Silicon-Photonic Electro-Optic Phase Modulators Integrating Transparent Conducting Oxides

Georgios Sinatkas, *Student Member, IEEE,* and Emmanouil E. Kriezis, *Senior Member, IEEE*

***Abstract*—Higher-order digital modulation formats are demonstrated by electrically inducing free-carrier concentration changes in thin films of transparent conducting oxides, integrated into well-established silicon-photonic waveguiding architectures. The proposed near-infrared modulators employ as physical platforms the silicon-rib and silicon-slot waveguides, exploiting the highly dispersive and carrier-dependent epsilon-near-zero behavior of transparent conducting oxides to modulate the optical carrier. Advancing the existing studies on conventional amplitude modulation, phase-shift keying formats are investigated in this work, using a rigorous and physically consistent modeling framework that seamlessly combines solid-state physics with Maxwell wave theory through carrier-dependent material models. The designed in-line modulators achieve $V_\pi L$ products in the order of 0.1 Vmm, two orders of magnitude lower than their respective all-silicon or lithium niobate counterparts, accompanied by an insertion loss of about 3 dB$/\pi$. Switching speeds in the order of 50 GHz are feasible along with a potential for sub-pJ/symbol energy consumption, meeting the demands for on-chip optical modulation.**

***Index Terms*—Silicon photonics, Electrooptic modulators, Charge carriers, Phase modulation, Indium tin oxide.**

I. Introduction

Recent advances in materials science provide new outlooks for contemporary research on high-speed optical modulators. Stepping forward from the mature and well-established lithium-niobate ($LiNbO_3$) technology [1], modern material classes have attracted considerable attention as potential candidates for providing dynamically configurable media, including highly nonlinear organic molecules [2], two-dimensional (2D) materials (e.g. graphene) [3] as well as the here employed transparent conducting oxides (TCOs) [4]. TCOs uniquely exhibit an epsilon-near-zero (ENZ) behavior in the near-infrared (NIR) spectrum, allowing the TCO permittivity to range from dielectric to metallic values through tuning their free-carrier concentration.

The carrier-dependent and highly dispersive TCO permittivity in the NIR has already been exploited in plenty of amplitude-shift keying (ASK) designs, implementing the fundamental on-off keying (OOK) modulation scheme for either free-space [5], [6] or on-chip applications. In the latter case, both silicon-photonic [7]–[11] and (hybrid-) plasmonic platforms [12]–[22] have been investigated. The electro-optic (EO) control is usually achieved through field-effect changes in the carrier concentration of thin TCO films.

Manuscript submitted June 28, 2018. This work was supported by the "Research Projects for Excellence IKY/Siemens".

G. Sinatkas, and E. E. Kriezis are with the School of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, GR, 54124 Greece (e-mail: gsinatka@auth.gr).

This work steps forward to study theoretically on-chip phase-shift keying (PSK) modulators with TCOs, which are more favorable in state-of-the-art telecom applications, having though attracted significantly less attention [23]. The investigated PSK modulators integrate TCO films in two of the most representative examples of silicon-photonic waveguides, the silicon-rib (Si-rib) and silicon-slot (Si-slot) designs. The modulators are studied using a rigorous and physically consistent framework, unifying both solid-state and electromagnetic physics on a common finite-element (FEM) platform [10]. This enables the accurate modeling of the EO effect, avoiding the unrealistic and approximate methods occasionally employed in the literature (Appendix).

The proposed configurations can efficiently demonstrate both binary (BPSK) and quadrature phase shift keying (QPSK) formats using in-line architectures of TCO-based waveguides. The necessary length does not exceed few hundreds of micrometers for both modulation schemes, with the BPSK insertion loss (IL) calculated in the order of 3 dB, increased to 5 dB for the QPSK scheme as an exchange to increased bit rates. The average energy consumption is rigorously calculated in the order of 1 pJ/symbol, with the Si-slot design outperforming its Si-rib counterpart in terms of efficiency due to enhancing the light-matter interaction. The proposed modulators additionally support switching speeds of several tenths of GHz, calculated through strict temporal studies, suggesting a viable solution for on-chip optical modulation.

The remainder of this paper is organized as follows: In Section II, the underlying physical platforms and the EO switching mechanism of the TCO-based modulators are presented. The study of a TCO-based phase shifter follows in Section III, employed subsequently in Section IV as the primary building block for implementing the proposed modulators. The work is concluded in Section V.

II. Physical Platforms

As underlying physical platforms, the Si-rib and Si-slot platforms are examined, Fig. 1. Both waveguides are loaded with a 10-nm thick indium tin oxide (ITO) layer, separated from the underlying n-doped Si by a 5-nm hafnium dioxide (HfO_2) spacer. A moderate 10^{18} cm^{-3} donor concentration is selected for Si, while a free-electron concentration equal to 10^{19} cm^{-3} is assumed for ITO, adequately high for ensuring its metallic state according to the Mott criterion [24], but sufficiently low for avoiding excessive optical losses. ITO is selected as the most widely studied and used TCO, providing mature fabrication and integration processes along with low

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JQE.2018.2852144, IEEE Journal of Quantum Electronics

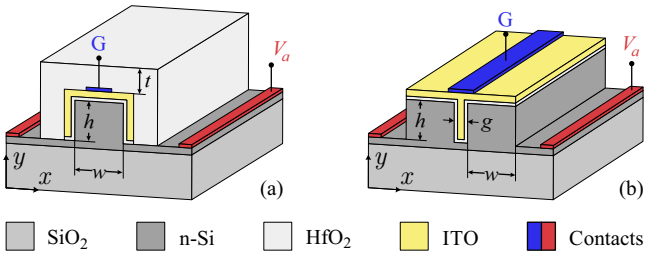

Fig. 1. (a) Si-rib waveguide, conformally coated by two successive layers of HfO_2 and ITO, cladded on top by a second HfO_2 layer to aid waveguiding ($t = 150$ nm). (b) Si-slot waveguide, with the slot ($g = 20$ nm) occupied by two layers of HfO_2 and ITO. In both designs, it is $w \times h = 180 \times 220$ nm^2, while the HfO_2, ITO, and silicon-slab thicknesses equal 5 nm, 10 nm, and 30 nm, respectively. An applied bias V_a between the n-doped Si and ITO induces the field-effect carrier-concentration changes in ITO.

resistance values [25]. Its properties are assumed thickness-independent, practically achieved by carefully controlling the deposition conditions [24].

The successive n-Si/HfO_2/ITO layers form a capacitor-like structure, which resembles the traditional metal-oxide-semiconductor (MOS) gate contact of field-effect transistors (MOSFETs). This structure is similarly used here for inducing carrier concentration changes in the waveguides by applying a bias difference between n-Si and ITO. Electrons are accumulated (depleted) in ITO under the application of a positive (negative) bias as shown in Fig. 5 of [10]. Equal and opposite electric charge is developed in the silicon layer in accordance with the charge conservation law. These carrier concentration changes modify in turn the NIR permittivity of the underlying materials, modulating eventually the complex effective refractive index, n_{eff}, of the guided mode. Due to the significantly higher dispersion of TCOs in the NIR, the effect is mainly dictated by the changes in the ITO layer.

Due to its high free-electron concentration, the ITO thickness can be well restricted to 10 nm, avoiding this way an unnecessarily excessive overlap with the guided wave, reducing as well the fabrication cost. The use of the high-k HfO_2 material, instead of the conventionally employed silicon dioxide (SiO_2), is strongly suggested by its higher dielectric constant ($\bar{\varepsilon}_{\text{HfO}_2}/\bar{\varepsilon}_{\text{SiO}_2} = 25/3.9$) [26], supporting the very same field effect at lower biasing values and, thus, reducing the necessary energy consumption [10]. At the same time, its breakdown field [27] is adequately high to sustain the biasing conditions employed in this work.

Both (quasi) transverse electric (TE) and transverse magnetic (TM) modes are supported by the Si-rib structure in Fig. 1(a) with the former exhibiting enhanced interaction with ITO due to its double-sided overlap [10], and thereby solely examined hereafter. By opting for a narrow silicon-core ($w = 180$ nm), weak waveguiding conditions are established, increasing the interaction with the superimposed ITO layer, while the use of the HfO_2 cladding prevents the mode from becoming leaky. Other high-index materials, such as chalgogenides glasses [28], can be equally preferred as dielectric claddings from an electromagnetic standpoint. For the Si-slot platform, the TE-polarized slot mode is of interest, being highly confined in the slot [10] and, thus, enhancing the light-matter interaction effects. Strong waveguide conditions are preferable in this case, achieved by reducing the slot width down to a few decades of nanometers using state-of-the-art lithography processes [29], [30]. Regarding the silicon-core height, the conventional 220 nm value is considered for both designs. The operating wavelength is $\lambda = 1.55$ μm and the respective material parameters are listed in Table I. The refractive indices for the semiconductor materials correspond to their equilibrium carrier concentration values.

TABLE I
MATERIAL REFRACTIVE INDICES AT 1.55 μm.

Material	Refractive Index
Silicon dioxide (SiO_2)	1.45
Hafnium dioxide [31] (HfO_2)	2.07
Silicon [32] (Si)	$3.48 - j1.11 \times 10^{-4}$
Indium tin oxide [24], [33] (ITO)	$1.98 - j1.12 \times 10^{-3}$

The effect of the metal electrodes will not be considered in the subsequent calculations. It would generally result in a moderate increase of the total IL due to the excitation of the highly lossy surface plasmon polariton modes (SPPs), reversible, though, by proper engineering. The top ITO electrode does not favor the excitation of SPPs due to both examined modes being in principal TE polarized. The n-Si contacts result in a more pronounced effect, especially in the case of the weakly-guided Si-rib mode. The latter can be mitigated by properly regulating the geometric and material contact properties.

The EO switching effect is modelled by coupling the physics of wave optics, macroscopically described by the Maxwell's equations, with the theory of solid-state physics, using classical, continuum medium models for the NIR permittivity. This multiphysics integration takes place on a finite-element method platform, COMSOL Multiphysics ®, avoiding the approximate techniques employed in the literature that occasionally overestimate the effect and result in spurious behavior (see Appendix). A detailed description of the modeling framework can be found in [10].

III. PHASE SHIFTER

The first step toward implementing phase modulation formats is the design of an elementary phase-shifting unit. A ϑ-radian phase shift ($\vartheta \geq 0$) can be achieved after a waveguide length L_ϑ equal to

$$L_\vartheta = \frac{\vartheta}{k_0 \left|\Delta \text{Re}\left\{n_{\text{eff}}\right\}\right|} = \frac{\vartheta}{\pi} \frac{\lambda/2}{\left|\text{Re}\left\{n_{\text{eff}}(V_a) - n_{\text{eff},0}\right\}\right|}, \quad (1)$$

inversely related to the absolute difference between the real parts of the effective indices at the reference unbiased state, $n_{\text{eff},0}$, and at a selected biased state, $n_{\text{eff}}(V_a)$. In (1), k_0 stands for the free-space wavenumber.

In Fig. 2(a), the signed difference $\Delta \text{Re}\left\{n_{\text{eff}}\right\}$ and the mode loss α in dB/μm are calculated for the Si-rib and Si-slot TE modes as functions of the applied bias V_a. The results are derived by studying an 1D horizontal waveguide cut, parallel

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JQE.2018.2852144, IEEE Journal of Quantum Electronics

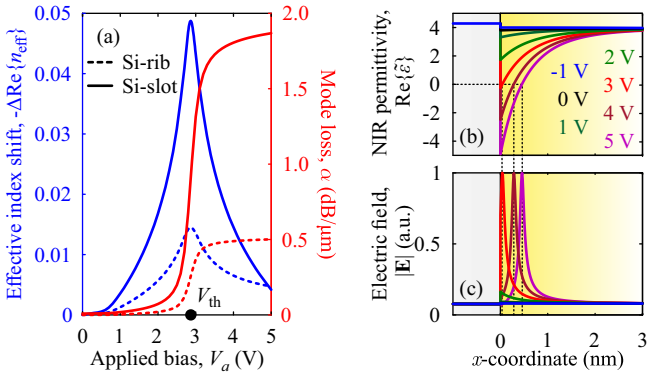

Fig. 2. (a) Signed difference $\Delta\mathrm{Re}\{n_{\mathrm{eff}}\}$ and mode loss α in $\mathrm{dB}/\mu\mathrm{m}$ as functions of the applied bias V_a for the 1D Si-rib (dashed lines) and Si-slot (solid lines) geometries. In both designs, the HfO_2 and ITO thicknesses equal 5 nm and 10 nm, respectively. (b) Spatial distribution of the near-infrared permittivity near the HfO_2/ITO interface for the 1D Si-rib waveguide under the application of seven biasing values, with the respective electric-field profiles given in (c). For $V_a \approx 3$ V, ITO enters the ENZ region, resulting in a field peak that shifts toward the inner ITO as the bias increases, closely following the respective ENZ-region shift in (b) (vertical dashed lines).

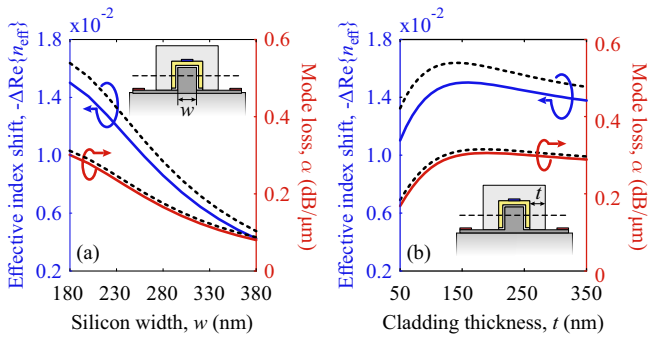

Fig. 3. One-dimensional evaluation of the geometry effect on the phase-shifting potential of the Si-rib waveguide with respect to its (a) silicon width w and (b) cladding thickness t, denoted in the respective insets. The effect is quantified by using as figures of merit the peak change of the real index part, $\Delta\mathrm{Re}\{n_{\mathrm{eff}}\}$, manifested between the unbiased state and $V_{\mathrm{th}} = 2.9$ V, as well as the respective IL penalty, calculated at $V_{\mathrm{th}} = 2.9$ V in $\mathrm{dB}/\mu\mathrm{m}$. The dashed lines incorporate additionally the free-carrier effect in silicon.

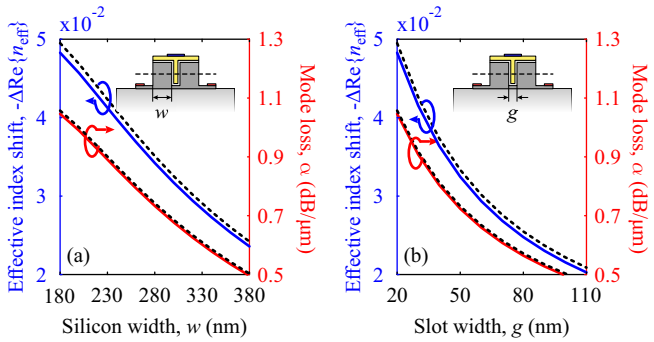

Fig. 4. One-dimensional evaluation of the geometry effect on the phase-shifting potential of the Si-slot waveguide with respect to its (a) silicon width w and (b) slot width g, denoted in the respective insets. The effect is quantified by using as figures of merit the peak change of the real index part, $\Delta\mathrm{Re}\{n_{\mathrm{eff}}\}$, achieved between the unbiased state and $V_{\mathrm{th}} = 2.9$ V, as well as the respective IL penalty, calculated at $V_{\mathrm{th}} = 2.9$ V in $\mathrm{dB}/\mu\mathrm{m}$. The dashed lines incorporate additionally the free-carrier effect in silicon.

to the polarization of the dominant electric-field component E_x. The downgraded 1D structure is solved for the TM mode, having the E_x component lying normal to ITO in order for the ENZ effect to manifest itself [10].

The study calculations show a sharp mode-loss increase around the threshold bias $V_{\mathrm{th}} = 2.9$ V for both designs, coinciding with the onset of the ENZ region at the HfO_2/ITO interface, as shown in Fig. 2(b). The high losses are maintained for $V_a > V_{\mathrm{th}}$ with the ENZ region simply being shifted further from the HfO_2/ITO interface, toward the inner ITO part. The electric-field peak is respectively shifted, Fig. 2(c), closely following the formation of the ENZ region and maintaining the highly lossy state.

At the same time, the real index part exhibits a non-monotonic behavior, with $|\Delta\mathrm{Re}\{n_{\mathrm{eff}}\}|$ peaking at V_{th} for both designs, decreasing nearly symmetrically either side. A shift in the second decimal place is evidenced with respect to the unbiased state, an order of magnitude larger than the plasma-dispersion effect in silicon [34], justifying the interest in TCO-based phase-shifting units [35]. Contrasting the performance between the two platforms, the effect in the Si-slot design is significantly enhanced compared to its Si-rib counterpart, justifying the interest in both platforms despite the possible fabrication challenges of the Si-slot design.

Both geometries evaluated in Fig. 2(a) have been carefully engineered using the parametric studies presented in Figs. 3-4. The interaction is quantified in terms of the maximum (in absolute value) attainable $\Delta\mathrm{Re}\{n_{\mathrm{eff}}\}$, achieved between the unbiased state and V_{th}, taking as well into account the respective loss penalty, calculated at V_{th} in $\mathrm{dB}/\mu\mathrm{m}$. In Fig. 3, the impact of both silicon width w and cladding thickness t on the Si-rib platform are parametrically examined, justifying setting them equal to 180 nm and 150 nm, respectively. Similarly, a silicon width w equal to 180 nm is selected for the Si-slot platform, Fig. 4(a). However, in contrast to the optimally weak waveguiding conditions of the TE Si-rib mode, strong waveguiding is preferable in this case, Fig. 4(b). Opting for a 20-nm wide slot may appear challenging from a fabrication standpoint, yet is totally feasible by the recent advances in the contemporary, state of the art lithography technology [29], with [30] suggesting a promising alternative for successively occupying the 20-nm slot with the HfO_2 and ITO layers. The dashed lines in Figs. 3 to 4 incorporate additionally the carrier concentration changes in silicon, not taken into account up to this point. It is evidenced that the silicon effect results just in small offsets to the solid-line results, beneficially contributing to the index change and introducing only a slightly higher loss penalty, mitigated eventually by the respective decrease in the necessary phase-shifting length. However, it is considerably weaker compared to the highly dispersive ITO behavior and it will not be further considered.

Eventually, the phase-shifting efficiency of the investigated TCO-loaded platforms is evaluated in Fig. 5, after rigorously solving for both 2D solid state and electromagnetic boundary value problems on the optimized waveguide cross-sections (solid lines). As a performance figure of merit, the normalized shifting length L_{rad} is employed, defined by dividing (1) with the targeted θ-radian phase shift, expressing the necessary

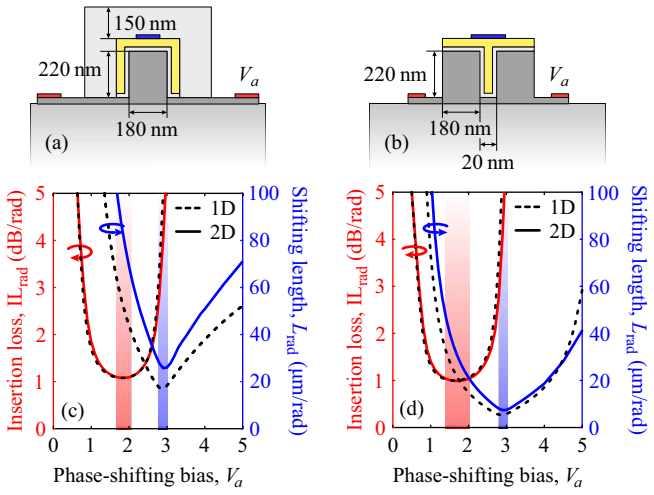

Fig. 5. (a) Si-rib and (b) Si-slot cross-sectional geometry, optimized for use in phase-shifting applications. In (c) and (d) the necessary interaction length $L_{\rm rad}$ for achieving a single radian phase shift is plotted as a function of the phase-shifting bias V_a for the Si-rib and Si-slot waveguides in (a) and (b), respectively, along with the resulting total loss $\mathrm{IL}_{\rm rad}$. The solid lines refer to rigorous 2D studies, being in close agreement to the respective 1D results, represented by the dashed lines.

TABLE II
COMPARISON OF PHASE MODULATION EFFICIENCY IN NANOPHOTONIC COMPONENTS.

Material platform	$V_\pi L_\pi$ (Vmm)	$\lvert V_\pi \rvert$ (V)	L_π (μm)	$\mathrm{IL}_{\rm rad}$ (dB/rad)
$LiNbO_3$ [1]*	36.0	9.0	4000	0.60
Carrier-effects in Si [36]*	8.50	10	1000	1.90
Silicon-organic hybrid [2]*	11.0	22	500	0.64
Single-layer graphene [37]	5.20	6.1	850	0.84
This work (Si-slot)	0.15	1.8	84	1.00

* Experimental works

length for achieving a single rad of phase shift. In the same figure, the respective $\mathrm{IL}_{\rm rad}$ penalty is plotted, calculated as the product between $L_{\rm rad}$ and the per-unit-length mode loss, providing the total loss per rad of phase shift. Both metrics are examined as functions of V_a for values up to 5 V. The dashed lines correspond to the respective 1D calculations, proven in relatively close agreement with the rigorous 2D solution, encouraging the preliminary 1D evaluation due to its considerably reduced computational cost.

In Fig. 5, two opposing design options are identified, each of which emphasizing either on compact or low-loss devices, with the respective biasing regions highlighted with blue and red shaded areas, respectively. Specifically, a minimum $L_{\rm rad}$ is achieved for both designs at $V_{\rm th} = 2.9$ V, coinciding with the peak change in $\mathrm{Re}\{n_{\rm eff}\}$, Fig. 2(a), but also with the onset of the highly lossy ENZ-region. As a result, even though a significantly reduced shifting length as short as 7 μm/rad (25 μm/rad) can be achieved for the Si-slot (Si-rib) platform, an excessive loss equal to 3.9 dB/rad (3.7 dB/rad) has to be tolerated. The latter falls down to the ultimate low of 1 dB/rad in the red shaded area, at

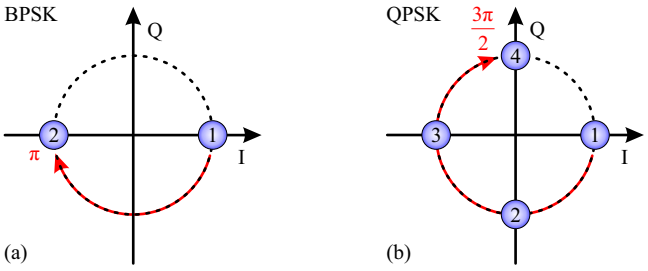

Fig. 6. Constellation diagrams for the (a) BPSK and (b) QPSK modulation format. A maximum of π- and $3\pi/2$-radian phase shift is necessary with respect to a reference state for implementing the BPSK and QPSK constellation diagram, respectively.

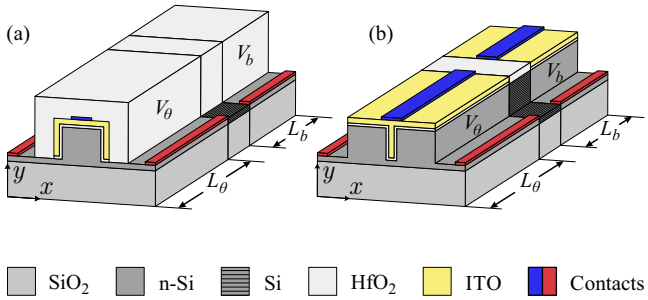

Fig. 7. Dual-segment modulation configurations implementing both BPSK and QPSK modulation formats, developed on the (a) Si-rib and (b) Si-slot platforms. Each segment is independently biased, being electrically insulated from one another by an intermediate, arbitrarily long, buffer waveguide of undoped silicon.

a phase-shifting bias equal to 1.8 V (2 V) for the Si-slot (Si-rib) platform, having to compromise though for a longer shifting length around 27 μm/rad (69 μm/rad). Targeting the design of low-loss phase shifters hereafter, biasing values around 2 V are selected, lying below the $V_{\rm th}$ threshold. This constitutes an essential deviation from the design principles of the TCO-based amplitude modulators [10], [22], which capitalize instead on the step-like loss behavior of the ENZ effect.

IV. PHASE MODULATION FORMATS

The designed phase-shifting units are favorable for several length-critical applications, including delay lines and interferometric configurations. An indicative comparison of their phase modulation efficiency to selective contemporary nanophotonic components is compiled in Table II; though a direct comparison may not equally apply to all cases, due to experimental data being included as well, the proposed designs demonstrate a strong potential in nearly all metrics. Most notably, a record low $V_\pi L_\pi$ product is predicted due to the simultaneous reduction in both V_π and L_π values. In this work, the designed phase-shifting units are used as building units for demonstrating PSK modulators, starting from the fundamental BPSK scheme and continuing to the more elaborate QPSK, Fig. 6.

TABLE III
DESIGN SPECIFICATIONS FOR THE SI-RIB AND SI-SLOT BPSK MODULATORS, INCLUDING THEIR SEGMENT LENGTHS L_θ AND L_b AND THEIR BIASING CONDITIONS AT THE j-TH STATE.

Design	Lengths (μm)		Biasing (V)			IL (dB)
	L_θ	L_b	j	$V_\theta{}^{(j)}$	$V_b^{(j)}$	
Si-rib	224	8.1	(1)	0.0	5.0	3.6
			(2)	2.0	0.0	3.6
Si-slot	85.3	2	(1)	0.0	5.0	3.2
			(2)	1.8	0.0	3.2

TABLE IV
DESIGN SPECIFICATIONS FOR THE SI-RIB AND SI-SLOT QPSK MODULATORS, INCLUDING THEIR SEGMENT LENGTHS L_θ AND L_b AND THEIR BIASING CONDITIONS AT THE j-TH STATE.

Design	Lengths (μm)		Biasing (V)			IL (dB)
	L_θ	L_b	j	$V_\theta{}^{(j)}$	$V_b^{(j)}$	
Si-rib	336	12.2	(1)	0.0	5.0	5.3
			(2)	1.1	3.1	5.3
			(3)	1.5	2.9	5.3
			(4)	2.0	0.0	5.3
Si-slot	128	2.9	(1)	0.0	5.0	4.8
			(2)	1.0	3.1	4.8
			(3)	1.4	2.9	4.8
			(4)	1.8	0.0	4.8

A. Binary phase-shift keying (BPSK) modulator

A BPSK modulation scheme requests a π-radian phase difference between its binary states. With reference to Fig. 5, a TCO-based phase shifter on the Si-slot (Si-rib) platform is capable of achieving a π-radian phase-shift under a V_π bias equal to 1.8 V (2 V) and after an interaction length L_π of 84 μm (216 μm), introducing 3.1 dB (3.4 dB) of IL. Thus, the $V_\pi L$ product, which is conventionally used for evaluating the phase-shifting performance, is calculated in the order of 0.1 Vmm, two orders of magnitude lower than the respective silicon-based counterparts reported in the literature [38].

To additionally ensure a balanced BPSK constellation, the principal π-radian phase-shifting unit is followed by a separately addressable balancing waveguide segment L_b in cascade architecture as illustrated in Fig. 7. The L_b segment manages to equalize the losses between the two states, equating them to the loss penalty paid for achieving the π-radian phase shift. The length of the phase-shifting unit is suitably adjusted to L_θ. Both lengths can be calculated by solving the linear equation system,

$$\begin{bmatrix} \mathrm{Re}\{{k_\theta}^{(2)} - {k_\theta}^{(1)}\} & \mathrm{Re}\{k_b^{(2)} - k_b^{(1)}\} \\ \mathrm{Im}\{{k_\theta}^{(2)} - {k_\theta}^{(1)}\} & \mathrm{Im}\{k_b^{(2)} - k_b^{(1)}\} \end{bmatrix} \begin{bmatrix} L_\theta \\ L_b \end{bmatrix} = \begin{bmatrix} \pi \\ 0 \end{bmatrix}, \quad (2)$$

with the first relation ensuring the π-radian phase shift between the two states and the second the IL equality. In (2), ${k_i}^{(j)} = \mathrm{Re}\{{k_i}^{(j)}\} - j\mathrm{Im}\{{k_i}^{(j)}\}$ stands for the bias-dependent wavenumber at the j-th modulation state of the i-th segment, with i and j obtaining values in $\{\theta, b\}$ and $\{1, 2\}$, respectively. Judiciously selecting the biasing conditions of the i-th segment at each j-th state allows for straightforwardly solving (2) with respect to the unknown length vector $[L_\theta, L_b]^{\mathrm{T}}$.

The design specifications for both Si-rib and Si-slot platforms are listed in Table III. Evidently, the L_θ segment undertakes the principal phase-shifting operation, while the few-micrometer long L_b segment contributes with a loss-balancing role. The designed BPSK modulators result in an overall IL around 3 dB, with the Si-slot platform reasonably allowing for more compact configurations. It is noted that even more compact designs could be achieved by opting for phase-shifting conditions in the blue shaded area of Fig. 5 instead of the red one, bound though to suffer from a significantly higher IL penalty.

B. Quadrature phase-shift keying (QPSK) modulator

Similar steps can be followed to demonstrate higher order phase-modulation schemes such as the QPSK format, equivalently referred to as four-quadrature amplitude modulation (4-QAM), Fig. 6(b). In this case, the peak phase shift with respect to the reference state equals $3\pi/2$ radians, accumulated according to Fig. 5 after an $L_{3\pi/2}$ interaction length of 127 μm (324 μm) and under a $V_{3\pi/2}$ bias of 1.8 V (2 V) for the Si-slot (Si-rib) waveguide, resulting in a minimal of IL equal to 4.7 dB (5.1 dB). This $3\pi/2$-radian phase shifter is subsequently used as the principal phase-shifting block in designing a QPSK modulator, similarly to the π-radian phase shifter employed for the BPSK modulator.

The conditions holding for the j-th state of the QPSK modulator, where j = 1-4, are similar to those of the BPSK implementation in (2), cumulatively summarized in the following equation system

$$\begin{bmatrix} \mathrm{Re}\{{k_\theta}^{(j)} - {k_\theta}^{(1)}\} & \mathrm{Re}\{k_b^{(j)} - k_b^{(1)}\} \\ \mathrm{Im}\{{k_\theta}^{(j)} - {k_\theta}^{(1)}\} & \mathrm{Im}\{k_b^{(j)} - k_b^{(1)}\} \end{bmatrix} \begin{bmatrix} L_\theta \\ L_b \end{bmatrix} = \begin{bmatrix} \phi \\ 0 \end{bmatrix}, \quad (3)$$

with the first relation representing the phase shift of the j-th state with respect to the reference state, $\phi = (j-1)\pi/2$, while the second equation ensures the IL equality among all four states.

Tailored solution options can be employed for solving (3) at each state, depending on the desired specifications. For minimum IL, it would be sufficient to maintain the peak loss, manifested at the maximum phase-shifted state $j = 4$, as low as possible by selecting its biasing conditions in the red shaded area of Fig. 5. The design specifications for the QPSK modulator are listed in Table IV for both Si-rib and Si-slot platforms. As expected, the QPSK modulators suffer from greater ILs (≈ 5 dB) compared to their BPSK counterparts, with higher order phase-modulation schemes being inevitably more lossy in exchange for increased bit rates.

C. Modulation speed and energy consumption

In Fig. 8, the mode-loss temporal behavior is illustrated under the application of a 20-ps rectangular pulse (dashed

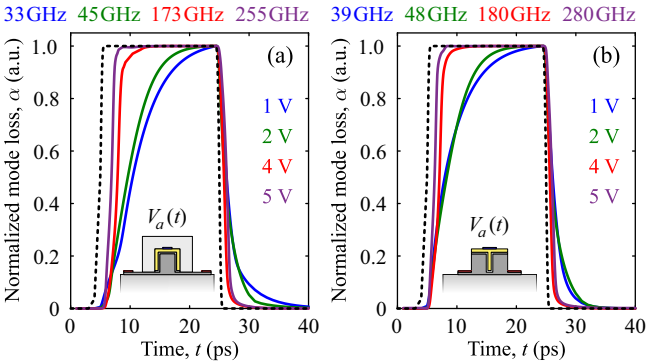

Fig. 8. Temporal behavior of the normalized mode loss for the (a) Si-rib and (b) Si-slot designs under a 20-ps rectangular electrical pulse (dashed line) of 1-ps rise time and parametrically changing amplitude. The supported optical bandwidth for each pulse amplitude is denoted above each subfigure.

line). The $10\% - 90\%$ rise time is set to 1 ps, while its peak amplitude is examined parametrically. To aid visualization, each loss curve has been normalized with respect to its peak value, after being reduced by the IL at the unbiased state.

In principle, due to the field-effect nature of the employed switching mechanism, the modulation speed is solely limited by the resistance-capacitance (RC) constant. The latter is bias-dependent as implied by the decreasing rise times of the loss curves for increasing amplitude values in Fig. 8. This means that the slowest rise time imposes a bottleneck to the ultimate modulation speed. With reference to Tables III and IV, pulse amplitudes approximately as low as 2 V and 1 V are used in the BPSK and QPSK designs, respectively, allowing for switching speeds as high as 48 GHz, even though ultra-high speeds, well-exceeding 100 GHz, are supported for pulses of higher amplitude, denoted in Fig. 8. The modulation bandwidth is calculated by $\mathrm{B} = 0.35/t_r$, where t_r stands for the 10%-90% rise time of the loss curve. The temporal performance of both Si-rib and Si-slot platforms is evaluated in Table V for both BPSK and QPSK schemes. An increase in the modulation speed can be achieved by reducing the silicon resistivity, principally contributing to the total structure resistance. Raising the dopant concentration to about 5×10^{18} cm^{-3}, modulation speeds above 50 GHz are feasible.

The energy demands of the proposed modulators are consistently estimated by summing the energy provided by the voltage source at each *charging* state transition, weighted by the respective occurrence frequency. This method is not based on calculating the CV^2 product, which generally introduces an ambiguity in the case of multi-level modulation formats. Specifically, the energy for switching between two states j and j' equals $W_{jj'} = V_{j'}\Delta Q_{jj'}$, expressing the necessary work for increasing the total charge by $\Delta Q_{jj'} = Q_{j'} - Q_j$ at a $V_{j'}$ potential. The $W_{jj'}$ energy incorporates both the energy dissipated during the charging process and the respective stored in the structure, dissipated eventually during discharge. Thus, $W_{jj'}$ can be considered as the upper intrinsic limit for switching between the j-th and j'-th state.

With reference to the dual-segment modulators in Fig. 7 and with respect to the notation adopted in Tables III and IV, the total energy consumption per symbol can be expressed as

TABLE V
MODULATION SPEED AND ENERGY CONSUMPTION FOR BOTH BPSK AND QPSK FORMATS, IMPLEMENTED ON EITHER SI-RIB OR SI-SLOT PLATFORM.

Design	GHz		pJ/symbol	
	BPSK	QPSK	BPSK	QPSK
Si-rib	45	33	3.8	4.0
Si-slot	48	39	1.4	1.5

$$W_e = \sum_i \sum_{j,j'} p_{jj'} V_i^{(j')} \Delta Q_i^{(jj')}, \tag{4}$$

where i denotes the modulation segment and j, j' the initial and final state, respectively. In (4), only the charging transitions $j \to j'$ should be considered, i.e., $\Delta Q_i^{(jj')} > 0$, with the rest transitions having either no energy penalty, $\Delta Q_i^{(jj')} = 0$, or resulting in the dissipation of the already considered stored energy, $\Delta Q_i^{(jj')} < 0$. The $p_{jj'}$ coefficient weights the contribution of each transition by its occurrence frequency, being equal to $1/4$ and $1/16$ in the case of equiprobable BPSK and QPSK modulation schemes, respectively. The potential $V_i^{(j')}$ coincides with the source bias at the final state j', while the charge difference $\Delta Q_i^{(jj')}$ is given by

$$\Delta Q_i^{(jj')} = L_i \iint_S \left(|\rho_{j'}(\mathbf{r})| - |\rho_j(\mathbf{r})|\right) \mathrm{d}S, \tag{5}$$

with $\rho(\mathbf{r})$ standing for the steady-state space charge density on the cross-section of each L_i-long modulation segment, integrated over either the silicon or ITO area S. The energy consumption is shown in Table V, estimated in the order of 1 pJ/symbol, with the Si-slot platform appearing about three times more efficient compared to its Si-rib counterpart due to providing reduced interaction lengths. Interestingly, the QPSK format exhibits only slightly higher energy demands compared to the BPSK scheme since both implementations share the same biasing range. Sub-pJ/symbol energy consumption can be achieved by opting for shorter designs (blue-shaded area in Fig. 5), having to tolerate though a larger penalty on the IL.

V. SUMMARY AND CONCLUSIONS

Higher-order phase modulators based on TCO-loaded silicon waveguides were rigorously studied, exploiting the highly dispersive and carrier-dependent permittivity of TCOs in the NIR. Both TCO-based Si-rib and Si-slot platforms yield a $V_\pi L$ product in the order of 0.1 Vmm, at the expense of an IL penalty around 3 dB/π, outperforming in terms of phase modulation efficiency existing material platforms. By employing in-line, dual-segment architectures, the design of BPSK and QPSK modulators of few hundreds of micrometers is demonstrated, exhibiting ILs around 3 and 5 dB, respectively. The proposed modulators support a 50 GHz speed potential and a sub-pj/symbol energy consumption, suggesting thereby a viable solution for on-chip optical modulation.

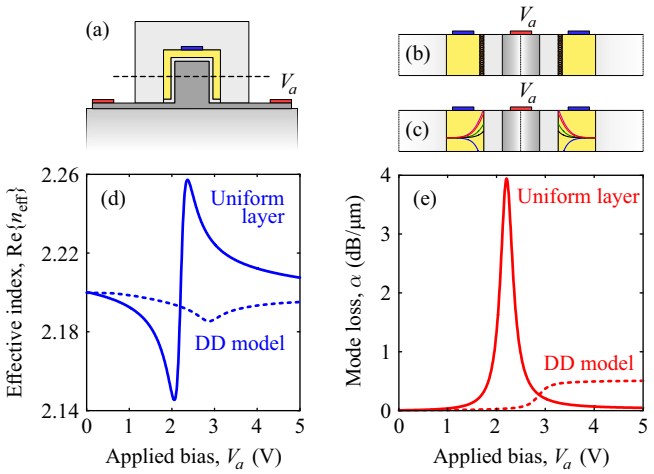

Fig. 9. With reference to the (a) Si-rib waveguide, an 1D comparison between (b) the uniform accumulation layer approach and (c) the rigorous DD model is presented by juxtaposing the respective calculations for (d) the real part of the effective index and (e) the mode loss.

Comparing to existing modulation architectures, the employed dual-segment configuration is much simpler, avoiding complex interferometric or bandwidth-limited resonant structures. For example, the conventional Mach-Zehnder modulators (MZMs) employ challenging Y-junctions and sensitive MZ interferometers, scaling also in complexity for higher-order formats. A QPSK MZM would require a nested configuration in push-pull operation [39], suffering additionally from large residual amplitude modulation in the case of TCO-based phase shifters. This effect is largely mitigated in the proposed dual-segment architecture, with the modulation performance depending rather on the reliability of the driving circuit for achieving precise phase swings.

In conclusion, this work contributes to the highly active field of TCO-based modulators by rigorously investigating the less studied, yet more favorable, PSK modulation schemes. The rigorous description of the underlying physical effects [10] rectifies the misconceptions stemming from the approximate and physically inconsistent models widely employed in the literature (Appendix) and, along with the strict calculations for both energy consumption and modulation speed, provides useful insight for future studies and experimental verifications.

Appendix
Comparison between the uniform layer approximation and the rigorous solid state theory approach

A direct comparison is presented between the rigorous solid-state model employed in this work, comprehensively described in [10], and the uniform layer approximation, extensively employed for modeling the electron-accumulation effect in the relevant literature [7]–[9], [14], [16], [17], [20], [23]. This comparison aims at highlighting the artificially overestimated and incorrect ENZ behavior described by the uniform layer approximation, eventually resulting in unrealistic designs.

As a working example is selected the Si-rib waveguide in its downgraded 1D geometry, Fig. 9(a), free of any corner effects to provide a straightforward comparison. The uniform layer approximation represents the electron-accumulation effect at the HfO_2/ITO interface by defining a few nanometer-thick layer, Fig. 9(b), of spatially constant, bias-dependent electron concentration, given by

$$n_{\rm acc} = n_0 + \frac{\varepsilon_0\ \bar{\varepsilon}_{\rm HfO_2}\ V_a}{e\ t_{\rm HfO_2}\ t_{\rm acc}}. \tag{6}$$

In (6), n_0 stands for the free-electron concentration at the flat-band region, equal to the ITO dopant concentration (10^{19} cm^{-3}), ε_0 is the vacuum dielectric constant, e the elementary charge, $\bar{\varepsilon}_{\rm HfO_2} = 25$ and $t_{\rm HfO_2} = 5$ nm are the relative dielectric constant and thickness of the thin insulating HfO_2 film, respectively, while V_a represents the applied bias. The uniform layer thickness $t_{\rm acc}$ is conservatively set equal to 1 nm, even though higher values (3-10 nm) are also frequent in the literature [7], [14], [16], [20].

In Fig. 9, the real part of the effective index and the loss per unit length for the 1D (TM) Si-rib mode are illustrated, calculated using both (6), Fig. 9(b), and the here employed rigorous model (drift-diffusion, DD, model), Fig. 9(c). Evidently, a significant inconsistency is observed between the two approaches, with the uniform layer approximation resulting in a large overestimate of the ENZ effect, predicting additionally a considerably contradicting behavior compared to the DD model for both $\mathrm{Re}\{n_{\rm eff}\}$ and mode-loss curves.

Specifically, (6) leads to a jump in $\mathrm{Re}\{n_{\rm eff}\}$ near the threshold bias $V_{\rm th}$ in contrast to the smoother behavior suggested by the DD model, which predicts instead a minimum for $\mathrm{Re}\{n_{\rm eff}\}$ at $V_{\rm th}$, equivalent to the $-\Delta\mathrm{Re}\{n_{\rm eff}\}$ peak in Fig. 2(a). Moreover, $V_{\rm th}$ appears downshifted by using the uniform layer approximation, due to (6) failing to account for the built-in potential $V_b = 0.63$ V, inherently developed across the heterogeneous n-Si/HfO_2/ITO junction. This mismatch can be readily rectified by considering in (6) the reduced bias $V_a - V_b$, preferred in relevant works [20].

The $\mathrm{Re}\{n_{\rm eff}\}$ profile exhibits a large, first-decimal-point change when calculated by (6), an order of magnitude greater than the respective peak change suggested by the DD model, Fig. 9(d). A similar mode-loss overestimate is reported in Fig. 9(e), where the peak value, approximately equal to 4 dB/μm, well-exceeds the rigorously calculated DD value of 0.5 dB/μm. In addition, the uniform layer approximation predicts a symmetrical decrease in mode loss either side of its peak, being in sheer contrast to the rigorously calculated step-like DD behavior, further corroborated by the results in [22]. Thus, apart from overestimating the ENZ behavior, the uniform layer approximation predicts artificial changes in the effective index, resulting in unrealistic designs. Opting for thicker accumulation layers would result in even more factitious and enhanced ENZ performance, highlighting the physical inconsistency of (6) for designing amplitude- and phase-modulation configurations.

References

[1] C. Wang, M. Zhang, B. Stern, M. Lipson, and M. Lončar, "Nanophotonic lithium niobate electro-optic modulators," *Optics Express*, vol. 26, no. 2, pp. 1547–1555, 2018.

[2] L. Alloatti, R. Palmer, S. Diebold, K. P. Pahl, B. Chen, R. Dinu, M. Fournier, J.-M. Fedeli, T. Zwick, W. Freude, C. Koos, and J. Leuthold, "100 GHz silicon-organic hybrid modulator," *Light Sci Appl*, vol. 3, p. e173, May 2014.

[3] M. Liu, X. Yin, E. Ulin-Avila, B. Geng, T. Zentgraf, L. Ju, F. Wang, and X. Zhang, "A graphene-based broadband optical modulator," *Nature*, vol. 474, no. 7349, pp. 64–67, Jun 2011.

[4] E. Feigenbaum, K. Diest, and H. A. Atwater, "Unity-order index change in transparent conducting oxides at visible frequencies," *Nano Letters*, vol. 10, no. 6, pp. 2111–2116, 2010, pMID: 20481480.

[5] J. Park, J.-H. Kang, X. Liu, and M. L. Brongersma, "Electrically tunable epsilon-near-zero (ENZ) metafilm absorbers," *Scientific reports*, vol. 5, p. 15754, 2015.

[6] D. C. Zografopoulos, G. Sinatkas, E. Lotfi, L. A. Shahada, M. A. Swillam, E. E. Kriezis, and R. Beccherelli, "Amplitude modulation in infrared metamaterial absorbers based on electro-optically tunable conducting oxides," *Applied Physics A*, vol. 124, no. 2, p. 105, Jan 2018.

[7] Z. Lu, W. Zhao, and K. Shi, "Ultracompact electroabsorption modulators based on tunable epsilon-near-zero-slot waveguides," *IEEE Photonics Journal*, vol. 4, no. 3, pp. 735–740, June 2012.

[8] A. P. Vasudev, J.-H. Kang, J. Park, X. Liu, and M. L. Brongersma, "Electro-optical modulation of a silicon waveguide with an "epsilon-near-zero" material," *Opt. Express*, vol. 21, no. 22, pp. 26 387–26 397, Nov 2013.

[9] H. Zhao, Y. Wang, A. Capretti, L. D. Negro, and J. Klamkin, "Broadband electroabsorption modulators design based on epsilon-near-zero indium tin oxide," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 21, no. 4, pp. 192–198, July 2015.

[10] G. Sinatkas, A. Pitilakis, D. C. Zografopoulos, R. Beccherelli, and E. E. Kriezis, "Transparent conducting oxide electro-optic modulators on silicon platforms: A comprehensive study based on the drift-diffusion semiconductor model," *Journal of Applied Physics*, vol. 121, no. 2, p. 023109, 2017.

[11] M. Badr, M. Abdelatty, and M. A. Swillam, "All-silicon directional coupler electro-optic modulator utilizing transparent conducting oxides," in *Laser Science*. Optical Society of America, 2017, pp. JW4A–41.

[12] A. Melikyan, N. Lindenmann, S. Walheim, P. M. Leufke, S. Ulrich, J. Ye, P. Vincze, H. Hahn, T. Schimmel, C. Koos, W. Freude, and J. Leuthold, "Surface plasmon polariton absorption modulator," *Opt. Express*, vol. 19, no. 9, pp. 8855–8869, Apr 2011.

[13] V. J. Sorger, N. D. Lanzillotti-Kimura, R.-M. Ma, and X. Zhang, "Ultra-compact silicon nanophotonic modulator with broadband response," *Nanophotonics*, vol. 1, no. 1, pp. 17–22, jan 2012.

[14] V. E. Babicheva and A. V. Lavrinenko, "Plasmonic modulator optimized by patterning of active layer and tuning permittivity," *Optics Communications*, vol. 285, no. 24, pp. 5500 – 5507, 2012.

[15] A. V. Krasavin and A. V. Zayats, "Photonic signal processing on electronic scales: Electro-optical field-effect nanoplasmonic modulator," *Phys. Rev. Lett.*, vol. 109, p. 053901, Jul 2012.

[16] C. Huang, R. J. Lamond, S. K. Pickus, Z. R. Li, and V. J. Sorger, "A sub-λ size modulator beyond the efficiency-loss limit," *IEEE Photonics Journal*, vol. 5, no. 4, pp. 2 202 411–2 202 411, Aug 2013.

[17] V. E. Babicheva, N. Kinsey, G. V. Naik, M. Ferrera, A. V. Lavrinenko, V. M. Shalaev, and A. Boltasseva, "Towards CMOS-compatible nanophotonics: Ultra-compact modulators using alternative plasmonic materials," *Opt. Express*, vol. 21, no. 22, pp. 27 326–27 337, Nov 2013.

[18] H. W. Lee, G. Papadakis, S. P. Burgos, K. Chander, A. Kriesch, R. Pala, U. Peschel, and H. A. Atwater, "Nanoscale conducting oxide plasmostor," *Nano Letters*, vol. 14, no. 11, pp. 6463–6468, 2014, pMID: 25302668.

[19] T. Amemiya, E. Murai, Z. Gu, N. Nishiyama, and S. Arai, "GaInAsP/InP-based optical modulator consisting of gap-surface-plasmon-polariton waveguide: theoretical analysis," *J. Opt. Soc. Am. B*, vol. 31, no. 11, pp. 2908–2913, Nov 2014.

[20] S. Zhu, G. Q. Lo, and D. L. Kwong, "Design of an ultra-compact electro-absorption modulator comprised of a deposited TiN/HfO_2/ITO/Cu stack for CMOS backend integration," *Opt. Express*, vol. 22, no. 15, pp. 17 930–17 947, Jul 2014.

[21] V. E. Babicheva, A. Boltasseva, and A. V. Lavrinenko, "Transparent conducting oxides for electro-optical plasmonic modulators," *Nanophotonics*, vol. 4, no. 1, pp. 165–185, 2015.

[22] U. Koch, C. Hoessbacher, J. Niegemann, C. Hafner, and J. Leuthold, "Digital plasmonic absorption modulator exploiting epsilon-near-zero in transparent conducting oxides," *IEEE Photonics Journal*, vol. 8, no. 1, p. 4800813, 2016.

[23] C. Lin and A. S. Helmy, "Dynamically reconfigurable nanoscale modulators utilizing coupled hybrid plasmonics," *Scientific Reports*, vol. 5, pp. 12 313 EP –, Jul 2015.

[24] I. Hamberg and C. G. Granqvist, "Evaporated Sn-doped In_2O_3 films: Basic optical properties and applications to energy-efficient windows," *Journal of Applied Physics*, vol. 60, no. 11, pp. R123–R160, 1986.

[25] X. Yu, T. J. Marks, and A. Facchetti, "Metal oxides for optoelectronic applications," *Nature materials*, vol. 15, no. 4, pp. 383–396, 2016.

[26] G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High- gate dielectrics: Current status and materials properties considerations," *Journal of Applied Physics*, vol. 89, no. 10, pp. 5243–5275, 2001.

[27] J. Yota, H. Shen, and R. Ramanathan, "Characterization of atomic layer deposition HfO_2, Al_2O_3, and plasma-enhanced chemical vapor deposition Si_3N_4 as metal–insulator–metal capacitor dielectric for GaAs HBT technology," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 31, no. 1, p. 01A134, 2013.

[28] J. Sanghera and I. Aggarwal, "Active and passive chalcogenide glass optical fibers for IR applications: a review," *Journal of Non-Crystalline Solids*, vol. 256, pp. 6–16, 1999.

[29] A. Grigorescu and C. Hagen, "Resists for sub-20-nm electron beam lithography with a focus on HSQ: state of the art," *Nanotechnology*, vol. 20, no. 29, p. 292001, 2009.

[30] K. Debnath, A. Z. Khokhar, G. T. Reed, and S. Saito, "Fabrication of arbitrarily narrow vertical dielectric slots in silicon waveguides," *IEEE Photonics Technology Letters*, vol. 29, no. 15, pp. 1269–1272, 2017.

[31] D. L. Wood, K. Nassau, T. Y. Kometani, and D. L. Nash, "Optical properties of cubic hafnia stabilized with yttria," *Appl. Opt.*, vol. 29, no. 4, pp. 604–607, Feb 1990.

[32] M. Nedeljkovic, R. Soref, and G. Z. Mashanovich, "Free-carrier electrorefraction and electroabsorption modulation predictions for silicon over the 1-14 μm infrared wavelength range," *IEEE Photonics Journal*, vol. 3, no. 6, pp. 1171–1180, Dec 2011.

[33] A. K. Kulkarni and S. A. Knickerbocker, "Estimation and verification of the electrical properties of indium tin oxide based on the energy band diagram," *Journal of Vacuum Science & Technology A*, vol. 14, no. 3, pp. 1709–1713, 1996.

[34] G. T. Reed, G. Mashanovich, F. Gardes, and D. Thomson, "Silicon optical modulators," *Nature photonics*, vol. 4, no. 8, pp. 518–526, 2010.

[35] G. Sinatkas, I. Skandalos, T. Christopoulos, and E. E. Kriezis, "Electro-optic phase modulators based on transparent-conducting-oxide loaded silicon waveguides," in *2017 IEEE 14th International Conference on Group IV Photonics (GFP)*. IEEE, 2017, pp. 147–148.

[36] A. Brimont, D. Thomson, F. Gardes, J. Fedeli, G. Reed, J. Martí, and P. Sanchis, "High-contrast 40 Gb/s operation of a 500 μm long silicon carrier-depletion slow wave modulator," *Optics letters*, vol. 37, no. 17, pp. 3504–3506, 2012.

[37] V. Sorianello, M. Midrio, and M. Romagnoli, "Design optimization of single and double layer graphene phase modulators in SOI," *Optics Express*, vol. 23, no. 5, pp. 6478–6490, 2015.

[38] L. Liao, A. Liu, R. Jones, D. Rubin, D. Samara-Rubio, O. Cohen, M. Salib, and M. Paniccia, "Phase modulation efficiency and transmission loss of silicon optical phase shifters," *IEEE journal of quantum electronics*, vol. 41, no. 2, pp. 250–257, 2005.

[39] P. Dong, L. Chen, C. Xie, L. L. Buhl, and Y.-K. Chen, "50-Gb/s silicon quadrature phase-shift keying modulator," *Opt. Express*, vol. 20, no. 19, pp. 21 181–21 186, Sep 2012.

Georgios Sinatkas was born in Kastoria, Greece, in 1990. He received the Diploma degree in electrical and computer engineering from the School of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2013, where he is currently a Ph.D. student. His research interests include the analysis and design of integrated nanophotonic devices, the study of free-carrier effects in semiconductors as well as the on-chip integration of novel materials.

Emmanouil E. Kriezis was born in Thessaloniki, Greece, in 1968. He received the Diploma degree in electrical engineering and the Doctorate degree from the School of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1991 and 1996, respectively. He joined the Department of Engineering Science, University of Oxford, Oxford, U.K., in 1998, initially as an Engineering and Physical Sciences Research Council Post-Doctoral Researcher. In 2001 he received the prestigious Royal Society University Research Fellowship to study light propagation in complex anisotropic media. Since 2014, he has been a Professor of optical and microwave communications with the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki. His current research interests include nonlinear effects at optical frequencies, plasmonics, silicon photonics, graphene devices and liquid crystal photonics.