Transparent conducting oxides for electro-optical plasmonic modulators

Babicheva, Viktoriia; Boltasseva, Alexandra; Lavrinenko, Andrei V.

Published in:
Nanophotonics

Link to article, DOI:
10.1515/nanoph-2015-0004

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Abstract: The ongoing quest for ultra-compact optical devices has reached a bottleneck due to the diffraction limit in conventional photonics. New approaches that provide subwavelength optical elements, and therefore lead to miniaturization of the entire photonic circuit, are urgently required. Plasmonics, which combines nanoscale light confinement and optical-speed processing of signals, has the potential to enable the next generation of hybrid information-processing devices, which are superior to the current photonic dielectric components in terms of speed and compactness. New plasmonic materials (other than metals), or optical materials with metal-like behavior, have recently attracted a lot of attention due to the promise they hold to enable low-loss, tunable, CMOS-compatible devices for photonic technologies. In this review, we provide a systematic overview of various compact optical modulator designs that utilize a class of the most promising new materials as the active layer or core—namely, transparent conducting oxides. Such modulators can be made low-loss, compact, and exhibit high tunability while offering low cost and compatibility with existing semiconductor technologies. A detailed analysis of different configurations and their working characteristics, such as their extinction ratio, compactness, bandwidth, and losses, is performed identifying the most promising designs.

Keywords: modulators; electro-optical materials; waveguide modulators; nanocircuits; plasmonics; surface plasmons; active plasmonics; transparent conducting oxides; epsilon-near-zero materials

1 Introduction

The advent of broadband optical signals in telecommunication systems led to unprecedented high bit rates that were unachievable in the electronic domain. Later on, photonics presented a new platform for high-speed data transfer by implementing hybrid electronic-photonic circuits, where data transmission using light instead of electric signals significantly boosted the data exchange rates on such photonic/electronic chips [1–3]. Different silicon-based photonic passive waveguides have recently been proposed and realized as possible optical interconnects for electronic chip communications [4]. Nevertheless, the main building block in the next generation networks architecture is an effective electro-optical modulator.

There are several different physical effects for light modulation as well as different material platforms, which can be grouped according to the three main mechanisms: thermo-optical, electro-optical, and electro-absorption.

The first case is connected with changes in the optical parameters with temperature. In spite of the principal speed limitation due to heat dissipation issues, some on-chip devices with a reasonable operation speed (predicted switching time of 1 µs) were demonstrated [5].

Much faster devices can be realized using the electro-optical mechanism, where variations in the complex refractive index are induced by an electric field or current. The nomenclature of electro-optical effects contains the linear electro-optic effect (Pockels effect), the quadratic electro-optic effect (Kerr effect), and carrier concentration change effect. The Pockels effect is a $\chi^{(2)}$ effect and cannot be exhibited by materials with centrosymmetric lattices, like silicon. However, under distortion of the lattice by strain, such an effect appears [6]. The Kerr effect, where changes in the refractive index are proportional to...
the square of electric field, has broader presentation in materials including silicon [7]. Although, both electro-optical effects are fast, they are rather weak. The strongest Pockels effect is exhibited by LiNbO$_3$, which is a classical material for electro-optical modulators, but hardly can be used in the subwavelength devices. The modulation speed of devices based on such effects is estimated to be about 100 GHz or more [8, 9].

The refractive index changes induced by a change in the carrier concentration have been also the focus of significant research attention for decades [3, 7, 10]. Gigahertz operating frequencies in silicon on-chip modulators were demonstrated a decade ago [11], and recently, the speed has been pushed up to 40 Gbit/s in a Mach–Zehnder interferometer scheme [12]. The carrier concentration can be altered in different ways, namely accumulation, injection, or depletion of free carriers, each method having its pros and cons [3].

The focus on the electro-absorption mechanism was created recently, when it was shown that the absorption factor (in other words, the imaginary part of the refractive index) can be quickly and efficiently altered by an external electric field. Such an effect appears due to the distortion of wave functions of holes and electrons in the bandgap followed by a slight red-shift of the bandgap. Named as the Franz–Keldysh effect in bulk semiconductors or the quantum-confined Stark effect in thin heterostructure layers, the effect was utilized in modulator designs [3, 13–16] exposing performance comparable (28 Gb/s in [13]) to photonic-based electro-optic devices. Estimates that a germanium quantum well can potentially operate with a modulation speed up to 100 GHz with the sufficiently high electric field intensity are presented in [14].

Naturally, the modulation mechanism can be projected onto a material platform to be employed in devices. Electronic fabrication lines prefer the group IV materials: silicon and germanium. III–V semiconductors exhibit even higher nonlinearities and larger bandgaps; however, their integration with the silicon-on-insulator platform still brings a lot of challenges. Polymers with nonlinear inclusions are considered as cheap alternatives to the conventional modulator materials, especially for filling nanoscaled gaps and slits in waveguides. Such small volumes intensify light-matter interactions due to the high-field concentration and allow fast-operation speed, at least 100 GHz as reported recently by Alloatti et al. [9]. Graphene is another cutting edge technology, and its modulation ability is attributed to the changes in the Fermi level driven by the applied voltage [17]. Moreover, a booming topic of electro-optical or all-optical modulation with graphene sheets both in the optical or terahertz ranges has gained massive attention [18–22].

Overall, from the material point of view, the integrated silicon platform is the most established. However, to compete with the compactness of conventional electronic elements, photonic solutions have to overcome the diffraction limit of light by implementing nanoscale optical components.

One avenue that offers subwavelength light confinement is to employ surface plasmon polaritons (SPPs) [23–34]. These are surface waves existing on metal-dielectric interfaces. SPPs are tightly confined to an interface and allow for the manipulation of light at the nanoscale, thus leading to a new generation of fast on-chip devices with unique capabilities. In particular, they could offer a higher bandwidth and reduced power consumption [35–37].

To provide the basic nanophotonic circuitry functionalities, elementary circuit components such as waveguides, modulators, sources, amplifiers, and photodetectors are required [35–41]. Among them, a waveguide is the key passive component connecting all elements in the circuit. Various designs of plasmonic waveguides have been proposed, aiming to achieve the highest mode localization while maintaining reasonable propagation losses [42–58]. However, because of the high intrinsic losses of plasmonic materials, signal propagation is highly damped even for the best (with the highest DC conductivity) metals such as silver and gold. Thus, the ultimate goal is to reduce material losses utilizing new plasmonic materials.

Recently, efforts have been directed towards developing new material platforms for integrated plasmonic devices [59–66]. Transparent conducting oxides (TCOs), for instance, indium tin oxide (ITO) and aluminum-doped zinc oxide (AZO), are oxide semiconductors conventionally used in optoelectronic devices such as flat panel displays and in photovoltaics. The optical response of TCOs is governed by free electrons, whose density is controlled through the addition of n-type dopants. The free carrier concentration in TCOs can be high enough so that TCOs exhibit metal-like behavior in the near-infrared (NIR) and mid-infrared (MIR) ranges, and can be exploited for subwavelength light manipulation (Figure 1). Moreover, in contrast to the optical properties of noble metals, which cannot be tuned or changed, the permittivity of TCOs can be adjusted via doping and/or fabrication process [59, 65, 67] (Figure 2), providing certain advantages for designing various plasmonic and nanophotonic devices.

The plasmonic properties of TCOs have been intensively studied over the past decade. The short-list of successful experimental demonstrations includes (i) possibility of high doping [67, 68] and achieving negative permit-
Figure 1: Optical properties of highly doped transparent conducting oxides: (A) real and (B) imaginary parts of permittivity. Shaded parts correspond to wavelengths where TCOs are metallic and allow plasmonic resonances. Crossover wavelength, that is ENZ region, is approximately at 1400 nm. Real values of permittivities of AZO, GZO, and ITO are similar. However, AZO possesses the lowest losses. The films were deposited using pulsed laser deposition and the deposition conditions were optimized to achieve the highest possible carrier concentration and lowest possible losses. During deposition, the oxygen partial pressure was 0.4 mTorr or lower and the substrate was heated to temperatures around 100°C. The films with thicknesses greater than about 50 nm exhibit very little thickness dependence in their optical properties. For experimental details, see [62] and [65]. Legend is the same on both plots. AZO: Aluminum-doped zinc oxide; GZO: Gallium-doped zinc oxide; ITO: Indium tin oxide.

Figure 2: Strong dependence of AZO permittivity [real (A) and imaginary (B) parts] on oxygen pressures upon deposition (0.1–50 mTorr). The value can be changed up to several units. The film thickness is approximately 120 nm. Losses in AZO is several times lower then losses of silver (Ag) at telecommunication wavelengths. Copyright (2013), AIP Publishing LLC, from Ref. [67], with permission of the authors. AZO: Aluminum-doped zinc oxide

LSPR [77, 84] or SPP guiding [85] with pattern films; and (vi) negative refraction [86] as well as near-perfect absorption in multilayer structures with ZnO components [87]. Theoretical studies of GZO [88] and first-principle theoretical calculations of optical properties of AZO [89] were reported as well.

Importantly, TCOs possess a unique property of having a small negative real part of the permittivity at the telecommunication wavelengths [59, 90]. Structures that exhibit such near-zero permittivity (epsilon-near-zero [ENZ]; [91, 92]) in the optical range can have a plasmonic resonance accompanied by slow light propagation. Thus, conditions for enhanced light–matter interactions are created [93].

An initial study of planar structures with silver layers showed that a plasmonic mode in such conditions ex-
Experiences an unexpected giant modal gain. The physical origin of such an increase is the slowing down of light propagation in the structure [94, 95]. However, this effect can only be realized at \( \lambda = 540 \) nm, close to the crossover point of the silver permittivity. It cannot be in any way tuned to the NIR range, where the real part of the silver permittivity is too high in magnitude. In contrast to conventional noble metals, TCO materials provide conditions for plasmonic resonances and slow light propagation [93, 96]. Implementing a TCO layer in such multilayer structures instead of silver opens up a possibility to achieve slow light propagation at technologically important telecommunication wavelength \( \lambda = 1.55 \) \( \mu \)m (Figure 3). Consequently, ENZ materials can drastically facilitate plasmonic nanolaser [97, 98] operation. According to numerical calculations, the extreme nonlinear dispersion and, correspondingly, extraordinarily high-group indices can be achieved with only a moderate increase of the optical losses [101]. It allows designing a cavity-free plasmonic nanolaser with an almost zero threshold and operational sizes far beyond the diffraction limit [102].

In addition to linear properties, nonlinearities at the ENZ point are of interest and are propelled by the pronounced local-field enhancement facilitated by metal nanostructures [103–106]. Recently, the enhancement of nonlinear properties in a hybrid TCO-gold structure due to the plasmon-induced hot carrier injection and tight-field confinement in the vicinity of a plasmonic nanostructure has been demonstrated [107–109]. The rapid energy transfer in the TCO produces a localized rise of the carrier temperature and increased carrier mobility, thus, providing a stronger nonlinear response. This ability to dramatically change the optical properties of TCOs from dielectric-like to metallic-like under an applied bias triggers new approaches for the realization of photonic modulators [110, 111].

To modulate and switch plasmonic signals in the waveguiding configuration, changes can be induced in either the real or imaginary parts of the permittivity by applying a bias. Traditional photonic approaches rely on small refractive index changes in the material forming the waveguides core. Consequently, it requires long-propagation distances to accumulate the sufficient phase or absorption changes, which result in increase in the size of devices (up to hundreds of micrometers). In contrast, TCOs are promising candidates for adding electro-optical capabilities to ultra-compact plasmonic devices. A large increase in the carrier concentration within a 10-nm thick ITO film of a metal-oxide-semiconductor (MOS) stack is demonstrated in Ref. [112]. Furthermore, additional functionalities can be gained by tuning a small absolute value of the TCO permittivity in and out of the plasmonic resonance [113]. The theoretically predicted extinction ratio (ER) is 50 dB/\( \mu \)m [113]. Such a high value allows for the 3 dB modulation depth in an extremely compact, only 60 nm long, device. The effect is based on the carrier concentration changes, and thus, the associated timescale is on the order of picoseconds. Therefore, around 100 GHz or even higher-frequency modulation speed is expected [112].
In this review, we discuss how the above-mentioned properties of TCOs can be exploited in plasmonic modulators. First, an overview of the absorption modulator concept and metrics of its performance is presented (Section 2). Further, we provide detailed analysis of TCOs’ properties (Section 3). As modulator examples, we discuss two particular designs, that is, ultra-compact plasmonic and silicon-waveguide-integrated modulators (Sections 4 and 5). Importance of their CMOS-compatibility is discussed in Section 6. It is followed by the comparison of performance, discussion of material platforms, and outlook (Section 7), as well as summary and conclusions (Section 8).

2 Absorption modulator concept and metrics of performance

A modulator is the prime component for digital signal encoding [3, 35–37]. The development of nanophotonic circuitry depends greatly on advance in modulation schemes configuration (Figure 4). Therefore, plasmonic switching devices form one of the most promising directions in nanophotonics, as it aims to employ SPP waves in semiconductor structures to control the properties of different nanodevices [114–127].

For a typical absorption modulator, it is convenient to introduce the logarithmic ER per unit length [128]:

\[ r = 10 \log_{10}(P_{on}/P_{off})/L = 8.68(\alpha_{off} - \alpha_{on}), \]  

(1)

where \( \alpha_{on} \) and \( \alpha_{off} \) are the imaginary parts of the effective propagation constants \( \beta_{eff} \) of waveguide modes (or losses) in the on-state and off-state, respectively. ER shows how strong one can vary the mode propagation through the waveguide. However, there is always a trade-off between the modulation depth and transmittance through the modulating system. For plasmonic switching devices that exhibit significant propagation losses, a modified parameter of performance or the figure of merit (FoM) can be defined as [128]:

\[ f = \frac{\alpha_{off} - \alpha_{on}}{\alpha_{on}}, \]  

(2)

where \( f \) describes how strong one can vary the mode propagation through the waveguide relative to the attenuation in the transmittive (on-) state. Thus, to optimize the modulator performance, one has to increase the ER of plasmonic devices while decreasing simultaneously their propagation losses. Equation (2) gives the length-independent characteristic of the structure. A proper length of the device can be found according to the signal level requirements and fabrication restrictions in particular modulator geometry.

![Figure 5: Schematic view of the horizontally arranged silicon-based plasmonic electro-absorption modulator. Copyright (2011), AIP Publishing LLC, from Ref. [134], with permission of the authors.](image)

![Figure 6: Schematic view of a multilayer stack: TCO and insulator layers sandwiched between electrodes. TCO: Transparent conducting oxides](image)

For plasmonic modulators, several other FoMs were proposed [113, 129]. For instance, in order to take into account the ultra-compact footprint of plasmonic devices, one can introduce the following FoM: 

\[ f^* = f_{\lambda_{eff, on}}/w_{on}, \]

where \( \lambda_{eff, on} = 2\pi/\text{Re}(\beta_{eff}) \) is the effective wavelength in the modulator in the on-state, and \( w_{on} \) is the on-state mode dimension [113]. Other important metrics include peak-to-peak voltage, modulation speed, and energy consumption [36, 129]. However, we refrain from introducing any linear or other combination of these characteristics, as it does not reflect the real trade-off in performance.
In general, for plasmonic modulators, two classes of active materials were suggested based on two different physical mechanisms for the refractive index control: carrier concentration change (used in e.g. silicon [40, 41, 130–136] and TCOs [112, 113, 128, 137–147]), and structural phase transitions (e.g. gallium [148–151] and vanadium dioxide [152–157]). In addition, plasmonic modulators based on thermo-optic [158–161] or nonlinear polymer [162, 163], graphene [164–166], indium phosphide [167, 168], barium titanate [169], bismuth ferrite [170], and germanium [171] were also reported.

The refractive index changes that accompany nanoscale material phase transitions are much higher than those caused by the carrier concentration change. In particular, an ER up to 2.4 dB/µm was demonstrated for a hybrid plasmonic modulator based on the metal–insulator phase transition in vanadium dioxide [154]. However, neither structural transitions nor thermo-optic responses provide the necessary bit rate, so only megahertz modulation operations are expected due to microsecond timescales of the aforementioned alternative processes.

An experimental study performed for a plasmonic latching switch, whose design is similar to the nanophotonic modulators based on conventional silicon waveguides, revealed the possibility of the creation and elimination of a conductive path in a gold/silicon dioxide/ITO-layered structure with a switching speed on the order of a few milliseconds [172, 173]. Similar to processes in an amorphous silicon layer, the changes in the optical transmission is attributed to the formation and annihilation of nanoscale metallic filaments [174].

The carrier concentration change in the accumulation layer is faster structural phase transitions, and thus optoelectronic modulators can theoretically achieve ultra-fast operational speeds in the hundreds of gigahertz being limited only by the RC time constant. It was shown that some plasmonic-based designs even outperform conventional silicon-based modulators [41, 112].

Effective modulation performance was obtained with horizontally arranged metal-insulator-silicon-insulator-metal slot waveguides [134–136]. The governing principle is based on inducing a high free electron density in the infinitesimal vicinity of the SiO$_2$/Si interface (Figure 5). Such electro-absorption CMOS-compatible modulation was characterized, and the 3-dB broadband modulation with just 3 µm-length of the device at approximately 6.5 V bias was reported [134]. Propagation losses varied between 1 and 2.9 dB/µm, which correspond to a FoM (2) $f = 1.9$. The advantage of this configuration is that it can be readily integrated in standard Si circuits. However, its high mode localization requires a very high aspect ratio of the waveguide core, which is challenging to fabricate.

Ultra-compact efficient plasmonic modulators based on strong-light localization in vertically arranged metal-silicon-insulator-metal waveguides were reported in Ref. [130]. The PlasMOStor consists of the semiconductor core with Si and SiO$_2$ layers, sandwiched between two silver plates [130]. It supports two modes: photonic and plasmonic, which interfere while propagating. The modulation principle is based on cutting off the photonic mode and thereby changing the integral transmittance. For the photonic mode, propagation losses can be changed from 2.4 to 28 dB/µm, providing the possibility to operate with high $f = 11$.

![Figure 7: Change of carrier concentration (A) and TCO permittivity (B) according to the Thomas–Fermi screening theory. 2.5 V per 5-nm thick SiO$_2$, or HfO$_2$ layer was taken for the calculations. 10$^{19}$, 5 × 10$^{20}$, and 10$^{21}$ cm$^{-3}$ correspond to different initial concentration of carriers in the TCO layer. Solid curves 1, 2, and 3 correspond to SiO$_2$ and dashed curves 4, 5, and 6 correspond to HfO$_2$. In (A), curve 7 shows approximation with the exponential decay defined by the Thomas–Fermi screening length. TCO: Transparent conducting oxides](image-url)
In contrast to silicon, changing the carrier concentration in TCOs can result in a much higher change of the effective refractive index. In particular, a striking unity order change $\Delta n = 1.5$ was experimentally demonstrated [110], which is three orders of magnitude higher than the maximal change of the effective refractive index in a silicon waveguide: $\Delta n = 10^{-3}$ [3]. Thus, TCOs can be efficiently exploited as active materials for light modulation providing extraordinary tuning of their complex refractive indices by changing the carrier concentration with the application of an electric field.

3 Transparent conducting oxides as highly tunable material

The first experimental demonstration of the carrier concentration change was made for a relatively thick TCO layer (thickness 300 nm) in a metal-insulator-metal structure. The carrier concentration growth from $1 \times 10^{21}$ to $2.83 \times 10^{22}$ cm$^{-3}$ was achieved in a 5-nm thick accumulation layer [110].

General concept of the carrier concentration change is shown in Figure 6. Applied bias causes either charge accumulation or its depletion in the TCO layer, depending on the direction of the electric field. In turn, variations in the carrier concentration $n_{c.c.}$ affect the TCO plasma frequency $\omega_{pl}$, and consequently, its permittivity $\varepsilon_{acum}$, which follows the Drude model $\varepsilon_{acum}(\omega) = \varepsilon_\infty - \omega_{pl}^2/\omega(\omega+i\gamma)$, where $\omega_{pl}^2 = n_{c.c.}e^2/m_{eff}$, $\varepsilon_\infty$ is the background permittivity, $\gamma$ is the collision frequency, $e$ is the charge of an electron, $m_{eff}$ is the effective mass of an electron, and $\varepsilon_0$ is the vacuum permittivity. The process can be described by the Thomas–Fermi screening theory [128, 138]. One can either find the spatial distribution of free carriers directly from the Thomas–Fermi screening theory, or use instead its exponential approximation with the Thomas–Fermi screening length (Figure 7A). Evidently, the TCO permittivity can reach relatively low values (Figure 7B).

Figure 7 shows that utilizing an HfO$_2$ layer, which possesses a high-static dielectric constant $k = 25$, instead of SiO$_2$ with ordinary $k = 4.5$, is very beneficial. For the same applied voltage, the carrier concentration induced in the accumulation layer can be several times higher than that in the case with SiO$_2$, providing larger change of the permittivity value. As the effect occurs only in a 5-nm accumulation layer, the rest of the material reacts passively for the applied voltage. It does not bring additional functionality, but rather increases optical losses significantly. Thus, later

![Figure 8](image-url): Influence of ITO permittivity on the absorption coefficient and FoM of the metal-insulator-metal-waveguide-based modulator. To obtain the complex propagation constant, the dispersion equation was solved for a four-layer structure, that is, 8-nm thick ITO and 70-nm thick silicon nitride insulation embedded in silver claddings, where the real part of the ITO permittivity was varied. Copyright (2012), Elsevier, from Ref. [139], with permission of the authors. ITO: Indium tin oxide; FoM: Figure of merit

![Figure 9](image-url): (A) SPPAM: the modulator section consists of a stack of silver, ITO, and silica layers, and can be integrated by means of a directional coupler with a silicon waveguide. The insets show how a photonic mode in the silicon waveguide excites the MIM waveguide mode in SPPAM via a hybrid mode in the directional coupler. (B) The absorption coefficient of the SPP is modulated by applying a voltage between two silver electrodes. (C) The MIM waveguide possesses a plasmonic resonance and is strongly affected by changes in the permittivity of the active layer. Copyright (2011), OSA, from Ref. [128], with permission of the authors. ITO: Indium tin oxide; MIM: Metal-insulator-metal; SPPAM: Surface plasmon polariton absorption modulator

Brought to you by | Georgia State University
Authenticated
Download Date | 8/25/15 5:35 PM
the focus of the study shifted towards ultra-thin (8–10 nm) active films [112, 128, 140].

Since the TCO’s properties depend strongly on fabrication conditions such as the doping level, annealing environment, pressure, and temperature ([59, 65]), it allows one to adjust TCO’s permittivity, and hence, to tailor the material for a specific device design ([139], Figure 8).

4 Ultra-compact plasmonic modulators

Metal-insulator-metal (MIM) configurations, where the dielectric core is sandwiched between two metal plates, were shown as one of the most compact and efficient waveguide layouts [42–45]. In such waveguides, light is localized in the gap with typical sizes of approximately 100 nm and less, that facilitate manipulation of light on the subwavelength scale. The SPP propagation length in the MIM waveguide is not extended and can be up to 10 µm only. However, the tight MIM mode confinement provides a base for designing extremely fast and ultra-compact plasmonic devices, including modulators, photodetectors, lasers, and amplifiers, where the long-propagation length is not required.

In the MIM configurations, an additional layer such as an ultrathin TCO film allows for control of the dispersion in the waveguide, and thus, a higher efficiency plasmonic modulator in comparison to other switchable materials. In addition to high speed and compactness of the MIM configurations, the metal plates can serve as intrinsic electrodes for electrical pumping of the active material, offering a simplified design.

One of the first proposed concepts of the TCO-based modulator includes layers of oxide and semiconductor, which are embedded in metal. For example, ITO was implemented in an MIM waveguide structure to demonstrate a subwavelength plasmonic modulator (Figure 9), in which a 5% change in the average carrier density (from $9.25 \times 10^{20}$ to $9.7 \times 10^{20} \text{ cm}^{-3}$) was assumed [128]. The authors employed a one-dimensional, four-layer structure model to find the SPP dispersion, and the Thomas–Fermi screening theory to determine the carrier density distribution. Numerical simulations with the finite element method were used to model performance of modulator with realistic dimensions. The structure supports an SPP having resonance conditions close to 1.55 µm due to...
the ITO layer, which has small permittivity values in the NIR range, though the resonance is broadened because of high losses in ITO. Consequently, this decreases performance of the device but increases the operational bandwidth at the same time. A similar structure based on a silicon-waveguide-integrated multilayer stack was fabricated and an ER up to \( r = 0.002 \, \text{dB/\mu m} \) was achieved. Theoretically, the Ag-ITO-Si\(_3\)N\(_4\)-Ag structure can have an ER up to \( r = 2 \, \text{dB/\mu m} \), but due to the high-mode confinement in the MIM structure and the high losses associated with the metal and ITO layers, the propagation length in this system is rather limited: losses are 24 \, \text{dB/\mu m} \) for the waveguide with the Si\(_3\)N\(_4\)-core and 9 \, \text{dB/\mu m} \) with the SiO\(_2\)-core. Overall, such a modulator is extremely compact.

The very high ER was achieved in a modulator based on an MIM waveguide with an ultrathin 5 \, \text{nm} \) gap [138]. Such an ultra-compact design leads to a significant change of the carrier concentration in a 2.5-\,\text{nm} thick ITO layer (Figure 10). The total length of the proposed nanoplasmonic modulator, including the coupling tapers, can be below 500 \, \text{nm}, providing a very small footprint acceptable for the on-chip integration.

The suggested plasmonic modulator can be a part of a metal–insulator–metal waveguide or coupled to another type of a plasmonic waveguide with high-mode localization. In addition, there are various possibilities to couple metal–insulator–metal waveguides to conventional photonic waveguides [175–177]. Similar concepts can be utilized for plasmonic waveguide modulator structures.

5 Silicon waveguide-based modulators

The MIM design provides a very high ER, and consequently allows an ultra-compact layout. However, device fabrication, its integration, and characterization are challenging. While Melikyan et al. [128] proposed the MIM design, the proof-of-concept demonstration was performed with a conventional silicon waveguide-based structure, where a MOS stack was deposited on top of a silicon photonic waveguide (Figure 11).

A similar approach was taken in Refs. [112, 137, 140, 141], where silicon photonic modulators used the TCO film as a dynamic layer. An ER \( r = 1 \, \text{dB/\mu m} \) was demonstrated for a plasmonic modulator utilizing a metal-oxide-ITO stack on top of a silicon photonic waveguide ([112], Figure 12). Under applied bias, the carrier concentration is increased from \( 1 \times 10^{19} \) to \( 6.8 \times 10^{20} \, \text{cm}^{-3} \), reducing the propagation length from 34 \, \text{\mu m} \) to 1.3 \, \text{\mu m}. Such waveguide-integrated plasmonic modulators are relatively easy to fabricate, their size is comparable to a single silicon waveguide, and propagation losses originate only from one metal interface. The configuration possesses low in-
Figure 15: (A) Modulator design [141] includes two additional metal layers: TiN and Cu. (B) Numerical modeling of the change in the carrier concentration. Cutoff frequency is approximately 11–54 GHz. Copyright (2014), OSA, from Ref. [141], with permission of the authors.

insertion loss (1 dB) and a broad operational band due to the non-resonant MOS-mode propagation.

One of the most important steps in developing TCO modulators is utilizing its small value of permittivity (Figure 13). It was pointed out in the first paper by Melikyan et al. [128] (Figure 9C) and the optimal operation was studied further [139], (Figure 8). However, because of the high initial carrier concentration in TCOs, the structure did not give high performance. Later, it was theoretically predicted that a very high ER (up to \( r = 20 \text{ dB/}\mu\text{m} \)) can be achieved utilizing ENZ properties of AZO [137]. For a pure photonic system (e.g. [137, 140], Figure 14), the ENZ condition \( \varepsilon = 0 \) causes a resonance. Due to the small absolute value of the permittivity, a large portion of the electric field is localized within the TCO layer providing efficient modulation. If plasmonic elements are involved, the resonance condition shifts towards negative values \( \varepsilon = -4 ... 0 \). The effect was discussed in detail in Ref. [113]. Later, in [141], it was mentioned that the ENZ point and absorption maximum do not coincide, which happens because metal layers (TiN and Cu) are put on a silicon waveguide (Figure 15). In general, recently proposed designs of silicon waveguide-integrated TCO modulators are similar but include a different number of additional layers, see, for example Refs. [141] and [145]. For comparison of performance in terms of different metrics, see table and discussion in Section 7.

6 CMOS-compatibility

Numerous plasmonic waveguides and modulators have been proposed and experimentally verified, but most of these structures incorporate noble metals such as gold or silver as plasmonic building blocks. However, noble metals are not compatible with the well-established CMOS manufacturing processes and pose challenges in fabrication and device integration. This fact limits the ultimate applicability of plasmonic structures based on noble metals in future optoelectronic devices. Thus, the next step towards practical implementation of plasmonic technologies requires the development of fully CMOS-compatible passive and active integrated-optics components. Recently, CMOS-compatible materials with adjustable/tailorable optical properties have been proposed as alternatives to noble metals in plasmonic devices [62, 65, 66, 178]. New configurations open up a way for dy-
Transparent conducting oxides for electro-optical plasmonic modulators

Figure 17: Suggested modulators designs are varied by compactness and ease of integration with low-loss waveguide (arrows show direction of more profound influence). Plasmonic multilayer structures offer additional advantages and possess strong ability to confine light [113, 128, 138, 139]. Although, a methodology of integration of MIM-based modulators into plasmonic or dielectric photonic networks has been developed, for example [138], the practical implementation of the proposed device is still challenging. Implementation of a TCO layer inside a dielectric waveguide [113, 137, 142, 145] is more complicated in comparison to just depositing it on top [112, 140, 141], but the mode overlap with the active TCO layer is also higher. Next possible designs can be based on slot- [143] or long-range dielectric-loaded plasmonic waveguides. MIM: Metal-insulator-metal; TCO: Transparent conducting oxides

Figure 18: (A) TCO modulator based on a plasmonic slot waveguide. (B) High modulation strength is achieved because of the good overlap of the propagating plasmonic mode and the ITO layer. Copyright (2014), American Chemical Society, from Ref. [143], with permission of the authors. ITO: Indium tin oxide; TCO: Transparent conducting oxides

Figure 19: Depending on the insulating material, permittivity of TCO in the accumulation layer varies significantly. Extremely high values can be reached using strontium titanate. Similar to Figure 7, data calculated for 2.5 V applied on per 5 nm thick HfO₂, La₂O₃, TiO₂, or SrTiO₃ layer, and 10¹⁹ cm⁻³ initial concentration in the 10-nm ITO film. TCO: Transparent conducting oxides

Dynamic switching and modulation capabilities offering low-cost and highly stable nonlinear properties. Similar to the advances in silicon technologies that led to the information revolution worldwide, the development of new CMOS-compatible plasmonic materials could revolutionize the field of hybrid photonic/electronic devices.

Plasmonic ceramics, such as titanium nitride and zirconium nitride, are among the best candidates that can replace conventional plasmonic metals [179, 180]. Titanium nitride is CMOS-compatible and provides higher mode confinement in comparison to gold [179]. Furthermore, titanium nitride is very thermally and chemically stable, extremely hard (one of the hardest ceramics), and biocompatible [181–184]. It can be grown epitaxially on many substrates including c-sapphire and [100]-silicon, forming ultra-smooth and ultra-thin layers [185]. It was theoretically shown that a TCO modulator integrated with a long-
7 Comparison metrics of performance and discussion

The summary of performance characteristics of the reported TCO-based plasmonic modulators is shown in the table below. Some numbers have been recalculated using parameters given in the papers. Based on the comparative analysis described above, one can draw several conclusions.

7.1 Design and basic waveguide structure

MIM-modulators offer high compactness (3 dB modulation can be achieved on a sub-100 nm scale), but the overall performance (e.g. FoM) is lower than that of silicon waveguides-based devices. Although, because conventional dielectric devices exploit photonic modes, the miniaturization level of such devices is limited. New configurations of hybrid plasmonic waveguides, such as slot-based or long-range dielectric-loaded plasmonic waveguides (Figure 17) provide the best trade-off between compactness and propagation losses. So hybrid photonic waveguides can serve as the most promising platform for plasmonic electro-optical modulators. Recently, a plasmonic slot waveguide, which is filled with low-doped TCO was suggested (Figure 18A) [143]. Such a waveguide has reasonably low losses and a small mode size forming the promising direction for further optimization. An ER \( r = 2.7 \ \mu \text{m/dB} \) and insertion losses 0.45 dB/\( \mu \text{m} \) are measured for a 300 \( \times \) 200 nm\(^2 \) slot (Figure 18B).

Furthermore, a more sophisticated design, which is based on directional coupling to conventional silicon waveguides, was suggested [142]. Another approach is to introduce a resonant structure along the waveguide [147, 186]. While higher ER can be achieved by utilizing directional coupling or resonant features, the operation bandwidth of the modulator is essentially limited. Similar to designs with ring resonators, these devices suffer from high sensitivity to fabrication imperfections and temperature-dependent functionality.

7.2 Constituent materials

Here we briefly discuss the main properties of materials employed in TCO-based electro-optical modulators.

TCO properties

As discussed in Section 2, and also illustrated in Figure 7, the rise of the carrier concentration in the accumulation layer is relatively steep, so ENZ or negative-index layer thicknesses and permittivity profiles are approximately the same for different initial concentrations (both high and low ones). As an obvious drawback of high initial carrier concentrations, the remaining unaffected and, therefore, passive part of the TCO film (approximately 8 nm out of 10 nm) brings additional high losses. Thus, low initial carrier concentrations in the TCO provide higher modulator FoMs and hence better performance.
Table 1: Comparison of TCO-based modulators. Notations “Th” and “Exp” in the column “Th/Exp”, signify what kind of analysis—theoretical or experimental was performed (in many cases there is a lack of experimental data). Designs that include “Si+” are based on conventional silicon waveguides. In parentheses, thickness of layers is in [nm]. Telecom wavelengths of signal propagation were considered, that is, $\lambda = 1.55$ or $1.31\ \mu m$). Different initial carrier concentrations (c.c.) in the TCO switchable layer were taken into account. Most of the studies either applied the Thomas–Fermi screening theory (notation “Th.–F.”) or took an average value of the carrier concentration. Software MEDICI was used in Ref. [141] and Silvaco ATLAS Device3D in Ref. [144]. Bias can be applied in both directions, that is, causing accumulation or depletion of carriers in TCOs (however, only in Ref. [141] depletion was employed). Only few works contain theoretical estimations of energy consumption as well as speed (sometimes RC constant is given). “NA” is notation in case if number is not mentioned and cannot be unambiguously estimated.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Date</th>
<th>Th/Exp</th>
<th>Design (thickness in nm)</th>
<th>Size, mm $\times$ mm</th>
<th>$\lambda$, nm</th>
<th>$r$, $\mu m$</th>
<th>$d_{\text{min}}$, dB/$\mu m$</th>
<th>$f_{\text{o}}$, GHz $\times 10^{13}$ cm$^{-3}$</th>
<th>C.c. distrib. in TCO</th>
<th>Bias, V</th>
<th>Energy</th>
<th>Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>[128]</td>
<td>04/11</td>
<td>Th</td>
<td>Ag/SiO$_2$(30)/ITO(8)/Ag</td>
<td>200 $\times$ 50</td>
<td>1.55</td>
<td>2</td>
<td>24</td>
<td>0.08</td>
<td>Th.–F., averaged</td>
<td>0 to 10</td>
<td>NA</td>
<td>RC 35 fs</td>
</tr>
<tr>
<td>[128]</td>
<td>04/11</td>
<td>Exp</td>
<td>Si$_3$N$_4$(200)/ITO(0)/Ag</td>
<td>240 $\times$ 220</td>
<td>0.002</td>
<td>NA</td>
<td>NA</td>
<td>82.7</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>[112]</td>
<td>04/12</td>
<td>Exp</td>
<td>Si+ITO(10)/SiO$_2$(20)/Au</td>
<td>800 $\times$ 340</td>
<td>1.31</td>
<td>1</td>
<td>0.1</td>
<td>0.06</td>
<td>NA</td>
<td>0 to 3</td>
<td>56 fl/bit (for 4 V)</td>
<td>300 GHz</td>
</tr>
<tr>
<td>[137]</td>
<td>05/12</td>
<td>Th</td>
<td>Si+AZO(10)/SiO$_2$(10)</td>
<td>450 $\times$ 140</td>
<td>1.31</td>
<td>17.4</td>
<td>1</td>
<td>17</td>
<td>averaged</td>
<td>NA</td>
<td>NA</td>
<td>RC 7.3 ps</td>
</tr>
<tr>
<td>[138]</td>
<td>07/12</td>
<td>Th</td>
<td>Al/ITO(2.5)/HfO$_2$(2.5)/Al</td>
<td>25 $\times$ 25</td>
<td>15.3</td>
<td>10.3</td>
<td>1.3</td>
<td>1</td>
<td>Th.–F.</td>
<td>0 to 2</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>[129]</td>
<td>07/13</td>
<td>Th</td>
<td>Si+ITO(8)/Si$_3$N$_4$(70)/Ag</td>
<td>200 $\times$ 78</td>
<td>1.55</td>
<td>3.2</td>
<td>8.7</td>
<td>0.37</td>
<td>92.5</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>[113]</td>
<td>10/13</td>
<td>Th</td>
<td>Si+Si$_3$N$_4$(10)/GZO(10)/Si</td>
<td>200 $\times$ 600</td>
<td>1.31</td>
<td>6</td>
<td>0.7</td>
<td>8.6</td>
<td>averaged</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>[140]</td>
<td>10/13</td>
<td>Th</td>
<td>Si+HfO$_2$(5)/ITO(10)</td>
<td>400 $\times$ 220</td>
<td>1.55</td>
<td>0.003</td>
<td>37</td>
<td>1</td>
<td>Th.–F.</td>
<td>0 to 2.3</td>
<td>1.3 pJ/bit</td>
<td>100 fl/bit</td>
</tr>
<tr>
<td>[141]</td>
<td>07/14</td>
<td>Th</td>
<td>Si+Si$_3$N$_4$(10)/HfO$_2$(5)/Cu</td>
<td>400 $\times$ 340</td>
<td>1.55</td>
<td>0.045</td>
<td>2.5</td>
<td>2.5</td>
<td>averaged</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>[142]</td>
<td>09/14</td>
<td>Th</td>
<td>Si+ITO(10)/SiO$_2$(10)</td>
<td>800 $\times$ 600</td>
<td>1.55</td>
<td>4.13</td>
<td>1.575</td>
<td>2.6</td>
<td>MEDICI, averaged</td>
<td>0 to 2</td>
<td>4.4 pJ/bit</td>
<td>11 GHz</td>
</tr>
<tr>
<td>[143]</td>
<td>10/14</td>
<td>Exp</td>
<td>Plasmonic slot waveguide</td>
<td>600 $\times$ 500</td>
<td>1.55</td>
<td>0.9</td>
<td>NA</td>
<td>NA</td>
<td>1.1</td>
<td>Th.–F.</td>
<td>0 to 2.4</td>
<td>78.8 fl/bit</td>
</tr>
<tr>
<td>[144]</td>
<td>11/14</td>
<td>Th</td>
<td>InP-GaInAsP+ITO(5)/GZO(10)/Au</td>
<td>500 $\times$ 1000</td>
<td>1.55</td>
<td>4.2</td>
<td>1.3</td>
<td>3.2</td>
<td>10</td>
<td>Silvaco</td>
<td>0 to 5</td>
<td>NA</td>
</tr>
<tr>
<td>[145]</td>
<td>12/14</td>
<td>Th</td>
<td>Si+Si$_3$N$_4$(7)/ITO(10)/Si$_3$N$_4$(7)</td>
<td>500 $\times$ 410</td>
<td>1.31</td>
<td>15.9</td>
<td>0.9</td>
<td>17.7</td>
<td>averaged</td>
<td>0 to 4</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
It is also important to mention that in theoretical analysis and modeling, direct averaging of concentration over the layer thickness (8–10 nm) does not properly describe the optical properties of the whole device, and more rigorous modeling approaches, which account for the non-uniform concentration, should be taken into account for providing accurate results.

**Insulator**

As an electrostatic material, HfO$_2$ performs better than SiO$_2$ (Section 2). The main advantages are that HfO$_2$ has a high-static dielectric constant ($k = 25$) as well as high-breakdown voltage limit. With the atomic layer deposition technique, smooth oxides films down to 5 nm thickness can be fabricated. Moreover, the permittivity of HfO$_2$ matches that of one of low-doped TCOs, which is desired for Si-waveguide-based modulators.

New high-dielectric constant (high-k) oxides, such as lanthanum oxide (La$_2$O$_3$), titanium dioxide (TiO$_2$), and strontium titanate (SrTiO$_3$) are currently under investigation [187]. Utilizing high-k materials can result in a much higher change of the TCO permittivity in the accumulation layer (Figure 19), and consequently can improve the performance.

**Dielectric materials in plasmonic waveguides**

Both high-index and low-index materials for waveguides were analyzed showing that the high-index cladding provides better performance [113, 141, 145]. For instance, for the design in Ref. [141], the ER is approximately $r = 20$ dB/µm for a silicon waveguide and only $r = 4$ dB/µm for a silicon nitride waveguide. Although there are two times higher propagation losses in the silicon-based waveguide, the high-index silicon waveguide-based modulator has higher FoM (see table). The similar conclusion was drawn in [128] comparing SPPAM performance with SiO$_2$ and Si$_3$N$_4$ cores.

**Other materials**

Including additional layers is suggested in Ref. [141] to provide better electric contacts. However, those layers of metal simultaneously increased insertion losses of the device (see table).

### 7.3 Outlook: Further directions

Despite the fact that there are many publications on TCO-based modulators, still many problems remain of fundamental interest. From the practical point of view, the experimental realization of TCO modulators is very challenging. The list includes, but not limited to, modulators fabrication, integration with compact low-loss waveguides, implementation of electrical contacts, elimination of leakage current through ultra-thin layers, and ultra-fast characterization of the modulated signal. The restriction on the film thickness (several nanometers) brings additional requirements on fabrication techniques. Only three of the whole publication list references [112], [128], and [143] provide the proof-of-principle modulation demonstration, so further experimental investigations are required.

#### Study of dynamic properties

Up-to-date, dynamic characterization of modulators in terms of speed limits is largely missing. Theoretically, the modulating speed in TCO modulators is limited only by the RC-circuit constrains, and it is expected to exceed hundreds of gigahertz [112, 143]. However, none of the experimental studies has reported the response of the modulators close to the upper limit (see table). A recently published study of dynamic tuning in multilayer structures shows the presence of a significant delay [111]. Thus, there is an urgent need for thorough experimental studies and advanced characterization of TCO-based devices, including the dynamic change of the applied voltage, high-frequency modulation of optical properties, determination of the upper limit of the modulation frequency, and the development of advanced theoretical models.

#### Study of carrier concentration

One of the aims is to develop a model to describe carriers’ dynamics and to investigate fundamental limitations on the increase in the carrier concentration. Thus, there is a need for further experimental studies of thin TCO films to go beyond the Thomas–Fermi screening theory. One should develop an analytical model of carrier distributions and their change under applied bias to find out how strong the carrier concentration in a thin TCO film can be varied by applied bias. Three-dimensional hybrid (micropscopic-electromagnetic) models of carrier dynamics can be very useful for real-time modeling of plasmonic modulator performance.
Ultra-thin and smooth layers
The TCO thickness should be as small as possible to reduce losses while keeping the main feature of the layer—the small negative real part of the permittivity. Thus, research efforts should be put towards the fabrication of ultra-thin, smooth continuous films of TCOs. It has been shown that reducing the thickness down to 8 nm is practical. Some theoretical studies consider thinner layers, for example, a few nanometers [138, 141], which will provide better performance (Figure 20). Yet, the feasibility of ultra-thin TCO films with optical properties similar to thicker layers has not been proven.

New TCO materials with superior properties
New intermediate carrier-density materials may offer the prospect of additional advantages going beyond conventional variations in optical properties, accompanied with lower losses and integration advantages. Such TCOs can provide extraordinary tuning and modulation of complex refractive indices, because their carrier concentrations can be changed over several orders of magnitude by applying an electric field. A broad variety of other TCO materials have been proposed, but most of them still have not been explored for plasmonic applications. Different techniques of high doping and loss mitigation have to be studied to obtain materials with better properties.

8 Conclusions
Plasmonics enables the merger between two major technologies: nanometer-scale electronics and ultra-fast photonics. As alternatives to conventional noble metals, new plasmonic materials (such as highly doped oxide semiconductors and plasmonic ceramics) offer many appealing advantages, including low-intrinsic loss, semiconductor-based design, compatibility with standard nanofabrication processes, and tunability. The demonstration of the plasmonic (metal-like) response of TCOs in the optical range was an important milestone, which opened up perspectives in employing this material for plasmonic applications. Implementing TCO layers in a waveguide pushes plasmonic resonances to lower frequencies, which are of interest for telecom applications.

It is expected that developing and optimizing highly-tunable materials such as TCOs will have a significant impact on nanophotonics technology providing CMOS-compatible passive and active devices. TCO-based plasmonic devices could also play an important role in footprint reduction and energy consumption, as well as increasing bandwidth, which are one of the main criteria in research in photonics technologies. Up to now, the lowest estimated power consumption of an un-optimized TCO-based plasmonic modulator is about 4 fJ/bit [143], which can further be decreased by engineering device design, for example, improving mode confinement, choosing more optimal geometry, and appropriate constituent materials. Thus, more experimental studies investigating the switching properties, that is, modulating strength and characteristic timescales, are required as well as the continued search for TCO materials with lower losses. The envisioned picture of new materials employment is complex hybrid circuitry bridging existing nanophotonic and nanoelectronic platforms and eventually, setting new standards in the levels of integration.

Acknowledgement: The authors are grateful to Nathaniel Kinsey, Jongbum Kim, and Marcello Ferrera for valuable discussions, experimental data, and help in manuscript preparation. V.E.B. acknowledges financial support from SPIE Optics and Photonics Education Scholarship and Kaj og Hermilla Ostenfeld foundation. A.B. acknowledges partial support from Air Force Office of Scientific Research grant (contract number FA9550-14-1-0138).

References
Appl. 3, doi:10.1038/lsa.2014.54 (2014).


[87] Zhang, Yun; Wei, Xiaoqing; Don, Wenjing; Huang, Chanyan; Zhang, Kenan; Sun, Yan; Chen, Xin; Dai, Ning, "Near-perfect infrared absorption from dielectric multilayer of plasmonic aluminum-doped zinc oxide." Applied Physics Letters 102, 231117 (2013).


