

A variational mode solver for optical waveguides based on quasi-analytical vectorial slab mode expansion

Alyona Ivanova¹, Remco Stoffer^{1,2}, and Manfred Hammer¹

¹MESA+ Institute for Nanotechnology, University of Twente, Enschede, The Netherlands

²PhoeniX Software, Enschede, The Netherlands

A flexible and efficient method for fully vectorial modal analysis of 3D dielectric optical waveguides with arbitrary 2D cross-sections is proposed. The technique is based on expansion of each modal component in some a priori defined functions defined on one coordinate axis times some unknown coefficient-functions, defined on the other axis. By applying a variational restriction procedure the unknown coefficient-functions are determined, resulting in an optimum approximation of the true vectorial mode profile. This technique can be related to both Effective Index and Mode Matching methods. A couple of examples illustrate the performance of the method.

Keywords: Integrated optics; Dielectric waveguides; Guided modes; Numerical modelling; Variational Methods

PACS: 42.82.m; 42.82.Et

1 Introduction

Three-dimensional optical channel waveguides are basic components of integrated optical devices such as directional couplers, wavelength filters, phase shifters, and optical switches. The successful design of these devices requires an accurate estimation of the modal field profiles and propagation constants. Over already some decades several classes of methods for the analysis of dielectric optical waveguides were developed: among these are techniques of more numerical character, like Finite Element and Finite Difference approximations, the Method of Lines, and Integral Equations Methods, but also more analytical approaches like Film Mode Matching (FMM) and the Effective Index Method (EIM). Detailed overviews of these techniques can be found in [1, 2, 3].

In the present paper we propose an extension of the scalar mode solver [4] to vectorial problems. Our method is based on individual expansions of each mode profile component into a set of a priori defined functions of one coordinate axis (vertical), here, field components of some slab waveguide mode. The expansion is global, meaning that the same basis functions are used at any point on the horizontal axis. The unknown expansion coefficients – in our case functions, defined on the horizontal axis – are found by means of variational methods [5, 6].

The present method can be viewed as some bridge between two popular approaches, namely the FMM on the one hand and the EIM on the other. In the standard EIM the 2D problem of finding modes of the waveguide is reduced to consecutive solving two 1D problems: at first, the 1D modes, and their propagation constants, of the constituting slab waveguides are found, and then their propagation constants are used to define effective refractive indices of a reduced 1D problem. In general this is a very quick and easy approach for a rough estimation of mode parameters. However, in case one of the constituting slabs doesn't support a guided mode (for example, some substrate material with air on top) it is impossible to uniquely define the effective refractive index in that particular region of the reduced problem. Should it be the refractive index of the cladding,

refractive index of the air, or something in between? The restriction of the present approach to one-term expansions will answer this question.

The validity of the method was checked on several structures, including waveguides with rectangular and non-rectangular piecewise-constant refractive index distributions, and a diffused waveguide. Comparison shows that the present method is a more consistent and accurate alternative to the standard EIM and also can be pushed to its limits and used for rigorous computations.

The paper is organized as follows. In section 2 the problem of finding vectorial modes of the dielectric waveguides is stated, then some properties of slab modes and the modal field ansatz are described in sections 3 and 4. The equations for the coefficient functions are derived in section 5. Section 6 outlines the numerical solution methods. The relation of the present method to the EIM and FMM is explained in more detail in sections 7 and 8. Then in section 9 numerical results for several waveguide configurations are presented. Finally some concluding remarks are made in section 10.

2 Variational form of the vectorial mode problem

Consider a z -invariant dielectric isotropic waveguide defined on its cross-section by a refractive index $n(x, y)$ or relative dielectric permittivity $\varepsilon(x, y) = n^2(x, y)$ distribution. Figure 1 shows two examples.

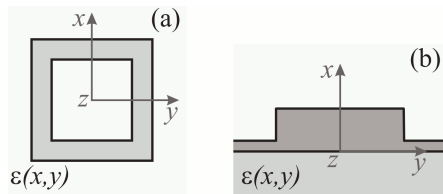


Figure 1: Examples for 3D dielectric waveguides defined on their cross-section by permittivity distribution $\varepsilon(x, y)$. The structures are invariant along the z -axis. (a) box-shaped hollow-core waveguide, a concept from [7], the subject of section 9.1, (b) a standard rib waveguide, investigated in section 9.2.

The propagation of monochromatic light, given by the electric $\bar{\mathbf{E}}$ and magnetic $\bar{\mathbf{H}}$ components of the optical field, with propagation constant β and frequency ω ,

$$\bar{\mathbf{E}}(x, y, z, t) = \mathbf{E}(x, y) e^{-i\beta z} e^{i\omega t}, \quad \bar{\mathbf{H}}(x, y, z, t) = \mathbf{H}(x, y) e^{-i\beta z} e^{i\omega t}, \quad (1)$$

is governed by the Maxwell equations for the mode profile components \mathbf{E} and \mathbf{H}

$$\begin{aligned} \omega\varepsilon_0\varepsilon\mathbf{E} + iC\mathbf{H} &= \beta R\mathbf{H}, \\ \omega\mu_0\mu\mathbf{H} - iC\mathbf{E} &= -\beta R\mathbf{E}, \end{aligned} \quad (2)$$

with

$$R = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & \partial_y \\ 0 & 0 & -\partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix}, \quad (3)$$

vacuum permittivity ε_0 , vacuum permeability μ_0 , relative permittivity $\varepsilon(x, y) = n^2(x, y)$. Here and further in this paper it is assumed that the relative permeability μ is equal to 1, as is the case for most materials at optical frequencies.

We will work with a variational formulation of the Maxwell equations. Solutions $(\beta, \mathbf{E}, \mathbf{H})$ of the equations (2) correspond to stationary points (\mathbf{E}, \mathbf{H}) of the functional [5]

$$\mathcal{F}(\mathbf{E}, \mathbf{H}) = \frac{\omega\varepsilon_0\langle\mathbf{E}, \varepsilon\mathbf{E}\rangle + \omega\mu_0\langle\mathbf{H}, \mathbf{H}\rangle + i\langle\mathbf{E}, C\mathbf{H}\rangle - i\langle\mathbf{H}, C\mathbf{E}\rangle}{\langle\mathbf{E}, R\mathbf{H}\rangle - \langle\mathbf{H}, R\mathbf{E}\rangle}, \quad (4)$$

with propagation constant $\beta = \mathcal{F}(\mathbf{E}, \mathbf{H})$ equal to the value of the functional at the stationary point. The inner product used is $\langle\mathbf{A}, \mathbf{B}\rangle = \int \mathbf{A}^* \cdot \mathbf{B} dx dy$. The natural interface conditions are the continuity of all tangential field components across the interfaces.

3 Slab modes

In this section we will consider modes of slab waveguides, which we will use in the next section as building blocks to construct approximations of the modes of waveguides with arbitrary 2D cross-sections. Furthermore, we introduce rotations of the slab modes; these rotations will be needed to provide a physical motivation for the particular form of the approximations that we will employ.

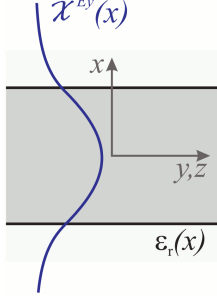


Figure 2: A slab waveguide with permittivity distribution $\varepsilon_r(x)$ and principal component χ^{E_y} of a corresponding TE slab mode (5).

A one dimensional TE mode, propagating in the z -direction with propagation constant $\beta_{r,TE}$, of the slab waveguide, given by the permittivity distribution $\varepsilon_r(x)$ (Figure 2) can be represented as

$$\begin{pmatrix} E_x, E_y, E_z \\ H_x, H_y, H_z \end{pmatrix} (x, y, z) = \begin{pmatrix} 0, & \chi^{E_y}(x), & 0 \\ \chi^{H_x}(x), & 0, & \chi^{H_z}(x) \end{pmatrix} e^{-i\beta_{r,TE}z}. \quad (5)$$

The principal electric component χ^{E_y} satisfies the equation

$$(\chi^{E_y}(x))'' + k^2 \varepsilon_r(x) \chi^{E_y}(x) = \beta_{r,TE}^2 \chi^{E_y}(x) \quad (6)$$

with vacuum wavenumber $k = 2\pi/\lambda$. The remaining two nonzero components of the mode profile can be derived directly from χ^{E_y} :

$$\chi^{H_x}(x) = -\frac{\beta_{r,TE}}{\omega\mu_0} \chi^{E_y}(x), \quad \chi^{H_z}(x) = \frac{i}{\omega\mu_0} (\chi^{E_y}(x))'. \quad (7)$$

The slab waveguide (Figure 2) is by definition invariant in the (y, z) -plane. So if a modal solution of Maxwell equations propagating in the z -direction will be rotated in the (y, z) -plane by an angle θ (Figure 3), it will still remain a modal solution of the Maxwell equations, but now propagating in the direction $(y, z) = (-\sin\theta; \cos\theta)$:

$$\begin{pmatrix} E_x, E_y, E_z \\ H_x, H_y, H_z \end{pmatrix} (x, y, z) = \begin{pmatrix} 0, & \chi^{E_y}(x) \cos\theta, & \chi^{E_y}(x) \sin\theta \\ \chi^{H_x}(x), & -\chi^{H_z}(x) \sin\theta, & \chi^{H_z}(x) \cos\theta \end{pmatrix} e^{-i\beta_{r,TE}(-\sin\theta y + \cos\theta z)}. \quad (8)$$

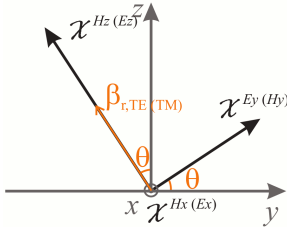


Figure 3: A slab TE (TM) mode propagating in the z -direction with propagation constant $\beta_{r,TE(TM)}$ is rotated around the x -axis by an angle θ . The rotated mode propagates with the same propagation constant, but in the direction $(y, z) = (-\sin\theta, \cos\theta)$.

Similarly a 1D TM slab mode, propagating in the z -direction with propagation constant $\beta_{r,TM}$

$$\begin{pmatrix} E_x, E_y, E_z \\ H_x, H_y, H_z \end{pmatrix} (x, y, z) = \begin{pmatrix} \chi^{E_x}(x), & 0, & \chi^{E_z}(x) \\ 0, & \chi^{H_y}(x), & 0 \end{pmatrix} e^{-i\beta_{r,TM}z}, \quad (9)$$

will still be a solution of the Maxwell equations after a rotation around the x -axis (Figure 3)

$$\begin{pmatrix} E_x, E_y, E_z \\ H_x, H_y, H_z \end{pmatrix} (x, y, z) = \begin{pmatrix} \chi^{E_x}(x), & -\chi^{E_z}(x) \sin\theta, & \chi^{E_z}(x) \cos\theta \\ 0, & \chi^{H_y}(x) \cos\theta, & \chi^{H_y}(x) \sin\theta \end{pmatrix} e^{-i\beta_{r,TM}(-\sin\theta y + \cos\theta z)}. \quad (10)$$

The principal magnetic component χ^{H_y} satisfies the equation

$$\left(\frac{1}{\varepsilon_r(x)} (\chi^{H_y(x)})' \right)' + k^2 \chi^{H_y(x)} = \beta_{\text{r,TM}}^2 \frac{1}{\varepsilon_r(x)} \chi^{H_y(x)}. \quad (11)$$

Again the remaining two nonzero components of the mode profile can be derived directly from χ^{H_y} :

$$\chi^{E_x(x)} = \frac{\beta_{\text{r,TM}}}{\omega \varepsilon_0 \varepsilon_r(x)} \chi^{H_y(x)}, \quad \chi^{E_z(x)} = -\frac{i}{\omega \varepsilon_0 \varepsilon_r(x)} (\chi^{H_y(x)})'. \quad (12)$$

4 Modal field ansatz

We now return to the vectorial modes of the 3D waveguides, as in section 2. Each field component $F \in \{E_x, E_y, E_z, H_x, H_y, H_z\}$ is represented individually as a superposition of m_F a priori known functions $X_j^F(x)$, defined on one coordinate axis, times some unknown coefficient-function $Y_j^F(y)$, defined on the other axis:

$$F(x, y) = \sum_{j=1}^{m_F} X_j^F(x) Y_j^F(y). \quad (13)$$

For the functions X we will take components of slab modes from some reference slice(s). Further in the paper two types of the expansion will be relevant, one which introduces 5 unknown functions Y per slab mode, and another one, which introduces only 3. These will be called five component approximation (VEIM5) and three component approximation (VEIM3), respectively.

In case of VEIM5, the TE basis mode (5) number j with mode profile components $\chi_j^{E_y}, \chi_j^{H_x}, \chi_j^{H_z}$ contributes to the expansion of components E_y, E_z, H_x, H_y and H_z with the form $\chi_j^{E_y} Y_j^{E_y}, \chi_j^{E_y} Y_j^{E_z}, \chi_j^{H_x} Y_j^{H_x}, \chi_j^{H_z} Y_j^{H_y}$ and $\chi_j^{H_z} Y_j^{H_z}$. Likewise, the TM basis mode (9) number l with mode profile components $\chi_l^{E_y}, \chi_l^{H_x}, \chi_l^{H_z}$ contributes to the expansion of components E_x, E_y, E_z, H_y and H_z with the form $\chi_l^{E_x} Y_l^{E_x}, \chi_l^{E_z} Y_l^{E_y}, \chi_l^{E_z} Y_l^{E_z}, \chi_l^{H_y} Y_l^{H_y}, \chi_l^{H_y} Y_l^{H_z}$, such that the complete expansion looks like

$$\begin{aligned} \begin{pmatrix} E_x, E_y, E_z \\ H_x, H_y, H_z \end{pmatrix} (x, y, z) &= \sum_{j \in \text{TE}} \begin{pmatrix} 0, & \chi_j^{E_y}(x) Y_j^{E_y}(y), & \chi_j^{E_y}(x) Y_j^{E_z}(y) \\ \chi_j^{H_x}(x) Y_j^{H_x}(y), & \chi_j^{H_z}(x) Y_j^{H_y}(y), & \chi_j^{H_z}(x) Y_j^{H_z}(y) \end{pmatrix} + \\ &+ \sum_{l \in \text{TM}} \begin{pmatrix} \chi_l^{E_x}(x) Y_l^{E_x}(y), & \chi_l^{E_z}(x) Y_l^{E_y}(y), & \chi_l^{E_z}(x) Y_l^{E_z}(y) \\ 0, & \chi_l^{H_y}(x) Y_l^{H_y}(y), & \chi_l^{H_y}(x) Y_l^{H_z}(y) \end{pmatrix}. \end{aligned} \quad (14)$$

This expansion has the drawback that the functions making up some of the components can become linearly dependent; for example, the full set of $\chi_l^{H_y}(x)$ components from TM modes form a complete set; thus any $\chi_j^{H_z}(x)$ from a TE mode can be expressed in that complete set of functions. When using a limited number of modes in the expansion, no problems result from this; however, increasing the number of modes will at some point make the problem ill-conditioned. Therefore, we introduce a different expansion, which we call VEIM3, in which we omit contributions of some modal components - making sure that each vector component is only represented by either TE or TM slab mode components. So a TE basis mode (5) number j with mode profile components $\chi_j^{E_y}, \chi_j^{H_x}, \chi_j^{H_z}$ contributes to the expansion of components E_y, E_z and H_x with the form $\chi_j^{E_y} Y_j^{E_y}, \chi_j^{E_y} Y_j^{E_z}, \chi_j^{H_x} Y_j^{H_x}$. Likewise a TM basis mode (9) number l with mode profile components $\chi_l^{E_y}, \chi_l^{H_x}, \chi_l^{H_z}$ contributes to the expansion of components E_x, H_y and H_z with the form $\chi_l^{E_x} Y_l^{E_x}, \chi_l^{H_y} Y_l^{H_y}, \chi_l^{H_y} Y_l^{H_z}$, such that the complete expansion looks like

$$\begin{aligned} \begin{pmatrix} E_x, E_y, E_z \\ H_x, H_y, H_z \end{pmatrix} (x, y, z) &= \sum_{j \in \text{TE}} \begin{pmatrix} 0, & \chi_j^{E_y}(x) Y_j^{E_y}(y), & \chi_j^{E_y}(x) Y_j^{E_z}(y) \\ \chi_j^{H_x}(x) Y_j^{H_x}(y), & 0, & 0 \end{pmatrix} + \\ &+ \sum_{l \in \text{TM}} \begin{pmatrix} \chi_l^{E_x}(x) Y_l^{E_x}(y), & 0, & 0 \\ 0, & \chi_l^{H_y}(x) Y_l^{H_y}(y), & \chi_l^{H_y}(x) Y_l^{H_z}(y) \end{pmatrix}. \end{aligned} \quad (15)$$

Note that in both expansions each contributing component χ^F of a 1D mode is used to represent the field not only in the slab segment where it belongs, as in EIM and FMM methods, but also in the whole waveguide. So even with a single slab mode in both expansions, (14) and (15), it is possible to construct an approximation of the field in the whole structure. In section 7 we will study in detail properties of such one-mode-expansions.

The form of the expansion (14) was inspired by the mode matching techniques that use the physically motivated approach of employing rotated modes (8), (10) to locally expand the total field [8, 9]. In the present approach though, we attribute those parts of the slab mode components that do not depend on x to the functions Y^F , treating them as unknowns – but the x -dependence of the y and z components is still the same. In the sections 7 and 8 we will study the behavior of these functions Y^F .

What concerns the choice of the reference slice(s), it seems that modal components from the slice, where the maximum power is expected to be localized, give the best results. Further in this paper VEIM5 will be used with a few modes only for rough and efficient approximations, while VEIM3 will be used with higher numbers of modes to obtain accurate, converged results. We do not restrict to using modes from one reference slice only; we observe that adding mode(s) from another slice can greatly improve accuracy for lower number of modes in the expansion. However, care must be taken; the problem can become ill-conditioned if the modes become nearly linear dependent.

In the following all the slab mode components χ , which are used to expand a field component F of the complete waveguide, we will denote as X^F (just like in eqn. (13)).

5 Reduced problem

The next question is how to find corresponding functions Y , such that the expansion (13) represents the true solution in the best possible way. For this purpose we apply variational restriction [2, 6] of the functional (4). In short it can be outlined as follows. As it was already mentioned the critical points of the functional (4), which satisfy some continuity conditions, are solutions of the Maxwell equations (2) and, vice versa, solutions of the Maxwell equations (2) are critical points of the functional (4).

After insertion of the expansions (14) or (15), variation of the functional (4) with respect to a function \mathbf{Y}^F , a vector function made up of all functions Y^F , results in the following system of first order differential equations for \mathbf{Y}^F with parameter β :

$$\begin{aligned}
\mathbf{A}_{11}\mathbf{Y}^{E_x} + \mathbf{A}_{12}(\mathbf{Y}^{H_z})' &= \beta\mathbf{A}_{13}\mathbf{Y}^{H_y} \\
\mathbf{A}_{21}\mathbf{Y}^{E_y} + \mathbf{A}_{22}\mathbf{Y}^{H_z} &= \beta\mathbf{A}_{23}\mathbf{Y}^{H_x} \\
\mathbf{A}_{31}\mathbf{Y}^{E_z} + \mathbf{A}_{32}(\mathbf{Y}^{H_x})' + \mathbf{A}_{33}\mathbf{Y}^{H_y} &= 0 \\
\mathbf{A}_{41}\mathbf{Y}^{H_x} + \mathbf{A}_{42}(\mathbf{Y}^{E_z})' &= \beta\mathbf{A}_{43}\mathbf{Y}^{E_y} \\
\mathbf{A}_{51}\mathbf{Y}^{H_y} + \mathbf{A}_{52}\mathbf{Y}^{E_z} &= \beta\mathbf{A}_{53}\mathbf{Y}^{E_x} \\
\mathbf{A}_{61}\mathbf{Y}^{H_z} + \mathbf{A}_{62}(\mathbf{Y}^{E_x})' + \mathbf{A}_{63}\mathbf{Y}^{E_y} &= 0.
\end{aligned} \tag{16}$$

The elements of the matrices \mathbf{A} include the overlap integrals (here: $\langle a, b \rangle = \int a^* b dx$) of the functions $X_j^F(x)$, their derivatives, and the local permittivity distribution of the waveguide:

$$\begin{aligned}
\mathbf{A}_{11}(p, j) &= \omega \langle X_p^{E_x}, \varepsilon X_j^{E_x} \rangle & \mathbf{A}_{12}(p, j) &= i \langle X_p^{E_x}, X_j^{H_z} \rangle & \mathbf{A}_{13}(p, j) &= \langle X_p^{E_x}, X_j^{H_y} \rangle \\
\mathbf{A}_{21}(p, j) &= \omega \langle X_p^{E_y}, \varepsilon X_j^{E_y} \rangle & \mathbf{A}_{22}(p, j) &= -i \langle X_p^{E_y}, (X_j^{H_z})' \rangle & \mathbf{A}_{23}(p, j) &= -\langle X_p^{E_y}, X_j^{H_x} \rangle \\
\mathbf{A}_{31}(p, j) &= \omega \langle X_p^{E_z}, \varepsilon X_j^{E_z} \rangle & \mathbf{A}_{32}(p, j) &= -i \langle X_p^{E_z}, X_j^{H_x} \rangle & \mathbf{A}_{33}(p, j) &= i \langle X_p^{E_z}, (X_j^{H_y})' \rangle \\
\mathbf{A}_{41}(p, j) &= \omega \mu \langle X_p^{H_x}, X_j^{H_x} \rangle & \mathbf{A}_{42}(p, j) &= -i \langle X_p^{H_x}, X_j^{E_z} \rangle & \mathbf{A}_{43}(p, j) &= -\langle X_p^{H_x}, X_j^{E_y} \rangle \\
\mathbf{A}_{51}(p, j) &= \omega \mu \langle X_p^{H_y}, X_j^{H_y} \rangle & \mathbf{A}_{52}(p, j) &= i \langle X_p^{H_y}, (X_j^{E_z})' \rangle & \mathbf{A}_{53}(p, j) &= \langle X_p^{H_y}, X_j^{E_x} \rangle \\
\mathbf{A}_{61}(p, j) &= \omega \mu \langle X_p^{H_z}, X_j^{H_z} \rangle & \mathbf{A}_{62}(p, j) &= i \langle X_p^{H_z}, X_j^{E_x} \rangle & \mathbf{A}_{63}(p, j) &= -i \langle X_p^{H_z}, (X_j^{E_y})' \rangle
\end{aligned} \tag{17}$$

Note that the permittivity appears only in \mathbf{A}_{11} , \mathbf{A}_{21} and \mathbf{A}_{31} , hence only these matrices are y -dependent.

If the permittivity exhibits discontinuities along the y -direction, the functions

$$\mathbf{Y}^{E_x} \quad \text{and} \quad \mathbf{Y}^{H_x}, \quad \mathbf{Y}^{E_z} \quad \text{and} \quad \mathbf{Y}^{H_z} \quad (18)$$

are required to be continuous at the respective positions.

It turns out that by algebraic operations the system of first order differential equations (16) can be reduced to a system of second order differential equations for the vector functions \mathbf{Y}^{E_x} and \mathbf{Y}^{H_x} only. Moreover, since the components E_y and E_z , H_y and H_z are approximated by the same functions χ in the representations (14) and (15), the matrices \mathbf{A} satisfy the following equalities:

$$\begin{aligned} \mathbf{A}_{13} &= -i \mathbf{A}_{12} \\ \mathbf{A}_{31} &= \mathbf{A}_{21}, \quad \mathbf{A}_{32} = i \mathbf{A}_{23}, \quad \mathbf{A}_{33} = -\mathbf{A}_{22} \\ \mathbf{A}_{43} &= -i \mathbf{A}_{42} \\ \mathbf{A}_{61} &= \mathbf{A}_{51}, \quad \mathbf{A}_{62} = i \mathbf{A}_{53}, \quad \mathbf{A}_{63} = -\mathbf{A}_{52}, \end{aligned} \quad (19)$$

and hence the system (16) reduces to

$$\mathbf{S}_1 \mathbf{u} + (\mathbf{S}_2 \mathbf{u}' + \beta \mathbf{S}_3 \mathbf{u})' = \beta^2 \mathbf{S}_2 \mathbf{u} + \beta \mathbf{S}_3 \mathbf{u}', \quad (20)$$

with $\mathbf{u}(y) = \begin{pmatrix} \mathbf{Y}^{E_x}(y) \\ \mathbf{Y}^{H_x}(y) \end{pmatrix}$ and (anti-)block-diagonal matrices \mathbf{S} of the following form:

$$\begin{aligned} \mathbf{S}_1 &= \begin{pmatrix} \mathbf{A}_{11} & 0 \\ 0 & \mathbf{A}_{41} \end{pmatrix}, \\ \mathbf{S}_2 &= \begin{pmatrix} -i \mathbf{A}_{12} (\mathbf{A}_{51} + \mathbf{A}_{52} \mathbf{A}_{21}^{-1} \mathbf{A}_{22})^{-1} \mathbf{A}_{53} & 0 \\ 0 & -i \mathbf{A}_{42} (\mathbf{A}_{21} + \mathbf{A}_{22} \mathbf{A}_{51}^{-1} \mathbf{A}_{52})^{-1} \mathbf{A}_{23} \end{pmatrix}, \\ \mathbf{S}_3 &= \begin{pmatrix} 0 & \mathbf{A}_{12} \mathbf{A}_{51}^{-1} \mathbf{A}_{52} (\mathbf{A}_{21} + \mathbf{A}_{22} \mathbf{A}_{51}^{-1} \mathbf{A}_{52})^{-1} \mathbf{A}_{23} \\ \mathbf{A}_{42} \mathbf{A}_{21}^{-1} \mathbf{A}_{22} (\mathbf{A}_{51} + \mathbf{A}_{52} \mathbf{A}_{21}^{-1} \mathbf{A}_{22})^{-1} \mathbf{A}_{53} & 0 \end{pmatrix}. \end{aligned} \quad (21)$$

Across the vertical interfaces continuity of

$$\mathbf{u} \quad \text{and} \quad \mathbf{S}_2 \mathbf{u}' + \beta \mathbf{S}_3 \mathbf{u} \quad (22)$$

is required.

As soon as the function \mathbf{u} , or in other words \mathbf{Y}^{E_x} and \mathbf{Y}^{H_x} , are known, the functions \mathbf{Y} corresponding to the four other components can be derived as follows:

$$\begin{aligned} \mathbf{Y}^{E_y} &= i \mathbf{A}_{21}^{-1} \mathbf{A}_{22} (\mathbf{A}_{51} + \mathbf{A}_{52} \mathbf{A}_{21}^{-1} \mathbf{A}_{22})^{-1} \mathbf{A}_{53} (\mathbf{Y}^{E_x})' + \beta (\mathbf{A}_{21} + \mathbf{A}_{22} \mathbf{A}_{51}^{-1} \mathbf{A}_{52})^{-1} \mathbf{A}_{23} \mathbf{Y}^{H_x}, \\ \mathbf{Y}^{E_z} &= \beta \mathbf{A}_{21}^{-1} \mathbf{A}_{22} (\mathbf{A}_{51} + \mathbf{A}_{52} \mathbf{A}_{21}^{-1} \mathbf{A}_{22})^{-1} \mathbf{A}_{53} \mathbf{Y}^{E_x} - i (\mathbf{A}_{21} + \mathbf{A}_{22} \mathbf{A}_{51}^{-1} \mathbf{A}_{52})^{-1} \mathbf{A}_{23} (\mathbf{Y}^{H_x})', \\ \mathbf{Y}^{H_y} &= \beta (\mathbf{A}_{51} + \mathbf{A}_{52} \mathbf{A}_{21}^{-1} \mathbf{A}_{22})^{-1} \mathbf{A}_{53} \mathbf{Y}^{E_x} + i \mathbf{A}_{51}^{-1} \mathbf{A}_{52} (\mathbf{A}_{21} + \mathbf{A}_{22} \mathbf{A}_{51}^{-1} \mathbf{A}_{52})^{-1} \mathbf{A}_{23} (\mathbf{Y}^{H_x})', \\ \mathbf{Y}^{H_z} &= -i (\mathbf{A}_{51} + \mathbf{A}_{52} \mathbf{A}_{21}^{-1} \mathbf{A}_{22})^{-1} \mathbf{A}_{53} (\mathbf{Y}^{E_x})' + \beta \mathbf{A}_{51}^{-1} \mathbf{A}_{52} (\mathbf{A}_{21} + \mathbf{A}_{22} \mathbf{A}_{51}^{-1} \mathbf{A}_{52})^{-1} \mathbf{A}_{23} \mathbf{Y}^{H_x}. \end{aligned} \quad (23)$$

Note that substituting Equation (23) into the continuity conditions (22) shows that (22) exactly implies the continuity of the relevant electromagnetic components (18).

6 Method of solution

In general the system (20) can be solved by the Finite Element method [10, 11]. It relies on a spatial discretization, i.e. divides the whole computational domain into a number of elements. On each of these elements the unknown function is represented as a superposition of some basis functions. The coefficients of the expansion are found using the weak form of eqn. (20). While this method is very general, it quickly introduces a large number of unknowns.

However, due to common techniques of fabrication many waveguides do not have a completely arbitrary refractive index distribution, but rather one which is piecewise constant along the horizontal axis. The waveguide then can be split in several vertical slices, where the refractive index does not change in the horizontal direction. In each of these layers the general solution of (20) can be written down analytically. Gluing them together across the vertical interfaces will give the desired mode profile.

Both of these methods can be applied to find not only the fundamental, but also higher order modes. In the following we will outline each of these methods in more detail.

6.1 Arbitrary refractive index distribution: Finite Element Method

In case of an arbitrary permittivity distribution $\varepsilon(x, y)$ (diffused waveguide, waveguide with slanted sidewalls) the matrices \mathbf{S} depend on y , as their elements include overlap integrals with the permittivity $\varepsilon(x, y)$. One of the ways to solve the differential equation (20) is by using the Finite Element Method.

By multiplying both sides of (20) from the left by some continuous test vector-function \mathbf{v} and integrating over y one gets the weak form of equation (20):

$$\int (-\mathbf{v}^\top \mathbf{S}_1 \mathbf{u} + (\mathbf{v}^\top)' \mathbf{S}_2 \mathbf{u}') dy + \beta \int ((\mathbf{v}^\top)' \mathbf{S}_3 \mathbf{u} + \mathbf{v}^\top \mathbf{S}_3 \mathbf{u}') dy + \beta^2 \int \mathbf{v}^\top \mathbf{S}_2 \mathbf{u} dy = 0. \quad (24)$$

Then we expand the solution \mathbf{u} into a finite combination of the basis functions φ_{ij} ,

$$\mathbf{u}(y) = \sum_{i=1}^{n_d} \sum_{j=1}^{n_g} a_{ij} \varphi_{ij}(y), \quad (25)$$

with n_d the dimension of the vector \mathbf{u} ; n_g the number of consecutive grid points y_j into which the y -axis has been divided, and

$$\varphi_{ij}(y) = \begin{pmatrix} 0 \\ \vdots \\ \hat{\varphi}_j(y) \\ \vdots \\ 0 \end{pmatrix} \leftarrow i^{\text{th}} \text{ position} \quad (26)$$

with, for example, linear basis functions

$$\hat{\varphi}_j(y) = \begin{cases} 0, & y < y_{j-1} \text{ or } y \geq y_{j+1}; \\ \frac{y-y_{j-1}}{y_j-y_{j-1}}, & y_{j-1} \leq y < y_j; \\ \frac{y_{j+1}-y}{y_{j+1}-y_j}, & y_j \leq y < y_{j+1}. \end{cases} \quad (27)$$

As eqn. (24) should hold for an arbitrary continuous \mathbf{v} , we choose it to be one of the basis functions φ_{ij} . For $i = 1, \dots, n_d$ and $j = 1, \dots, n_g$ this results in the system of exactly $n_d \cdot n_g$ linear equations

$$(-\hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2) \mathbf{a} + \beta (\hat{\mathbf{S}}_3 + \hat{\mathbf{S}}_5) \mathbf{a} + \beta^2 \hat{\mathbf{S}}_4 \mathbf{a} = 0, \quad (28)$$

where $\mathbf{a}^\top = (a_{ij}) = ([a_{11}, \dots, a_{n_d 1}], [a_{12}, \dots, a_{n_d 2}], \dots, [a_{1 n_g}, \dots, a_{n_d n_g}])$ (the subscript ij here refers to the j^{th} element of the i^{th} subvector). Since for any square matrix \mathbf{M} of dimension $n_d n_g \times n_d n_g$

$$\varphi_{pm}^\top \mathbf{M} \varphi_{ij} = \hat{\varphi}_m(y) \cdot M_{pi}(y) \cdot \hat{\varphi}_j \quad (29)$$

holds, the matrices \mathbf{S} turn to be of the following form

$$\begin{aligned} \hat{\mathbf{S}}_1(pm, ij) &= \int \hat{\varphi}_m S_{1pi} \hat{\varphi}_j dy, \\ \hat{\mathbf{S}}_2(pm, ij) &= \int (\hat{\varphi}_m)' S_{2pi} (\hat{\varphi}_j)' dy, \\ \hat{\mathbf{S}}_3(pm, ij) &= \int (\hat{\varphi}_m)' S_{3pi} \hat{\varphi}_j dy, \\ \hat{\mathbf{S}}_4(pm, ij) &= \int \hat{\varphi}_m S_{2pi} \hat{\varphi}_j dy, \\ \hat{\mathbf{S}}_5(pm, ij) &= \int \hat{\varphi}_m S_{3pi} (\hat{\varphi}_j)' dy, \end{aligned} \quad (30)$$

where the indices pm and ij have the same meaning as in the definition of the vector \mathbf{a} .

The solution of the quadratic eigenvalue problem (28) with β as an eigenvalue can be found by introducing an auxiliary vector $\mathbf{b} = \beta\mathbf{a}$. (28) can then be transformed into the following linear eigenvalue problem

$$\begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\hat{\mathbf{S}}_4^{-1}(-\hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2) & -\hat{\mathbf{S}}_4^{-1}(\hat{\mathbf{S}}_3 + \hat{\mathbf{S}}_5) \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} = \beta \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}. \quad (31)$$

This is a quite straightforward, but expensive approach, as the dimension of the transformed problem is doubled in comparison to the original one. Other more involved approaches to tackle a quadratic eigenvalue problem can be found e.g. in [12]. We apply standard general eigenvalue solvers as embedded within the LAPACK [13] package. Specialized solvers could be employed, provided that an initial guess for the propagation constant, or a range of possible eigenvalues, are available for the problem at hand. On the other hand, there are situations where all the propagation constants β and corresponding functions \mathbf{u} need to be found together, e.g. if one wants to expand a 3D field in terms of vectorial modes of some channel waveguide, as required for the implementation of transparent boundary conditions [14].

While the entire 2D problem could also be solved directly by means of a Finite Element method, the number of degrees of freedom in such cases would be much higher than when solving the 1D equations (20) using the Finite Element Method; instead of having to use a triangulation of the entire 2D domain, only 1D finite elements are needed; furthermore, the number of degrees of freedom on each node is equal to the number of modes in the expansion, which is typically a small number.

6.2 Piecewise constant refractive index distribution

If a waveguide has a piecewise constant rectangular refractive index profile, it can be divided by vertical lines into slices with constant refractive index distribution along the y -direction. In each of these slices the matrices \mathbf{S} do not depend on y . Then (20) can be rewritten in a more familiar manner: Inside each of the slices \mathbf{u} should satisfy a system of second order differential equations with constant coefficients \mathbf{S} and a parameter β^2

$$\mathbf{S}_1\mathbf{u} + \mathbf{S}_2\mathbf{u}'' = \beta^2\mathbf{S}_2\mathbf{u}, \quad (32)$$

together with the continuity conditions (22). Moreover the matrices \mathbf{S}_1 and \mathbf{S}_2 are block-diagonal in such a way that the equations for the functions \mathbf{Y}^{E_x} and \mathbf{Y}^{H_x} decouple inside each of the slices; coupling occurs only across the vertical interfaces.

Inside each slice a particular solution of the system (32) can be readily written as

$$\mathbf{u} = c e^{\alpha y} \mathbf{p} \quad (33)$$

with some constants c, α and a vector \mathbf{p} . By substituting (33) into (32) we find a generalized eigenvalue problem with $\eta^2 = \beta^2 - \alpha^2$ as an eigenvalue:

$$\mathbf{S}_1\mathbf{p} = \eta^2\mathbf{S}_2\mathbf{p}. \quad (34)$$

So inside each of the slices with uniform permittivity along the y -axis the function \mathbf{u} can be represented as

$$\mathbf{u} = \sum_j \left(c_{1j} e^{\sqrt{\beta^2 - \eta_j^2} y} + c_{2j} e^{-\sqrt{\beta^2 - \eta_j^2} y} \right) \mathbf{p}_j \quad (35)$$

with eigenvalues η_j and corresponding eigenvectors \mathbf{p}_j from (34).

By matching the solutions of the each individual slab across the vertical interfaces using (22) and looking only for exponentially decaying solutions for $y \rightarrow \pm\infty$, one can obtain an eigenvalue problem

$$\mathbf{M}(\beta)\mathbf{c} = 0. \quad (36)$$

The vector \mathbf{c} consists of all unknown coefficients c_{1j} and c_{2j} from the representations of \mathbf{u} (35) on all individual slices. The matrix \mathbf{M} depends on β in a non-linear, even non-polynomial way. One of the strategies to tackle this is at first to specify a range of admissible values $\beta \in [I_1, I_2]$, where solutions β are sought. As we are looking only for propagating modes, with decaying field (35) at $y \rightarrow \pm\infty$, I_1 should be not smaller than the biggest eigenvalue η_j of (34) in the left-most and the right-most slabs. At the same time we require that there exists at least one oscillating function in at least one vertical slab. So I_2 should be smaller than the biggest eigenvalue η_j of (34) of all the constituting slabs, except the left- and the right-most ones. Once this interval is at hand, we scan through it looking for a β such that the matrix $\mathbf{M}(\beta)$ has at least one zero eigenvalue. Obviously, to find a non-trivial solution with certain accuracy requires some iterations. Moreover a large step size might lead to missing some roots while scanning the interval.

Once a nontrivial solution β , \mathbf{c} of (36) is at hand, \mathbf{u} can be reconstructed using (35). And then all field components can be obtained according to expressions (23) together with (14) or (15).

7 Relation with the Effective Index Method

In the following section we are going to show what happens if only a single, TE or TM, slab mode is taken into account in VEIM5 (14). Using the variational reasoning we will rigorously derive an analog to the Effective Index method.

7.1 TE polarization

Let us take only one TE slab mode with propagation constant β_r from a reference slice r with permittivity distribution $\varepsilon_r(x)$, and use it to represent the vectorial field profile of the complete waveguide as in eqn. (14). Due to the fact that $X^{E_x} \equiv 0$, according to (20) the unknown function Y^{H_x} satisfies the eqn.

$$A_{41}Y^{H_x} + \left(-i A_{42}(A_{21} + A_{22}A_{51}^{-1}A_{52})^{-1}A_{23}(Y^{H_x})' \right)' = \beta^2 \left(-i A_{42}(A_{21} + A_{22}A_{51}^{-1}A_{52})^{-1}A_{23} \right) Y^{H_x}. \quad (37)$$

After some manipulations, using the relations between the modal components χ^{E_y} , χ^{H_x} and χ^{H_z} of the slab mode the above relation can be rewritten as follows

$$\left(\frac{1}{\varepsilon_{\text{eff}}} (Y^{H_x})' \right)' + k^2 Y^{H_x} = \beta^2 \frac{1}{\varepsilon_{\text{eff}}} Y^{H_x} \quad (38)$$

with

$$\varepsilon_{\text{eff}}(y) = \frac{\beta_r^2}{k^2} + \frac{\langle \chi^{E_y}, (\varepsilon(x, y) - \varepsilon_r(y)) \chi^{E_y} \rangle}{\langle \chi^{E_y}, \chi^{E_y} \rangle}. \quad (39)$$

This looks exactly as a TM mode equation, similar to the standard Effective Index Method. In the reference slice one has $\varepsilon = \varepsilon_r$, and the effective permittivity ε_{eff} is equal to the squared effective index of the mode of the reference slice β_r^2/k^2 . In other slices this squared effective index is modified by the difference between the local permittivity and that of the reference slice, weighted by the local intensity of the fundamental component of the reference mode profile. Hence, on the contrary to the EIM, even in slices where no guided mode exist, the effective permittivity can still be rigorously defined.

Now it is instructive to see how the mode profile adjusts both in the reference slabs and elsewhere. Inside a slice with constant permittivity ε_{eff} , eqn. (38) permits solutions of the form

$$Y^{H_x} = c_+ e^{i\alpha y} + c_- e^{-i\alpha y} \quad (40)$$

for arbitrary constants c_+ and c_- and with

$$\beta^2 + \alpha^2 = k^2 \varepsilon_{\text{eff}}. \quad (41)$$

With the abbreviation $\rho^2 = k^2 \varepsilon_{\text{eff}}$ from (23) it follows that

$$\begin{aligned} Y^{H_z} &= \frac{\beta_r \beta}{\rho^2} (c_+ e^{i\alpha y} + c_- e^{-i\alpha y}), \\ Y^{E_y} &= Y^{H_z}, \\ Y^{E_z} &= \frac{\beta_r \alpha}{\rho^2} (c_+ e^{i\alpha y} - c_- e^{-i\alpha y}), \\ Y^{H_y} &= -Y^{E_z}. \end{aligned} \quad (42)$$

By introducing an angle θ such that $\cos \theta = \beta/\rho$, one can write

$$\begin{aligned} \begin{pmatrix} Y^{E_x}, Y^{E_y}, Y^{E_z} \\ Y^{H_x}, Y^{H_y}, Y^{H_z} \end{pmatrix}(y) &= c_+ \frac{\beta_r}{\rho} e^{i\rho \sin \theta y} \begin{pmatrix} 0, & \cos \theta, & \sin \theta \\ \rho/\beta_r, & -\sin \theta, & \cos \theta \end{pmatrix} + \\ &+ c_- \frac{\beta_r}{\rho} e^{-i\rho \sin \theta y} \begin{pmatrix} 0, & \cos \theta, & -\sin \theta \\ \rho/\beta_r, & \sin \theta, & \cos \theta \end{pmatrix}. \end{aligned} \quad (43)$$

If we use the principal square roots of α^2 and ρ^2 for α and ρ , and the principal inverse cosine for θ eq. 43 can be interpreted as follows. In the slice where the reference slab mode lives $\rho = \beta_r$, and we find that functions Y act as a rotation of the slab mode, such that the projection of the propagation constant of this mode onto the z -axis will match the global propagation constant β . In other slices, in addition to the rotation of the y and z components of the slab mode, the x component is scaled by ρ/β_r .

7.2 TM polarization

Analogously, the eqn. (20) can be rewritten for a single TM mode, with a field template as in (14). We now have $X^{H_x} \equiv 0$ in eqn. (20) and using the properties of the TM slab mode, the original equation for the unknown function Y^{E_x} ,

$$\begin{aligned} A_{11} Y^{E_x} + (-i A_{12} (A_{51} + A_{52} A_{21}^{-1} A_{22})^{-1} A_{53} (Y^{E_x})')' = \\ \beta^2 (-i A_{12} (A_{51} + A_{52} A_{21}^{-1} A_{22})^{-1} A_{53}) Y^{E_x}, \end{aligned} \quad (44)$$

can be rewritten as

$$\left(\frac{1}{\varepsilon_{1\text{eff}}} (Y^{E_x})' \right)' + k^2 \varepsilon_2 Y^{E_x} = \beta^2 \frac{1}{\varepsilon_{1\text{eff}}} Y^{E_x}, \quad (45)$$

with

$$\begin{aligned} \varepsilon_{1\text{eff}}(y) &= \frac{\beta_r^2}{k^2} \frac{\langle \chi^{E_z}, \varepsilon_r(x) \chi^{E_z} \rangle}{\langle \chi^{E_z}, \varepsilon(x, y) \chi^{E_z} \rangle} + \frac{\langle \chi^{H_y}, \chi^{H_y} \rangle}{\langle \chi^{H_y}, \frac{1}{\varepsilon_r(x)} \chi^{H_y} \rangle} \frac{\langle \chi^{E_z}, (\varepsilon(x, y) - \varepsilon_r(x)) \chi^{E_z} \rangle}{\langle \chi^{E_z}, \varepsilon(x, y) \chi^{E_z} \rangle}, \\ \varepsilon_2(y) &= \frac{\langle \chi^{E_x}, \varepsilon(x, y) \chi^{E_x} \rangle}{\langle \chi^{E_x}, \varepsilon_r(x) \chi^{E_x} \rangle}. \end{aligned} \quad (46)$$

This appears to be neither a standard TE nor a TM mode equation, but something in between, with the local refractive index distribution appearing both in the terms with and without derivative. In the reference slice with $\varepsilon = \varepsilon_r$, the effective permittivity $\varepsilon_{1\text{eff}}$ is equal to the squared effective index β_r^2/k^2 of the mode of the reference slice and $\varepsilon_2 = 1$. Contrary to the EIM, even in slices where no guided mode exists quantities that act like effective indices can still be rigorously defined.

What concerns the mode profile, in intervals along the y -axis with constant $\varepsilon_{1\text{eff}}$ and ε_2 , local solutions of eqn. (45) are of the form

$$Y^{E_x} = c_+ e^{i\alpha y} + c_- e^{-i\alpha y} \quad (47)$$

with

$$\beta^2 + \alpha^2 = k^2 \varepsilon_{1\text{eff}} \varepsilon_2. \quad (48)$$

Let us denote the right hand side of eqn. (41) as $\rho^2 = k^2 \varepsilon_{1\text{eff}} \varepsilon_2$ and $\varepsilon_3(y) = \frac{\langle \chi^{Ez}, \varepsilon(x, y) \chi^{Ez} \rangle}{\langle \chi^{Ez}, \varepsilon_r(x) \chi^{Ez} \rangle}$, then according to eqn. (23) one obtains

$$\begin{aligned} Y^{Hz} &= \frac{\beta_r \alpha \varepsilon_2}{\rho^2} (c_+ e^{i\alpha y} - c_- e^{-i\alpha y}), \\ Y^{Ey} &= -\frac{1}{\varepsilon_3} Y^{Hz}, \\ Y^{Hy} &= \frac{\beta_r \beta \varepsilon_2}{\rho^2} (c_+ e^{i\alpha y} + c_- e^{-i\alpha y}), \\ Y^{Ez} &= \frac{1}{\varepsilon_3} Y^{Hy}. \end{aligned} \quad (49)$$

By introducing an angle θ such that $\cos \theta = \beta/\rho$, one can write

$$\begin{aligned} \begin{pmatrix} Y^{Ex}, Y^{Ey}, Y^{Ez} \\ Y^{Hx}, Y^{Hy}, Y^{Hz} \end{pmatrix} (y) &= c_+ \frac{\beta_r \varepsilon_2}{\rho} e^{i\rho \sin \theta y} \begin{pmatrix} \rho/\beta_r \varepsilon_2, & -\varepsilon_3^{-1} \sin \theta, & \varepsilon_3^{-1} \cos \theta \\ 0, & \cos \theta, & \sin \theta \end{pmatrix} + \\ &+ c_- \frac{\beta_r \varepsilon_2}{\rho} e^{-i\rho \sin \theta y} \begin{pmatrix} \rho/\beta_r \varepsilon_2, & \varepsilon_3^{-1} \sin \theta, & \varepsilon_3^{-1} \cos \theta \\ 0, & \cos \theta, & -\sin \theta \end{pmatrix}. \end{aligned} \quad (50)$$

In the reference slice $\rho = \beta_r$ and we find that the functions Y also in this case act as a rotation of the slab mode. In all other slices, while the y - and z -components of the magnetic field are just rotated by the angle θ , the electric y and z components are not only rotated, but also scaled by ε_3^{-1} . In addition to this the x -component is scaled by $\rho/\beta_r \varepsilon_2$.

8 Relation with the Film Mode Matching Method

As we could see in the previous section if only one, TE or TM, slab mode is used to expand the total field profile using the 5 component expansion (14), the variational procedure leads to functions Y that act as a rotation. Then the field representation inside the slice where the slab mode lives replicates the field ansatz of the FMM (cf. section 4). In the following we look at the the case when multiple TE and TM slab modes appear in the 5 component expansion (14).

Let us rewrite the second, third, fifth and sixth equations of (16) as

$$\begin{aligned} \mathbf{I}_{\text{TE}}(\mathbf{Y}_{\text{TE}}^{Ey} - \mathbf{Y}_{\text{TE}}^{Hz}) + \mathbf{I}_{\text{TM}}(\mathbf{Y}_{\text{TM}}^{Ey} + \mathbf{Y}_{\text{TM}}^{Hz}) &= \mathbf{A}_{21}^{-1} \mathbf{A}_{23} (\beta \mathbf{Y}^{Hx} - \mathbf{G}_{\text{TE}} \mathbf{Y}_{\text{TE}}^{Hz}) \\ -\mathbf{I}_{\text{TE}}(\mathbf{Y}_{\text{TE}}^{Ey} - \mathbf{Y}_{\text{TE}}^{Hz}) + \mathbf{I}_{\text{TM}}(\mathbf{Y}_{\text{TM}}^{Ey} + \mathbf{Y}_{\text{TM}}^{Hz}) &= \mathbf{A}_{51}^{-1} \mathbf{A}_{53} (\mathbf{G}_{\text{TM}} \mathbf{Y}_{\text{TM}}^{Ey} - i(\mathbf{Y}^{Ex})') \\ \mathbf{I}_{\text{TE}}(\mathbf{Y}_{\text{TE}}^{Ez} + \mathbf{Y}_{\text{TE}}^{Hy}) + \mathbf{I}_{\text{TM}}(\mathbf{Y}_{\text{TM}}^{Ez} - \mathbf{Y}_{\text{TM}}^{Hy}) &= \mathbf{A}_{21}^{-1} \mathbf{A}_{23} (\mathbf{G}_{\text{TE}} \mathbf{Y}_{\text{TE}}^{Hy} - i(\mathbf{Y}^{Hx})') \\ \mathbf{I}_{\text{TE}}(\mathbf{Y}_{\text{TE}}^{Ez} + \mathbf{Y}_{\text{TE}}^{Hy}) - \mathbf{I}_{\text{TM}}(\mathbf{Y}_{\text{TM}}^{Ez} - \mathbf{Y}_{\text{TM}}^{Hy}) &= \mathbf{A}_{51}^{-1} \mathbf{A}_{53} (\beta \mathbf{Y}^{Ex} - \mathbf{G}_{\text{TM}} \mathbf{Y}_{\text{TM}}^{Ez}) \end{aligned} \quad (51)$$

with functions \mathbf{Y}_b^a corresponding to a vector of all the functions Y related to modal component of polarization b , used to expand component a of the total field. \mathbf{G}_b is a diagonal matrix with propagation constants $\beta_{r,j}$ of the slab modes of polarization b sitting on the diagonal. Matrices

$$\mathbf{I}_{\text{TE}} = \begin{pmatrix} \mathbf{1}_{n_{\text{TE}} \times n_{\text{TE}}} \\ \mathbf{0}_{n_{\text{TM}} \times n_{\text{TE}}} \end{pmatrix}, \quad \mathbf{I}_{\text{TM}} = \begin{pmatrix} \mathbf{0}_{n_{\text{TE}} \times n_{\text{TM}}} \\ \mathbf{1}_{n_{\text{TM}} \times n_{\text{TM}}} \end{pmatrix}, \quad (52)$$

have been introduced to increase the readability of the equations. Here $\mathbf{1}_d$ and $\mathbf{0}_d$ denote correspondingly the unity- and zero-matrix of a dimension d , and symbols n_{TE} and n_{TM} denote the number of slab modes of respectively TE and TM polarization included in the expansion (14).

Obviously, functions \mathbf{Y} that satisfy

$$\begin{aligned} \mathbf{Y}_{\text{TE}}^{Ey} &= \mathbf{Y}_{\text{TE}}^{Hz}, & \mathbf{Y}_{\text{TM}}^{Ey} &= -\mathbf{Y}_{\text{TM}}^{Hz}, \\ \beta \mathbf{Y}^{Hx} &= \mathbf{G}_{\text{TE}} \mathbf{Y}_{\text{TE}}^{Hz}, & \mathbf{G}_{\text{TM}} \mathbf{Y}_{\text{TM}}^{Ey} &= i(\mathbf{Y}^{Ex})', \\ \mathbf{Y}_{\text{TE}}^{Ez} &= -\mathbf{Y}_{\text{TE}}^{Hy}, & \mathbf{Y}_{\text{TM}}^{Ez} &= \mathbf{Y}_{\text{TM}}^{Hy}, \\ \mathbf{G}_{\text{TE}} \mathbf{Y}_{\text{TE}}^{Hy} &= i(\mathbf{Y}^{Hx})', & \beta \mathbf{Y}^{Ex} &= \mathbf{G}_{\text{TM}} \mathbf{Y}_{\text{TM}}^{Ez} \end{aligned} \quad (53)$$

are solutions of (51). Using these relations together with the first and the fourth equations of (16) result in

$$\begin{aligned} (\mathbf{Y}^{E_x})'' + (\mathbf{G}_{\text{TM}})^2 \mathbf{Y}^{E_x} &= \beta^2 \mathbf{Y}^{E_x} \\ (\mathbf{Y}^{H_x})'' + (\mathbf{G}_{\text{TE}})^2 \mathbf{Y}^{H_x} &= \beta^2 \mathbf{Y}^{H_x}. \end{aligned} \quad (54)$$

According to eqns. (51) and (54) all the functions Y decouple inside the slice where the slab modes belongs to. So we can solve these equations for all the components of \mathbf{Y}^{E_x} and \mathbf{Y}^{H_x} separately. Solutions of (54) have the form

$$Y_j^{H_x} = c_{+,j} e^{i\alpha_j y} + c_{-,j} e^{-i\alpha_j y} \quad (55)$$

with

$$\beta^2 + \alpha_j^2 = \beta_{r,j}^2. \quad (56)$$

Other components can be derived from (51) as

$$\begin{aligned} \begin{pmatrix} Y_j^{E_x}, Y_j^{E_y}, Y_j^{E_z} \\ Y_j^{H_x}, Y_j^{H_y}, Y_j^{H_z} \end{pmatrix} (y) &= c_{+,j} e^{i\beta_{r,j} \sin \theta_j y} \begin{pmatrix} 0, & \cos \theta_j, & \sin \theta_j \\ 1, & -\sin \theta_j, & \cos \theta_j \end{pmatrix} + \\ &+ c_{-,j} e^{-i\beta_{r,j} \sin \theta_j y} \begin{pmatrix} 0, & \cos \theta_j, & -\sin \theta_j \\ 1, & \sin \theta_j, & \cos \theta_j \end{pmatrix}, \end{aligned} \quad (57)$$

where $\cos \theta_j = \beta / \beta_{r,j}$.

Hence the functions Y_j corresponding to TE slab mode number j rotate the original slab mode around the x -axis such that the projection of its propagation constant $\beta_{r,j}$ onto the direction of propagation z will be precisely the propagation constant β of the mode of the complete waveguide structure. The same is true for TM slab modes.

We showed that the field ansatz of rotated slab modes, as used locally in the Film Mode Matching method [9, 8] can be found also by the present approach where it appears to be optimal. While in itself it might seem rather pointless to reinvent the method, the idea behind the present technique might be used in deriving some sort of analogue of the FMM for full 3D scattering problems, in which the structure varies in all 3 directions [14].

9 Numerical results

We will illustrate the method with four examples. The first two deal with waveguides with piecewise constant rectangular refractive index distribution. The third example is a waveguide with slanted sidewall and the fourth is an indiffused waveguide. We will use the acronym VEIM (variational effective index method) for results of the technique as introduced in sections 2 - 8.

9.1 Box-shaped waveguide

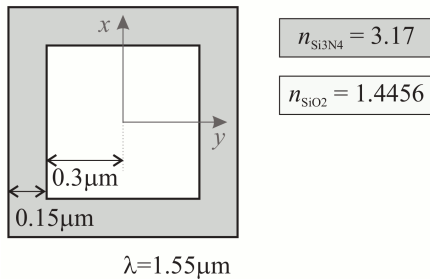


Figure 4: Structure of the Box-Shaped Waveguide. The vertical extents of the computational window range from $-2.5\mu\text{m}$ to $2.5\mu\text{m}$.

Consider the box-shaped waveguide of Figure 4, originating from [7]. It can be divided into five vertical slices with three distinct cross-sections (Figure 5, left). We take slab modes from the side walls of the box (Figure 5, middle) to approximate the modal field in the entire cross-section (Figure 5, right). The waveguide will be analyzed with both 3 (Equation (15)) and 5 (Equation (14))

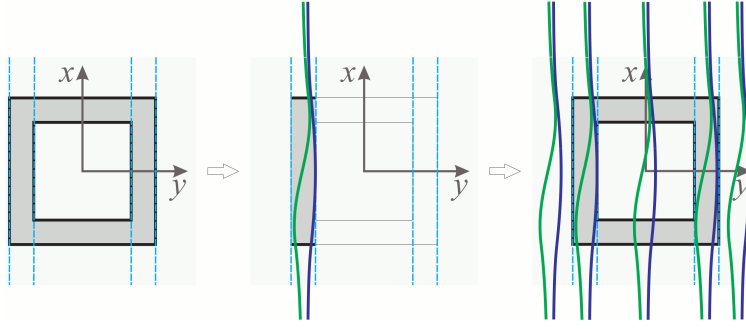


Figure 5: Subdivision of the waveguide into slices. Slab modes of the side walls are used to approximate the field of the mode everywhere.

) component approximations, denoted by $\text{VEIM3}_{a,b}$ and $\text{VEIM5}_{a,b}$, where a and b are the number of TE and TM slab modes taken into account.

In Figure 6 $\text{VEIM5}_{1,0}$ approximation of the vectorial mode profile of the fundamental TE-like mode is shown. In this case χ^{E_y} is multiplied by Y^{E_y} and Y^{E_z} to get E_y and E_z respectively; χ^{H_x} is multiplied by Y^{H_x} to get H_x ; and χ^{H_z} is multiplied by Y^{H_y} and Y^{H_z} to get H_y and H_z respectively. The figure contains plots of all contributing functions. Consistent with the observation in sec. 7.1, we see that $Y^{E_y} = Y^{H_z}$ and $Y^{E_z} = -Y^{H_y}$. Note that, contrary to the EIM, the field profile can still be visualized even when no local guided slab mode exists.

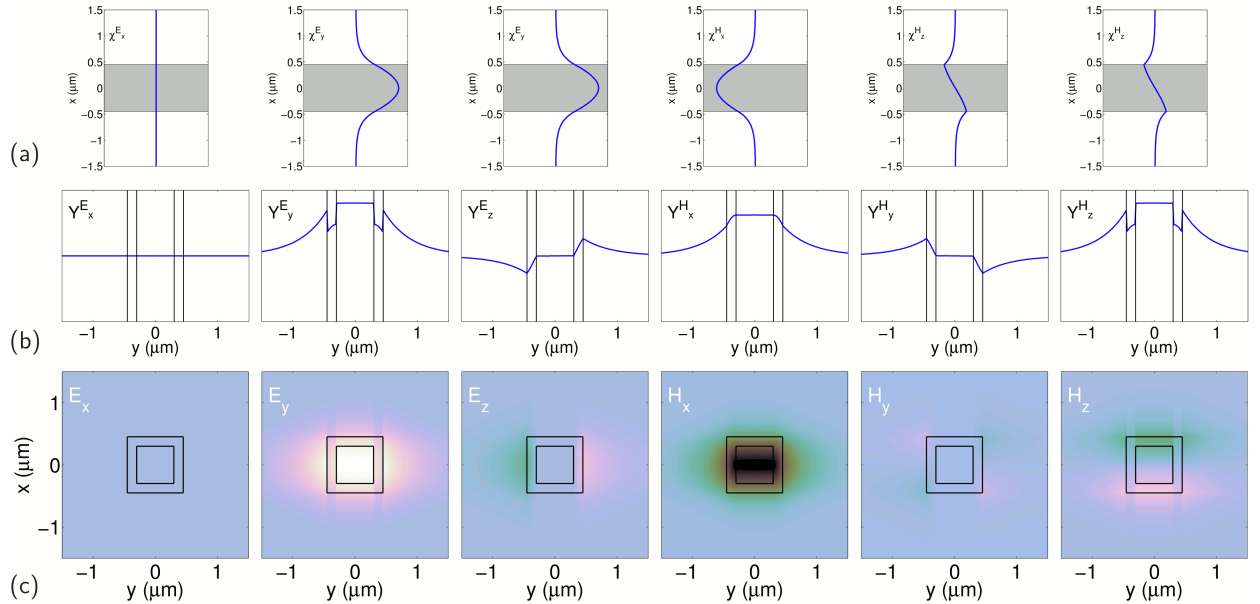


Figure 6: Square waveguide: (a) Functions χ in expansion $\text{VEIM5}_{1,0}$; (b) Functions Y in expansion $\text{VEIM5}_{1,0}$; (c) Vectorial field profile $\text{VEIM5}_{1,0}$.

Next, Figure 7 gives an impression of the "converged" field profile obtained using $\text{VEIM3}_{30,30}$. The slab mode basis has been discretized by Dirichlet boundary conditions on the boundaries of the vertical computational window as given in Figure 4. Comparison with Figure 6 shows that even with a single mode in the representation, the main features of the true field profile are already visible. So the present method with one mode in the expansion can very well serve as a quick tool for qualitative analysis of the waveguide structures, while also being able to quantitatively analyze the waveguide by using more modes in the expansion.

Figure 8 shows the propagation constant of the fundamental modes of the waveguide versus the number m of TE and TM modes in the expansion $\text{VEIM3}_{m,m}$ for both the present method and a commercial FMM solver [15]. Both methods converge to the same value with comparable convergence speed.

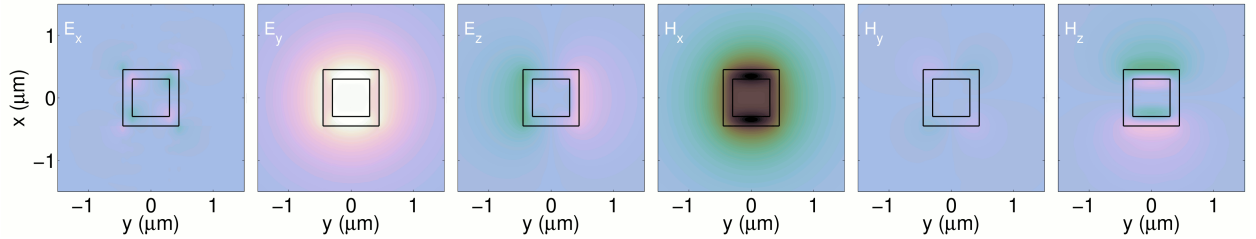


Figure 7: "Converged" (VEIM_{330,30}) vectorial field profiles of the fundamental TE-like mode.

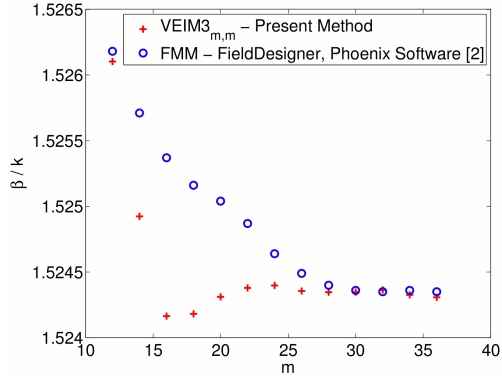


Figure 8: Convergence of the effective index of the fundamental TE-like mode of the box-shaped waveguide Figure 4.

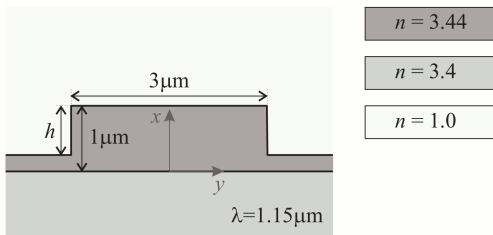


Figure 9: Structure of the Rib Waveguide. Vertical extents of the computational window are $[-2, 2]\mu m$.

9.2 Rectangular rib waveguide

In this section we consider the rib structure from Figure 9, which is used as a benchmark waveguide in [2, 16, 17, 4]. The structure supports a fundamental TE and TM mode for all etch depths h in the range we look in, which is $[0.2, 1]$. The modes are strongly polarized, and thus it may be expected that an expansion using only TE or only TM modes (similar to a semi-vectorial calculation) will give good results.

At etch depths greater than $0.5\mu m$ guided modes do not exist outside the central slice, so the EIM fails to uniquely determine the effective refractive index of those regions. We analyze this structure with both 3- (15) and 5-component (14) approximations. In the following figures we will refer to them as VEIM_{3 a,b,c,d} and VEIM_{5 a,b,c,d} correspondingly. The subscript letters stand for number of slab modes used in the current approximation: a – number of TE modes from the central slice, b – number of TE modes from the outer slice, c – number of TM modes from the central slice, d – number of TM modes from the outer slice.

The slab modes are calculated using Dirichlet boundary conditions on the upper and lower computational domain boundaries. Because of this, the outer slice mode is still defined when the guided mode of that slice goes below cut-off.

Figure 10 and Figure 11 show plots of the TE and TM effective indices correspondingly using these different expansions versus the etch depth. The figures also show the corresponding EIM results, and, as reference, FMM results obtained by the commercial mode solver [15].

Comparing the results of our method with only one TE (VEIM_{5 $1,0,0,0$}) or TM (VEIM_{5 $0,0,1,0$}) mode of the central slice in the expansion (14) to the EIM results, shows that for larger etch depths, our results are much closer to the reference results - especially after the outer slice has gone below cut-off and the EIM uses the substrate refractive index as (constant) outer slice effective index.

Adding one outer slice mode to the VEIM expansion greatly improves its accuracy, especially

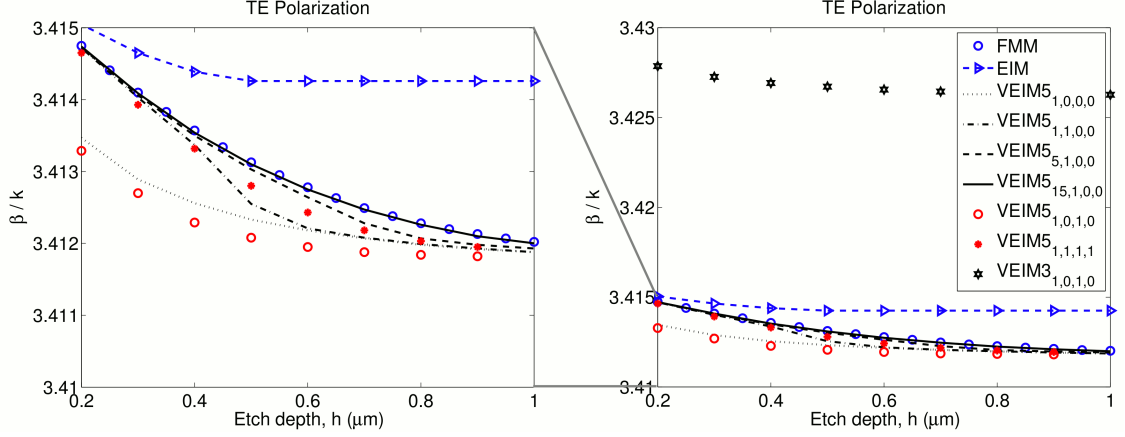


Figure 10: Convergence of the effective index (β/k) of the fundamental TE mode of the rib waveguide.

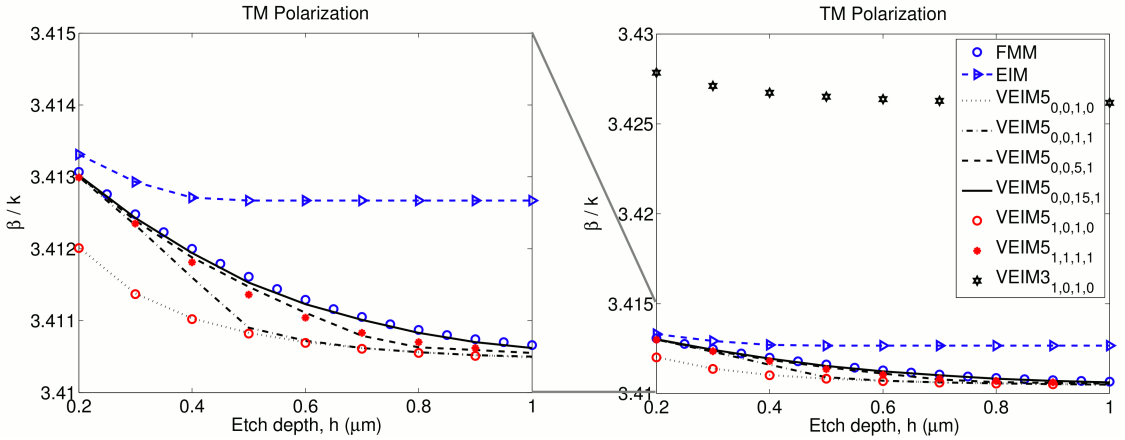


Figure 11: Convergence of the effective index (β/k) of the fundamental TM mode of the rib waveguide.

if it is a guided slab mode; the $\text{VEIM5}_{1,1,0,0}$ curves are much closer to the reference results than the $\text{VEIM5}_{1,0,0,0}$ curves, especially at etch depths below $0.5\mu\text{m}$.

Taking five inner and one outer slice mode $\text{VEIM5}_{5,1,0,0}$ moves the results closer to the reference curve, while fifteen inner and one outer slice modes $\text{VEIM5}_{15,1,0,0}$ yield results that almost coincide with the reference. Note that these results use only TE or only TM modes in the 5-component expansion (14), i.e. the resulting fields are semi-vectorial; apparently a semivectorial approximation is sufficient for an accurate estimation of the effective indices of this structure.

The present method when using just one central slice TE and TM mode simultaneously with the three-component-per-mode approximation $\text{VEIM3}_{1,0,1,0}$ (15) yields results that are quite far from the reference data. Moving to the five-component-per-mode approximation $\text{VEIM5}_{1,0,1,0}$ (14), on the other hand, gives much better results. Moreover, adding outer slice TE and TM modes $\text{VEIM5}_{1,1,1,1}$ greatly improves the estimation of propagation constant for both, TE and TM, polarizations.

9.3 Waveguide with non-rectangular piecewise constant cross-section

The waveguide cross-section of Figure 12 is part of a polarization rotator in InP/InGaAsP, proposed in [18]. Due to its slanted sidewall, the modes of this structure are highly hybrid.

Because of the slanted sidewall, the finite element scheme is more suitable to calculate the modes of this structure; the semi-analytical method requires a rather large number of slices, while the finite elements automatically take the slant into account.

Figure 13 shows the convergence of the effective index of the fundamental mode of the waveguide versus the number of modes in the 3-component expansion $\text{VEIM3}_{a,b}$ (15), with a and b being

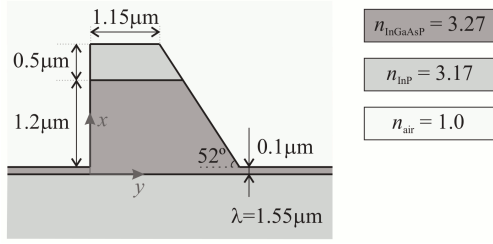


Figure 12: Structure of polarization converter from [18]. The computational window in the calculations is $(x, y) \in [-2, 2.5] \times [-2, 3.5] \mu m^2$; 50 elements are used in the finite element scheme.

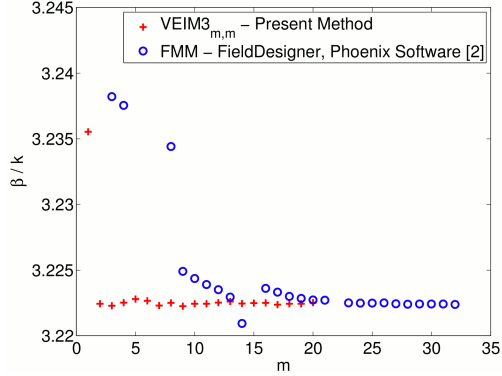


Figure 13: Convergence of the effective index of the fundamental mode of the polarization converter.

numbers of TE and TM slab modes from the central ($y \in (0, 1.15) \mu m$) slab. It also shows the convergence of the commercial FMM mode solver [15], in which the structure is subdivided into 50 slices. Remarkably, starting from just 2 TE and TM modes in the 3-component expansion (15) VEIM_{3,2,2}, the effective index is stable and close to the converged value of the FMM solver; 320 modes in the FMM solver lead to an effective index of 3.2225, while with just 7 TE and TM modes the current method predicts already an effective index of 3.2223. The field profiles also converge rapidly; Figure 14 shows the vectorial fields for (a) one (VEIM_{3,1,1}), (b) two (VEIM_{3,2,2}), and (c) seven (VEIM_{3,7,7}) TE and TM modes in the expansion (15).

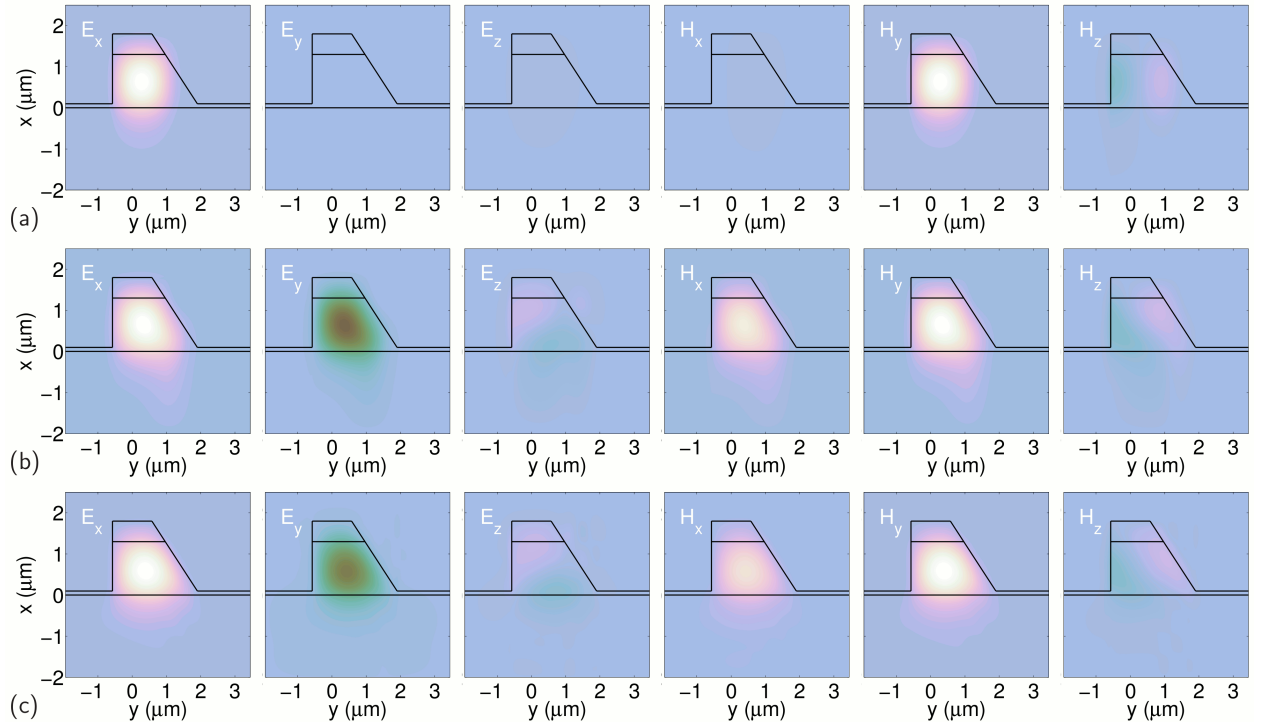


Figure 14: Vectorial field profile of the fundamental mode of the polarization converter. (a) VEIM_{3,1,1}, (b) VEIM_{3,2,2}, (c) VEIM_{3,7,7}.

9.4 Indiffused waveguide

To show the flexibility of the present method we apply it to a diffused waveguide [19] with a refractive index distribution given by

$$n^2(x, y) = \begin{cases} n_s^2 + n_s^2(1.05^2 - 1) \exp(-x^2/16) \exp(-y^2/4), & \text{if } x > 0; \\ n_c^2, & \text{if } x < 0, \end{cases} \quad (58)$$

with $n_s^2 = 2.1$, $n_c^2 = 1.0$ and $\lambda = 1.3\mu\text{m}$. Similar to the slanted sidewall waveguide described above, the finite element implementation of the presented method is the more suitable, since it takes into account the nonuniform distribution in the y -direction of the refractive index automatically. Vertically, the structure is subdivided into 7 layers; horizontally, 20 finite elements are used. The computational window used in the calculations is defined as $(x, y) \in [-1, 8] \times [-6, 6]\mu\text{m}^2$.

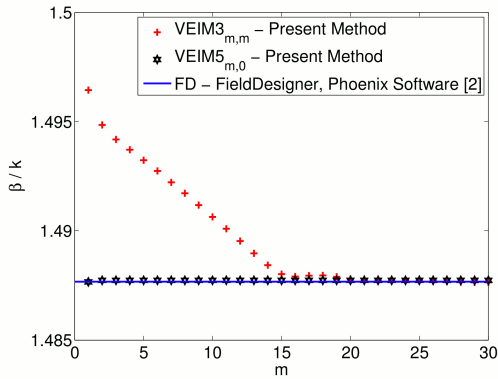


Figure 15: Convergence of the effective index of the fundamental mode of the diffused waveguide.

Figure 15 shows the convergence of the effective index of the fundamental mode of the indiffused waveguide versus the number of modes in the 3- (VEIM3_{a,b}) and 5-component (VEIM5_{a,b}) approximations, with a and b being numbers of TE and TM slab modes of the central ($y = 0\mu\text{m}$) slab. The results are compared to the rigorous Finite Difference simulation (with 129×129 grid points) [15]. Since the fundamental mode is strongly polarized, the semi-vectorial approximation appears to converge much faster.

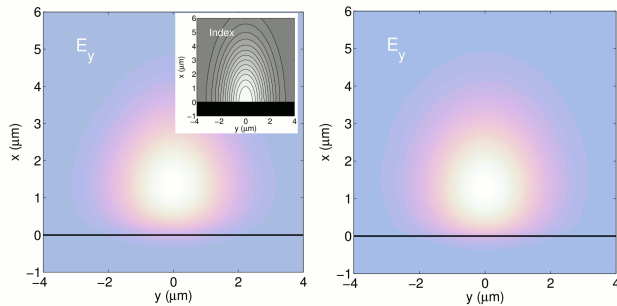


Figure 16: Field profiles of the dominant electric component E_y of the fundamental TE mode: left – VEIM3_{1,1}, right – VEIM3_{15,15}.

On Figure 16 field profiles of the dominant electric component E_y of the fundamental TE mode are shown. The effective index $N_{\text{eff}} = \beta\lambda/2\pi$ of the fundamental mode on the left picture is 1.4965 and on the right – 1.48802, which compares well with the Finite Difference simulation – 1.48797.

10 Concluding remarks

A variational method for the fully vectorial analysis of arbitrary isotropic dielectric waveguides was developed. Similar to the scalar approach [4] this method gives rather accurate estimates of the propagation constants, sometimes even with only a few terms in the expansion.

When applying the present method with only one slab mode in the expansion of the modal field of the complete waveguide, this mode is transformed in all different slices to fit the true solution there the best. Together with the shape transformation, the effective index of this mode is uniquely

transformed. Additionally, the expression for the transformed propagation constant is quite simple and is certainly not more complicated, than the calculation of a slab mode. In this way the present procedure turns out to be a simple and still a more rigorous way to obtain a first intuitive guess for the propagation constant and field profile, than the standard Effective Index Method.

It turns out that in case a TE mode is used in the expansion, the reduced equation appears to be a TM mode equation. At the same time when a TM mode is used, the reduced equation appears to be neither TE, nor TM mode equation, but something in between, with the effective refractive indices appearing both under the derivative sign and in the right part of the equation.

While in the Film Mode Matching method, rotated modes of each slice are used to locally expand the field, VEIM uses only one set of modes everywhere. We showed that in the reference slice, where the 1D modes are calculated, VEIM predicts exactly the same rotations as the Film Mode Matching method uses. In the reference slice the total field profile is a superposition of these rotated 1D TE and TM modes; in other slices, however, the components of all the 1D modes mix.

Of course the question remains - would some other combination of slab mode components lead to faster convergence? For example, one could imagine that in a certain case a superposition of e.g. explicitly selected profiles of specific slices would lead to similar results as if one would take functions related to more, let's say, five but consecutive modes - from the fundamental to the fourth order. However, adding field profiles from different slices may lead to a (near) linear dependency of functions X , and result in non-unique functions Y . Obviously, the safe choice is to use in the approximation of a component of the total field only profiles from a single slice. Nevertheless, when only a few modal components are used, it may, as our calculations show, be beneficial to use one or two modes from other slice(s).

Similar ideas can be applied to optical scattering problems in 2D and 3D. A preliminary account of corresponding simulations has been given in [20], [14].

Acknowledgements

This work was supported by the Dutch Technology Foundation (BSIK / NanoNed project TOE.7143)

References

- [1] K. S. Chiang. Review of numerical and approximate methods for the modal analysis of general optical dielectric waveguides. *Optical and Quantum Electronics*, 26:S113–S134, 1994.
- [2] C. Vassallo. 1993-1995 Optical mode solvers. *Optical and Quantum Electronics*, 29:95–114, 1997.
- [3] R. Scarmozzino, A. Gopinath, R. Pregla, and S. Helfert. Numerical techniques for modeling guided-wave photonic devices. *IEEE Journal of Selected Topics in Quantum Electronics*, 6(1):150–162, 2000.
- [4] O. V. Ivanova, M. Hammer, R. Stoffer, and E. van Groesen. A variational mode expansion mode solver. *Optical and Quantum Electronics*, 39(10–11):849–864, 2007.
- [5] C. Vassallo. *Optical Waveguide Concepts*. Elsevier, Amsterdam, 1991.
- [6] E. W. C. van Groesen and J. Molenaar. *Continuum Modeling in the Physical Sciences*. SIAM publishers, Philadelphia, USA, 2007.
- [7] F. Morichetti, A. Melloni, M. Martinelli, R. G. Heidemann, A. Leinse, D. H. Geuzebroek, and A. Borreman. Box-shaped dielectric waveguides: A new concept in integrated optics? *Journal of Lightwave Technology*, 25(9):2579–2589, 2007.
- [8] P. Bienstman. Two-stage mode finder for waveguides with a 2D cross-section. *Optical and Quantum Electronics*, 36(1–3):5–14, 2004.

- [9] A. S. Sudbø. Film mode matching: a versatile numerical method for vector mode fields calculations in dielectric waveguides. *Pure and Applied Optics*, 2:211–233, 1993.
- [10] COMSOL. www.comsol.com .
- [11] J. van Kan, G. Segal, and F. Vermolen. *Numerical Methods in Scientific Computing*. Delft Academic Press, VSSD, Delft, The Netherlands, 2005.
- [12] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Review*, 43(2):235–286, 2001.
- [13] LAPACK. www.netlib.org/lapack/ .
- [14] O. V. Ivanova, R. Stoffer, L. Kauppinen, and M. Hammer. Variational effective index method for 3D vectorial scattering problems in photonics: TE polarization. In *Progress In Electromagnetics Research Symposium PIERS 2009, Moscow, Proceedings*, pages 1038–1042, 2009.
- [15] Phoenix Software, P.O. Box 545, 7500 AM Enschede, The Netherlands; <http://www.phoenixbv.com/> .
- [16] G. R. Hadley and R. E. Smith. Full-vector waveguide modeling using an iterative finite-difference method with transparent boundary conditions. *Journal of Lightwave Technology*, 13(3):465–469, 1995.
- [17] M. Lohmeyer. *Guided waves in rectangular integrated magneto-optic devices*. Cuvillier Verlag, Göttingen, 1999. Dissertation, Universität Osnabrück.
- [18] S.S.A. Obayya, N. Somasiri, B.M.A. Rahman, and K.T.V. Grattan. Full vectorial finite element modeling of novel polarization rotators. *Optical and Quantum Electronics*, 35(4–5):297–312, 2003.
- [19] A. Sharma and P. Bindal. An accurate variational analysis of single-mode diffused channel waveguides. *Optical and Quantum Electronics*, 24(12):1359–1371, 1992.
- [20] O. V. Ivanova, R. Stoffer, and M. Hammer. A dimensionality reduction technique for scattering problems in photonics. 1st International Workshop on Theoretical and Computational Nano-Photonics TaCoNa-Photonics, Conference Proceedings, 47 (2008).