Rigorous modal analysis of silicon strip nanoscale waveguides

D. M. H. Leung*, N. Kejalakshmy, B. M. A. Rahman and K. T. V. Grattan

School of Engineering and Mathematical Sciences,
City University London, Northampton Sq, London, EC1V 0HB England UK
*D.M.H.Leung@city.ac.uk

Abstract: A full-vectorial **H**-field Finite Element Method has been used for the rigorous modal analysis of silicon strip waveguides. The spatial variation of the full-vectorial **H** and **E**-fields are also discussed in details and further, the Poynting vector is also presented. The modal area, hybridness, single mode operation and birefringence are also described for such silicon strip waveguides.

©2010 Optical Society of America

References and links

1. M. Lipson, "Guiding, Modulating, and Emitting Light on Silicon-Challenges and Opportunities," J. Lightwave Technol. **23**(12), 4222–4238 (2005).
2. Y. Vlasov, and S. J. McNab, "Losses in single-mode silicon-on-insulator strip waveguides and bends," Opt. Express **12**(8), 1622–1631 (2004).
3. D. Dai, and S. He, "Ultrasmall overlapped arrayed-waveguide grating based on Si nanowire waveguides for dense wavelength division multiplexing," IEEE J. Sel. Top. Quantum Electron. **12**(6), 1301–1305 (2006).
4. L. Liao, D. Samara-Rubio, M. Morse, A. Liu, D. Hodge, D. Rubin, U. D. Keil, and T. Franck, "High speed silicon Mach-Zehnder modulator," Opt. Express **13**(8), 3129–3135 (2005).
5. O. Boyraz, and B. Jalali, "Demonstration of a silicon Raman laser," Opt. Express **12**(21), 5269–5273 (2004).
6. L. Yin, Q. Lin, and G. P. Agrawal, "Soliton fission and supercontinuum generation in silicon waveguides," Opt. Lett. **32**(4), 391–393 (2007).
7. B. M. A. Rahman, and J. B. Davies, "Finite element solution of integrated optical waveguides," J. Lightwave Technol. **2**(5), 682–688 (1984).
8. B. M. A. Rahman, and J. B. Davies, "Penalty Function Improvement of Waveguide Solution by Finite Elements," IEEE Trans. Microw. Theory Tech. **32**(8), 922–928 (1984).
9. B. M. A. Rahman, and J. B. Davies, "Vector-H finite element solution of GaAs/GaAlAs rib waveguides," Optoelectronics, IEE Proceedings J **132**, 349–353 (1985).
10. ISO 11146, Laser and laser related equipment – Test methods for laser beam widths, divergence and beam propagation ratios, International Organization for Standardization, Geneva, Switzerland, 2005.
11. N. Somasiri, and B. M. A. Rahman, "Polarization crosstalk in high index contrast planar silica waveguides with slanted sidewalls," J. Lightwave Technol. **21**(1), 54–60 (2003).
12. B. M. A. Rahman, S. S. A. Obayya, N. Somasiri, M. Rajarajan, K. T. V. Grattan, and H. A. El-Mikathi, "Design and characterization of compact single-section passive polarization rotator," J. Lightwave Technol. **19**(4), 512–519 (2001).

1. Introduction

In recent years, there has been an increasing interest in the design of ultra-small and compact photonic devices. However most photonic devices have been fabricated using exotic materials, such as Gallium Arsenide (GaAs), Lithium Niobate (LiNbO$_3$) and Indium Phosphide (InP) that are complicated to process and unless their critical material properties, such as their excellent electro-optic properties or emissive capabilities are explicitly needed, are excessively expensive to produce. Recent research [1] has suggested that Silicon (Si) can be the material of choice to reduce the cost of fabrication and to increase the functionality of the components in building photonic integrated circuit (PIC). Si waveguides based on the silicon-on-insulator (SOI) platform are the most basic important building blocks used in many PIC systems, such as ring resonators [2], arrayed waveguide filters [3], modulators [4] and lasers [5]. In photonic-based telecommunication systems, light is strongly confined in the Si waveguide because of the high refractive index contrast between the cladding and the core,

and therefore, it allows for the waveguide core to be shrunk down to a submicron cross-section. It also permits the minimum bending radius of waveguides to be reduced, consequently also allowing an overall reduction in the dimensions of the PIC, which ultimately decreases the cost of Si photonics devices. Previous research [6] has also shown that due to the strong electric field in Si waveguides of small dimension, an ultra-high optical power density can be achieved in which the longitudinal field, E_z has a large effect on the waveguide nonlinearity.

It is the aim of this paper to provide a novel comprehensive analysis defining the modal characteristics of a silicon strip waveguide with a nanoscale cross-section, and in doing so, the effects of the critical size of such waveguide are also presented. To undertake such analysis, an accurate and numerically efficient vector-**H**-field finite-element method (VFEM) [7] is used to calculate the propagation constant and the full-vectorial modal field profiles of the waveguide. The full-vectorial electric field (**E**) is also derived from the vector **H**-field obtained to characterize modal properties of such a waveguides.

2. Theory

A full-vectorial Finite Element Method (FEM) is used in this study to obtain the modal solutions of a Si nanowire optical waveguide. The FEM, based on the vector-**H**-Field formulation has been established as one of the most accurate and numerically efficient approaches to obtaining the modal field profiles and propagation constants of the fundamental and higher-order quasi-TE and TM modes. The full-vectorial formulation is based on the minimization of the following functional in terms of the nodal values of the full **H**-field vector,

$$\omega^2 = \frac{\int \left[\left(\nabla \times H \right)^* \cdot \varepsilon^{-1} \left(\nabla \times H \right) + p \left(\nabla \cdot H \right)^* \left(\nabla \cdot H \right) \right] dxdy}{\int H^* \cdot \mu H \, dxdy} \tag{1}$$

where **H** is the full vectorial magnetic field, * denotes a complex conjugate and transpose, ω^2 is the eigenvalue where ω is the angular frequency of the wave and ε and μ are the permittivity and permeability, respectively. The penalty function approach has also been incorporated to impose divergence-free condition of the magnetic field to reduce the appearance of spurious modes and p is the dimensionless penalty parameter. It has been shown [8] that a 'p' value of the order of $1/n_g^2$ would give a satisfactory result, where n_g is the refractive index of the core. This full-vectorial FEM modal solution may also be used to determine the beat lengths between the fundamental and the higher order modes and also the polarization beat length between the two quasi-TE and TM polarized fundamental modes.

3. Results

3.1 Waveguide structure

The structure considered in the rigorous investigation carried out is a typical silicon strip optical waveguide or Si photonic wire waveguide. This particular type of waveguide consists in general of a silicon (Si) core with small rectangular cross-section, surrounded by either silica (SiO_2) or air. This waveguide can be fabricated by using a SOI (Silicon-On-Insulator) wafer on a Si substrate. A resist mask can be used on the surface of the Si layer and the Si waveguide core may be formed by etching down to the SiO_2 buffer layer by use of an Inductively Coupled Plasma (ICP) dry etcher. The Si core can then either be buried under a thick SiO_2 over-layer or surrounded with air. In this study, the thickness of the core waveguide is taken as 260 nm and that of the lower SiO_2 buffer layer as 1.50 μm. The refractive index of the rectangular Si core at 1.55 μm wavelength is taken as 3.50. The refractive indices for the SiO_2 cladding or air cladding are taken as 1.50 and 1.0, respectively.

3.2 Modal solution

In the study of modal field profile, the **H**-field based VFEM is used to obtain the modal solutions of such a waveguide. For this study, due to the availability of two-fold symmetry of the waveguide structure, only a quarter of the structure is considered, in which more than 80,000 irregular sized first order triangular elements have been employed to represent the waveguide structure. It takes about 2 minutes cpu time on a dual-core Pentium processor computer running solaris platform.
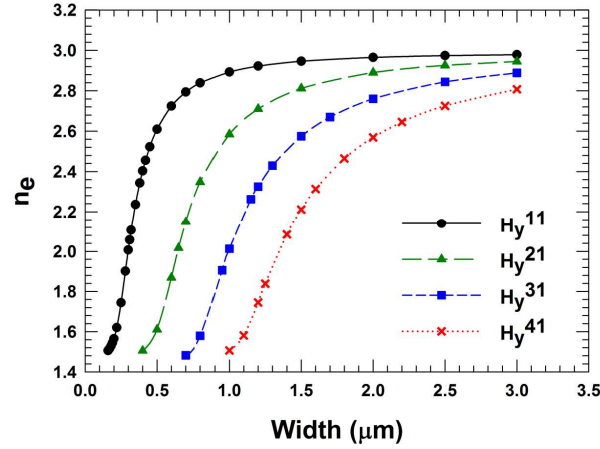
Fig. 1. Variations of the effective index, n_e with the waveguide width, W, for different quasi-TE modes.

In the simulations carried out, the waveguide height (H) was kept constant at 260 nm, while the waveguide width (W) was decreased slowly from 3 µm to nano-dimensions to determine the single-mode-to-multimode-transition. In Fig. 1, variations of the effective index (n_e) with the width (W) for the fundamental H_y^{11} and higher order H_y^{21}, H_y^{31} and H_y^{41} quasi-TE modes are presented for the SiO_2-cladding. The effective index, n_e, of a given mode is a normalized propagation parameter, which can be defined by $n_e = \beta_0/k_0$, where β_0 is the propagation constant of that mode and k_0 is the free space wavenumber defined as $k_0 = \omega(\varepsilon_0\mu_0)^{1/2} = 2\pi/\lambda$. It can be observed that when the width of the waveguide is large in comparison to the height of the structure, the waveguide supports many modes and n_e is closer to the refractive index of Si. As W, decreases, n_e also decreases and gets closer to the refractive index of SiO_2 where the mode reaches its cut-off. It is shown here that the cut-off widths for the H_y^{41}, H_y^{31}, H_y^{21} and H_y^{11} modes are 1000 nm, 700 nm, 400 nm and 200 nm, respectively. This simulation data generated suggest that the single-mode operation occurs when the waveguide width lies approximately between 200 nm to 400 nm when the operating wavelength, λ, is 1550 nm with a height of 260 nm. Accuracy of the solution for the H_y^{11} mode was identified by using Aitken extrapolation [9] to be better than 0.065%, when W = 500 nm.

3.3 Effective area

Mode size area or effective area is an important design parameter for various applications and Fig. 2 shows the variation of the effective area (A_{eff}) with the waveguide width for different quasi-TE modes. Following the second moment of intensity distribution (recommended by ISO Standard 11146), the definition of the effective area [10] can be given by:

$$A_{eff} = \frac{\left(\iint_{\Omega} |E_t|^2 \, dxdy \right)^2}{\iint_{\Omega} |E_t|^4 \, dxdy} \tag{2}$$

where E_t is the transverse electric field vector and the integration is carried out over the whole cross-section of the waveguide, Ω.

It can be observed that as W reduces, A_{eff} reduces to a minimum value and any further reduction of W, results in the sudden increases of the A_{eff} as the mode approaches its cut-off. The minimum effective area, A_{min}, for the H_y^{11} mode is obtained as 0.0956 μm^2 when the value of W = 320 nm, which means that the mode is more confined when the effective area is a minimum. Further rigorous simulations were carried out for the quasi-TE H_y^{21}, H_y^{31} and H_y^{41} modes. It was found that the widths for the minimum A_{eff} values for the H_y^{21}, H_y^{31} and H_y^{41} modes are 650 nm, 1000 nm and 1300 nm, respectively. It can be noted that all the A_{eff} values are very similar when the waveguide width, W, is large but it is only slightly higher for the higher order modes.
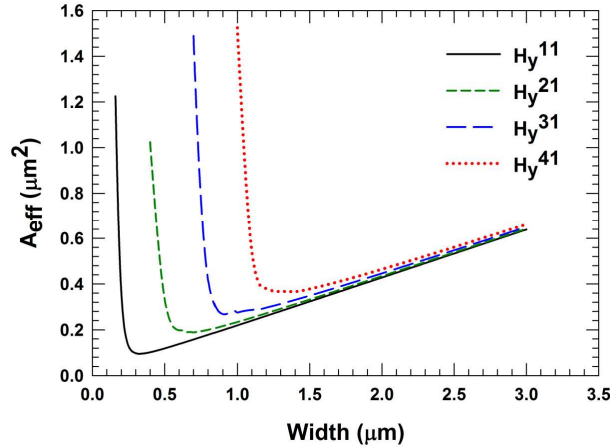

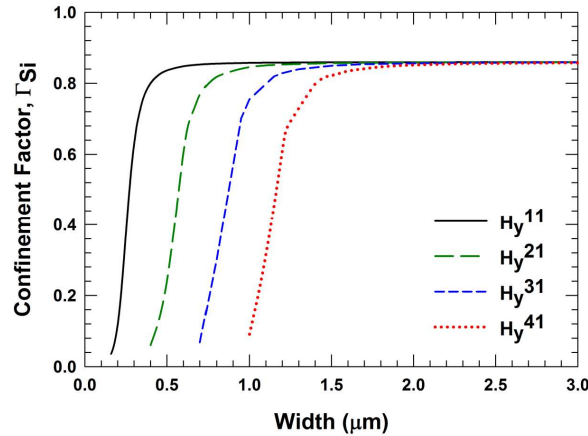Fig. 2. Variations of effective area, A_{eff}, with the width for different quasi-TE modes.

Fig. 3. Variations of power confinement factor in silicon, Γ_{Si}, with the waveguide width, W, for different quasi-TE modes.

3.4 Power confinement

The variations of the power confinement with the width, W, are shown in Fig. 3 for the H_y^{11}, H_y^{21}, H_y^{31} and H_y^{41} modes. The confinement factor in any particular area normalized to the total power, which is obtained by integrating the Poynting vector, from the **H**-and **E**-fields as given below:

$$S_z = \iint_\Omega \left\{ \mathbf{E}^* \times \mathbf{H} \right\}_z \, dx dy \qquad (3)$$

It is expected that as the waveguide dimension becomes large, most of the power would be confined in the Si core and that, Γ_{Si} would be close to 1.0. However, it can be noted that, the maximum power confinement in this case is closer to 0.85, because the height of the core was restricted to 260 nm. If the height of the core also becomes larger, then the power confinement in the Si core could approach 1.0. It can be observed here that as the width is reduced, the power confinement in the Si core also reduces. It can be also observed that although, for a wider waveguide, the power confinement for all the four modes shown here are similar, but for a narrower waveguide, the power confinements for the higher order modes are smaller.

The variations of the effective index, n_e, the effective area, A_{eff} and the power confinement in the SiO$_2$ cladding region, Γ_{SiO2} with the width, W, of the fundamental H_y^{11} mode are shown together in Fig. 4, for their comparison. It can be observed that for the single mode operation, in the case of a 260 nm thick waveguide, the value of the width should lie between 200 nm and 400 nm and this may be chosen optimally to be around 320 nm, when the spot-size is also the smallest. In this case, the power confinement in SiO$_2$ is 0.31 and the remaining power will be in the Si core. From this figure, it can be seen that the power confinement range in SiO$_2$ for single mode operation range could be between 0.89 and 0.20 for those width values varying from 200 nm to 400 nm respectively.
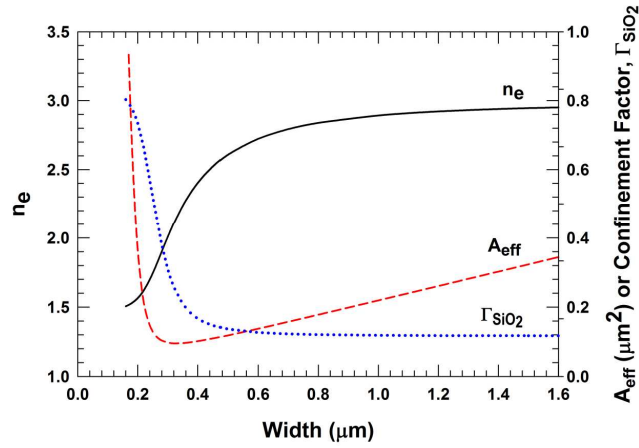

Fig. 4. Variations of n_e, A_{eff} and Γ_{SiO2} with the waveguide width for the H_y^{11} mode.

3.5 Modal **H**-field profiles

For the quasi-TE mode, the H_y field component is dominant, and H_x and H_z are the non-dominant components. The dominant H_y field component of the H_y^{11} mode is shown as an inset in Fig. 5 for the waveguide width, W = 300 nm and height, H = 260 nm.
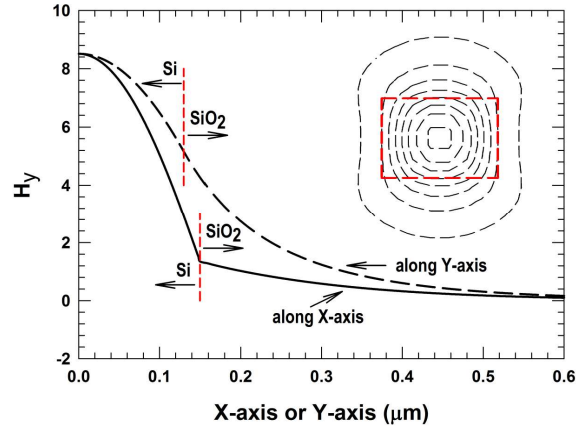
Fig. 5. Variations of H_y along X-axis and Y-axis for the H_y^{11} mode

The field profile shown in Fig. 5 clearly identifies the maximum intensity occurring at the center of the core. It is also shown that there is symmetry along the vertical and horizontal axes, and the mode extends considerably more into the top and bottom SiO$_2$ cladding region. To illustrate its variation more clearly, the variation of the H_y field along the Y-axis is also shown by a dashed line in Fig. 5. In this case, the H_y field is monotonically decreasing along the Y-axis and when y = 0.13 μm (as H = 260 nm), at the boundary between the Si and SiO$_2$, the magnitude of the H_y field is 60% of its maximum value. The variations of the H_y field along the X-axis is also shown by a solid line, which also decreases monotonically but with a rapid reduction inside the Si core until at the boundary interface, following which its slope reduces in the SiO$_2$ cladding region. At the interface, (at x = 0.15 μm, as W = 300 nm) the magnitude of the H_y field is only about 16% of its maximum value, which is significantly lower than field values at the upper and lower interfaces. In its contour plot it is clearly visible that the modal confinement in the horizontal direction is much stronger. The effective index, n_e, of the H_y^{11} mode was found to be 2.00711 with its effective area A_{eff}, equal to 0.098 μm^2, when W = 300 nm.
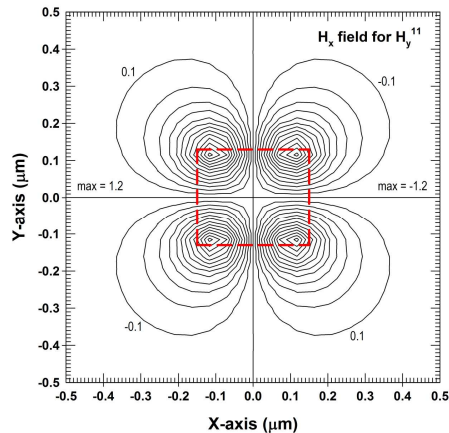

Fig. 6. Non-dominant H_x field profile of the H_y^{11} mode.

As this is a high index contrast Si waveguide, the non-dominant H_x field component of H_y^{11} mode has a significant magnitude, which is shown in Fig. 6. In the case where H = 260 nm and W = 300 nm, the maximum magnitude of H_x field is found to be 19% of maximum H_y field. As the H_x field is related to the spatial derivative of the dominant H_y field, it clearly

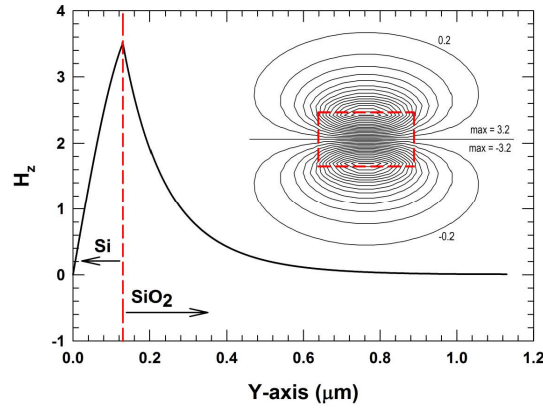shows four peak values at the four corners of the waveguide with alternate positive and negative signs.

Fig. 7. Variations of the H_z field along the Y-axis for the H_y^{11} mode.

The contour of non-dominant H_z field component of the H_y^{11} mode is shown as an inset in Fig. 7. The maximum magnitude of H_z is found to be 41% of the maximum H_y field and it is significantly higher than the non-dominant of H_x field. The H_z field is zero along the X-axis and it can be observed that the maximum intensity occurs at the two horizontal interfaces between the Si and SiO_2. This is because the value of H_z is proportional to the derivative $\partial E_x/\partial y$ and therefore the H_z value is shown to peak in the y-direction. It can also be noted that all the three **H**-field components of the H_y^{11} mode are continuous across the Si/SiO_2 interfaces.

3.6 Hybridness

Modes in optical waveguides with two-dimensional confinement are not truly TE or TM, but hybrid in nature, which means that all the six components of the vector magnetic and electric fields are always present. For the quasi-TE H_y^{11} mode, the H_y component is the dominant and H_x and H_z each are non-dominant. The variations of the modal hybridness with the width for the H_y^{11}, H_y^{21}, H_y^{31} and H_y^{41} modes are shown in Fig. 8. The modal hybridness in this case is defined as the ratio of the maximum value of non-dominant H_x field to the maximum value of dominant H_y field. This is an important parameter to calculate polarization cross-talk [11] and also in the design of polarization rotators [12].
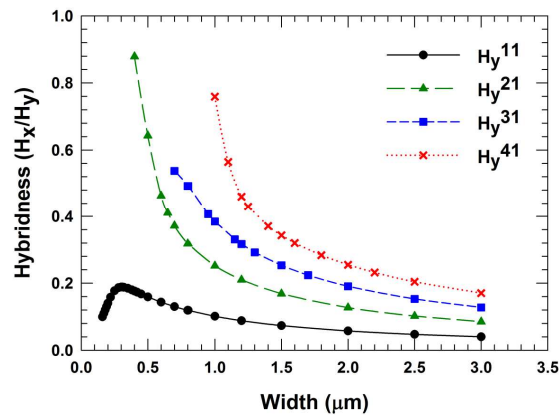

Fig. 8. Variations of H_y hybridness with the width for the H_y^{11}, H_y^{21}, H_y^{31} and H_y^{41} modes.

For the H_y^{11} mode its hybridness is shown by a solid line in Fig. 8. This is low when the width is large, and as W reduces, the hybridness reaches its maximum, then slowly reduces as the fundamental mode approaches its cut-off region when the A_{eff} also increases. The maximum hybridness for the fundamental H_y^{11} mode is found to be 0.19 at W = 320 nm when its effective area was also the smallest. Furthermore, the behavior for the higher order modes (H_y^{21}, H_y^{31} and H_y^{41}) are shown to be similar to that of the fundamental mode, in which, as the width decreases, the hybridness increases until each higher order modes approaches its cut-off regions.
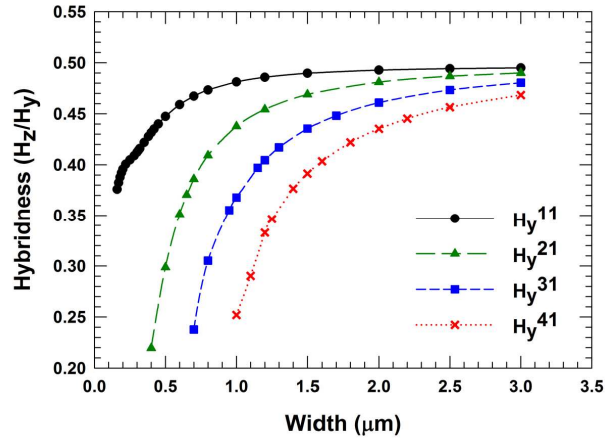

Fig. 9. Variations of H_z hybridness with the width for the H_y^{11}, H_y^{21}, H_y^{31} and H_y^{41} modes.

For the non-dominant H_z field, the hybridness can also be defined as the ratio of H_z/H_y, and this parameter is a function of the width, is shown in Fig. 9. It can be observed that the hybridness increases as the width increases. It is also shown that the H_z hybridness of the fundamental H_y^{11} mode is significantly larger than its H_y hybridness. The H_z hybridness of the fundamental and higher order modes follows a similar behavior, in which, they all decreases as width decreases. In this case, with W = 300 nm and H = 260 nm, the Si core waveguide is asymmetric. However, it was observed that in the case of when height varied with a fixed width, the H_z hybridness increases as the waveguide height decreases (but this is not shown in here).
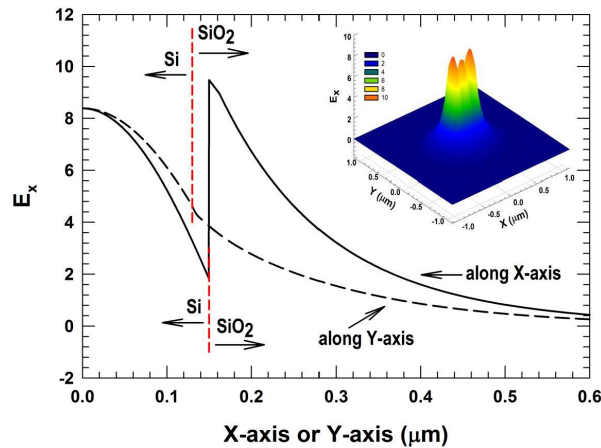

Fig. 10. Variations of the E_x field along the X and Y-axes for the H_y^{11} mode.

*3.7 Modal **E**-field profiles*

Once all the three components of the vector **H**-field are obtained, Maxwell's $\nabla \times \mathbf{H}$ equation is used to calculate the three components of the **E**-field vector. The fundamental quasi-TE H_y^{11} mode contains all the three components of the electric field, E_x, E_y and E_z: the E_x field is dominant, and the E_y and E_z field components are non-dominant. Variations of the dominant E_x field along the X and Y-axes for the fundamental H_y^{11} mode are shown in Fig. 10 for the situation when W = 300 nm and H = 260 nm. The variation of the E_x field along the Y-axis, shown here as a dashed-line, reduces monotonically from the center of the waveguide core and is continuous at the interface between Si/SiO$_2$, as required by the boundary condition. The SiO$_2$ and Si interfaces are shown by two vertical dashed lines. However, the E_x field along the X-axis, shown here as a solid line, reduces more quickly in the core and at the Si/SiO$_2$ interface increases abruptly with a step change in the ratio of $(3.5/1.5)^2 = 5.44$ in the SiO$_2$ region. Therefore the magnitude of the E_x field in the SiO$_2$ region can be significantly higher than that in the core region and this behavior is also shown in the 3D-contour of E_x field, this being shown as an inset in Fig. 10.
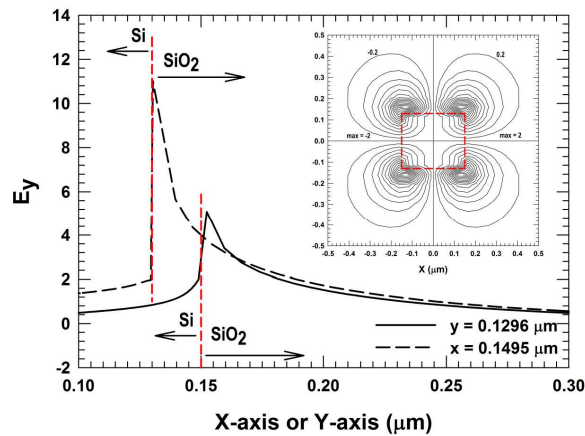

Fig. 11. Variations of the E_y field along the X and Y-axes for the H_y^{11} mode.

The contour of the non-dominant E_y field is illustrated here as an inset in Fig. 11. From the E_y contour, it can be observed that the **E**-field is zero along the X and Y-axes. So the variations of E_y inside the core area, but slightly away from the boundary interfaces are shown in Fig. 11. The variations of E_y along the Y-direction at x = 0.1495 µm is shown by a dashed-line and the variations of E_y along the X-direction at y = 0.1296 µm is shown further as a solid line. It can also be seen that the E_y field is small inside the core and increases along the Y-direction. There is a step change noticeable at the SiO$_2$ and Si interface in the Y-direction, with E_y being normal to this interface and seen as discontinuous here. On the other hand, although E_y also peaks at the Si/SiO$_2$ interface along the X-direction, but this was more gradual. It can be noted that while the E_x field was more concentrated in the core region (shown in Fig. 10), the E_y field is predominantly present around the Si/SiO$_2$ interfaces, as can be seen from Fig. 11.
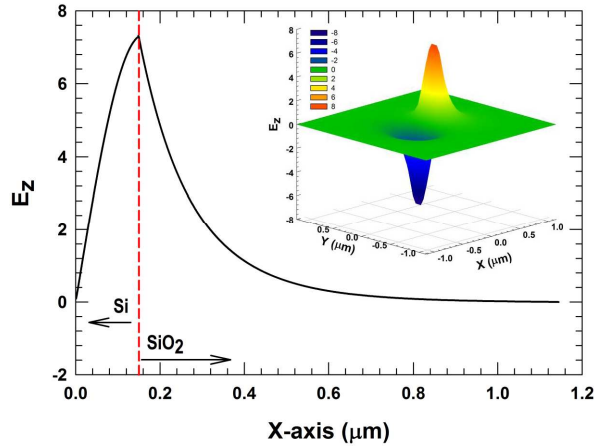
Fig. 12. Variations of the E_z field along the X-axis for the H_y^{11} mode.

The 3D-contour of non-dominant E_z field component of the H_y^{11} mode is shown as an inset in Fig. 12. It can be noted that E_z is zero along the Y-axis. The variation of the E_z field along the X-axis is also shown here. It can be observed that the maximum value of the longitudinal E_z field is 80% of the maximum E_x at the Si/SiO$_2$ interface and then is noted to reduce gradually in the SiO$_2$ region. Such a high longitudinal field can be used in the beam shaping, but this can also influence its polarization properties.
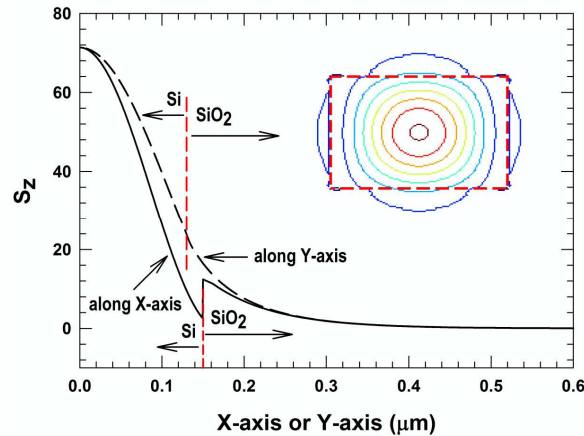

Fig. 13. Variations of the S_z intensity along the X and Y-axes for the H_y^{11} mode.

3.8 Poynting vector profiles

From the full vectorial **E** and **H** fields, the Poynting Vector (S_z) may be calculated. The contour of the S_z intensity profile is shown as an inset in Fig. 13, where the core of the rectangular Si waveguide is also outlined by dashed lines. The variations of S_z along the X and Y-axes are also shown here. Variations of the S_z along the Y-axis are shown by a dashed-line, which clearly indicates that it reduces monotonically and continuous at the interface and extends more along the Y-axis. However, the variation of the S_z along the X-axis is shown by solid line which reveals a small discontinuity step at the Si/SiO$_2$ interface and it decays more quickly in the Si core. It was observed, (but not shown here), that for a smaller waveguide width, the S_z discontinuity step can be even larger with a significant S_z field extending into the SiO$_2$ cladding in the horizontal X-direction.
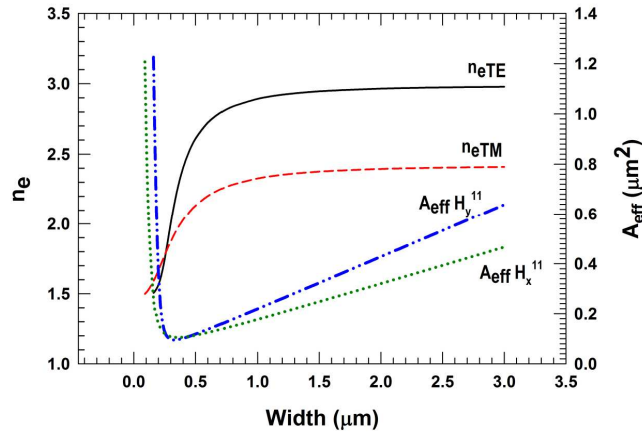
Fig. 14. Variations of the n_e and A_{eff} with the width for the H_y^{11} and H_x^{11} modes.

The variations of effective indices and effective areas for the fundamental H_y^{11} and H_x^{11} modes are shown in Fig. 14. In this study, mostly the effective index for the quasi-TE H_y^{11} mode is larger than the quasi-TM H_x^{11} mode: however when the waveguide's width became smaller, the difference between the effective indices of the H_y^{11} and H_x^{11} modes becomes smaller, and they are equal when the waveguide becomes a square nano-scale waveguide with $W = H = 260$ nm.
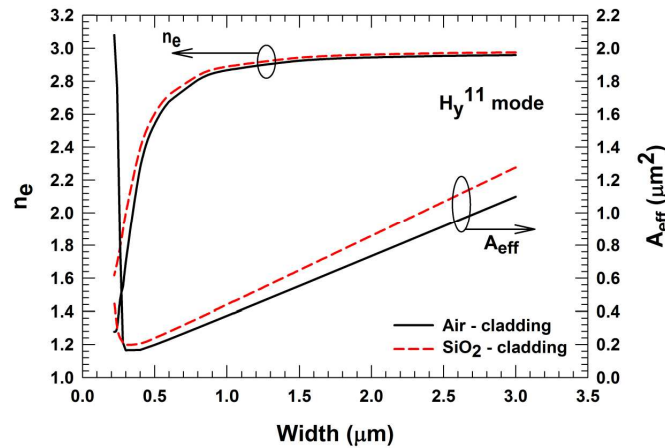

Fig. 15. Variations of the n_e and A_{eff} with SiO$_2$ or Air cladding for the H_y^{11} mode

In the design of PIC, some silicon guided-wave components may not be surrounded by SiO$_2$ cladding but be in air and the variations of their effective indices and the effective areas with the width are rigorously investigated. Here, the air-clad structure has only got one-fold symmetry, so half of the waveguide is used in the simulation. Figure 15 shows the variations of effective indices and effective areas for the fundamental H_y^{11} mode with the width for both air and SiO$_2$ cladding. It can be observed that when the waveguide is covered with SiO$_2$ both the effective indices and the effective areas are slightly higher than that with the air-cladding.
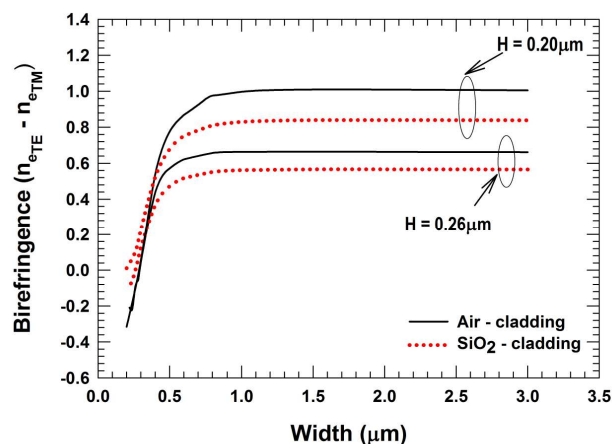
Fig. 16. Variations of the modal birefringence with the width for the fundamental modes with silica or air cladding.

Finally, the variations of modal birefringence are shown in Fig. 16, for the waveguide heights, 200 nm and 260 nm. Here, the modal birefringence is defined as the difference between the effective indices of the fundamental quasi-TE and TM modes. It is shown here that modal birefringence increases as W increase and then saturates. It can also be observed that with the air-cladding waveguide, the modal birefringence is higher than that with the SiO$_2$-cladding due to a stronger refractive index contrast between the Si core and upper cladding. It can be noted that the modal birefringence difference is larger when the height of the waveguide is smaller as in this case the waveguide asymmetry is increased. It is also shown here that the birefringence is zero when the waveguide's width and height are the same.

4. Conclusion

There is an impetus to develop low-cost photonic devices that can be furthered by exploiting the well-developed, low-cost CMOS technology. The higher index contrast, Δn, of silicon also allows for smaller waveguides and compact bend designs which will allow more reliable PIC to be developed, with increased functionality. It is shown here that for a silicon waveguide with a strong index contrast, the single mode operation region is 200 nm to 400 nm, when the waveguide height is 260 nm and the operating wavelength is 1550 nm. However, it is shown here that non-dominant field components are significantly higher, (often more than 50%), in which case the simpler scalar approaches often considered would be totally inadequate. It is also shown that, for this case (H = 260 nm), the effective area not only is minimum, when W = 320 nm, but also these values are similar for both the polarizations. In this case, the modal birefringence is also small and spot-size is near circular and could be a suitable dimension for many applications.

In this study, the spatial variations of all the six components of the **E** and **H**-fields are shown, together with the S_z profile. The modal hybridness, the power confinement and the modal birefringence for silica and air-clad Si nanowires are also shown here. It should be noted that the performance of these highly birefringence guided-wave devices are polarization dependent. However, in the presence of slanted side walls and bends, there may be additional polarization conversion that occurs due to the presence of higher magnitude of the non-dominant field components and this can only be studied by using a rigorous full-vectorial approach, as was done in this work.