,n}(\mathbf{m}) \approx (P_{1,n}(\mathbf{m} + h\mathbf{e}^l) - P_{1,n}(\mathbf{m}))/h$ , where  $\mathbf{e}^l$  is the  $l$ -th euclidian unit vector.

## 4. EXAMPLES

**4.1. Linear taper.** The first example is used to validate our method with the results of Hermansson et al. [HYD83]. We considered two radially symmetric waveguides with different diameter, which were connected by a linear taper with transition length  $L$ . For the computation of the input wave, we solved an eigenvalue problem (2.7) for  $\beta_{\text{in}}$ ,  $U^{\text{in}}$ , and the corresponding one for  $\beta_{\text{out}}$ ,  $U^{\text{out}}$ , with  $N = 41730$  degrees of freedom on the computational domain  $\Omega = [-35 \mu\text{m}, 35 \mu\text{m}]^2 \times [-0.625 \mu\text{m}, 0.625 \mu\text{m}]$ . This domain was surrounded by a perfectly matched layer (PML) of thickness  $5 \mu\text{m}$  in  $xy$ -direction. The taper was discretised with  $N_z = 160$  nodes in  $z$ -direction. Moreover, we considered the first guided mode with wavelength  $\lambda = 1 \mu\text{m}$  in a waveguide with a parabolic index of refraction profile  $n := \sqrt{\varepsilon_r}$ ,

$$n(x, y, z) = \begin{cases} n_{\text{co}} \left( 1 - 0.06 \left( \frac{\sqrt{x^2 + y^2}}{r(z)} \right)^2 \right)^{1/2} & \text{for } \sqrt{x^2 + y^2} \leq r(z), \\ n_{\text{cl}} & \text{for } \sqrt{x^2 + y^2} > r(z). \end{cases}$$

The refractive indices are  $n_{\text{cl}} = 1.5$ ,  $n_{\text{co}} = n_{\text{cl}}/\sqrt{0.94} \approx 1.547$  and  $r(z)$  is the radius of the waveguide at  $z$ . Hence the waveguide is weakly guiding since (1.1) is fulfilled. The radius of the input and output waveguide was  $r_{\text{in}} = 5 \mu\text{m}$  and  $r_{\text{out}} = 30 \mu\text{m}$ . In what follows we show results for two different transition lengths  $L = 100 \mu\text{m}$  and  $L = 200 \mu\text{m}$ .

In Figure 2 the power of the first mode propagating in the taper is shown. It is in good agreement with the corresponding Figures 5 and 6 in [HYD83]. The transmitted power in the first mode at the end of the taper can be found in Table 1. The reference values were taken from the figures cited.

transition length $L$ in $\mu\text{m}$	transmitted power	extrapolated reference values
100	88.52%	87%
200	99.24%	99%

TABLE 1. Transmitted power of the first mode in a linear taper in comparison to reference values obtained from [HYD83, Figs. 5, 6].

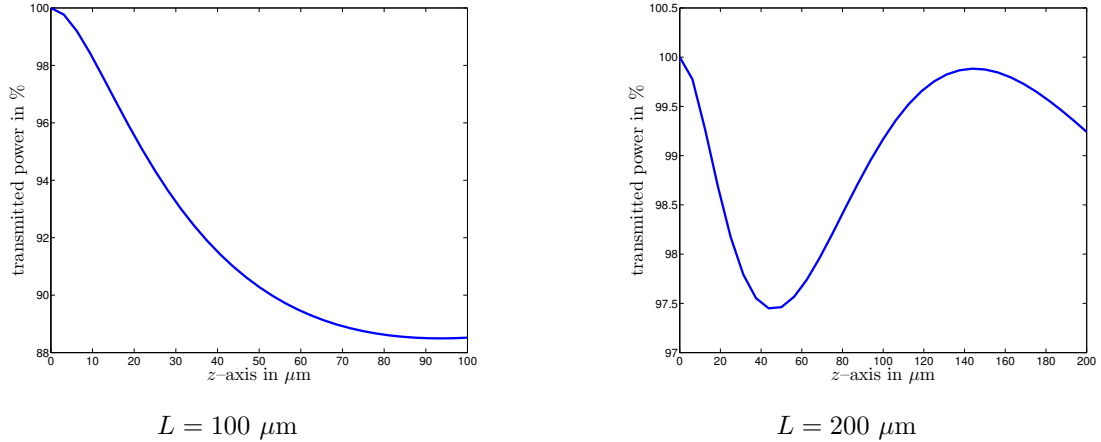


FIGURE 2. Transmitted power in the first guided mode in a linear taper.

**4.2. S-shaped waveguides.** In this example the transition between two identical waveguides is considered. However, there is an offset between the input and output waveguide. For that reason they must be bent and losses occur. As in the previous example we consider the first guided mode in a weakly guiding waveguide with a parabolic index of refraction profile

$$n(x, y, z) = \begin{cases} n_{\text{co}} \left( 1 - d \left( \frac{\sqrt{(x-f(z))^2 + y^2}}{r(z)} \right)^2 \right)^{1/2} & \text{for } \sqrt{(x-f(z))^2 + y^2} \leq r, \\ n_{\text{cl}} & \text{for } \sqrt{(x-f(z))^2 + y^2} > r, \end{cases}$$

with parameters  $n_{\text{co}} = 2.2$  and  $n_{\text{cl}} = 2.1$  and  $d = 1 - (n_{\text{cl}}/n_{\text{co}})^2$ . In contrast to the first example the radius  $r = 1.5 \mu\text{m}$  remains constant. These values are typical for waveguides made from lithium niobate ( $\text{LiNbO}_3$ ). The vacuum wavelength is given by  $\lambda = 0.6328 \mu\text{m}$  and the shape function

$$f(z) = -\frac{b}{2} \sin\left(\frac{2z-L}{2L}\pi\right) \quad \text{for } z \in [0, L], \quad (4.1)$$

describes a transition between the input and output waveguide. Both waveguides are shifted against each other by  $b = 50 \mu\text{m}$  along the  $x$ -axis. The input wave  $\psi_{\text{in}}^{(1)}$  is computed with 41730 degrees of freedom on the computational domain  $\Omega = [-28 \mu\text{m}, 28 \mu\text{m}] \times [-4 \mu\text{m}, 4 \mu\text{m}] \times [-0.25 \mu\text{m}, 0.25 \mu\text{m}]$ , again surrounded by a PML of thickness  $4 \mu\text{m}$  in  $xy$ -direction. We use a step size of  $\Delta z = 1 \mu\text{m}$  in the beam propagation method (2.5). The computed losses for different transition lengths  $L$  are shown in Figure 3 together with some reference values computed by [Mar78, Fig. 6] using an analytic approximation formula. For large  $L$  the loss is low but increases very fast when  $L$  becomes smaller. In Figure 4 the absolute value of the E-field of the first mode is shown for  $L = 300 \mu\text{m}$  and  $L = 600 \mu\text{m}$ . In the first case the higher radiation is clearly visible.

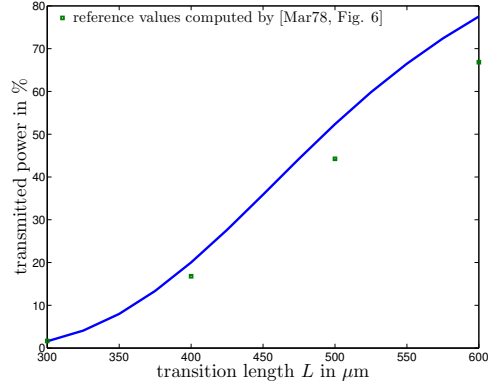


FIGURE 3. Losses in a sinusoidal shaped transition profile and different transition lengths  $L$

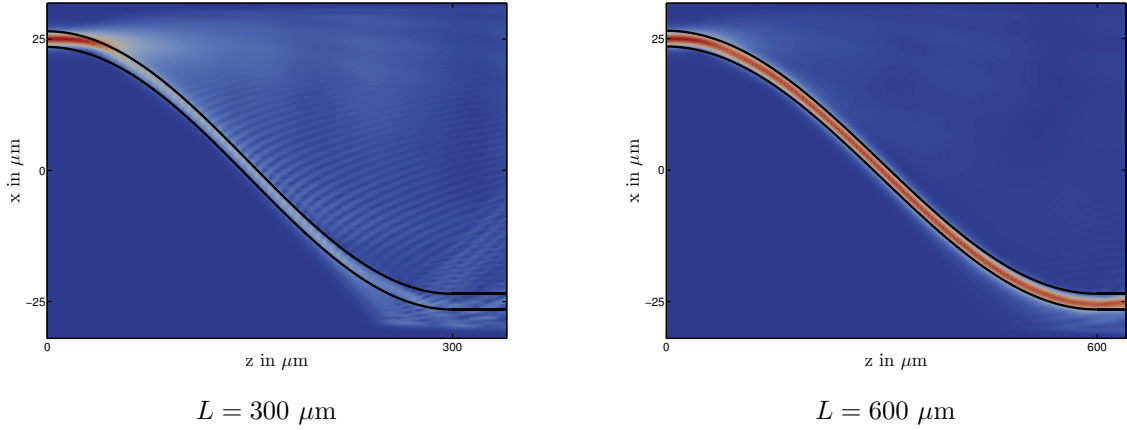


FIGURE 4. Absolute value of the E-field in a sinusoidal transition and transition lengths  $L = 300 \mu\text{m}$  and  $L = 600 \mu\text{m}$ .

**4.3. Optimized S-shaped waveguides.** Using a sinusoidal transition profile is not the only possibility to connect two waveguides with an offset. In this example two other common profiles are considered and compared with an optimized profile using our optimization method, Section 3, and the previous sinusoidal profile. The transition length is given by  $L = 600 \mu\text{m}$ . The first new profile is a modified sinusoidal profile

$$f(z) = -\frac{b}{2\pi} \sin\left(\frac{2z-L}{L}\pi\right) - \frac{z}{L}b + \frac{1}{2}b \quad \text{for all } z \in [0, L] \quad (4.2)$$

and the second one consists of two connected arcs

$$f(z) = \begin{cases} \sqrt{R^2 - z^2} - R + \frac{1}{2}b & \text{for } z \in [0, \frac{1}{2}L], \\ -\sqrt{R^2 - (z-L)^2} + R - \frac{1}{2}b & \text{for } z \in (\frac{1}{2}L, L], \end{cases} \quad (4.3)$$

with radius  $R = L^2/(4b) + b/4$ , see [BL83]. For the optimization we used a model  $\mathbf{m} = [m_1, m_2]$ , where  $m_i$  denotes the  $x$  component of the waveguide center depending on  $z$ ,  $m_0 = b/2$  and  $m_3 = -b/2$  are

prescribed and the  $z$ -positions are  $z = [0, 200, 400, 600]$ . The model values are interpolated by cubic splines  $\Phi$ . The algorithm stops if the stopping criterion  $|\mathbf{m}^{(k+1)} - \mathbf{m}^{(k)}| \leq \text{tol} = 10^{-3}$  is fulfilled.

Using different initial values of the model  $\mathbf{m}$ , the Gauß–Newton method terminates after about 7 to 10 iterations. However, only 4 to 5 iterations are needed to reduce the loss significantly, see Figure 5. The evolution of the optimized profile during the optimization process is shown in Figure 6. The computed power losses for all profiles are given in Table 2. As one can see the sinusoidal profile produces already a low loss for this configuration and can be hardly improved. But there is a great difference of about 15% between the best and the worst profile. These results are also consistent with those of [MBB93] where a Gauß–Newton like method had been used. In contrast to our work, the losses were computed by using an approximation formula which depended on the slope of the profile and experimental data.

transition profile	transmitted power
sinusoidal (4.1)	77.52 %
modified sinusoidal (4.2)	62.87 %
connected arcs (4.3)	70.95 %
GN-Spline-optimized	78.39 %

TABLE 2. Power losses for different transition profiles.

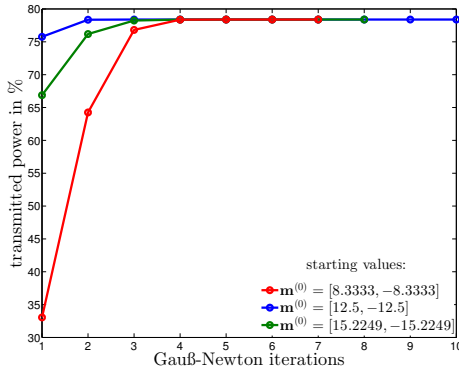


FIGURE 5. Transmitted power in the first mode dependent on Gauß–Newton iterations using different starting values.

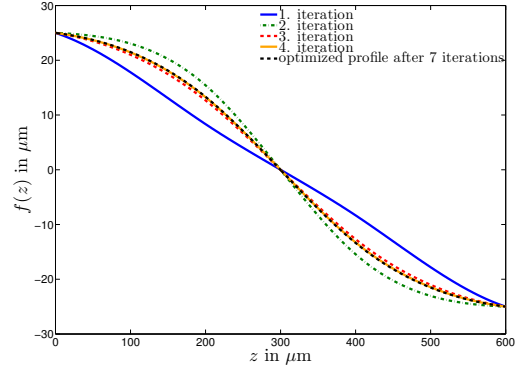


FIGURE 6. Evolution of the optimized profile during the Gauß–Newton optimization process.

**Acknowledgement.** S. Findeisen is financially supported by the DFG Research Training Group 1294 ”Analysis, Simulation and Design of Nanotechnological Processes“ at the KIT.

## REFERENCES

- [BBB<sup>+</sup>12] S. Balay, J. Brown, K. Buschelman, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang. PETSc Web page, 2012. <http://www.mcs.anl.gov/petsc>.
- [BDD<sup>+</sup>00] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*. SIAM, Philadelphia, PA, 2000.
- [BHK07] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II – A general-purpose object-oriented finite element library. *ACM Trans. Math. Softw.*, 33, 2007.

- [BL82] R. Baets and P. E. Lagasse. Calculation of radiation loss in integrated-optic tapers and Y-junctions. *Appl. Opt.*, 21(11):1972–1978, 1982.
- [BL83] R. Baets and P. E. Lagasse. Loss calculation and design of arbitrarily curved integrated-optic waveguides. *J. Opt. Soc. Amer.*, 73(2):177–182, 1983.
- [CL04] S. L. Chui and Y. Y. Lu. Wide-angle full-vector beam propagation method based on an alternating direction implicit preconditioner. *J. Opt. Soc. Amer. A*, 21(3):420–425, 2004.
- [CRRT12] C. Campos, J. E. Román, E. Romero, and A. Tomás. SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems, 2012. <http://www.grycap.upv.es/slepc>.
- [DS96] J. E. Dennis jun. and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, volume 16 of *Classics in Applied Mathematics*. SIAM, Philadelphia, PA, 1996.
- [EF01] H. W. Engl and T. Felici. On shape optimization of optical waveguides using inverse problem techniques. *Inverse Problems*, 17:1141–1162, 2001.
- [Ged96] S. D. Gedney. An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices. *IEEE Trans. Antennas and Propagation*, 44(12):1630–1639, 1996.
- [Had92] G. R. Hadley. Wide-angle beam propagation using padé approximant operators. *Opt. Lett.*, 17(20):1426–1428, 1992.
- [HYD83] B. Hermansson, D. Yevick, and P. Danielsen. Propagating beam analysis of multimode waveguide tapers. *IEEE Journal of Quantum Electronics*, 19(8):1246–1251, 1983.
- [LB09] K. Q. Le and P. Bienstman. Wide-angle beam propagation method without using slowly varying envelope approximation. *J. Opt. Soc. Amer. B Opt. Phys.*, 26(2):353–356, 2009.
- [Lif03] G. Lifante. *Integrated photonics: Fundamentals*. John Wiley & Sons Ltd, Chichester, 2003.
- [Mar78] D. Marcuse. Length optimization of an S-shaped transition between offset optical waveguides. *Appl. Opt.*, 17(5):763–768, 1978.
- [MBB93] F. J. Mustieles, E. Ballesteros, and P. Baquero. Theoretical s-bend profile for optimization of optical waveguide radiation losses. *IEEE Photonics Technology Letters*, 5:551–553, 1993.
- [Mon03] P. Monk. *Finite element methods for Maxwell's equations*. Numerical mathematics and scientific computation, Oxford science publications. Clarendon Press, Oxford, 2003.
- [NW99] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer, New York, NY, 1999.
- [RDL81] J. van Roey, J. van der Donk, and P. E. Lagasse. Beam-propagation method: analysis and assessment. *J. Opt. Soc. Amer.*, 71(7):803–810, 1981.
- [SA04] A. Sharma and A Agrawal. New method for nonparaxial beam propagation. *J. Opt. Soc. Amer. A*, 21(6):1082–1087, 2004.
- [SGBV98] D. Schulz, C. Glingener, M. Bludszuweit, and E. Voge. Mixed finite element beam propagation method. *J. Lightwave Technol.*, 16(7):1336–1342, 1998.

## IWRMM-Preprints seit 2011

- Nr. 11/01 Tobias Jahnke, Derya Altintan : Efficient simulation of discret stochastic reaction systems with a splitting method
- Nr. 11/02 Tobias Jahnke : On Reduced Models for the Chemical Master Equation
- Nr. 11/03 Martin Sauter, Christian Wieners : On the superconvergence in computational elasto-plasticity
- Nr. 11/04 B.D. Reddy, Christian Wieners, Barbara Wohlmuth : Finite Element Analysis and Algorithms for Single-Crystal Strain-Gradient Plasticity
- Nr. 11/05 Markus Bürg: An hp-Efficient Residual-Based A Posteriori Error Estimator for Maxwell's Equations
- Nr. 12/01 Branimir Anic, Christopher A. Beattie, Serkan Gugercin, Athanasios C. Antoulas: Interpolatory Weighted-H2 Model Reduction
- Nr. 12/02 Christian Wieners, Jiping Xin: Boundary Element Approximation for Maxwell's Eigenvalue Problem
- Nr. 12/03 Thomas Schuster, Andreas Rieder, Frank Schöpfer: The Approximate Inverse in Action IV: Semi-Discrete Equations in a Banach Space Setting
- Nr. 12/04 Markus Bürg: Convergence of an hp-Adaptive Finite Element Strategy for Maxwell's Equations
- Nr. 12/05 David Cohen, Stig Larsson, Magdalena Sigg: A Trigonometric Method for the Linear Stochastic Wave Equation
- Nr. 12/06 Tim Kreutzmann, Andreas Rieder: Geometric Reconstruction in Bioluminescence Tomography
- Nr. 12/07 Tobias Jahnke, Michael Kreim: Error bound for piecewise deterministic processes modeling stochastic reaction systems
- Nr. 12/08 Haojun Li, Kirankumar Hiremath, Andreas Rieder, Wolfgang Freude: Adaptive Wavelet Collocation Method for Simulation of Time Dependent Maxwell's Equations
- Nr. 12/09 Andreas Arnold, Tobias Jahnke: On the approximation of high-dimensional differential equations in the hierarchical Tucker format
- Nr. 12/10 Mike A. Botchev, Volker Grimm, Marlis Hochbruck: Residual, Restarting and Richardson Iteration for the Matrix Exponential
- Nr. 13/01 Willy Dörfler, Stefan Findeisen: Numerical Optimization of a Waveguide Transition Using Finite Element Beam Propagation

Eine aktuelle Liste aller IWRMM-Preprints finden Sie auf:

[www.math.kit.edu/iwrmm/seite/preprints](http://www.math.kit.edu/iwrmm/seite/preprints)

## **Kontakt**

Karlsruher Institut für Technologie (KIT)  
Institut für Wissenschaftliches Rechnen  
und Mathematische Modellbildung

Prof. Dr. Christian Wieners  
Geschäftsführender Direktor

Campus Süd  
Engesserstr. 6  
76131 Karlsruhe

E-Mail: [Bettina.Haindl@kit.edu](mailto:Bettina.Haindl@kit.edu)

---

[www.math.kit.edu/iwrmm/](http://www.math.kit.edu/iwrmm/)

## **Herausgeber**

Karlsruher Institut für Technologie (KIT)  
Kaiserstraße 12 | 76131 Karlsruhe

*Januar 2013*

---

[www.kit.edu](http://www.kit.edu)