

Extension of Marcatili's analytical approach for rectangular silicon optical waveguides

Wouter J. Westerveld, Suzanne M. Leinders, Koen W.A. van Dongen, H. Paul Urbach, and Mirvais Yousefi

Abstract—Marcatili's famous approximate analytical description of light propagating through rectangular dielectric waveguides, published in 1969, gives accurate results for low-index-contrast waveguides. However, photonic integrated circuit technology has advanced to high-index-contrast (HIC) waveguides. In this paper, we improve Marcatili's model by adjusting the amplitudes of the components of the electromagnetic fields in his description. We find that Marcatili's eigenvalue equation for the propagation constant is also valid for HIC waveguides. Our improved method shows much better agreement with rigorous numerical simulations, in particular for the case of HIC waveguides. We also derive explicit expressions for the effective group index and the effects of external forces on the propagation constant. Furthermore, with our method the phenomenon of avoided crossing of modes is observed and studied.

Index Terms—Optical waveguides, Electromagnetic propagation, Electromagnetic fields, Integrated optics, Silicon on insulator technology, Optical sensors.

I. INTRODUCTION

The propagation of light through rectangular dielectric optical waveguides cannot be described in closed analytical form. Marcatili's famous approximate analytical approach [1] has been used since the 1970's and is treated in many textbooks on optical waveguides theory [2]–[5]. His method is, however, derived for waveguides with a low refractive index contrast, while nowadays technology has advanced to high-index-contrast (HIC) waveguides. Silicon-on-insulator (SOI) technology has, for example, become one of the focus platforms for integrated optics over the last decade. The large refractive index contrast of the materials allows for small device footprint. High-yield mass fabrication is provided using CMOS processes from the electronics industry, that have been tailored to photonic applications [6]. Behavior of integrated optical components, such as ring resonator filters or arrayed waveguide grating (AWG) based multiplexers depend critically on the exact knowledge of the propagating modes in the

waveguide [7], [8]. Although numerical solutions such as the circular harmonics method, the film mode matching (FMM) method, the variational mode expansion method (VMEM), or the finite element method (FEM) are available [9]–[12], we believe that an analytical model is useful in order to gain insight in the physics of the devices, and also for fast explorative simulations of photonic integrated circuitry [13].

In this paper, we extend the range of waveguides for which Marcatili's approximate approach can be applied, in particular to high-index-contrast waveguides. Similar to Marcatili, we use an *ansatz* for the form of the modal fields that is based on separation of variables in the waveguide core. The large index contrast causes, in Marcatili's original approach, a severe mismatch of the electromagnetic fields just inside and just outside the core of the waveguide. We show that Marcatili's eigenvalue equation for the propagation speed of the light through waveguides is, in fact, more general and we obtain improved modal electromagnetic fields for the same eigenvalue equation, which have a much lower mismatch. An analytical description is presented, and is compared with the fundamental mismatch of this *ansatz*, which is found by means of an optimization algorithm.

Next to this, explicit equations are derived for the effective group index and for the linearized influence of external effects on the effective index of the modes. As an example, we analytically calculate the influence of temperature on the effective index of the modes in the waveguide. Also, results are presented on photonic evanescent field sensors, where the refractive index of the medium in the vicinity of the waveguide is probed with the evanescent tail of the waveguide mode [14], [15].

Throughout this paper, we test the analysis with the first three modes in a typical SOI waveguide with a guiding layer height of 300 nm, with infrared light that has a free-space wavelength $\lambda_0 = 1550$ nm. These guides consist of a thin monocrystalline silicon layer ($n = 3.476$) on top of a thick silicon dioxide (BOX) layer ($n = 1.444$) [16]. The influence of the silicon substrate below the BOX layer is neglected.

In the next section, we present our extension of Marcatili's approach. In Sec. III, we apply the eigenvalue equation for the effective index and derive explicit equations for the effective group index and for the effects of external forces. In Sec. IV, we compare our theory with rigorous simulations of typical SOI waveguides. Section V concludes this paper.

This article appeared in IEEE/OSA Journal of Lightwave Technology, vol. 30, no. 14, July 15, 2012. This work was supported by TNO, Delft, The Netherlands, and the IOP Photonic Devices Program of NL Agency.

W. J. Westerveld is with the Optics Research Group, Faculty of Applied Sciences, Delft University of Technology, 2628CJ Delft, The Netherlands, and also with TNO, 2628 CK Delft, The Netherlands.

S.M. Leinders and K. W. A. van Dongen are with the Laboratory of Acoustical Wavefield Imaging, Faculty of Applied Sciences, Delft University of Technology, 2628 CJ Delft, The Netherlands.

H. P. Urbach is with the Optics Research Group, Faculty of Applied Sciences, Delft University of Technology, 2628 CJ Delft, The Netherlands (e-mail: h.p.urbach@tudelft.nl).

M. Yousefi is with Photonic Sensing Solutions, 1013 EN Amsterdam. Digital Object Identifier 10.1109/JLT.2012.2199464

II. THEORY

In this section, we describe an approximate model of how light travels through rectangular high-index-contrast (HIC) dielectric waveguides. In Sec. II-A, we make an *ansatz* for modes in a rectangular waveguide, and briefly investigate the implications of this *ansatz* at the interfaces of the waveguide core. In Sec. II-B, we describe Marcatili's choice of parameters for the *ansatz* and Sec. II-C presents our improved method. In Sec. II-D, we present a quantification of the error that arises from the discontinuities of the electromagnetic field at the interfaces, and we propose an optimization method based on this quantification. In Sec. II-E, we use this quantification to discuss the different methods.

A. Rectangular waveguides

In regular lossless dielectric waveguides, the core of the waveguide has a higher refractive index (n_1) than the surrounding media (n_2 - n_5), as depicted in Fig 1. The refractive indices for the outer quadrants, i.e. in the corner regions, are not specified because we neglect these regions. We consider a monochromatic wave with angular frequency ω , propagating in the waveguide direction (z) with a propagation constant β . For a two-dimensional refractive index profile $n(x, y)$, solutions of Maxwell's equations can be found in the form

$$\mathcal{E}(x, y, z, t) = \text{Re} \{ \mathbf{E}(x, y) \exp[i(\omega t - \beta z)] \}, \quad (1)$$

and a similar description of \mathcal{H} , with \mathcal{E} and \mathcal{H} being the electric and magnetic field, respectively. The free-space propagation constant is $k_0 = \omega/c$, where c is the speed of light in vacuum. When the propagation constant β is larger than the propagation constant that is allowed in the regions outside the waveguide core, due to spectral cut-off (i.e. $\beta > k_0 n_j$, with $j = 2, \dots, 5$), the light is confined in the core of the waveguide. The lateral confinement of such guided waves dictates the light to exist in the form of certain modes, or "standing waves".

Using Maxwell's equations, it is now possible to describe the full electromagnetic fields in terms of the longitudinal field components [2], i.e. in region j

$$E_x = \frac{-\iota}{K_j^2} \left(\beta \frac{\partial E_z}{\partial x} + \omega \mu_0 \frac{\partial H_z}{\partial y} \right), \quad (2)$$

$$E_y = \frac{-\iota}{K_j^2} \left(\beta \frac{\partial E_z}{\partial y} - \omega \mu_0 \frac{\partial H_z}{\partial x} \right), \quad (3)$$

$$H_x = \frac{-\iota}{K_j^2} \left(\beta \frac{\partial H_z}{\partial x} - \omega \epsilon_0 n_j^2 \frac{\partial E_z}{\partial y} \right), \quad (4)$$

$$H_y = \frac{-\iota}{K_j^2} \left(\beta \frac{\partial H_z}{\partial y} + \omega \epsilon_0 n_j^2 \frac{\partial E_z}{\partial x} \right), \quad (5)$$

where μ_0 and ϵ_0 are the permeability and the permittivity of vacuum, and K_j is defined by

$$K_j = \sqrt{n_j^2 k_0^2 - \beta^2}. \quad (6)$$

All components, and hence in particular the longitudinal components E_z and H_z , satisfy the reduced wave equation (here given for E_z) [2]:

$$\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_z}{\partial y^2} + K_j^2 E_z = 0. \quad (7)$$

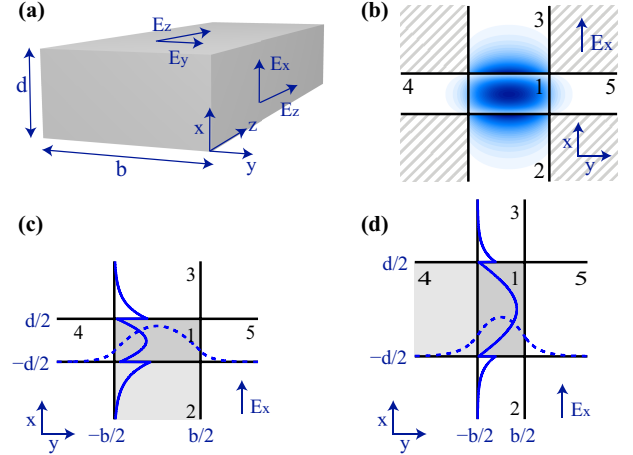


Fig. 1. Schematics of a rectangular waveguide with a width of 600 nm and a height of 300 nm, used to guide light with a free-space wavelength of 1550 nm. (a) 3-D sketch of the waveguide, also displaying the tangential electric field components. (b) Cross-section of the waveguide at $z = 0$. Regions 1 to 5 are indicated. The shaded corner regions are neglected in this analysis. The color plot represents E_x , the dominant electric field component, of the fundamental TM-like mode. (c) Sketch of the cross-section of the waveguide with the coordinate frame for a typical TM-like mode. Shape of E_x in the x direction (solid line) and in the y -direction (dashed line). Refractive indices $n_1 = 3.476$, $n_2 = 1.444$, $n_3 = n_4 = n_5 = 1$. (d) Sketch of the cross-section of the waveguide with the coordinate frame for a typical TE-like mode. The waveguide geometry is rotated such that E_x is tangential to the "upper" surface of the waveguide. Refractive indices $n_1 = 3.476$, $n_4 = 1.444$, $n_2 = n_3 = n_5 = 1$. The mode profiles in this figure are calculated using the *amplitude optimization method*.

Furthermore, when Eqs. (2)-(5) are satisfied, the electric and magnetic fields satisfy $\nabla \cdot (\epsilon_0 n_j^2 \mathcal{E}) = 0$ and $\nabla \cdot \mathcal{H} = 0$, respectively.

In this section, we adopt a description of the behavior of the light in a rectangular waveguide that is based on separation of spatial variables (x, y) in the core region, similar to Marcatili's *ansatz*. We neglect the effect of the corners based on the observation that the field is small in those areas. The modal field then consists of standing waves in the core of the waveguide and an exponentially decaying field outside the core (see Fig 1). We will show that the form we propose here can not provide an exact solution of Maxwell's equations, and thus provides an approximate description of the physics. The proposed solution obeys Maxwell's equations in regions 1-5, so that all errors that arise from the approximation show themselves at the interfaces between the waveguide core and its cladding. At these interfaces, continuity of the electromagnetic field components tangential to these interfaces is required (referred to as the electromagnetic boundary conditions), but we allow for discontinuities that are small compared to the field strength in the core of the waveguide. When all tangential components are continuous, the normal components automatically obey Maxwell's equations.

In our analysis, we propose a form where the electric field is predominantly polarized in the x -direction. From symmetry, the field could as well be chosen predominantly polarized in the y -direction as there is no further discrimination between the x - and the y -direction. Sketches (a)-(c) in Fig. 1 show the fundamental TM-like mode in a waveguide with a width

of 600 nm and a height of 300 nm. Sketch (d) shows the fundamental TE-like mode in this waveguide. Instead of rotating the coordinate-frame, the waveguide itself was rotated such that the “top” surface of the waveguide is tangential to E_x .

We make the *ansatz* on the form of the longitudinal components of the modal electromagnetic field in region 1:

$$E_z = A_1 \sin[k_x(x + \xi)] \cos[k_y(y + \eta)], \quad (8)$$

$$H_z = A_2 \cos[k_x(x + \xi)] \sin[k_y(y + \eta)]. \quad (9)$$

Then the transversal electromagnetic field components inside region 1 follow from Eqs. (2)-(5) as

$$E_x = \frac{(\beta A_1 k_x + \omega \mu_0 A_2 k_y)}{i K_1^2} \cos[k_x(x + \xi)] \cos[k_y(y + \eta)], \quad (10)$$

$$E_y = \frac{(-\beta A_1 k_y + \omega \mu_0 A_2 k_x)}{i K_1^2} \sin[k_x(x + \xi)] \sin[k_y(y + \eta)], \quad (11)$$

$$H_x = \frac{(-\beta A_2 k_x + \omega \epsilon_0 n_1^2 A_1 k_y)}{i K_1^2} \sin[k_x(x + \xi)] \sin[k_y(y + \eta)], \quad (12)$$

$$H_y = \frac{(\beta A_2 k_y + \omega \epsilon_0 n_1^2 A_1 k_x)}{i K_1^2} \cos[k_x(x + \xi)] \cos[k_y(y + \eta)]. \quad (13)$$

In region 2, we set

$$E_z = A_3 \exp[\gamma_2(x + d/2)] \cos[k_y(y + \eta)], \quad (14)$$

$$H_z = A_4 \exp[\gamma_2(x + d/2)] \sin[k_y(y + \eta)], \quad (15)$$

in region 3

$$E_z = A_5 \exp[-\gamma_3(x - d/2)] \cos[k_y(y + \eta)], \quad (16)$$

$$H_z = A_6 \exp[-\gamma_3(x - d/2)] \sin[k_y(y + \eta)], \quad (17)$$

in region 4

$$E_z = A_7 \sin[k_x(x + \xi)] \exp[\gamma_4(y + b/2)], \quad (18)$$

$$H_z = A_8 \cos[k_x(x + \xi)] \exp[\gamma_4(y + b/2)], \quad (19)$$

and in region 5

$$E_z = A_9 \sin[k_x(x + \xi)] \exp[-\gamma_5(y - b/2)], \quad (20)$$

$$H_z = A_{10} \cos[k_x(x + \xi)] \exp[-\gamma_5(y - b/2)]. \quad (21)$$

Here, amplitudes $A_1 - A_{10}$, spatial frequencies k_x , k_y , spatial shifts ξ , η , and exponential decay strengths $\gamma_j > 0$, $j = 2, \dots, 5$ are still to be determined. The transversal components are calculated from Eqs (2)-(5). From the wave equation (7) in region 1, we arrive at

$$K_1^2 = n_1^2 k_0^2 - \beta^2 = k_x^2 + k_y^2, \quad (22)$$

or, with positive β ,

$$\beta = \sqrt{n_1^2 k_0^2 - k_x^2 - k_y^2}. \quad (23)$$

The effective (refractive) index of the mode in the waveguide is defined by $n_{\text{eff}} = \beta/k_0$. From the wave equation (7), together with Eq. (22), we find

$$\gamma_2^2 = (n_1^2 - n_2^2)k_0^2 - k_x^2, \quad (24)$$

$$\gamma_3^2 = (n_1^2 - n_3^2)k_0^2 - k_x^2, \quad (25)$$

$$\gamma_4^2 = (n_1^2 - n_4^2)k_0^2 - k_y^2, \quad (26)$$

$$\gamma_5^2 = (n_1^2 - n_5^2)k_0^2 - k_y^2. \quad (27)$$

The sine and cosine dependency on x and y , of the transversal electric (E_z) and magnetic (H_z) field components, was chosen such that the dominant electric field component (E_x) is described by a cosine function in both x and y direction. For most common waveguides $n_2 \approx n_3$ and $n_4 \approx n_5$, from which follows that spatial shifts ξ and η are small. From the observation that the field of the fundamental mode of a waveguide has its highest energy density in the center of the guide, we expect that the field components with a cosine behavior in both x - and y -directions (E_x and H_y) carry the majority the field's energy, next the components with a sine and a cosine dependence in x or y (E_z and H_z), and finally the least energy is expected in the field components that have a sinusoidal profile in both x - and y -direction (E_y and H_x). These components have antinodes (i.e. high energy density) in the corners of the waveguide core.

This description has $A_1 - A_{10}$, ξ , η , k_x and k_y as free parameters. There are 4 interfaces, with 4 tangential field components to be matched per interface, adding up to 16 equations from the electromagnetic boundary conditions. The tangential field components are depicted in Fig. 1(a). Field amplitude A_1 is used as normalization factor, so we end up with an overdetermined system of only 13 free parameters for 16 equations.

In summary, we proposed an *ansatz* on the form of the electromagnetic fields of the modes in a rectangular dielectric waveguide, such that Maxwell's equations are obeyed in all regions 1-5. This *ansatz* has 13 free parameters. From continuity of the tangential electromagnetic field components at the interfaces, 16 boundary condition equations follow. In the remainder of Sec. II-A, the requirements that follow from continuity at either the surface normal (Sec. II-A1) or parallel (Sec. II-A2) to the dominant electric field component, are given.

1) *Obeying the electromagnetic boundary conditions at the interfaces normal to the dominant electric field component:*

In this section, we derive the requirements that follow from continuity of the fields at the 1-2 and 1-3 interfaces. The dominant electric field component, E_x , is orthogonal to these interfaces, so an infinitely wide, in the y -direction ($b \rightarrow \infty$), rectangle will describe a TM mode in a slab waveguide, see Fig. 1(c). From all eight electromagnetic boundary conditions at these interfaces, we find

$$A_2 = \frac{\omega \epsilon_0 n_1^2 k_y}{\beta k_x} A_1, \quad (28)$$

$$A_3 = A_1 \sin[k_x(\xi - d/2)], \quad (29)$$

$$A_4 = A_2 \cos[k_x(\xi - d/2)], \quad (30)$$

$$A_5 = A_1 \sin[k_x(\xi + d/2)], \quad (31)$$

$$A_6 = A_2 \cos[k_x(\xi + d/2)], \quad (32)$$

together with

$$\tan[k_x(\xi - d/2)] = -\frac{n_1^2 \gamma_2}{n_2^2 k_x}, \quad (33)$$

$$\tan[k_x(\xi + d/2)] = \frac{n_1^2 \gamma_3}{n_3^2 k_x}. \quad (34)$$

Equations (29)-(32) follow directly from the continuity of E_z and H_z . The continuity of E_y and H_y is most easily verified by substituting Eqs. (28)-(34) into the remaining electromagnetic boundary conditions. With these field amplitudes $A_2 - A_6$, the magnetic field component H_x is zero in regions 1, 2 and 3, as follows from Eq. (4).

The last two equations, (33) and (34), can be recognized as the eigenvalue equations for a TM mode in a slab waveguide [2], [4]. These eigenvalue equations thus do not only hold for a slab solution where $\partial/\partial y = 0$ and H_y , E_x and E_z are the non-zero field components, but also for our *ansatz* where there is a variation in the y-direction. We eliminate ξ from these two equations, and arrive at the functional

$$G(k_x, k_0, n_1, n_2, n_3, d) \equiv \tan[k_x d] - \frac{n_1^2 k_x (n_3^2 \gamma_2 + n_2^2 \gamma_3)}{n_2^2 n_3^2 k_x^2 - n_1^4 \gamma_2 \gamma_3} = 0. \quad (35)$$

2) *Obeying the electromagnetic boundary conditions at the interface parallel to the dominant electric field component:* When we obey all eight boundary conditions at the 1-4 and 1-5 interface, to which the dominant electric field component, E_x , is parallel, we find

$$A_2 = \frac{\beta k_y}{\omega \mu_0 k_x} A_1, \quad (36)$$

$$A_7 = A_1 \cos[k_y(\eta - b/2)], \quad (37)$$

$$A_8 = A_2 \sin[k_y(\eta - b/2)], \quad (38)$$

$$A_9 = A_1 \cos[k_y(\eta + b/2)], \quad (39)$$

$$A_{10} = A_2 \sin[k_x(\xi + b/2)], \quad (40)$$

together with

$$\tan[k_y(\eta - b/2)] = -\gamma_4/k_y, \quad (41)$$

$$\tan[k_y(\eta + b/2)] = \gamma_5/k_y. \quad (42)$$

Eliminating η from latter two equations gives

$$F(k_y, k_0, n_1, n_4, n_5, b) \equiv \tan[k_y b] - \frac{k_y(\gamma_4 + \gamma_5)}{k_y^2 - \gamma_4 \gamma_5} = 0. \quad (43)$$

Equations (37)-(40) follow from the continuity of E_z and H_z . Continuity of E_x and H_x is checked by substituting Eqs. (36)-(42) into the four boundary conditions corresponding to these field components at the two interfaces. It follows from Eq. (3) that with these field amplitudes A_2 , $A_7 - A_{10}$, the electric field component E_y is zero in regions 1, 4 and 5. Equations (41) and (42) are identical to the eigenvalue equations for a TE mode in a slab waveguide.

3) *Conclusion:* The eigenvalue equations that follow from the analysis of slab waveguides, in which invariance of the field in one direction is assumed, are in fact more general. Identical equations follow from the *ansatz* for E_z and H_z , i.e. Eqs. (8)-(21), and from imposing the boundary conditions on the horizontal or vertical interfaces of the rectangular waveguide.

Obeying all electromagnetic boundary conditions at the interfaces normal to E_x (Sec. II-A1) demand a different amplitude coefficient of the magnetic field in the core (A_2) then the conditions that follow from the interfaces parallel to E_x (Sec. II-A2). Therefore the *ansatz* has no solutions that exactly obeys the electromagnetic boundary conditions at all interfaces simultaneously. The next sections are devoted to different possibilities for choosing the free parameters such that a low mismatch of the fields at the boundaries is achieved.

4) *Normalization:* Throughout this work, A_1 is normalized such that the power flux through waveguide regions 1-5 equals unity, i.e.

$$P = \frac{1}{2} \text{Re} \left\{ \iint_{\text{regions 1-5}} (E_x H_y^* - E_y H_x^*) dx dy \right\} = 1, \quad (44)$$

where the integral runs over the regions 1-5.

B. Marcatili's approach

Marcatili has developed a widely used analytical approach for low-index-contrast waveguides [1]. For propagating modes in these guides, $k_0 n_j \approx \beta$ because modes are not guided otherwise, so $k_x/k_0 n_j$ and $k_y/k_0 n_j$ are much smaller than unity. Therefore those quantities are neglected in second order. This is often referred to as "far from cutoff", while, "close to core material spectral cutoff" would be more appropriate.

In Marcatili's work, H_x is set to zero in all regions. With this requirement, all electromagnetic boundary conditions at the 1-2 and 1-3 interface can be satisfied. This was shown in Sec. II-A1, where the requirement that $H_x = 0$ in regions 1, 2 and 3 followed from the boundary conditions at these interfaces. At the 1-4 and 1-5 interfaces approximations are necessary since not all boundary conditions can be satisfied at these interfaces. H_x is set to zero and continuity follows trivially. In the approximate matching of E_x across these interfaces, quantities on the order of $(k_x/k_0 n_j)^2$ are neglected, which is reasonable for low-index-contrast waveguides. For these guides, $\sqrt{\mu_0} H_z$ is larger than $\sqrt{\epsilon_0} E_z$, and is matched across the horizontal interfaces, while E_z is not matched. From the requirements above, it is found that the eigenvalue equation of this waveguide is identical to the slab eigenvalue equations (33), (34), (41), and (42). We will refer to this approach as *Marcatili's $H_x = 0$ method*.

Although neglecting terms on the order of $(k_x/k_0 n_j)^2$ is valid for low-index-contrast waveguides, this quantity is even larger than unity outside the core region of high-index-contrast waveguides. This approximation introduces a large mismatch in the continuity of E_x , which is the dominant electric field component.

Similarly, another approximate solution is obtained by setting $E_y = 0$ [1], [3]. With this demand, all boundary conditions at the interface parallel to E_x can be satisfied, and mismatches occur at the 1-2 and 1-3 interfaces. Analogue to the approach above, E_y is trivially matched and E_z is matched, but H_z is not. H_y , which is the dominant magnetic field component, is matched while neglecting terms on the order of $(k_x/k_0n_j)^2$. Although the field amplitudes are different to the case where $H_x = 0$ in all regions, the eigenvalue equations of this waveguide description are identical and thus also given by Eqs. (33), (34), (41), and (42). We will refer to this approach as *Marcatili's $E_y = 0$ method*.

C. Improved method

In this section, we present two improvements of Marcatili's methods that give a better description of high-index-contrast waveguides. Two choices for the matching of the boundary conditions are presented, one where the fields at the interfaces normal to the dominant electric field component, E_x , are continuous, and one where the fields at the interfaces parallel to E_x are continuous. In Sec. II-E, we show that the latter is more accurate for the cases considered in this paper.

We argue that the dominant boundary conditions for determining k_x and ξ are at the 1-2 and 1-3 interfaces, and therefore Eqs. (33) and (34) are used to determine k_x and ξ . Similarly, continuity at the 1-4 and 1-5 interfaces is used to determine k_y and η using Eq. (41) and (42). This is supported by the argument that the two different approximations of the modal electromagnetic field as presented by Marcatili both yield these eigenvalue equations.

1) *Improved $H_x \approx 0$ method*: In comparison with Marcatili's approach, we remove the discontinuity of the dominant electric field at the cost of the weak magnetic field component, H_x , being not continuous across the 1-4 and 1-5 interfaces. In this approach, we demand continuity of the tangential electromagnetic fields at the 1-2 and 1-3 interfaces, which are normal to E_x . This determines the electromagnetic fields in regions 1, 2 and 3 by amplitudes A_2 - A_6 , analogue to Marcatili's $H_x = 0$ method. At the 1-4 and 1-5 interfaces, we choose to match E_x and H_z perfectly. The first is the dominant electric field component. The latter is chosen in favor of matching E_z because an infinitely high ($d \rightarrow \infty$) waveguide is identical to a slab waveguide in which $E_z = 0$. These requirements determine the electromagnetic field in regions 4 and 5, in which H_x is not necessary zero. We will refer to this method as the *Improved $H_x \approx 0$ method*.

This gives the slab eigenvalue equations (35) and (43), and field amplitude Eqs. (28)-(32), together with

$$A_7 = A_1 \left(1 + \frac{k_0^2(n_1^2 - n_4^2)}{\beta^2} \right) \cos[k_y(\eta - b/2)], \quad (45)$$

$$A_8 = A_2 \sin[k_y(\eta - b/2)], \quad (46)$$

$$A_9 = A_1 \left(1 + \frac{k_0^2(n_1^2 - n_5^2)}{\beta^2} \right) \cos[k_y(\eta + b/2)], \quad (47)$$

$$A_{10} = A_2 \sin[k_y(\eta + b/2)]. \quad (48)$$

2) *Improved $E_y \approx 0$ method*: This improvement removes the discontinuity of the dominant magnetic field component H_y , that was present in Marcatili's $E_y = 0$ method. We match all tangential field components at the 1-4 and 1-5 interface, which are parallel to the dominant electric field component. From Sec. II-A2 follows that $E_y = 0$ in regions 1, 4, and 5. At the 1-2 and 1-3 interfaces, H_y is matched because it is the dominant magnetic field component, and E_z is matched in favor of H_z because an infinitely wide waveguide ($b \rightarrow \infty$) is identical to a slab waveguide in which $H_z = 0$.

We find the slab eigenvalue equations (35) and (43), and field amplitude Eqs. (36)-(40), together with

$$A_3 = A_1 \sin[k_x(\xi - d/2)], \quad (49)$$

$$A_4 = A_2 \left(1 + \frac{k_0^2(n_1^2 - n_2^2)}{\beta^2} \right) \cos[k_x(\xi - d/2)], \quad (50)$$

$$A_5 = A_1 \sin[k_x(\xi + d/2)], \quad (51)$$

$$A_6 = A_2 \left(1 + \frac{k_0^2(n_1^2 - n_3^2)}{\beta^2} \right) \cos[k_x(\xi + d/2)]. \quad (52)$$

D. Least-discontinuity optimization of the ansatz parameters

In Sec. II-A, we presented an *ansatz* on the form of the electromagnetic field for modes in a rectangular waveguide. This *ansatz* was chosen such that Maxwell's equations are satisfied in regions 1-5, so that all errors manifest themselves at the four interfaces of the waveguide core. This error, which is the discontinuity of the tangential electromagnetic field components at these interfaces, is referred to as the mismatch. The measure we adopt to quantify this mismatch, or error, is the average energy density that is associated with these discontinuities. In Sec. IV, we show that this intuitive quantity excellently agrees with rigorous numerical results. This analysis is performed in a cross-section of the waveguide at $z = 0$, at time $t = 0$, as further longitudinal and temporal behavior follows trivially from Eq. (1). We define:

$$U_{\text{mm}} = \frac{\epsilon_0}{4l} \oint (n^+ + n^-)^2 \cdot |\hat{\nu} \times (\mathbf{E}^+ - \mathbf{E}^-)|^2 dl \quad (53)$$

$$+ \frac{\mu_0}{l} \oint |\hat{\nu} \times (\mathbf{H}^+ - \mathbf{H}^-)|^2 dl.$$

The four interfaces of the waveguide are simultaneously described by the integral. The line integral runs along the circumference of the waveguide in the (x,y)-plane, and $l = 2(b+d)$ is the length of this circumference. \mathbf{E}^+ and \mathbf{E}^- are the electric fields just outside and inside the waveguide core region 1, so that $(\mathbf{E}^+ - \mathbf{E}^-)$ represents the discontinuity of this field, and $\hat{\nu}$ is a unit vector orthogonal to the waveguide surface. The cross product of $\hat{\nu}$ with the discontinuity in the field just selects the tangential components. n^+ and n^- are the refractive indices just outside and inside the waveguide. At the interface, an average refractive index $(n^+ + n^-)/2$ is assumed to calculate the energy density of the electric field components.

Although U_{mm} can be intuitively interpreted as an energy density, we cannot attach a rigorous physical meaning to this quantity. The mismatch in the fields only occurs at interfaces, which have no physical volume. Therefore the energy density cannot be integrated over volume in order to obtain a total energy.

In this section, we use this quantity to propose a new (semi-) analytical method. In the next section, we use this quantity to investigate the methods of the two previous sections (II-B and II-C) without resorting to numerical simulations.

In what we call the *full optimization method*, the mismatch in the fields at the boundaries, as given by Eq. (53), is minimized using an unconstrained nonlinear optimization as implemented in MATLAB [17]. The initial parameters are calculated from the improved method as presented in Sec. II-C2.

The *full optimization method* may be compared with the variational mode expansion method (VMEM) [11], which is also based on a separation of variables. However, the methods differ substantially because the modal field *ansatz* differs and also the object function used in the variational principle of VMEM differs from the object function that is minimized in our approach. The VMEM is applicable to more general structures.

It is interesting to know how accurate k_x and k_y , and thus the effective refractive index, are calculated using the analytical methods. The effective indices that follow from the eigenvalue equations on the one hand, and from the least-mismatch optimization on the other hand, are presented in Fig. 2. It can be seen that the difference between the two methods is small. The influence of this difference in propagation speed on the mismatch of the fields at the interfaces is investigated by optimizing the field amplitudes $A_2 - A_{10}$ with fixed k_x , k_y , ξ , and η as calculated from the analytical eigenvalue equations. As Eq. (53) is quadratic in the amplitudes $A_2 - A_{10}$, this minimum can be found analytically. This method is referred to as the *amplitudes optimization method*. As can be seen in Fig. 3, the mismatch of the method with the fitted k 's (k_x and k_y) and method with the analytical calculated k 's is almost identical. Therefore we conclude that, with the *ansatz* for the fields as described in Sec. II-A, and the error of the model described by the field mismatch, Eq. (53), the values of k_x and k_y are very accurately calculated from analytical slab eigenvalue equations, for the typical SOI waveguides as considered in this paper.

E. Discussion of the different methods

In Fig. 3, all six different methods are compared. The energy associated with mismatch of the electromagnetic fields at the boundaries of the core of the waveguide, U_{mm} , is plotted for three types of waveguide modes. U_{mm} can be interpreted as the energy density of the error at the interfaces. In order to get a feeling for the magnitude of U_{mm} , the average energy density in the core region of the waveguide is also plotted in this figure (black dashed line).

For all geometries, the improved methods are better than Marcatili's original methods. Considering the fundamental modes in a 400 nm wide waveguide, and the 1st order TE-like mode in a 650 nm wide waveguide, we see that the mismatch in Marcatili's $E_y \approx 0$ method is 1.5 to 4 times larger than the mismatch in the improved $E_y \approx 0$ method. In Marcatili's method, the power flux through regions 2, 3, 4, or 5, i.e. in the evanescent tail of the mode, is sometimes negative (or in backwards propagating direction). This is an indication of the

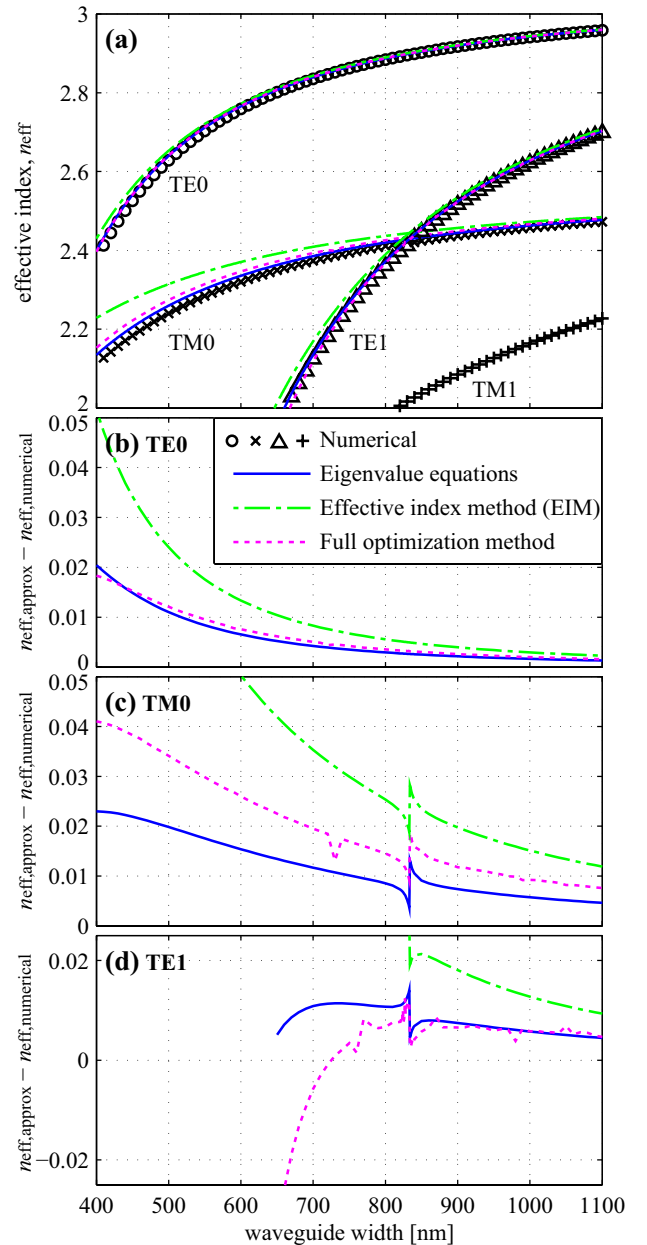


Fig. 2. Effective refractive indices calculated using four different methods. Plot (a) presents the first 3 modes in the waveguide core (TE0, TM0, TE1). In comparison with conventional notation, e.g. TE₀₀, we dropped a zero, as in our waveguide geometries all higher-order modes have higher-order standing waves only in the direction of the width of the waveguide. The numerically calculated effective index of the TM1 mode is included for completeness. Plots (b)-(d) show the difference in the effective index as calculated by the analytical method with respect to the numerical method.

inaccuracy of the method, and this behavior was not observed for the other methods.

For the waveguide geometries as considered in this paper, the improved $E_y \approx 0$ method has smaller errors in the continuity of the fields at the boundaries than the improved $H_x \approx 0$ method. However, this is geometry dependent, as is expected from the trend in the TM-like mode. For TM-like modes in waveguides wider than 1765 nm ($b > 6d$), the improved $H_x \approx 0$ method has a lower mismatch than the improved $E_y \approx 0$ method. This behavior of the asymptotes is

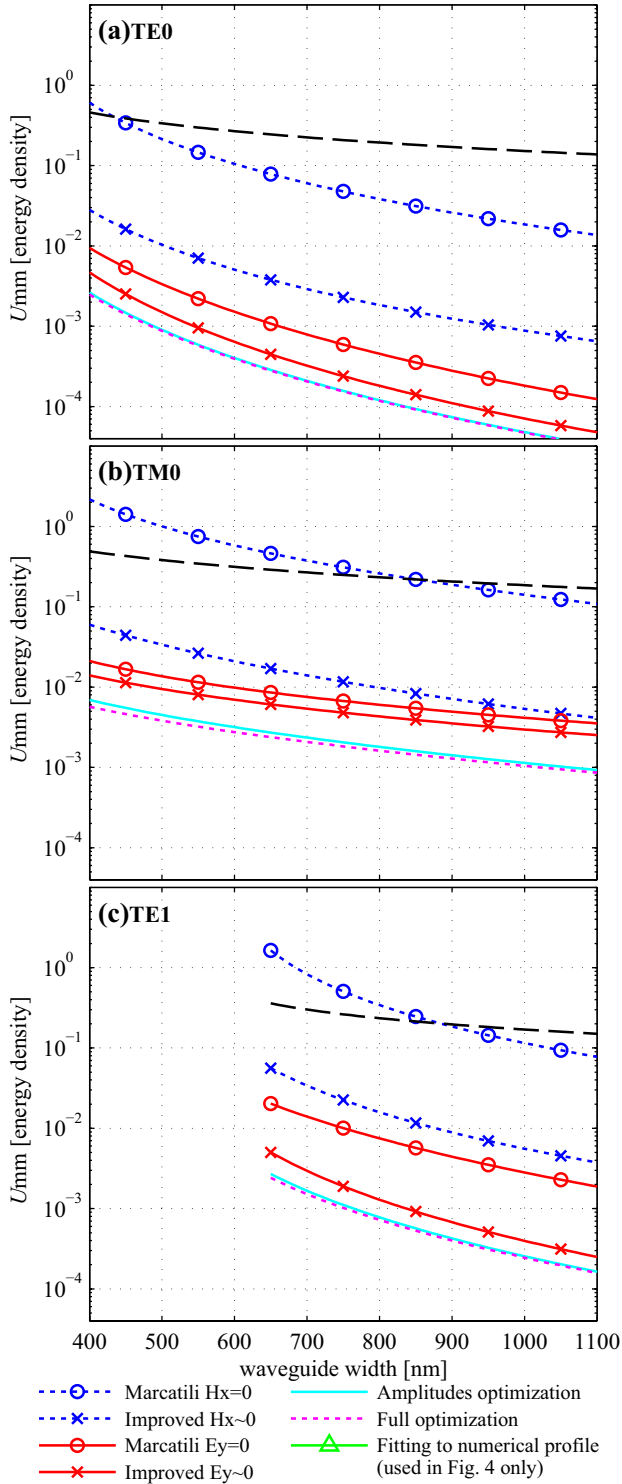


Fig. 3. Mismatch of the tangential field components at the interfaces of the waveguide core, U_{mm} , see Eq. (53). Typical, 300 nm high, SOI waveguide. Plots (a)-(c) represent the TE0, TM0, and TE1 modes in the waveguide. In each plot, 6 different methods to calculate the free parameter in the analytical description of the waveguide mode profile are presented. The black dashed line indicates the average energy density in the core (calculated using the *full optimization method*).

expected from theory of slab waveguides. In the limit $b \rightarrow \infty$, the waveguide is a slab waveguide with interfaces 1-2 and 1-3, so the *improved $H_x \approx 0$ method* is expected to give better results, because the fields described by this methods are continuous across these interfaces. In the limit $d \rightarrow \infty$, only the 1-4 and 1-5 interfaces play a role, so we expect the *improved $E_y \approx 0$ method* to give the most accurate results. For the cases considered in this plot, average energy density associated with the mismatch (U_{mm}) of the *improved $E_y \approx 0$ method* is below 3% of the average energy density in the core of the waveguide.

In conclusion, the effective index of the mode is accurately calculated by the eigenvalue equations. The amplitudes A_2 to A_{10} are calculated more accurately with the improved methods than with Marcatili's original methods. A better quantitative evaluation can be made by a comparison of the methods with rigorous numerical calculated modal profiles, which is presented in Sec. IV.

III. NOVEL APPLICATIONS OF THE EIGENVALUE EQUATION OF THE PROPAGATION CONSTANT

In this section, we derive explicit expressions for the effective group index and for the influence of a change in the waveguide.

A. Modal dispersion

The modal dispersion can be described by the (effective refractive) group index:

$$n_g \equiv n_{\text{eff}} - \lambda_0 \frac{\partial n_{\text{eff}}}{\partial \lambda_0} = \frac{\partial \beta}{\partial k_0}. \quad (54)$$

The group index is often used to design photonic integrated circuits. For example to calculate of the free spectral range (FSR) of ring resonators. From Eq. (23), we find

$$\frac{\partial \beta}{\partial k_0} = \frac{1}{\beta} \left(k_0 n_1^2 + k_0^2 n_1 \frac{\partial n_1}{\partial k_0} - k_x \frac{\partial k_x}{\partial k_0} - k_y \frac{\partial k_y}{\partial k_0} \right). \quad (55)$$

The 1st and 2nd term on the right-hand-side of this equation are specified by the material refractive indices. The refractive indices $n_j(k_0)$ may depend on frequency and thus on $k_0 = \omega/c$. The 3rd term is calculated from Eq. (35). Although k_x is only given as implicitly, $\partial k_x / \partial k_0$ can be calculated explicitly. The total derivative of the left-hand-side of Eq. (35) with respect to k_0 , dG/dk_0 , equals zero for solutions of $G = 0$. The height d does not depend on frequency. So we get

$$\frac{dG}{dk_0} = \frac{\partial G}{\partial k_0} + \frac{\partial G}{\partial k_x} \frac{\partial k_x}{\partial k_0} + \frac{\partial G}{\partial n_1} \frac{\partial n_1}{\partial k_0} + \frac{\partial G}{\partial n_2} \frac{\partial n_2}{\partial k_0} + \frac{\partial G}{\partial n_3} \frac{\partial n_3}{\partial k_0}, \quad \text{or,}$$

$$\frac{\partial k_x}{\partial k_0} = - \frac{\frac{\partial G}{\partial k_0} + \frac{\partial G}{\partial n_1} \frac{\partial n_1}{\partial k_0} + \frac{\partial G}{\partial n_2} \frac{\partial n_2}{\partial k_0} + \frac{\partial G}{\partial n_3} \frac{\partial n_3}{\partial k_0}}{\frac{\partial G}{\partial k_x}}. \quad (56)$$

Similarly, the 4th term of the right-hand-side of Eq. (55) is calculated from Eq. (43) as

$$\frac{\partial k_y}{\partial k_0} = - \frac{\frac{\partial F}{\partial k_0} + \frac{\partial F}{\partial n_1} \frac{\partial n_1}{\partial k_0} + \frac{\partial F}{\partial n_4} \frac{\partial n_4}{\partial k_0} + \frac{\partial F}{\partial n_5} \frac{\partial n_5}{\partial k_0}}{\frac{\partial F}{\partial k_y}}. \quad (57)$$

The partial derivatives in Eqs. (56) and (57) are straightforward to calculate, as shown in appendix A.

B. Influence of a linear external effect

Analogue to the derivation of the effective group index, it is also possible to calculate the linearized influence of an external effect on the propagation constant of the mode. For example, the influence of a temperature change in the waveguide or the influence of a change in the refractive index of the surrounding medium can be calculated. We describe the external effect by a parameter χ . In this theory, the waveguide properties n_j , b and d are assumed to be known, as well as the first-order influence of the external effect on these properties, i.e. $\partial n_j/\partial\chi$, $\partial d/\partial\chi$ and $\partial b/\partial\chi$. Taking the derivative of Eq. (23) with respect to χ gives

$$\frac{\partial\beta}{\partial\chi} = \frac{1}{\beta} \left(n_1 k_0^2 \frac{\partial n_1}{\partial\chi} - k_x \frac{\partial k_x}{\partial\chi} - k_y \frac{\partial k_y}{\partial\chi} \right). \quad (58)$$

As in the previous section, $\partial k_x/\partial\chi$ is calculated by taking the derivative of Eq. (35), $G = 0$, and solving $\partial k_x/\partial\chi$:

$$\frac{\partial k_x}{\partial\chi} = - \frac{\frac{\partial G}{\partial n_1} \frac{\partial n_1}{\partial\chi} + \frac{\partial G}{\partial n_2} \frac{\partial n_2}{\partial\chi} + \frac{\partial G}{\partial n_3} \frac{\partial n_3}{\partial\chi} + \frac{\partial G}{\partial d} \frac{\partial d}{\partial\chi}}{\frac{\partial G}{\partial k_x}}. \quad (59)$$

And for $\partial k_y/\partial\chi$, we find from Eq. (43) that

$$\frac{\partial k_y}{\partial\chi} = - \frac{\frac{\partial F}{\partial n_1} \frac{\partial n_1}{\partial\chi} + \frac{\partial F}{\partial n_4} \frac{\partial n_4}{\partial\chi} + \frac{\partial F}{\partial n_5} \frac{\partial n_5}{\partial\chi} + \frac{\partial F}{\partial b} \frac{\partial b}{\partial\chi}}{\frac{\partial F}{\partial k_y}}. \quad (60)$$

Equations (59) and (60) are given explicitly in Appendix B.

1) *Influence of temperature on SOI waveguides:* In the case of a temperature change, $\partial n_j/\partial\chi$ is given by the thermo-optic effect. The change in cross-section of the waveguide, $\partial d/\partial\chi$ and $\partial b/\partial\chi$ are described by linear thermal expansion. The thermo-optic coefficients of silicon and silicon dioxide, as well as the linear thermal expansion coefficients, are reported in Refs. [16] and [18], respectively.

2) *Evanescent field sensor:* In evanescent field sensors, liquids with varying refractive indices flow over the SOI waveguides. In this case, χ describes the liquid refractive index, so that for TM-like modes $n_3 = n_4 = n_5 = \chi$ and $\partial n_3/\partial\chi = \partial n_4/\partial\chi = \partial n_5/\partial\chi = 1$, while the other waveguide properties are constant. For TE-like modes, $n_2 = n_3 = n_5 = \chi$ and $\partial n_2/\partial\chi = \partial n_3/\partial\chi = \partial n_5/\partial\chi = 1$ (see Fig. 1).

IV. COMPARISON WITH RIGOROUS NUMERICAL METHODS

In this section, we compare the analytical results of the previous section with rigorous numerical results. To find the transverse modes of the waveguide, we use the film mode matching (FMM) method as implemented in the FimmWave software package [10], [12]. In this method the cross-section of the ridge waveguide is split in vertical slices, and 1-dimensional modes are computed analytically for each slice. The 2-D modes are now found by simultaneously solving the modal amplitude factors of the 1-D modes in all slices such that they construct a field obeying Maxwell's equations. The area of the numerical simulation extends $2 \mu\text{m}$ from the waveguide, and 200 1-D modes are used per slice. The width of the waveguide is varied in steps of 10 nm, and in steps of 1 nm between 820 nm and 840 nm. For verification, we

compared the FMM method with the finite element method (FEM) that is also implemented in the FimmWave software, using ~ 210 gridpoints in both the x- and the y-direction. For all datapoints presented in this work, the difference (between FMM and FEM) in effective index is below 10^{-3} and the relative energy in the difference field is below 10^{-4} .

The measure that is used to compare two electromagnetic fields is the relative energy in the difference field, i.e.

$$\Delta U = \frac{\iint_{\text{regions 1-5}} (n^2 \epsilon_0 |\mathbf{E}^A - \mathbf{E}^N|^2 + \mu_0 |\mathbf{H}^A - \mathbf{H}^N|^2) dx dy}{\iint_{\text{regions 1-5}} (n^2 \epsilon_0 |\mathbf{E}^N|^2 + \mu_0 |\mathbf{H}^N|^2) dx dy} \quad (61)$$

where \mathbf{E}^A and \mathbf{E}^N are the analytically and rigorous numerically calculated fields, respectively. This integral runs over all regions that are described by the analytical solution. The integral is calculated analytically with the assumption that the numerical field is piecewise-constant in space.

A. Improved analytical methods

The effective refractive indices are calculated using four different methods, including the FMM method, and are presented in Fig. 2. For completeness, results of the effective index method (EIM) [2] are also included. The analytical eigenvalue equations, described in Sec. II-A, give, for most geometries, the most accurate effective index. The method where the effective index is calculated using least-mismatch fitting in the boundary conditions as presented in Sec. II-D is less accurate for TM-like modes. The error in the EIM becomes significantly large for strong confined modes. The discontinuities in plots (c) and (d) at a width of 833 nm are discussed in Sec. IV-B. The other small ‘‘wiggles’’ in the *full optimization method* are due to numerical instabilities. It can be seen that the error in the effective index calculated using the analytical eigenvalue equation remains below 2% for all cases. The error is, for all methods, lower for the weaker-confined modes.

Figure 4 presents the energy of the difference field between the analytical and the numerical FMM method. First of all, it is clear that something special happens for the fundamental TM-like (TM0) and the 1st TE-like (TE1) modes in the waveguide with a width of ~ 830 nm. We will explain this in Sec. IV-B. The very small ‘‘wiggles’’ in the fundamental TE-like mode between waveguide widths of 820 nm and 840 nm are due to our numerical discretization of the electromagnetic fields, which not exactly match the refractive index distribution for features below 10 nm.

The fundamental limit of the *ansatz* is found by optimizing all free parameters such that the energy in the difference field compared to the numerical result (the quantity plotted in this figure) is minimal.

The method that gives the least difference with the rigorous numerical result is the method where k_x , k_y , ξ , and η follow from the slab eigenvalue equations, while the field amplitudes A_2 - A_{10} are optimized by minimizing the mismatch, U_{nm} , in Eq. (53) (*amplitude optimization method*). For this method, the energy in the difference field is below 2% for all waveguide

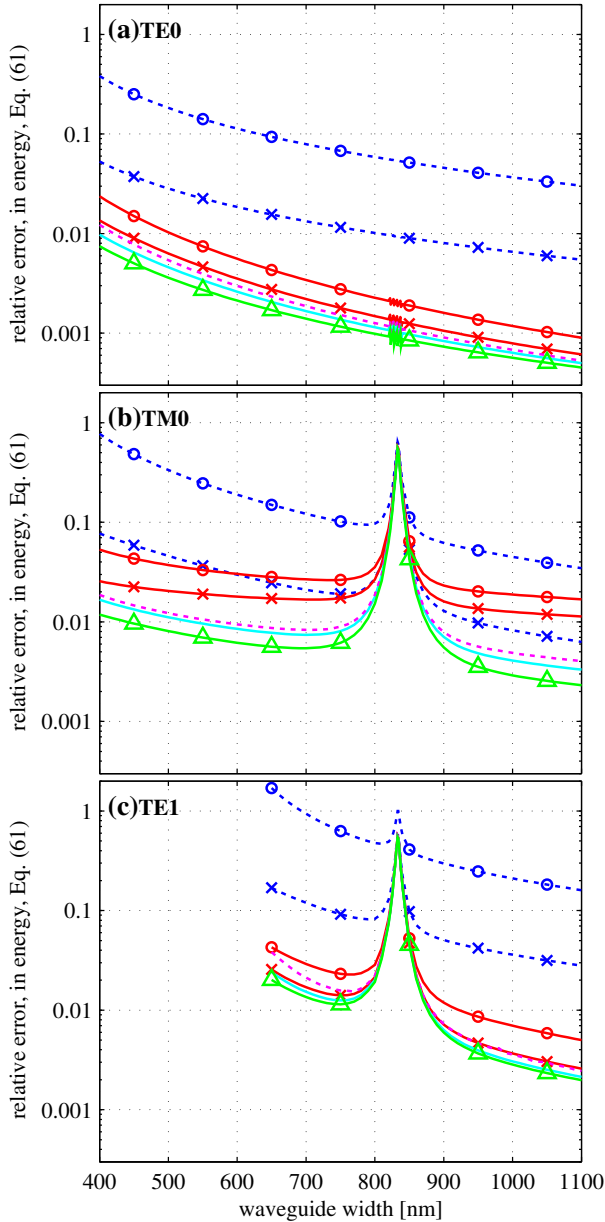


Fig. 4. Energy in the electromagnetic difference field between the analytically calculated fields and the rigorous numerically calculated fields, normalized to the energy in the numerically calculated field. The fields are only considered in regions 1-5. Legend and layout of the figure is identical to Fig. 3.

modes considered here (except for the crossing at 830 nm). Also, this method is very close to the fundamental limit of the *ansatz*.

The fact that the eigenvalue equations work so well, justifies the assumption that the boundary conditions at the 1-2 and 1-3 interfaces can be best used to calculate k_x and the boundary conditions on the 1-4 and 1-5 interfaces can be best used to calculate k_y . Deviating from these k -values to minimize the total mismatch in the boundary conditions only slightly reduces this mismatch and increases the error in the k -values. Still, the quantification of the mismatch as done in Sec. II-D gives surprisingly accurate results. The difference between the fundamental limit of the *ansatz* and both the *amplitude*

optimization method and the *full optimization method* is very small, especially considering that there is a fundamental error in the *ansatz*, and that the error of the fields only manifests itself at the interfaces of the waveguide core, which is difficult to interpret from a physical perspective.

The improved approaches are, for all cases, more accurate than Marcatili's original approaches. For these methods, the full field-profiles are easily and intuitively calculated. The fields of the SOI waveguides considered in this work are accurately described by the *improved $E_y \approx 0$ method*, with less than 3% energy in the difference field (except for the crossing at 830 nm).

The *improved $E_y \approx 0$ method* is more accurate than the *improved $H_x \approx 0$ method* for ($b < 3d$), i.e. for TE-like modes and for TM-like modes in waveguides that are not very wide. From the mismatch in the boundary conditions, the same trend was expected and this crossing was expected at ($b \approx 6d$). Also, the difference between the *improved $E_y \approx 0$ method*, the *improved $H_x \approx 0$ method* and the optimization methods follows the same trend as expected from Sec. II-D.

In conclusion, the improved Marcatili's method describes the effective indices and the field distributions of typical SOI waveguides sufficiently accurately for many applications, except when the effective indices of two modes are similar. In the next section, the modes found from our approximation are used to explain what happens in these regions. When compared to rigorous numerical solutions, the error in the effective index was below 2% and the relative error in terms of the energy of the fields was less than 3%. In addition, the optimization method gives even more accurate results for the fields.

B. The avoided crossing phenomenon

Figures 2 and 4 suggest that something interesting happens at the apparent crossing of the effective indices of the TM0-like and the TE1-like modes when the width of the guide is changed. A detailed inspection of the waveguides with widths around 833 nm is presented in Fig. 5. In plot (a), it can be seen that the numerically computed effective indices of the 2nd and 3rd mode in the waveguide (counted from high to low effective index) actually do not cross each other, but show a behavior that is known in quantum mechanics as *avoided crossing* [19]. We investigate the modes that were found numerically, E_2^N and E_3^N , in terms of the analytically computed approximate modes of the waveguide.

We will verify that the actual modes E_2^N and E_3^N can in good approximation be written as a superposition of the TM0-like, E_{TM0} , and TE1-like, E_{TE1} modes that were calculated using the approximate theory (*amplitude optimization method*), i.e.

$$\mathbf{E}_i^N \approx a\mathbf{E}_{\text{TM0}} + b\mathbf{E}_{\text{TE1}}, \quad (62)$$

for some real a and b , and $i = 2$ or 3 . The phase of mode \mathbf{E}_i is chosen such that coefficient b is positive. The coefficient a of the TM0-like mode can be either positive or negative. The approximate calculated modes \mathbf{E}_{TM0} and \mathbf{E}_{TE1} , are in good approximation orthonormal such that normalization of \mathbf{E}_i in the norm of Eq. (44) implies $b = \sqrt{1 - a^2}$.

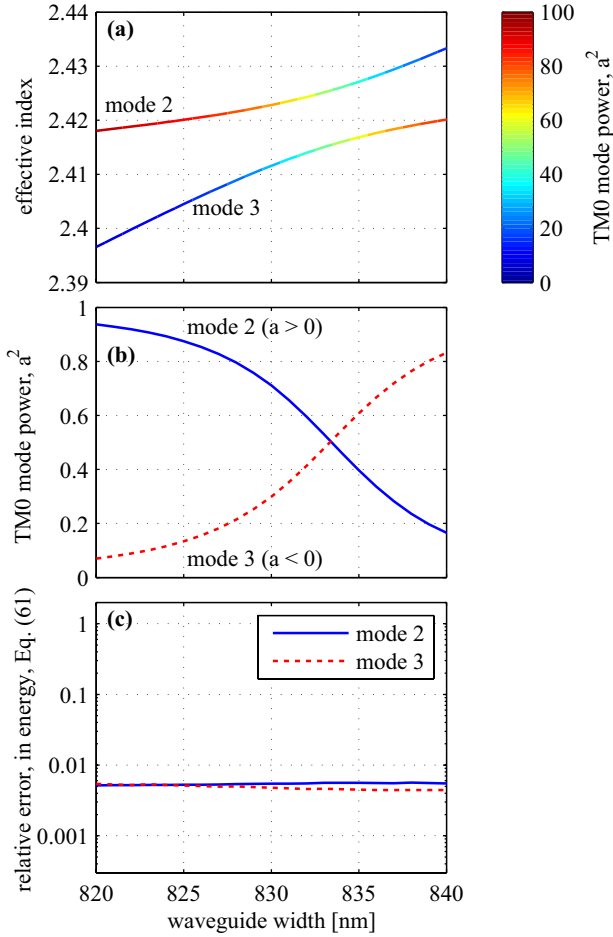


Fig. 5. Investigation of the avoided crossing of effective indices of two modes. Plot (a): numerically calculated effective indices of the 2nd and 3rd mode, zoom-in of Fig. 2(a). Plot (b): power in the TM0-like mode when the fields of the modes in plot (a) are written as a superposition of a TM0-like and a TE1-like mode. The curves in plot (a) are color-coded accordingly. Plot (c): Relative energy in the electromagnetic difference field between the superposition and the rigorous numerically calculated fields, to be compared with Fig 4 (b,c).

The coefficient a of the TM0-like mode is optimized such that the difference measured using Eq. (61) between the left- and right-hand sides of Eq. (62) is minimum. The result is plotted in Fig. 5(b), where it can be seen that mode 2 looks like a TM0-mode at the left of the crossing, while it looks like a TE1-like mode on the right-hand-side of the crossing, whereas close to the crossing the modes are an equal mixture of E_{TM0} and E_{TE1} . In Fig. 5(c), it can be seen that the error between the superposition E_i and the rigorous numerically calculated field E_i^N close to the apparent crossing is small and similar to the error that was found away from the crossing (see Fig. 4). Therefore we may indeed conclude that the field around the crossing can be written as a superposition of modes of the types that are present away from the crossing. Figure 6 presents the electric fields E_{TM0} , E_{TE1} , E_2^N and E_3^N for a 833 nm wide by 300 nm high waveguide, where $a^2 \approx b^2 \approx 0.5$.

Using this observation, we will derive a qualitative description of this *avoided crossing*. The exact description of propagation of light through waveguides that are invariant in

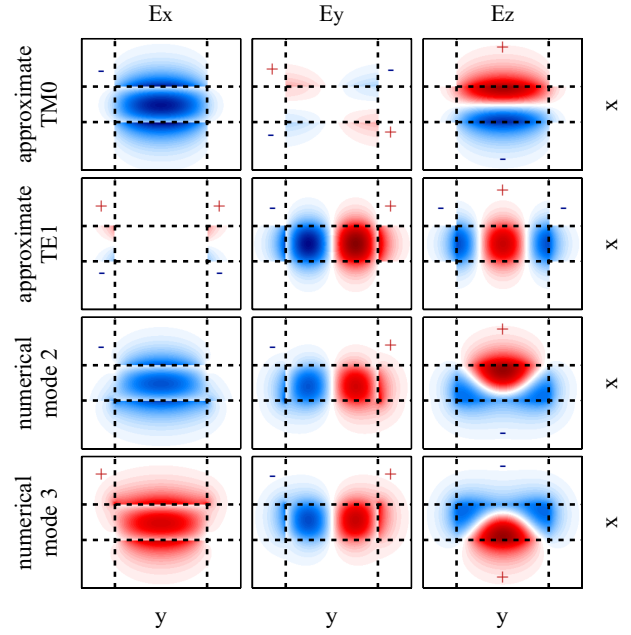


Fig. 6. Electric fields of different modes in a 833 nm wide by 300 nm high waveguide. The dashed lines separate the different regions (see Fig. 1). Modes E_{TM0} , E_{TE1} , E_2^N and E_3^N are plotted from top to bottom. Color indicates the field strength, white regions have a low field strength. Regions with positive field strength are indicated with a plus (+) sign and red. Regions with a negative field strength are indicated with a minus sign (-) and blue.

the z -direction can be formulated as an eigenvalue problem with the propagation constant as eigenvalue. In fact, by eliminating the longitudinal components E_z and H_z from Maxwell's equations, one obtains the following eigenvalue problem for the transverse components only

$$\hat{O} \begin{pmatrix} E_x(x, y) \\ E_y(x, y) \\ H_x(x, y) \\ H_y(x, y) \end{pmatrix} = \beta_i \begin{pmatrix} E_x(x, y) \\ E_y(x, y) \\ H_x(x, y) \\ H_y(x, y) \end{pmatrix} \quad (63)$$

with the propagation constant β_i as eigenvalue and with \hat{O} a second order partial differential operator with respect to transverse variables x and y . We consider forward propagating waves with positive β_i . The operator \hat{O} is not symmetric, however it can be shown that solutions of Eq. (63) are orthogonal with respect to the bilinear form derived from the power flux [3]

$$\langle i|j \rangle \equiv \iint_{-\infty}^{\infty} (\mathbf{E}_i \times \mathbf{H}_j^*) \cdot \hat{z} \, dx dy = \delta_{ij} \quad (64)$$

where the inner product between two solutions i and j is defined, δ_{ij} is the Kronecker delta function, and the eigenfields have been normalized. We adopt a *bra-ket* notation and write Eq. (63) as

$$\hat{O}|i \rangle = \beta_i|i \rangle \quad (65)$$

We now apply the aforementioned observation that in good approximation the electromagnetic fields of the modes in the

waveguide, also around the crossing, can be written as a superposition of the approximate fields, hence

$$\mathbf{E}_i \approx a\mathbf{E}_{\text{TM0}} + b\mathbf{E}_{\text{TE1}}, \quad \text{or} \quad |i\rangle \approx a|a\rangle + b|b\rangle, \quad (66)$$

where $|a\rangle$ and $|b\rangle$ represent the TM0-like and TE1-like modes in the waveguide, respectively, while $|i\rangle$ represents the exact solutions of Eq. (65). We only consider the 2nd and 3rd approximate solutions here. As will become clear, only modes with similar propagation constants have to be taken into account around the crossing, whereas the other modes are already accurately calculated by the approximate methods presented in this paper. Substituting Eq. (66) in Eq. (65) and taking the inner product with $\langle a|$ gives

$$a\langle a|\hat{O}a\rangle + b\langle a|\hat{O}b\rangle \approx \beta_i (a\langle a|a\rangle + b\langle a|b\rangle). \quad (67)$$

If we also take the inner product of Eq. (65) with $\langle b|$ we arrive at the (2x2)-system:

$$\begin{pmatrix} \langle a|\hat{O}a\rangle & \langle a|\hat{O}b\rangle \\ \langle b|\hat{O}a\rangle & \langle b|\hat{O}b\rangle \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \approx \beta_i \begin{pmatrix} \langle a|a\rangle & \langle a|b\rangle \\ \langle b|a\rangle & \langle b|b\rangle \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}, \quad (68)$$

or,

$$M \begin{pmatrix} a \\ b \end{pmatrix} \approx \beta_i \begin{pmatrix} a \\ b \end{pmatrix}, \quad (69)$$

where $M =$

$$\frac{1}{D} \begin{pmatrix} \langle b|b\rangle\langle a|\hat{O}a\rangle - \langle a|b\rangle\langle b|\hat{O}a\rangle & \langle b|b\rangle\langle a|\hat{O}b\rangle - \langle a|b\rangle\langle b|\hat{O}b\rangle \\ -\langle b|a\rangle\langle a|\hat{O}a\rangle + \langle a|a\rangle\langle b|\hat{O}a\rangle & -\langle b|a\rangle\langle a|\hat{O}b\rangle + \langle a|a\rangle\langle b|\hat{O}b\rangle \end{pmatrix}$$

and

$$D = \langle a|a\rangle\langle b|b\rangle - \langle a|b\rangle\langle b|a\rangle.$$

The modes that we found in our approximate analysis are almost orthonormal, so $\langle a|a\rangle$ and $\langle b|b\rangle$ are approximately unity and $\langle a|b\rangle$ and $\langle b|a\rangle$ are approximately zero. Away from the crossing, we found that the approximate solutions $|a\rangle$ and $|b\rangle$ obey relation Eq. (65) so that $\langle a|\hat{O}a\rangle \approx \beta_a$, $\langle b|\hat{O}b\rangle \approx \beta_b$, while $\langle a|\hat{O}b\rangle$ and $\langle b|\hat{O}a\rangle$ are small. This allows us to write the relation Eq. (69) as

$$\begin{pmatrix} \beta_a + \delta_a & \delta_{ab} \\ \delta_{ba} & \beta_b + \delta_b \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \approx \beta_i \begin{pmatrix} a \\ b \end{pmatrix}, \quad (70)$$

where δ_a , δ_b , δ_{ab} , and δ_{ba} are quantities that are much smaller than the β 's. This system has the eigenvalues [19]

$$\beta_{2,3} = \frac{\beta'_a + \beta'_b}{2} \pm \frac{\sqrt{(\beta'_a - \beta'_b)^2 + 4\delta_{ab}\delta_{ba}}}{2}, \quad (71)$$

with corresponding eigenvectors $\mathbf{v}_{2,3}$ (not normalized)

$$\begin{pmatrix} -\beta'_a + \beta'_b \pm \frac{2\delta_{ab}}{\sqrt{(\beta'_a - \beta'_b)^2 + 4\delta_{ab}\delta_{ba}}} \\ 1 \end{pmatrix}, \quad (72)$$

where $\beta'_a \equiv \beta_a + \delta_a$ and $\beta'_b \equiv \beta_b + \delta_b$. The two propagation constants are closest when $\beta'_a = \beta'_b$ but are always separated by a minimum distance $4\sqrt{\delta_{ab}\delta_{ba}}$, so that they never intersect. For small $\delta_a, \delta_b, \delta_{ab}, \delta_{ba} \ll |\beta_a - \beta_b|$, we find the eigenvector for $\beta_a > \beta_b$ to be $\mathbf{v}_2 \approx (1, 0)$ and $\mathbf{v}_3 \approx (0, 1)$. The upper propagation constant, β_2 , has a TM0-like mode in this limit, while the lower propagation constant, β_3 , has a TE1-like mode. For $\beta_b > \beta_a$ we find $\mathbf{v}_2 \approx (0, 1)$ and $\mathbf{v}_3 \approx (1, 0)$, so that the upper propagation constant now has a TE1-like mode while the lower propagation constant has a TM0-like mode. An

interesting case occurs when $\beta'_a = \beta'_b$ and $\delta_{ab} = \delta_{ba}$. Then the normalized eigenvectors of this system are $\mathbf{v}_2 = \frac{1}{\sqrt{2}} \cdot (1, 1)$ and $\mathbf{v}_3 = \frac{1}{\sqrt{2}} \cdot (1, -1)$, i.e. they are either an addition or subtraction of the eigenvectors far from the crossing.

This simple description of the *avoided crossing* agrees with the observations of the numerically computed modal profiles as presented in Figs. 5 and 6.

C. Novel applications of the eigenvalue equation of the propagation constant

We applied the analytical eigenvalue equations for the effective index of the modes in rectangular waveguide to explicitly calculate of the effective group index (Sec. III-A), as well as the linearized influence of an external effect (Sec. III-B). In this section, we compare those analytical results with rigorous numerical simulations. For numerical calculation of the linearized influence of an external effect on the effective index of the modes in the waveguide, we use the following procedure: first, we calculated the effective refractive index (n_{eff}) for different values of the external effect (χ) that are around the starting value (χ_0), then, a linear fit is used to obtain the $\partial n_{\text{eff}}/\partial \chi$.

1) *Effective group index*: The method presented in Sec. III-A is investigated in this section. Equations (54), (77) and (78) are used to calculate the effective group index of the modes in the waveguides. The silicon dispersion is taken into account, with $\partial n_1/\partial k_0 = 3.147 \cdot 10^8 \text{ m}^{-1}$ [16]. These results are compared with the group index that was numerically calculated using the FMM method as implemented in FimmWave. Results are presented in Fig. 7 where can be seen that the error remains below 4%.

2) *Temperature-induced effective index change*: To rigorously compute the linearized influence of temperature variations on the effective index of the waveguide, the temperature in the simulation is varied from 20 °C to 28 °C in steps of 2 °C. The influence $\partial n_{\text{eff}}/\partial T$ is compared with the analytically calculated influence in Fig. 8. Two effects are taken into account: the thermo-optic effect using $\partial n_1/\partial \chi = 1.83 \cdot 10^{-4} \text{ }^\circ\text{C}^{-1}$, and the thermal expansion using $\partial b/\partial \chi = \alpha b \text{ m}^\circ\text{C}$ and $\partial d/\partial \chi = \alpha d \text{ m}^\circ\text{C}$, with linear thermal expansion coefficient $\alpha = 2.6 \cdot 10^{-6} \text{ }^\circ\text{C}^{-1}$ [16], [18]. The effects are investigated simultaneous (net effect) and the effect of the expansion only is also shown separately. This separate investigation does, of course, not simulate a physical measurement but it gives a good test case for the method. From Fig. 8, it can be concluded that the effects are well estimated with a relative error below 7%. The error is much smaller for modes that are less confined, and the effect of the geometrical change is better estimated than the effect of the silicon index change.

3) *Cladding-index induced effective index change*: In Sec. III-B2, we described the linearized influence of a change in refractive index of the cladding of the waveguide. This is, for example, used in evanescent field sensors, where a liquid flows over the waveguide surface. For example, aqueous solutions with different concentrations of sucrose or sodium chloride are used in Refs. [14] and [15], respectively. The rigorous numerical results are obtained by changing the cladding refractive

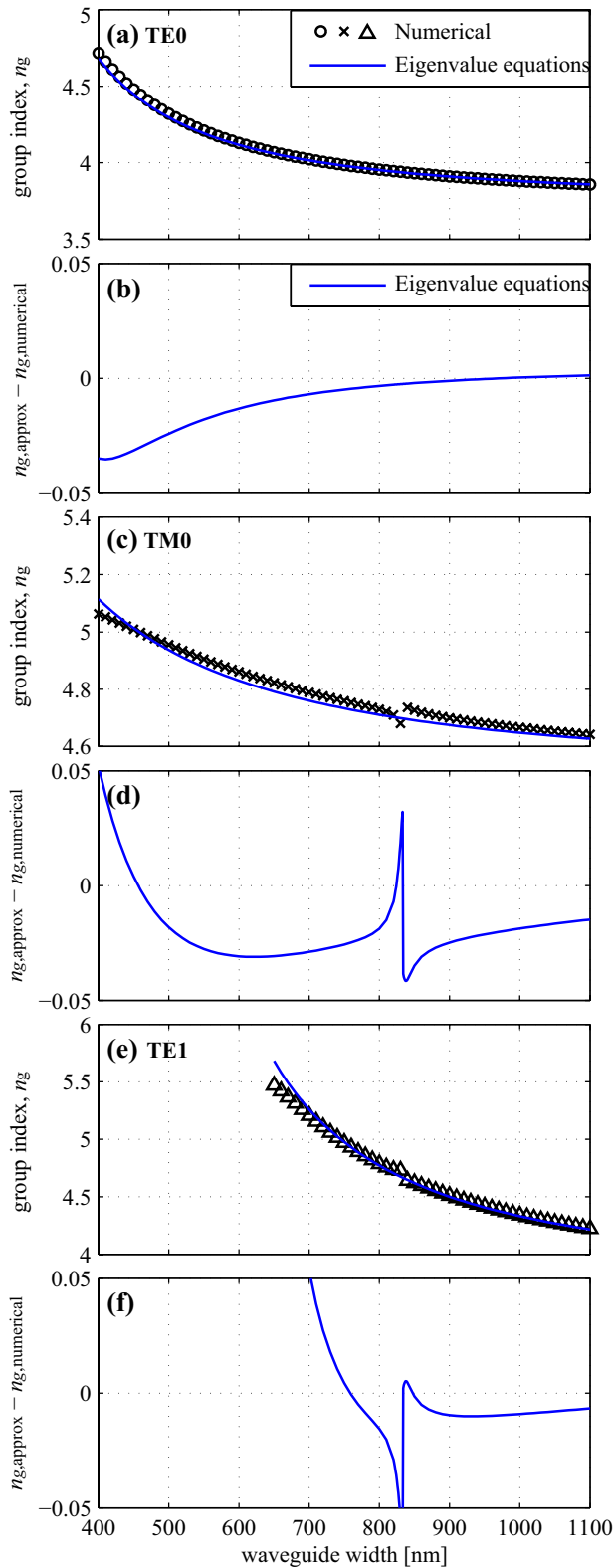


Fig. 7. Effective group indices. Plots (a), (c), (e) present the numerical result, as well as the result of the extension of Marcattili's method (eigenvalue equation). Plots (b), (d), (f) show the difference between the two methods.

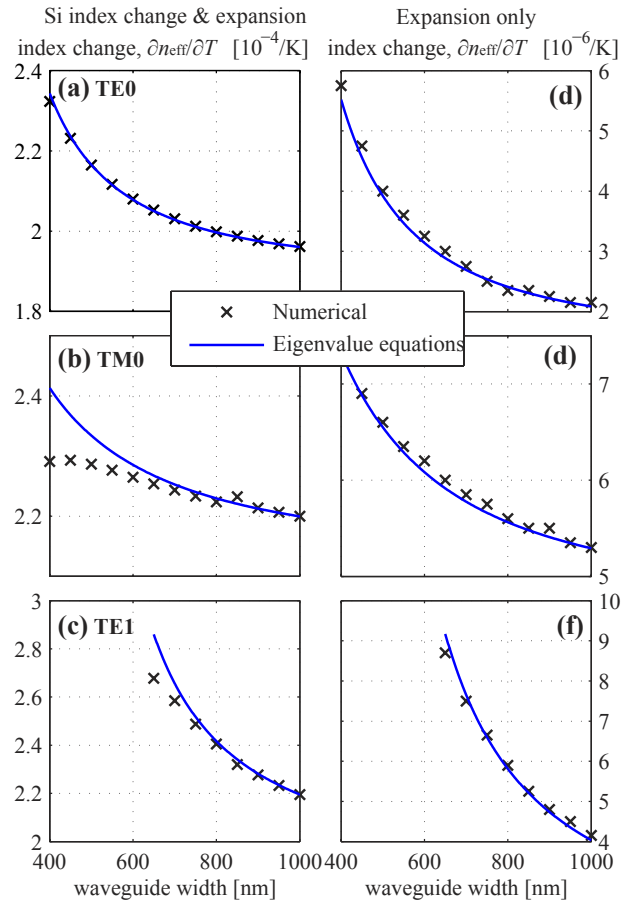


Fig. 8. Shift in effective refractive index due to temperature variation. (Left plots) Shift in effective refractive index due to change in silicon refractive index (thermo-optic effect) as well as expansion of the waveguide (thermal expansion) is depicted. (Right plots) Only the geometry of the waveguide is varied (thermal expansion), with the silicon refractive being constant. The result of the extension of Marcattili's method, and the result of the numerical FMM method are presented.

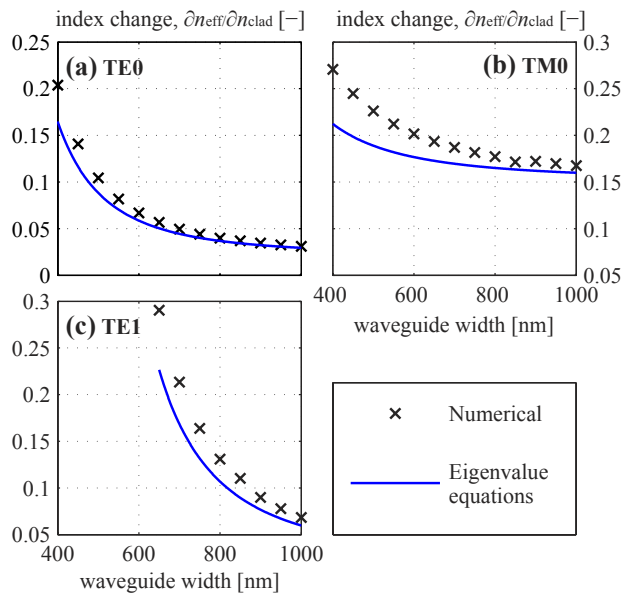


Fig. 9. Shift in effective index due to change in cladding refractive index. The unit is refractive index change per cladding material index change [-/-]. The initial cladding is water ($n = 1.31$).

index from 1.310 (water) to 1.320 in steps of 0.002. From this, $\partial n_{\text{eff}}/\partial n_{\text{clad}}$ is extracted. Results are shown in Fig. 9, where it can be seen that relative errors are up to 23%. These errors rapidly decrease for wider waveguides, i.e. for less-confined modes.

V. CONCLUSION

We derived a model for high-index-contrast rectangular dielectric waveguides, based on an *ansatz* that describes the fields in the core of the waveguide as standing waves. As in Marcatili's work, we also arrive at eigenvalue equations for the spatial frequencies of the standing waves, i.e. k_x and k_y , that are identical to the eigenvalue equations of slab waveguides. In order to obtain k_x , we use the limit that the waveguide extends to infinity in the y-direction, and order to obtain k_y , we use the limit that the waveguide extends to infinity in the x-direction. The effective refractive index of the mode in the waveguide, which directly follows from k_x and k_y , agrees excellently with rigorous numerical mode solvers (relative error < 2%). In comparison with Marcatili's original work, our novel choice of electromagnetic field amplitudes in the *ansatz* removed the discontinuity of the dominant electromagnetic field components, which was severe for high-index-contrast waveguides. This improvement led to better agreement with numerical simulations, with a relative energy of the difference field below 3%, except when the effective indices of two modes in the waveguide are similar. We quantified the error that arises from the discontinuity of the fields in terms of energy density. With this quantification, we accurately predicted the trend in the error of the different analytical methods. In addition, this quantification allowed us to optimize the amplitudes in the *ansatz* such that minimal mismatch is achieved. Our *amplitude optimization method* accurately describes both TE-like as well as TM-like modes in the rectangular waveguides (except when two modes have similar effective indices), and the error of this method with respect to the numerical results is very close to the fundamental error in the *ansatz*.

Next to this, we derived explicit expressions for the effective group index of the waveguide, taking dispersion into account. The error in comparison with rigorous numerical methods was below 4%. Also, explicit expressions are derived for the linearized influence of an external effect. We predicted the influence of temperature, as well as the influence of a change of the refractive index of the cladding of the waveguide.

We applied our method to interpret the results of a rigorous numerical mode solver, and showed that the modes in the waveguide show *avoided crossing* behavior when the effective indices of a TE-like and a TM-like mode are similar. A quick analysis of waveguides can be performed with this method, while the results of this method can be used as starting point for rigorous mode solvers. Marcatili's method has long been used for educational purposes, as is indicated by its frequent occurrence in textbooks. With our work, the development of intuitive analytical methods now follows the technological developments to high-index-contrast waveguides.

APPENDIX A

EXPLICIT EQUATION FOR MODAL DISPERSION

In Sec. III-A, partial derivatives $\partial k_x/\partial k_0$ and $\partial k_y/\partial k_0$ were given by Eqs. (56) and (57). In this section, we give $\partial k_x/\partial k_0$ and $\partial k_y/\partial k_0$ for the case when only the core material is dispersive, i.e. $n_1(k_0)$ and where the other refractive indices do not depend on the frequency, thus also not on k_0 . We define:

$$\alpha_2 \equiv \left(\frac{k_x^2}{n_1^4} + \frac{\gamma_3^2}{n_3^4} \right) \frac{1}{n_2^2 \gamma_2}, \quad (73)$$

$$\alpha_3 \equiv \left(\frac{k_x^2}{n_1^4} + \frac{\gamma_2^2}{n_2^4} \right) \frac{1}{n_3^2 \gamma_3}, \quad (74)$$

$$\alpha_4 \equiv \frac{k_y^2 + \gamma_5^2}{\gamma_4}, \quad (75)$$

$$\alpha_5 \equiv \frac{k_y^2 + \gamma_4^2}{\gamma_5}. \quad (76)$$

And find that Eqs. (56) and (57) are explicitly given by Eqs. (77) and (78), which are shown on the next page.

APPENDIX B

EXPLICIT EQUATION FOR EXTERNAL EFFECT

In this appendix, we give the explicit form of Eqs. (59) and (60). Let:

$$\alpha_a \equiv \frac{\gamma_2}{n_2^2} + \frac{\gamma_3}{n_3^2}, \quad (79)$$

$$\alpha_b^2 \equiv \frac{k_x^2}{n_1^4} - \frac{\gamma_2 \gamma_3}{n_2^2 n_3^2}. \quad (80)$$

Then Eqs. (59) and (60) are identical to Eqs. (81) and (82), see next page, where α_2 - α_5 are defined in Eqs. (73) - (76).

ACKNOWLEDGMENT

The authors would like to thank dr. Omar El Gawhary, dr. Jose Pozo, mr. Vincent Brulis, dr. Jos Thijssen, dr. Ad Verbruggen, and ir. Kevin van Hoogdalem for fruitful discussions.

REFERENCES

- [1] E. Marcatili, "Dielectric rectangular waveguide and directional coupler for integrated optics," *The Bell System Technical Journal*, vol. 48, pp. 2071–2121, Mar. 1969.
- [2] D. Marcuse, *Theory of dielectrically optical waveguides*. San Diego: Academic Press Inc, 1991.
- [3] C. Yeh and F. I. Shimabukuro, *The Essence of Dielectric Waveguides*. New York: Springer Science+Business Media, LCC, 2008.
- [4] C. Pollock and M. Lipson, *Integrated Photonics*. Boston: Kluwer Academic Publishers, 2003.
- [5] R. G. Hunsperger, *Integrated optics: theory and technology*. Berlin: Springer-Verlag, 2002.
- [6] P. Dumon, W. Bogaerts, A. Tchelnokov, J.-M. Fedeli, and R. Baets, "Silicon nanophotonics," *Future Fab International*, vol. 25, pp. 29–36, Apr. 2008.
- [7] A. Yariv, "Universal relations for coupling of optical power between microresonators and dielectric waveguides," *Electronics Letters*, vol. 36, no. 4, pp. 321–322, feb 2000.
- [8] M. Smit, "New focusing and dispersive planar component based on an optical phased array," *Electronics Letters*, vol. 24, no. 7, pp. 385–386, mar 1988.
- [9] J. Goell, "A circular-harmonic computer analysis of rectangular dielectric waveguides," *The Bell System Technical Journal*, vol. 48, pp. 2133–2160, 1969.

$$\frac{\partial k_x}{\partial k_0} = \frac{k_x k_0 \left(\alpha_2 (n_1^2 - n_2^2) + \alpha_3 (n_1^2 - n_3^2) + (\alpha_2 + \alpha_3) n_1 k_0 \frac{\partial n_1}{\partial k_0} \right) + \left(\frac{\gamma_2}{n_2^2} + \frac{\gamma_3}{n_3^2} \right) \left(\frac{4k_x^3}{n_1^5} - \frac{2k_x}{n_1} \left(\frac{k_x^2}{n_1^4} - \frac{\gamma_2 \gamma_3}{n_2^2 n_3^2} \right) \right) \frac{\partial n_1}{\partial k_0}}{\left(\frac{\gamma_2}{n_2^2} + \frac{\gamma_3}{n_3^2} \right) \left(\frac{k_x^2}{n_1^4} + \frac{\gamma_2 \gamma_3}{n_2^2 n_3^2} \right) + k_x^2 (\alpha_2 + \alpha_3) + n_1^2 d \left(\frac{k_x^2}{n_1^4} - \frac{\gamma_2 \gamma_3}{n_2^2 n_3^2} \right)^2 \sec^2[k_x d]} \quad (77)$$

$$\frac{\partial k_y}{\partial k_0} = \frac{k_y k_0 \left(\alpha_4 (n_1^2 - n_4^2) + \alpha_5 (n_1^2 - n_5^2) + (\alpha_4 + \alpha_5) n_1 k_0 \frac{\partial n_1}{\partial k_0} \right)}{(\gamma_4 + \gamma_5)(k_y^2 + \gamma_4 \gamma_5) + k_y^2 (\alpha_4 + \alpha_5) + b(k_y^2 - \gamma_4 \gamma_5)^2 \sec^2[k_y b]} \quad (78)$$

$$\begin{aligned} \frac{\partial k_x}{\partial \chi} = & k_x \left\{ \left((\alpha_2 + \alpha_3) k_0^2 n_1 + \alpha_a \left(\frac{4k_x^2}{n_1^5} - \frac{2\alpha_b^2}{n_1} \right) \right) \frac{\partial n_1}{\partial \chi} - \left(\alpha_2 k_0^2 n_2 + 2 \frac{\gamma_2}{n_2^2} \left(\alpha_b^2 + \alpha_a \frac{\gamma_3}{n_2^2} \right) \right) \frac{\partial n_2}{\partial \chi} \right. \\ & - \left. \left(\alpha_3 k_0^2 n_3 + 2 \frac{\gamma_3}{n_3^2} \left(\alpha_b^2 + \alpha_a \frac{\gamma_2}{n_2^2} \right) \right) \frac{\partial n_3}{\partial \chi} - \alpha_b^4 n_1^2 k_x \sec^2[k_x d] \frac{\partial d}{\partial \chi} \right\} \\ & \cdot \left\{ \alpha_a \left(\frac{k_x^2}{n_1^4} + \frac{\gamma_2 \gamma_3}{n_2^2 n_3^2} \right) + k_x^2 (\alpha_2 + \alpha_3) + n_1^2 \alpha_b^4 d \sec^2[k_x d] \right\}^{-1} \end{aligned} \quad (81)$$

$$\frac{\partial k_y}{\partial \chi} = \frac{k_y k_0^2 \left((\alpha_4 + \alpha_5) n_1 \frac{\partial n_1}{\partial \chi} - \alpha_4 n_4 \frac{\partial n_4}{\partial \chi} - \alpha_5 n_5 \frac{\partial n_5}{\partial \chi} \right) - k_y (k_y^2 - \gamma_4 \gamma_5)^2 \sec^2[k_y b] \frac{\partial b}{\partial \chi}}{(\gamma_4 + \gamma_5)(k_y^2 + \gamma_4 \gamma_5) + k_y^2 (\alpha_4 + \alpha_5) + b(k_y^2 - \gamma_4 \gamma_5)^2 \sec^2[k_y b]}, \quad (82)$$

-
- [10] A. S. Sudbo, "Improved formulation of the film mode matching method for mode field calculations in dielectric waveguides," *Pure and Applied Optics: Journal of the European Optical Society Part A*, vol. 3, pp. 381–388, 1994.
- [11] O. Ivanova, M. Hammer, R. Stoffer, and E. van Groesen, "A variational mode expansion mode solver," *Optical and Quantum Electronics*, vol. 39, pp. 849–864, 2007.
- [12] "Fimmwave - numerical waveguide mode solver," Photon Design Ltd, Oxford, United Kingdom, Jan. 2011.
- [13] A. Melloni, D. Roncelli, F. Morichetti, A. Canciamilla, and A. Bakker, "Statistical design in integrated optics," in *CLEO/Europe and EQEC 2009 Conference Digest*. Optical Society of America, 2009, p. JS11_4.
- [14] A. Densmore, D.-X. Xu, P. Waldron, S. Janz, P. Cheben, J. Lapointe, A. Delage, B. Lamontagne, J. Schmid, and E. Post, "A silicon-on-insulator photonic wire based evanescent field sensor," *Photonics Technology Letters, IEEE*, vol. 18, no. 23, pp. 2520–2522, dec.1, 2006.
- [15] K. D. Vos, I. Bartolozzi, E. Schacht, P. Bienstman, and R. Baets, "Silicon-on-insulator microring resonator for sensitive and label-free biosensing," *Opt. Express*, vol. 15, no. 12, pp. 7610–7615, Jun 2007.
- [16] "Material database and material models," distributed with FimmWave software package, Photon Design Ltd, Oxford, United Kingdom, Jan. 2011.
- [17] "Matlab - the language of technical computing," The MathWorks, Inc., Natick, Massachusetts, USA, 2010.
- [18] Y. Okada and Y. Tokumaru, "Precise determination of lattice parameter and thermal expansion coefficient of silicon between 300 and 1500 k," *Journal of Applied Physics*, vol. 56, no. 2, pp. 314–320, 1984.
- [19] L. Landau and E. Lifshitz, *Quantum Mechanics*, 3rd ed. Pergamon Press Ltd., 1977, sec. 79.