Attojoule-efficient graphene optical modulators

Rubab Amin,[1] (iD) Zhizhen Ma,[1] Rishi Maiti,[1] Sikandar Khan,[1] Jacob B. Khurgin,[2] Hamed Dalir,[3] and Volker J. Sorger[1,*]

[1]Department of Electrical and Computer Engineering, George Washington University, Washington, DC 20052, USA
[2]Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, Maryland 21218, USA
[3]Omega Optics, Inc., 8500 Shoal Creek Blvd., Bldg. 4, Suite 200, Austin, Texas 78757, USA
*Corresponding author: sorger@gwu.edu

Electro-optic modulation is a technology-relevant function for signal keying, beam steering, or neuromorphic computing through providing the nonlinear activation function of a perceptron. With silicon-based modulators being bulky and inefficient, here we discuss graphene-based devices heterogeneously integrated. This study provides a critical and encompassing discussion of the physics and performance of graphene. We provide a holistic analysis of the underlying physics of modulators including graphene's index tunability, the underlying optical mode, and discuss resulting performance vectors for this novel class of hybrid modulators. Our results show that reducing the modal area and reducing the effective broadening of the active material are key to improving device performance defined by the ratio of energy-bandwidth and footprint. We further show how the waveguide's polarization must be in-plane with graphene, such as given by plasmonic-slot structures, for performance improvements. A high device performance can be obtained by introducing multi- or bi-layer graphene modulator designs. Lastly, we present recent results of a graphene-based hybrid-photon-plasmon modulator on a silicon platform and discuss electron beam lithography treatments for transferred graphene for the relevant Fermi level tuning. Being physically compact, this 100 aJ/bit modulator opens the path towards a novel class of attojoule efficient opto-electronics. © 2018 Optical Society of America

OCIS codes: (250.7360) Waveguide modulators; (160.2100) Electro-optical materials; (250.5403) Plasmonics.

https://doi.org/10.1364/AO.57.00D130

1. INTRODUCTION

The functionality of optical modulation bears versatility in both optical communication and information processing [1,2]. Some key applications requiring optical signal modulation (directly or indirectly) include fast signal keying [3,4], beam steering such as in phased arrays for LIDAR [5,6], and more recently, neuromorphic computing by providing the nonlinear activation function of a perceptron [7,8]. Bringing optical modulators into technological context, in recent years both communication and computing have experienced increased challenges to deliver performance, while demand for higher data handling capacity at lower energy-per-bit functions is yet rising. The latter is driven by the rapid increase of internet and data-heavy applications (i.e., media streaming and cloud computing), and is expected to continue to rise with progressing digitalization driven by densification of connected networks and edge-computing known as the Internet of things [9]. It is understood that electronic interconnects limit data bandwidth, while optics enables natural data-handling parallelism given by the bosonic nature of photons and the weak interaction between light and matter, respectively [10–12]. Both communication and

emerging computing systems improve performance with increased modulator speed, wide spectral bandwidth, and reduced footprint and energy-per-bit functions [8,13,14]. The performance of a modulator, being a three-terminal device, is mainly driven by electrostatics (similar to a transistor), material index change capability, effective material broadening (state-smearing), and optical mode-engineering. We recently showed that reducing both the mode area and effective broadening of the active (i.e., modulated) material are critical to improve modulation performance [8,15]. This can be achieved in the former by ensuring a maximum field overlap of the optical mode with the active material by "squeezing" the optical mode to sub-diffraction-limited modes. In the latter (broadening), by trivially cooling the device increased material quality can be achieved [i.e., reduced (in)homogeneous broadening], or selecting an active material with a naturally lower broadening such as low-dimensional materials. These include predominantly quantum dots and wells, but also van der Waals layered two-dimensional (2D) materials. We previously found that the material Q-factor is generally lower than that of a cavity ("photonic-Q"), and thus another design rule for modulators is to avoid material resonances since they introduce loss. An optimized approach, however, is to utilize

photonic enhancement such as via cavity or waveguide dispersion engineering. Here we caution high cavity Q approaches, as the long photon lifetime limits the re-modulation of the active material; a Q of mid-10^4, for instance, results in a resistive-capacitive (RC)-limited speed of about 10 GHz, which is not considered fast under current signal keying standards. Modulators are classified by as electro-absorptive ($\Delta\kappa$) or electro-refractive (Δn) depending on the optical property of the material changed due to the applied electric field. A wide variety of material-index change mechanisms exist to include state-filling, Pauli blocking, exciton modulation, or free carriers. Material-wise, the contestant to production-level photonic integration modulators are silicon-based devices [16–18]. Electro-optically, silicon shows weak ($\Delta n \sim 10^{-4}$–10^{-3}) free-carrier-based dispersion [19]. As a result, silicon modulators are bulky ($\sim$millimeters long), require relatively high voltage levels (2–5 V), and have a relatively high insertion loss ($\sim$3–5 dB) given the long length. Here, we consider graphene as an active material for modulation due to its electro-optic properties originating from its unique band structure where charge carriers are governed by the relativistic Dirac equation [20,21]. The high mobility of graphene allows reduction of the RC delay time via a lowered contact resistance [8,22–24]. Graphene has the potential to fulfill the current state-of-the-art modulator performance due to following special features: first, it exhibits the highest charge carrier mobility ($\sim 10^5$ cm^2/V-s) at room temperature since the electrons in graphene behave as massless Dirac fermions, which is essential to obtain faster modulation (>tens of gigahertz) [22,25]. Second, due to its linear band structure, it can absorb the electromagnetic wave over a spectrally broad wavelength band, i.e., from ultraviolet (UV) to terahertz (THz) even in the microwave region, which provides opportunities for operating at Telecom windows [26]. Third, graphene exhibits strong light–matter interaction (LMI; $\sim$2.3% absorption) in spite of being a single-atom-thick material ($\sim$0.34 nm), which can offer exceptionally high nonlinearity for ultrafast light [27,28]. Fourth, the interband transition of graphene can be controlled by tuning the Fermi level (E_F) by means of electrical gating, optical excitation, or chemical doping [29–31]. Here, graphene becomes opaque when the Fermi level falls outside of $\pm h\nu_0/2$ due to the Pauli blocking; otherwise, graphene is transparent for a photon energy up to $2E_F$ [32]. Fifth, graphene additionally fortifies 2D plasmons due to its free-carrier absorption (i.e., intraband transitions), which exhibits strong optical confinement in the infrared wavelength region [33,34]. Finally, graphene can be integrated on any other substrate and selectively on photonic devices such as

waveguides, fibers, optical micro-cavities, and most importantly, with a mature Si CMOS platform due to its mechanical flexibility and chemical robustness [35–37]. In this work, we assess the state-of-the-art of graphene-based electro-optic modulators (EOMs). We provide a synthesized analysis starting from a device evolution benchmark over graphene optical models, waveguide mode considerations, and to recent device demonstrations showcasing the state of the art.

2. EOM PERFORMANCE MATRIX

EOMs, being the key building block for photonic integrated circuits (PICs), have attracted attention in recent years due to their ability to provide signal encoding. As mentioned earlier, the important performance matrix for EOMs includes modulation speed, power consumption, and (less importantly) the footprint area (possibly 3D volume). In this section, we summarize selected studies on EOM development, focusing on active material selection and photonic platform design using a unified performance benchmark, where we define the figure of merit (FOM) as $\text{FOM} = \frac{\text{Data capacity}}{\text{Power} \times \text{Cost}} = \frac{f_{3\text{-dB}}}{(E/\text{bit}) \times \text{Area}}$ [16]. For seamless integration with CMOS technology, extensive investigation has been sought to utilize the classic semiconductor, silicon, for electro-optic modulation by using free-carrier injection/depletion. However, due to the intrinsically low carrier concentration, and weak free-carrier dispersion of silicon, either a long arm Mach–Zehnder interferometer (MZI) or a low-loss, high-Q resonator is needed to increase the LMI. Since the RC delay of the device and power consumption both suffer from larger device volume, pioneering work on silicon MZI and micro-ring modulators have a rather low overall FOM, which sets our baseline for the performance matrix (Table 1 and Fig. 1) [17,18]. With the breakthrough on low dimensional materials, graphene has shown great tunability as an absorptive material, as demonstrated first in 2011 by integrating this atomic thin material onto a conventional photonic silicon waveguide [26]. The silicon waveguide used is rather compact (tens of micrometers), benefiting from the drastic change of graphene's absorption compared to silicon. However, although the waveguide dimension was optimized to maximize the mode overlap between the propagating mode and the active material (graphene), at telecom wavelength, the conventional photonic mode still has a certain physical dimension cutoff (220 nm for Si at 1550 nm wavelength), which is usually much larger compare to the 0.34 nm thickness of single-atom-layer

Table 1. Performance Comparison Table for Highlighted EOM Works[a]

Year	Mechanism	Bandwidth (GHz)	Footprint (μm²)	E/bit (fJ)	FOM	Refs.
2004	Si MZI	3	10,000	50,000	6×10^{-9}	[18]
2005	SI Ring	1.5	150	20,000	5×10^{-7}	[17]
2011	Graphene/P	1	40	1000	2.5×10^{-6}	[26]
2012	2× Graphene/P	10	20	2000	2.5×10^{-5}	[38]
2012	HPP ITO	25	5	56	8.9×10^{-2}	[39,40]
2015	Poly/SPS	72	20	25	1.4×10^{-1}	[41]
2015	Graphene Ring	30	6400	80	5.8×10^{-5}	[25]
2017	Graphene HPP	20*	20	3	3.3×10^{-1}	[42]
2017	Graphene SPS*	100	3	0.4	83	[43]

[a]P, Photonic; SPP, Surface Plasmon; SPS, SPP Slot; HPP, Hybrid Photon Plasmon.
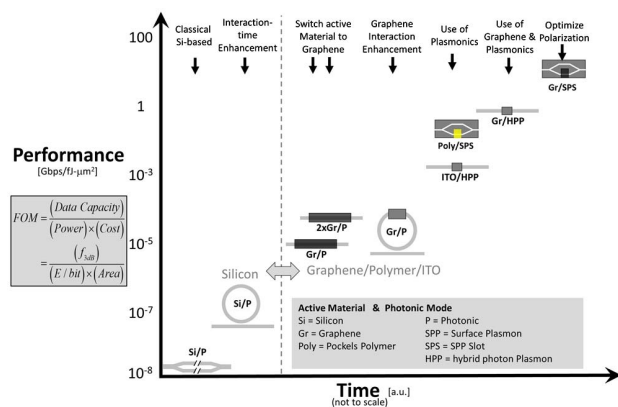
Fig. 1. Roadmap of EOM development using the pre-defined performance matrix. Graphene showed predominant advantage compared to silicon as the active material selection, while plasmonic mode (HPP/SPP/SPS) provide a much higher light-active material interaction by squeezing the mode into the active material region. Thus the integration of unity index change material and plasmonic mode could reduce the device size and bring the performance to the upper side sweet spot.

active material graphene. Thus, for either graphene EOM work from 2011 or 2012 in Table 1, the device dimension is still bulky due to the low LMI [26,38]. Nevertheless, in 2015 the Lipson group demonstrated a graphene modulator with 30-GHz-fast modulation using a low-loss Si_3N_4 micro-ring resonator as the underlying photonic platform. Yet, this device still suffers from the bulky dimensions and reduced thermal stability from the ring resonator [25].

On the other hand, increased optical confinement has been explored, such as in plasmonic modulators [39–41]. Here, the plasmonic modes increase the modal overlap factor of the active material with that of the optical mode and provide a free metallic contact that has a smaller series resistance in the device, hence reducing the RC delay compared to the photonic platform, which usually requires doped semiconductors as the electrical contact. By squeezing the mode into the active material region to increase the LMI and thus the FOM, modulators have been demonstrated by using other active materials such as transparent conductive oxides (e.g., ITO), but also Pockels-effect field-driven (non-charge-driven) materials such as polymers [39–41]. Following this hybrid-integration scheme (e.g., silicon plus novel material) Leuthold's group has shown that the latter enables a platform for ultra-fast modulation approaching 200 GHz. Lately, benefiting from the platform design and material selection, studies on integrating graphene with plasmonic modes have shown promising results as the FOM increases [42,43]. This is enabled by (a) reducing the optical mode area via high-plasmonic-mode confinement from slot-like waveguide modes, (b) the use of unity-strong tunable index materials such as graphene, (c) improved electrostatics by using thin high-k dielectrics, and (d) low-contact resistances supported by metallic synergistic contacts being part of the design. In the following sections, we will review these designs and compare different plasmonic modes for graphene integration. With optimized designs in mind, we discuss how the modulator FOM can be further improved (Fig. 1).

3. DIFFERENT MODELS FOR GRAPHENE'S OPTICAL PROPERTIES

While several graphene conductivity models are being discussed in the literature, including the random phase approximation (RPA) [44], the Kubo formulism is also being used [45]. We note that the discussion about the relative contributions of free-carrier (intraband) versus band-to-band (interband) transitions is still ongoing in the field.

A. Kubo Model and Different Approximations

Graphene shows anisotropic material properties given its dimensions; in its honeycomb-like lattice plane, the in-plane permittivity ($\varepsilon_\parallel$) can be tuned by varying its chemical potential μ_c, whereas the out-of-plane permittivity is reported to remain constant around 2.5, resembling that of graphite [46]. Here, we model graphene by a surface conductivity model $\sigma(\omega, \mu_c, \gamma, T)$, where ω is the angular frequency, μ_c is the chemical potential, γ is the scattering rate, and T is the temperature. The Kubo formula describing the conductivity is given by Eq. (1),

$$\sigma(\omega, \mu_c, \gamma, T)$$
$$= \frac{iq^2(\omega - i2\gamma)}{\pi\hbar^2} \left[\frac{1}{(\omega - i2\gamma)^2} \int_0^\infty E\left(\frac{\partial f_d(E)}{\partial E} - \frac{\partial f_d(-E)}{\partial E}\right) dE \right.$$
$$\left. - \int_0^\infty \frac{f_d(-E) - f_d(E)}{(\omega - i2\gamma)^2 - 4(E/\hbar)^2} dE \right], \quad (1)$$

where $f_d(E) = 1/[\exp((E - \mu_c)/k_B T) + 1]$ is the Fermi–Dirac distribution function, and $\hbar$ is the reduced Planck's constant $\hbar = h/2\pi$; γ is taken as 0.001 eV. We can calculate the in-plane conductivity $\sigma_\parallel$ from the Kubo formula followed by estimation of the in-plane permittivity as $\varepsilon_\parallel = 1 - \frac{\sigma_\parallel}{i\omega\varepsilon_0\Delta}$, where $\Delta = 0.35$ nm is the thickness of a single graphene sheet. The corresponding refractive index and extinction coefficient are found from the complex in-plane permittivity [43]. The first integral in Eq. (1) describes intraband motion of free electrons near energy E or holes near energy $-E$, while the second integral is due to momentum-conserving ("vertical") transitions between the states with energies $-E$ in the valence and E in the conduction band. The intraband contribution is similar to the Drude response in Si and ITO, and, since at optical frequencies in telecom region $\omega \gg \gamma$, the intraband conductivity in this region is mostly imaginary, i.e., its contribution to the dielectric constant is mostly real. In contrast, the interband contribution has a pole in the vicinity of $\omega = 2E/\hbar$, and hence it can contribute to both real and imaginary parts of optical conductivity, i.e., to both the extinction coefficient and refractive index. Both intra- and interband terms in Eq. (1) depend on the temperature, T. In the end, every carrier-driven modulator will be limited by a voltage sharpness limited by the carrier-distribution smearing width, which can be estimated by the Boltzmann approximation to be $3 - 4k_B T < 100$ meV. However, we note that all experimental EOMs operate far from this limit, requiring at least $10\times$ that amount (i.e., usually a few volts).

With applied gate voltage, the chemical potential μ_c can be tuned, and, according to Eq. (1), the real part of the conductivity (imaginary part of the permittivity) at frequencies close to $\omega_0 = |\mu_c|/\hbar$ experiences sharp change. The real part of

permittivity also experiences change, mostly within few γs from ω_0, but as shown in the literature, it is the large change of absorption that is the most attractive feature of graphene. It is important to note that only the in-plane component of the electric field experiences large change in absorption, which has consequences for the modal overlap factor, Γ. The voltage bias for the metal and silicon are denoted as V_f and V_{Si}, respectively. From Maxwell's equations, the carrier density n (i.e., $\pm$ for $e-/h+$) in graphene has to be equal to the change in the normal component of the electric displacement current across graphene, $\varepsilon_{ox}E_{ox} - \varepsilon_{Si}E_{Si} = en$, where ε_{ox} and ε_{Si} are the electric permittivity of the oxide and silicon layer, respectively; E_{ox} is the electric field across the oxide; and E_{Si} the electric field in the silicon adjacent to graphene. The Dirac point energy in graphene E_D is related to the carrier density n via $n = \int_{E_D}^{\infty} f(E)D(E - E_D)dE \approx \text{sgn}(-E_D)\frac{E_D^2}{\pi\hbar^2 v_F^2}$, where $f(E)$ is the Fermi–Dirac distribution and $D(E - E_D)$ is the density of states in graphene. Notice that the limit of $[E_D] \gg k_B T$ was used; $T = $ Temperature, $v_F = $ Fermi velocity of graphene. The electric field across the oxide is given by $E_{ox} \approx \frac{1}{et_{ox}}(\Phi_G - \Phi_m + ev_F)$, where Φ_G and Φ_m are the work functions of graphene ($\Phi_G = 4.6$ eV) and the metal (e.g., $\Phi_{Cu} = 4.65$ eV), respectively. With the electron affinity of silicon, $\chi_{Si} = 4.05$ eV, a Schottky junction exists. Here, we consider a p-doped silicon-oxide-insulator (SOI)-based waveguide core with doping n_A. The electric field in the silicon adjacent to graphene is given by $E_{Si} = (2en_A(V_{bi} - V_{Si})/\varepsilon_{Si})^{1/2}$, where V_{bi} is the built-in potential in silicon. Solving these equations relates the graphene's chemical potential μ_c, to the applied voltage, V_f.

The Kubo model can have various approximations, and we have used the Hanson approximation for the interband transitions at 0 K in previous works [15,47]. Here we want to compare the diversity among different results operating at $\lambda = 1550$ nm. All of the results depicted show the similar trend in n and κ (Fig. 2). κ decreases as Pauli blocking takes effect

Fig. 2. Different n and κ results for graphene based on Kubo formula at $\lambda = 1550$ nm. Based on results from Gusynin et al. [48], Stauber et al. ($T = 300$ K) [49], Hanson ($T = 0$ K for interband and $T = 300$ K for intraband transitions) [45], Ooi et al. ($T = 300$ K) [50], and Gosciniak and Tan [51].

and starts to grow again around 0.5 eV due to free-carrier intraband absorption. Theoretical models for calculating the optical conductivity of graphene differ from each other since different assumptions were made; for instance, Gusynin et al. developed a frequency-dependent electrical conductivity tensor using the Kubo formulation [48]. The scattering rate is assumed to be $\Gamma = 17$ K, and the Fermi velocity in graphene is taken as $v_F = 1.1 \times 10^6$ ms^{-1} [48]. Stauber et al. also developed a method based on the Kubo formulation to calculate the optical conductivity of graphene at 300 K by taking into account its full density of states and found that in the optical regime the corrections to the Dirac cone approximation are quite small, where the hopping parameter is assumed as 2.7 eV [49]. The Hanson approach approximates the interband transitions at 0 K, but the intraband transitions take the temperature dependence into account [45]. Ooi et al. also report graphene complex index data based on different approximations for the Kubo model at $T = 300$ K [50]. The results from Gosciniak et al. are also derived from the Kubo formula [51].

All these n and κ data reported follow similar trends with n decreasing in the Pauli blocking range and staying at a minimum value for the entire chemical potential range, and κ decreases to a minimum near 0.4–0.45 eV and grows again due to intraband absorption. All the data sets have n and κ intersect each other near 0.5 eV where a transition from dielectric graphene to metallic graphene can be observed. The start points in chemical potential i.e., 0 eV data are different for each result, and as a result all the plots are shifted with similar trends. This can be understood from the fact that many other factors such as quality of the graphene sheet, growth/transfer techniques, impurity, trap states, measurement conditions/criteria, etc. impact the complex graphene conductivity. Some other notable approaches towards characterizing graphene optical properties, beyond those shown in Fig. 2, are also available and can be utilized. Peres et al. studied the electronic properties of graphene in the presence of defects and electron–electron interaction as a function of temperature, external frequency, gate voltage, and magnetic field [52]. Wunsch et al. and Hwang and Das Sarma developed similar methods to calculate the conductivity of graphene from the polarization based on the Dirac cone approximation for a finite chemical potential and arbitrary radian frequency, respectively [53,54]. Simsek generated a closed-form approximate expression based on Stauber's method, which generates results with less than 0.8% maximum absolute error for $\lambda > 250$ nm [55]. The photons with energies larger than $2\mu_c$ experience absorption, while the photons with smaller energies do not; therefore, when the chemical potential moves across $\mu_c = \hbar\omega/2 \approx 0.4$ eV, the extinction coefficient undergoes rapid downward change as evidenced by the sharp negative peak of $\partial\kappa/\partial\mu_c \sim -\delta(\mu_c - 0.4)$, where $\delta(x)$ is the delta function. According to the Kramers–Kronig transform, the real part of the refractive index also experiences rapid change in the vicinity of $\mu_c \approx 0.4$ eV as $\partial n/\partial\mu_c \sim -1/\delta(\mu_c - 0.4)$. In practice, a large change in the extinction coefficient occurs in the range of μ_c values near 0.35–0.45 eV as the Pauli blocking effect can be smeared for ~ 0.1 eV at room temperature. The modulation technique we investigate here involves interband transitions, whereas intraband free-carrier absorption can rise
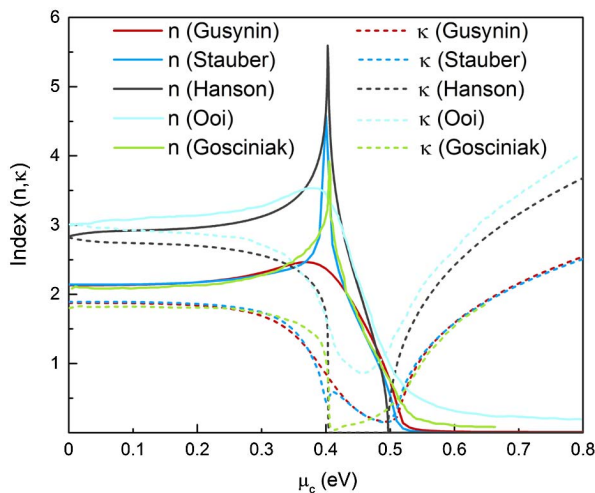
even after the Pauli blocking range of ~0.1 eV and can be approximated by the Drude model. Depending on these interband transitions of Pauli blocking and temperature-dependent smearing of the transition, n-dominance is present for μ_c ranges past the transition until the free-carrier absorptions start around 0.5 eV due to very low κ values there and abruptly changing n values. κ-dominance is on either side of the n-dominant region.

B. RPA Model

The optical response of graphene is modeled using local random phase approximation (local RPA) in the literature [44]. At temperature T, the 2D conductivity of graphene is given by [56]

$$
\begin{aligned}
\sigma_{\mathrm{RPA}}(\omega) \\
= \frac{2q^2 k_B T}{\pi \hbar^2} \frac{i}{\omega + i/\tau} \ln \left| 2 \cosh \left(\frac{\pi}{2 k_B T} \right) \right| \\
+ \frac{q^2}{4\hbar} \left\{ H(\omega/2, T) + \frac{4i\omega}{\pi} \int_0^\infty \frac{H(\zeta, T) - H(\omega/2, T)}{\omega^2 - 4\zeta^2} \mathrm{d}\zeta \right\},
\end{aligned}
$$
(2)

where $H(\omega, T) = \sinh(\hbar \omega / k_B T) / \{\cosh(\mu_c / k_B T) + \cosh(\hbar \omega / k_B T)\}$. The first term in Eq. (2) represents the intraband contribution, and the remaining terms are contributions of the interband transitions to the total graphene conductivity. Here, τ is the electron relaxation time. While Landau damping itself is already included in the conductivity model, the relaxation time typically has other contributions from: (a) impurity scattering, (b) scattering with phonons ($\hbar \omega_{\mathrm{Ph}} = 0.2$ eV) in graphene and phonon modes of polar substrates, (c) higher-order processes such as phonon coupled to $e - h$ excitations (which have to be treated separately), etc. [57–59]. In the literature, relaxation times as long as 1000 fs have been reported

Fig. 3. Graphene electro-optical properties dictated by the local random phase approximation (RPA) model; (a) single-layer graphene conductivity, σ (real and imaginary parts), and (b) real and imaginary index, n and κ versus chemical potential, μ_c.

[60,61]. However, for frequencies larger than the optical phonon frequency of graphene, typically $\tau \approx 50$ fs [57]. In this work, we use a graphene DC mobility μ_{DC} of 10^4 cm^2 V^{-1} s^{-1}, which is close to most experimental reported values, but lower than the record-high measurements. The mobility and relaxation time are related by $\tau = \mu_{\mathrm{DC}} E_F / q v_F^2$ [62]. The relaxation times considered in this work are in the same order of magnitude range as in [58].

The RPA model closely follows the Drude approximations for the intraband transitions, i.e., free-carrier absorption in graphene. But responses from the interband transitions are dissimilar to the Kubo model (Fig. 3). Hence, the optical index of graphene, both real and imaginary parts, in the local-RPA model are noticeably different from the values generated by the Kubo model. Both sets of results for graphene's optical index (Kubo and RPA) are achieved from conductivity models at $\lambda = 1550$ nm, but are quite dissimiliar in the relevant chemical potential region. It is understandable that different models will have different limitations and trade-offs in accuracy, and since there is this dissimilarity in the optical index results for both the aforementioned models and in ongoing discussions in the field to articulate accurate results from models, we choose to derive expressions for graphene interband absorption relevant to optical modulators from the first principles in the following section.

C. Ab Initio Approach for Graphene Interband Absorption

In order for us to derive expressions for graphene absorption relevant to optical modulators, we first define relevant waveguide modal parameters such as effective area and thickness. A generic schematic of a current-driven electro-absorptive modulator (EAM) is given in Fig. 4(a), showing the modulator length L and cross-sectional width W. For electrostatic gating, the active layer (thickness, d_a) is accompanied with two cladding layers on either side. The drive current I_d flows injecting charges ($-Q$ in the active layer and $+Q$ in the gate) into the active layer with applied drive voltage V_d. Beneficial to the modulator's performance is a strong LMI, which suggests a small effective mode area, S_{eff}, leading to a short effective thickness t_{eff}, which is useful towards achieving performance advantages. We evaluate the Poynting vector $S(x) = E_y H_x - E_x H_y$ for defining the effective area. Applying the Maxwell equations for the waveguide geometry, we find $S(x) = \frac{\omega n^2(x,y) \varepsilon_0 (E_x^2 + E_y^2)}{2\beta} = \frac{n^2(x,y)(E_x^2 + E_y^2)}{2 n_{\mathrm{eff}} \eta_0}$, where we introduce the effective index as $n_{\mathrm{eff}} = \beta c / \omega$. The total power flow is then $P = \iint S(x,y) \mathrm{d}x \mathrm{d}y = \frac{1}{2 n_{\mathrm{eff}} \eta_0} \iint n^2(x,y)(E_x^2 + E_y^2) \mathrm{d}x \mathrm{d}y = \frac{n_{\mathrm{eff}} E_{a0}^2}{2 \eta_0} S_{\mathrm{eff}}$, where E_{a0} is the magnitude of the transverse electric field in the active layer (Fig. 4). The effective area of the waveguide is given by

$$
S_{\mathrm{eff}} = \frac{1}{n_{\mathrm{eff}}^2} \iint n^2(x,y)(E_x^2 + E_y^2) \mathrm{d}x \mathrm{d}y / E_{a0}^2.
$$
(3)

This definition of the effective area may be different from others in the literature, but the difference is negligible and involves the distinctions between arranging of different indices, i.e., the effective and group indices. Note that the active layer is not necessary at the center of the waveguide, i.e., E_a may not be
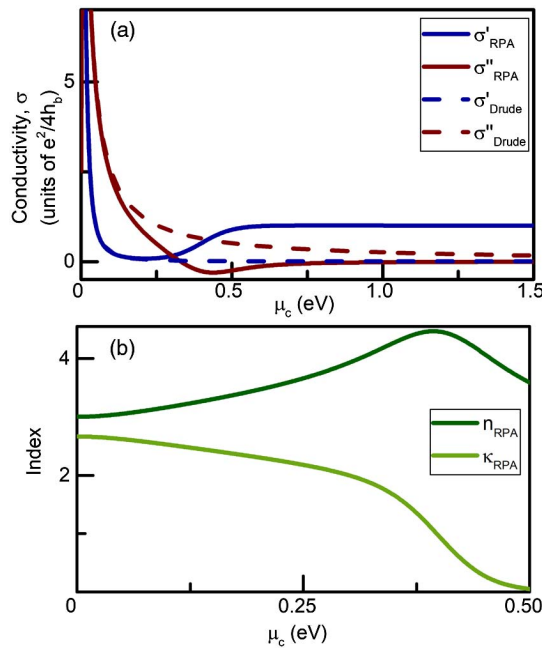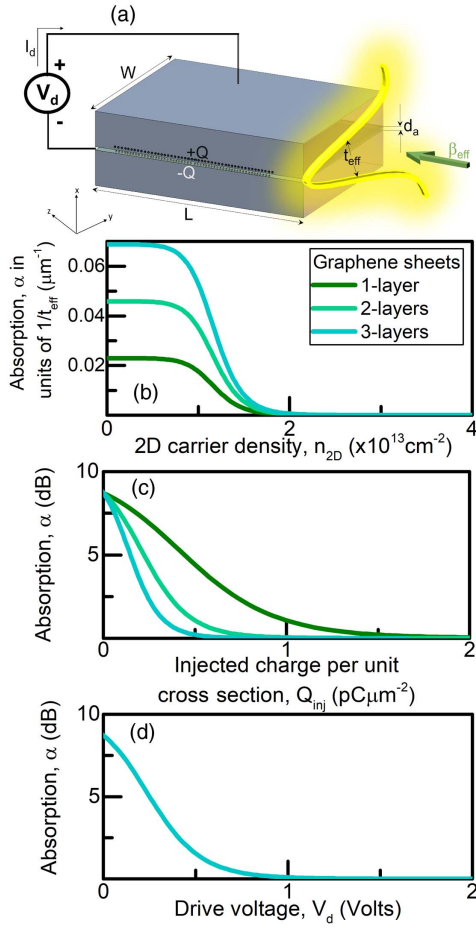
Fig. 4. (a) Schematic of the waveguide structure with a biasing scheme to the active region [drive voltage, V_d and current (charge) flow into the active layer, I_d], the charge $-Q$ in the active layer and $+Q$ in the gate is induced, and the propagation direction (indicated with green arrow, β_{eff}). The associated electric field in the y direction is shown. d_a is the width of the graphene sheet, and t_{eff} is the effective thickness. The relevant coordinate system used in this work is also included to accompany the text. (b) Normalized absorption, α in units of $1/t_{\text{eff}}$ (µm^{-1}) as a function of carrier concentration, n_{2D} (cm^{-2}). Optical absorption versus (c) injected charge, Q_{inj} (pCµm^{-2}), and (d) drive voltage, V_d (Volts), respectively.

the peak electric field. Simply in 1D, the effective thickness of the waveguide is $t_{\text{eff}} = S_{\text{eff}}/W$.

Even though there have been many derivations of graphene's optical absorption [48,49], it is useful to re-derive the interband (band-to-band) absorption in graphene from the first principles on an *ab initio* basis. This can be useful as it allows us to see how closely related the nature of the absorption in graphene is to any other 2D structure [e.g., quantum wells (QWs)] not-withstanding the presence of any Dirac electrons. In fact, when the Pauli blocking electrons are located roughly 0.4 eV away from the Dirac point (assuming $\lambda = 1550$ nm $= 0.8$ eV), all the peculiarity of graphene becomes irrelevant and its performance as an EAM is conceptually no different from any other semiconductor material. The matrix element of momentum between the valence and conduction bands is $P_{\text{cv,gr}} = m_0 v_F$, where $v_F \sim 10^8$ cm/s is the Fermi velocity. The matrix element

of the Hamiltonian in the $p \cdot A$ gauge then becomes $H_{\text{cv}} = \frac{1}{2} e v_F \cdot E/\omega$. The joint density of states can be written as

$$\rho_{2D}(\hbar\omega) = \frac{1}{2\pi} \frac{\omega}{\hbar v_F^2}, \tag{4}$$

where the dispersion relation for the transition frequency $\omega = 2kv_F$ has been implemented. We also take into account the importance of both spin and valley (K, K') degeneracies in graphene. Substituting our results into Fermi's golden rule, we obtain

$$\frac{dn_{2D}(y)}{dt} = \frac{2\pi}{\hbar} \frac{e^2}{4} E_{\text{eff}}^2 \frac{1}{2} \frac{v_F^2}{\omega^2} \frac{1}{2\pi} \frac{\omega}{\hbar v_F^2} [1 - f_c] = \frac{e^2 E_{\text{eff}}^2}{8\hbar^2 \omega} [1 - f_c]. \tag{5}$$

Here, we formulated each graphene sheet populated with n_{2D} carriers that are distributed according to the Fermi–Dirac distribution $f_c(E_c) = \{1 + \exp[(E_c - E_F)/k_B T]\}^{-1}$. Subsequently, we can acquire the expression for the graphene interband absorption coefficient in the waveguide as

$$\alpha_{gr} = \frac{\pi \alpha_0}{n_{\text{eff}} t_{\text{eff}}} [1 - f_c]. \tag{6}$$

Next, we consider the changes in graphene absorption with density of electrons to achieve modulation characteristics. Since the maximum change in absorption occurs when the Fermi level approaches $E_F = \hbar\omega/2 \gg k_B T$, the relation between density of electrons and Fermi level can be expressed as $n_{2D} \approx E_F^2/\pi\hbar^2 v_F^2$, and differentiating it with respect to the Fermi level, we obtain $dn_{2D}/dE_F \approx 2E_F/\pi\hbar^2 v_F^2$. This leads us to the absorption change as

$$\frac{d\alpha_{gr}}{dn_{2D}} = \frac{\partial \alpha}{\partial f_c} \frac{\partial f_c}{\partial E_{Fc}} \frac{dE_F}{dn_{2D}} = -\sigma_{gr}(\omega) t_{\text{eff}}^{-1}, \tag{7}$$

where the differential absorption cross section is

$$\sigma_{gr}(\omega) = \frac{\pi \alpha_0}{n_{\text{eff}}} \frac{1}{k_B T} \frac{\exp[(\hbar\omega/2 - E_{Fc})/k_B T]}{\{1 + \exp[(\hbar\omega/2 - E_{Fc})/k_B T]\}^2} \frac{\hbar^2 v_F^2}{2E_F}. \tag{8}$$

Equation (8) obviously peaks at $E_F = \hbar\omega/2$, and for maximum absorption variation we can obtain

$$\sigma_{gr}(\omega) = \frac{\pi^2 \alpha_0}{n_{\text{eff}}} \frac{1}{k_B T} \frac{1}{4} \frac{\hbar^2 v_F^2}{\hbar\omega} \approx 9.7 \times 10^{-17} \text{ cm}^2 \frac{1}{n_{\text{eff}} k_B T}. \tag{9}$$

Finally, we can introduce any additional broadening (temperature, strain, etc.), and also in order to account for the fact that the absorption does not change linearly with the Fermi level, the effective broadening can be defined as $\hbar\gamma_{\text{eff}} = ((\hbar\gamma)^2 + (3k_B T)^2)^{1/2}$, and Eq. (9) becomes

$$\sigma_{gr}(\omega) = \frac{\pi^2 \alpha_0}{\hbar\gamma_{\text{eff}}} \frac{1}{4} \frac{\hbar^2 v_F^2}{\hbar\omega} \approx 9.7 \times 10^{-17} \text{ cm}^2 \frac{1}{\hbar\gamma_{\text{eff}}}. \tag{10}$$

Compared with the results for a QW modulator, we see that the results are strikingly similar for graphene and QW modulators. This is easy to understand, since graphene and many typical III-V semiconductors have roughly similar matrix transition elements, and for a given transition energy have similar joint densities of states. Similarly, following Eqs. (4) and (6), we find the absorption coefficient dependency for graphene with 2D carrier density as

$$\alpha_{gr} = \frac{\pi\alpha_0}{n_{\text{eff}}t_{\text{eff}}} \frac{1}{\exp\left[\frac{\hbar v_F \sqrt{\pi n_{2D}} - \frac{\hbar\omega}{2}}{k_B T}\right] + 1}. \tag{11}$$

Next, we add the intraband conductivity. When the electric field $\frac{1}{2}Ee^{-i\omega t}$ + c.c. is applied according to the second law of Newton, $\frac{d}{dt}\hbar k = -\gamma\hbar k - eE$. The Fermi distribution shifts by $\Delta k = -\frac{eE/2}{\hbar(-i\omega+\gamma)} = \frac{eE/2}{\hbar}\frac{\gamma+i\omega}{\omega^2+\gamma^2}$. To calculate surface current, we assume that the temperature is much smaller than the Fermi energy E_F, and the sum up over the states then becomes $\frac{J}{2} = -e\frac{2\times2}{4\pi^2}\int_0^\pi v_F \cos\theta k_F \Delta k \cos\theta d\theta = \frac{e^2}{2\pi\hbar}v_F k_F \frac{E}{2}\frac{\gamma+i\omega}{\omega^2+\gamma^2}$, and we obtain $\sigma = \frac{e^2}{2\pi\hbar}\frac{E_F}{\hbar}\frac{\gamma+i\omega}{\omega^2+\gamma^2}$. The real part of conductivity is $\sigma' = \frac{e^2}{2\pi\hbar}\frac{E_F}{\hbar\gamma_{\text{eff}}}$ where $\gamma_{\text{eff}} = \frac{\omega^2+\gamma^2}{\gamma}$. The energy dissipation inside the graphene layer is $\frac{dU}{dt} = \frac{1}{2}\sigma'E^2 = \sigma'\eta_0 I$. Therefore, the absorption coefficient for intraband absorption is $\sigma_{\text{FCA,gr}} = \sigma'\eta_0 = \frac{e^2\eta_0}{2\pi\hbar}\frac{E_F}{\hbar\gamma_{\text{eff}}} = 2\alpha_0\frac{E_F}{\hbar\gamma_{\text{eff}}}$. This value is much less than interband absorption of graphene since $E_F \approx \hbar\omega/2$, $\alpha_{\text{FCA,gr}} \approx \alpha_0\gamma/\omega \ll \pi\alpha_0$. Therefore, the intraband absorption can be neglected for our intended purpose.

Additionally, we have evaluated the change of total absorption $\alpha_{gr}L$ as a function of the injected charge per unit waveguide cross-sectional area, $Q_{\text{inj}} = en_{2D}Lt_{\text{eff}}W/S_{\text{eff}}t_{\text{eff}} = en_{2D}L/t_{\text{eff}}$ [Fig. 4(c)]. Only the AC charge $Q_{\text{inj}} = e(n_{2D} - n_{2D,0})L/t_{\text{eff}}$ is shown [Fig. 4(c)], where $n_{2D,0} = 0.9 \times 10^{13}$ cm^{-2} is the electron density that fetches the Fermi level within $3k_BT$ from the 0.4 eV transition point. We then obtain the drive voltage via $V_d = Q_{\text{inj}}/C_g$, where C_g is the gate capacitance [Fig. 4(d)].

4. DIFFERENT WAVEGUIDE MODES FOR GRAPHENE MODULATORS

We study three different mode structures with the aim to explore modulator-suitable material/mode combinations for both electro-absorption (EA) and electro-optic (EO) modulation mechanisms. The aim is to increase the LMIs towards ultra-compact modulators while preserving a high extinction ratio (ER), i.e., modulation depth, and we consider plasmonics as a spatial mode compression tool towards increasing the LMI and compare two distinct plasmonic modes with a bulk case for comparison. The two plasmonic modes analyzed are the slot waveguide in a metal-insulator-metal (MIM) configuration [63–65], and a hybrid plasmonic polariton (HPP) design in a metal-insulator-semiconductor (MIS) configuration [66,67]. All mode structures are chosen on top of a SiO$_2$ substrate, thus providing a leveled playing field. The bulk graphene structure consists of placing a single layer graphene on a gate oxide on top of a Si waveguide, thus forming an electrical capacitor. The Si waveguide was chosen to have a height of 200 nm, which supports the second-order transverse-magnetic (TM) mode, resulting in an improved modal overlap (i.e., in-plane electric field) with the active graphene sheet [26]. The underlying silicon waveguide width is taken as diffraction limited λ/n. It is important to point out that all gate oxides in this work are fixed with an oxide thickness such that we can compare them in a similar standard (Fig. 5).

The graphene slot consists of placing a single layer of graphene on top of the SiO$_2$ substrate separated by a gating oxide of 5 nm. Then two metal pads form the slot structure. Here the gap is 20 nm, which we have found previously to deliver high modulation performance [43]. We note that broader gap dimensions lead to higher-order modes, lower optical confinement, and hence lower ER. This value (20 nm) can be understood from two aspects both relating to the fact that metallic confinement beyond 20 nm is not favorable: (a) the skin depth of plasmons at telecom wavelengths is about 20–30 nm, and (b) the Purcell factor reduces dramatically beyond tens of nanometers small plasmonic cavities due to high losses and field leakage [14]. Our results indeed confirm a modal confinement to the gap and a high field strength (peak $|\mathbf{E}|^2$), which is about 4500-fold higher compared to the bulk case [Fig. 5(e)]. The hybrid graphene mode is comprised of a metal layer on top of a 10 nm oxide layer, and the graphene single layer is sandwiched inside the oxide. These are stacked on top of a Si waveguide with 200 nm thickness [68]. The hybrid graphene structure shows reasonably high confined modes and a field enhancement of 63 times compared to the bulk case. It is worth mentioning that the slot and HPP modes are comparably lossy without accounting for the material loss to contribute as a byproduct of tuning. As such, one would intuitively expect these to be suitable for EAM devices. However, there are also regions where the EOMs (real part index tuning) via phase shifting outperform bulk cases despite the high losses. Lastly,
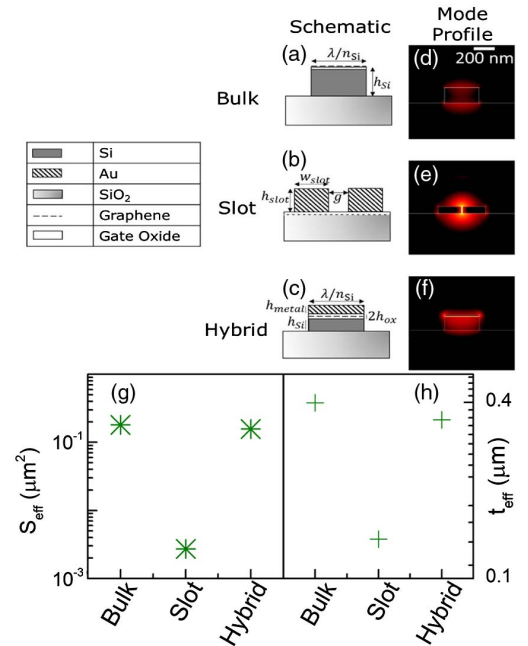

Fig. 5. (a)–(f) Schematics and FEM simulated mode profiles for different photonic and plasmonic modes for consideration in optical waveguide-based modulators. The relevant parameters are $\lambda/n_{\text{Si}} = 451$ nm, $h_{\text{Si}} = 200$ nm, $h_{\text{slot}} = 100$ nm, $w_{\text{slot}} = 300$ nm, $g = 20$ nm, $h_{\text{metal}} = 20$ nm. All gate oxides have thickness $h_{\text{ox}} = 5$ nm to ensure similar electrostatics. (g) and (h) Effective area and effective thickness for the modes in consideration. The slot graphene mode exhibits both lowest S_{eff} and t_{eff} due to the polarization matching of the E-field with the active graphene sheet in this mode, implying high LMI.

only gold (Au) is used for the modal simulations, which has a reasonably low ohmic loss at near infrared wavelengths. However, thermal softness of gold at elevated temperatures (especially for nanostructures) and the incompatibility with CMOS technology may warrant a future study to explore other plasmonic materials such as TiO_2 [69,70].

Using multiple layers of active material, i.e., multiple graphene sheets, can be an interesting approach to increase the maximum absorption—that will reduce the length, increase voltage, and increase speed. We have calculated modulator lengths and ER for up to four layers of actively biased graphene sheets [8]; beyond a few layers the biasing schemes can become tricky in the slot structure. This multilayer approach is not only limited to the slot structure by default; we only chose to showcase it since (a) it performs well among the chosen waveguide designs (Fig. 5), and (b) due to graphene's selectivity of interacting only with the in-plane components of the relevant electric field in the mode and the slot's favorable support in achieving such criteria. The ER scales linearly with the addition of extra graphene layers in the slot structure as the maximum absorption into the graphene layers increases and, as such, the device length also shortens. Our results indicate ER values of 1.2, 2.1, and 3.9 dB/µm for 1, 2, and 4 layers of graphene in the slot structure, respectively [8]. Similarly, the necessary length to achieve 3 dB modulation of the optical signal scales gradually with the number of graphene layers in the slot, i.e., device length for 1, 2, and 4 layered graphene slots are 2.5, 1.7, and 0.8 µm, respectively for a modulation depth of 3 dB [8]. With graphene, only few layers are enough as the bandgap opens up, and we need fewer carriers per layer for desired modulation. Also, bi-layer or multilayer graphene could be used, which has a higher density of states (DOS), and thus higher absorption that could be beneficial for EAMs. The introduced small bandgap is a by-product of a reduced number of electrons being wallowed around below Fermi level and not contributing to the modulation. Adding more layers demonstrates monotonic improvements in both ER and scaling, but biasing the active layers can become challenging after a few.

5. DEVICE PERFORMANCE PARAMETERS

Next we combine our details and results from above to analyze exerted modulator device performance metrics. Starting by finding the device length L, we assume an arbitrarily chosen $ER = 10$ dB for the off-state absorption, normalized by units of t_{eff} [Fig. 6(a)]. From Fig. 4(c), we can determine the switching charge necessary to obtain this 10 dB on-off ratio, based on a simple perturbative estimation, $Q_{sw} \sim 2.2e/\sigma \sim 10^{-13}$–$10^{-12}$ C/µm².

Next, we calculate the capacitance, $C_g = \varepsilon_0 \varepsilon_{gate} L / t_{eff} d_{gate}$ [Fig. 6(b)]. Now one can finally determine the switching energy-per-bit as $U_{sw} = \frac{1}{2} Q_{sw} V_{sw} = \frac{1}{2} Q_{sw}^2 / C_g$ [Fig. 6(c)]. Further, we calculate the 3-dB cutoff frequency $f_{3-dB} = 1/2\pi R C_g$, assuming $R = 50$ Ω [Fig. 6(d)].

Such a low resistance may not be realistic for photonic bulk modes, where partial and selective doping has to be used in order to keep both the carrier density and hence optical loss low. In contrast, the metal deployed in plasmonics can also serve as a low-resistance contact [8]. Graphene devices implementing the state-of-the-art advanced processes have experimentally
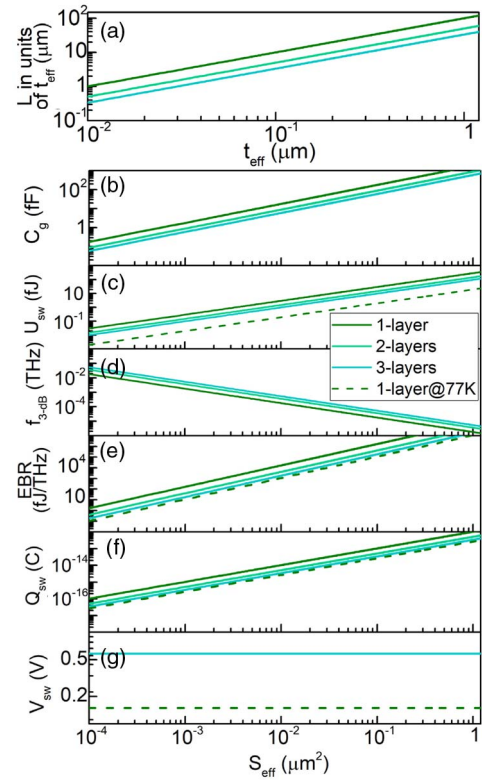

Fig. 6. (a) 10 dB absorption modulator length, L versus effective thickness, t_{eff}. (b)–(g) Relevant device parameters versus effective mode areas, S_{eff} for varying number of graphene layers and thermal cooling at 77 K to reduce the effective broadening, γ_{eff}; (b) switching charge, Q_{SW} (C); (c) capacitance, C (fF); (d) switching energy (E/bit function), U_{SW} (fJ); (e) 3-dB modulation bandwidth, f_{3-dB} (THz); (f) energy-bandwidth ratio, EBR (fJ/THz); and (g) switching voltage, V_{SW} (V). The varying amount of broadening corresponding to room temperature, i.e., 300 K for the single layer and thermal cooling down to 77 K; $d_{gate} = 5$ nm.

demonstrated the series resistance considerably reduced to below 50 Ω [4,25,71,72]. As such, the contact resistance may vary by 1 order of magnitude between photonic bulk designs and plasmonic waveguide cases. The requirement of a micrometer-tight integration of optoelectronic devices has also been mentioned recently as a viable path for attojoule-per-bit efficient devices [8,73]. The relevant FOM for modulators is the ratio of the switching energy and cutoff frequency ("energy bandwidth ratio" or EBR), $EBR = U_{sw}/f_{3-dB} = \pi Q_{sw}^2 R$, where evidently lower EBR is desired.

To study the dependence on the effective broadening, γ_{eff} for graphene with change in the operating temperature the device parameters adjusts with the scaling factor $\varsigma = T/300$ K, as our results are calculated based on room temperature performances. As such, the switching charge Q_{SW} scales as ς, while the capacitance C_g stays constant and so does the 3-dB modulation bandwidth, f_{3-dB}. The modulator length L is unchanged from the variation in broadening; the switching voltage V_{SW} scales as ς^2; switching energy, i.e., the energy-per-bit function U_{SW} and energy-bandwidth ratio EBR both scale as ς^2.

Our results demonstrate rationale for cryogenic operation for graphene modulators to cut down energy overhead and achieve better EBR. It would certainly be a viable option if operating temperatures would be near cryogenic already, as this does not take into account the energy overhead needed to cool the entire system down to near-cryogenic temperatures. Evidently, using multilayers of active graphene can help to lessen the switching energy even lower. Our results show indeed attojoule-efficient graphene modulators are feasible, and the slot structure can fetch tens of attojoule-efficient switching at room temperature. We have recently experimentally demonstrated a hybrid graphene absorption modulator requiring hundreds of attojoules per bit, discussed below.

6. EXPERIMENTAL DEVICE RESULTS

We have recently demonstrated a modulator using a HPP mode [8] that electrically resembles a metal-oxide-semiconductor (MOS) capacitor with an additional oxide-graphene sandwiched inside the MOS oxide [66]. Probing the modulator's transfer function at DC bias results in steep switching. A charge-driven modulator's switching voltage is given by $V = Q/C \propto (t_{ox}/\varepsilon_{gate})n_{eff}A_m(\gamma_{eff}/f_{12})$, where Q and C are the physical charge and gate capacitance, respectively, t_{ox} and ε_{gate} are the gate's oxide thickness and permittivity, n_{eff} is the effective mode index of the waveguide, A_m is the mode area, f_{12} is the oscillator strength, γ_{eff} is the material broadening, which is the sum over (in)homogeneous and temperature broadening (the latter ~75 meV at room temperature) [74]. Thus, low energy-per-bit functions (E/bit ~CV^2) can be achieved by improving the electrostatics (low t_{ox}, high-k dielectrics), increasing the effective mode index, reducing the mode area (i.e., sub diffraction limited modes), or improving the material quality or temperature-reducing broadening and enabling a steeper transition across the Fermi–Dirac function. In our device, we used a thin 5 nm oxide of relatively high-k (Al_2O_3, ALD deposited). In general, the optical interconnect or photonic circuit designer has voltage-choice options; for instance, one could drive the modulator across 0.75 V for a 0.4 dB/μm steep modulation, resulting in a compact $L = 8$ μm device to achieve an ER of 3 dB for small signal modulation corresponding to an E/bit of 3.2 fJ. Alternatively, a bias of only 0.1 V or ER = 0.2 dB/μm increases the capacitance due to the longer required length $L = 15$ μm, but over-proportionally reduces the energy efficiency, resulting in a low 110 aJ/bit.

The electrostatic Pauli blocking of graphene requires capacitive gating. An engineering challenge here is to minimize the voltage drop from the contacts to the device region. In photonic (non-plasmonic) devices, the weak LMI requires up to millimeter-long modulation lengths known from silicon photonics foundries. This extended length requires careful loss management, challenging insertion loss; typical methods are to highly dope selective regions near (but not including) the active device area. Such bias configuration results in typical contact resistance to the semiconductor terminal of about hundreds of ohms. With graphene's contact resistance being in the order of 10–100 Ω, a photonic device's total contact resistance is about 1000 Ω and requires 5–6 V of drive voltage [4]. In contrast, plasmonic modulators are micrometer-compact, which makes

their loss-per-length less of a factor. Here, the metal forming the compressed optical mode can be used synergistically as a low-resistance contact right at the active device region. Together, this allows the total resistance of both contacts to be on the order of 100 Ω. This means that plasmonics fundamentally allows a higher efficient usage of voltage, which particularly impacts the energy efficiency due to the ~V^2 scaling. We note that a lower contact resistance will also reduce the power dissipation of monolithically integrated plasmon lasers [11,75], 2 × 2 switches, [76] and photonic routers [77].

Indeed, experimentally, graphene's doping level can be tuned via post-transfer processes and probed in a three-terminal field-effect structure. The Fermi energy of graphene layer can be significantly altered by metal electrode deposition on graphene. The graphene's Fermi energy is shifted upon metal deposition by testing a back-gated field effect transistor design. The graphene layer is set on 50 nm silicon oxide, and the back gate is silicon. The Au contact pads are deposited by electron beam deposition. After the deposition of the Au contact pads, the Fermi energy of the graphene channel shifts to –1.80 eV, which becomes p-type doped. We found that the electrical properties of graphene can also be shifted back to n-type by exposure to the electron beam. To execute this in a controlled fashion, we selected the writing function from an electron beam system. After the electron beam treatment, we observe an $I - V_{gate}$ curve right shift corresponding to a p-type doping. The Fermi energy shifted to about 0.96 eV, which changed to n-type doped. This method shows a simple hands-on option to control the Fermi energy of graphene. In our experiment, we used a source-drain voltage, $V_{SD} = 20$ mV. Graphene channel dimensions are 10 μm × 10 μm. Exposure is done in a dose-controlled method (320 μC/cm^2) using an electron beam lithography (EBL) system. We note that the ideal chemical potential for a graphene-based modulator may not be at $\mu_c = 0$ eV; that would necessarily be at the steepest index change region in the graphene's transfer function (Figs. 2 and 3), which is near 0.4 eV (for $\lambda = 1550$ nm = 0.8 eV). Thus, in order to reduce static parasitic power dissipation, it would be elegant to permanently bias graphene to this point without applied bias for modulation. The latter could be done either chemically or via a post-processing electron doping as demonstrated here.

7. SUMMARY AND OUTLOOK

In this work, we discussed all the aspects of graphene modulators, starting from the optical properties with various theoretical models, design of different device configurations, and fabrication and analysis of their performances. We have developed an *ab initio* approach for graphene interband absorption for modulation since different approximations in the literature predict ambiguous modulation results. We have shown pathways to optimizing the modulation response of graphene modulators by varying the number of active graphene layers in the design and varying the operating temperature to achieve sharper transitions, i.e., reduced broadening. We also predict device performance based on our theoretical model and follow up with experimental device results closely following the perturbative theoretical predictions. We show pathways to achieving attojoule-efficient switching in graphene using plasmonic

modes, both theoretically and experimentally. However, many challenges still remain to be solved in terms of modulation speed, insertion loss, and energy consumption in order to use them as an optical interconnect for future data communication and computing technology. One of the main limiting factors of modulation bandwidth is the value of the RC constant, which mainly originates due to the high contact resistance of graphene devices. Several other 2D materials, such as transitional metal dichalcogenides (TMDs) and black phosphorus, emerged as more optically active beyond graphene. The interesting fact about TMDs is that the optical transition is governed by the excitons, where the exciton binding energy can be very high (~few hundreds of millielectronvolts). Like graphene, the refractive index of TMDs monolayers also can be tuned around exciton resonances by using CMOS compatible electrical gating. Henceforth, the large variety of available 2D materials and their heterostructures may outperform the current competing technologies, which could enable some new modulation mechanism.

REFERENCES

1. A. Alduino and M. Paniccia, "Wiring electronics with light," Nat. Photonics **1**, 153–155 (2017).
2. H. Subbaraman, X. Xu, A. Hosseini, X. Zhang, Y. Zhang, D. Kwong, and R. T. Chen, "Recent advances in silicon-based passive and active optical interconnects," Opt. Express **23**, 2487–2511 (2015).
3. J. Leuthold, C. Koos, W. Freude, L. Alloatti, R. Palmer, D. Korn, J. Pfeifle, M. Lauermann, R. Dinu, S. Wehrli, M. Jazbinsek, P. Gunter, M. Waldow, T. Wahlbrink, J. Bolten, H. Kurz, M. Fournier, J.-M. Fedeli, H. Yu, and W. Bogaerts, "Silicon-organic hybrid electro-optical devices," IEEE J. Sel. Top. Quantum Electron. **19**, 114–126 (2013).
4. H. Dalir, Y. Xia, Y. Wang, and X. Zhang, "Athermal broadband graphene optical modulator with 35 GHz speed," ACS Photon. **3**, 1564–1568 (2016).
5. S. M. Jackson, G. T. Reed, C. K. Tang, A. G. R. Evans, J. Clark, C. Aveyard, and F. Namavar, "Optical beamsteering using integrated optical modulators," J. Lightwave Technol. **15**, 2259–2263 (1997).
6. D. Kwong, A. Hosseini, J. Covey, Y. Zhang, X. Xu, H. Subbaraman, and R. T. Chen, "On-chip silicon optical phased array for two-dimensional beam steering," Opt. Lett. **39**, 941–944 (2014).
7. M. A. Nahmias, B. J. Shastri, A. N. Tait, and P. R. Pruncal, "A leaky integrate-and-fire laser neuron for ultrafast cognitive computing," IEEE J. Sel. Top. Quantum Electron. **19**, 1–12 (2013).
8. V. J. Sorger, R. Amin, J. B. Khurgin, Z. Ma, H. Dalir, and S. Khan, "Scaling vectors for attojoule per bit modulators," J. Opt. **20**, 014012 (2017).
9. S. J. B. Yoo, "Future prospects of silicon photonics in next generation communication and computing systems," Electron. Lett. **45**, 584–588 (2009).
10. D. Miller, "Device requirements for optical interconnects to silicon chips," Proc. IEEE **97**, 1166–1185 (2009).
11. N. Li, K. Liu, D. K. Sadana, and V. J. Sorger, "Monolithic III-V on silicon nanolaser structure for optical interconnects," Sci. Rep. **5**, 14067 (2015).
12. S. Sun, A.-H. A. Badawy, V. Narayana, T. El-Ghazawi, and V. J. Sorger, "The case for hybrid photonic plasmonic interconnects (HyPPIs): low-latency energy-and-area- efficient on-chip interconnects," IEEE Photon. J. **7**, 4801614 (2015).
13. A. N. Tait, T. Ferreira de Lima, E. Zhou, A. X. Wu, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Neuromorphic photonic networks using silicon photonic weight banks," Sci. Rep. **7**, 7430 (2017).
14. K. Liu, S. Sun, A. Majumdar, and V. J. Sorger, "Fundamental scaling laws in nanophotonics," Sci. Rep. **6**, 37419 (2016).
15. R. Amin, C. Suer, Z. Ma, J. Khurgin, R. Agarwal, and V. J. Sorger, "Active material, optical mode and cavity impact on electro-optic modulation performance," Nanophotonics **7**, 455–472 (2017).
16. G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Omson, "Silicon optical modulators," Nat. Photonics **4**, 518–526 (2010).
17. Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, "Micrometre-scale silicon electro-optic modulator," Nature **435**, 325–327 (2005).
18. A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. Paniccia, "A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor," Nature **427**, 615–618 (2004).
19. R. Soref and B. Bennett, "Electro-optical effects in silicon," IEEE J. Quantum Electron. **23**, 123–129 (1987).
20. A. K. Geim and K. S. Novoselov, "The rise of graphene," Nat. Mater. **6**, 183–191 (2007).
21. F. Bonaccorso, Z. Sun, T. Hasan, and A. Ferrari, "Graphene photonics and optoelectronics," Nat. Photonics **4**, 611–622 (2010).
22. W. Li, B. Chen, C. Meng, W. Fang, Y. Xiao, X. Li, Z. Hu, Y. Xu, L. Tong, H. Wang, W. Liu, J. Bao, and Y. R. Shen, "Ultrafast all-optical graphene modulator," Nano Lett. **14**, 955–959 (2014).
23. Z. Sun, T. Hasan, F. Torrisi, D. Popa, G. Privitera, F. Wang, F. Bonaccorso, D. M. Basko, and A. C. Ferrari, "Graphene mode-locked ultrafast laser," ACS Nano **4**, 803–810 (2010).
24. X. Gan, R.-J. Shiue, Y. Gao, I. Meric, T. F. Heinz, K. Shepard, J. Hone, S. Assefa, and D. Englund, "Chip-integrated ultrafast graphene photo-detector with high responsivity," Nat. Photonics **7**, 883–887 (2013).
25. C. T. Phare, L. Y.-H. Daniel, J. Cardenas, and M. Lipson, "Graphene electro-optic modulator with 30 GHz bandwidth," Nat. Photonics **9**, 511–514 (2015).
26. 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 **474**, 64–67 (2011).
27. E. Hendry, P. J. Hale, J. Moger, A. K. Savchenko, and S. A. Mikhailov, "Coherent nonlinear optical response of graphene," Phys. Rev. Lett. **105**, 097401 (2010).
28. S. Bhattacharya, R. Maiti, A. C. Das, S. Saha, S. Mondal, S. K. Ray, S. N. B. Bhaktha, and P. K. Dutta, "Efficient control of ultrafast optical nonlinearity of reduced graphene oxide by infrared reduction," J. Appl. Phys. **120**, 013101 (2016).
29. F. Wang, Y. Zhang, C. Tian, C. Girit, A. Zettl, M. Crommie, and Y. R. Shen, "Gate-variable optical transitions in graphene," Science **320**, 206–209 (2008).
30. R. Maiti, S. Haldar, D. Majumdar, A. Singha, and S. K. Ray, "Hybrid opto-chemical doping in Ag nanoparticle decorated monolayer CVD graphene probed by Raman spectroscopy," Nanotechnology **28**, 075707 (2017).
31. J. Lee, K. S. Novoselov, and H. S. Shin, "Interaction between metal and graphene: dependence on the layer number of graphene," ACS Nano **5**, 608–612 (2011).
32. Z. Li, E. Henriksen, Z. Jiang, Z. Hao, M. Martin, P. Kim, H. Stormer, and D. Basov, "Dirac charge dynamics in graphene by infrared spectroscopy," Nat. Phys. **4**, 532–535 (2008).
33. A. N. Grigorenko, M. Polini, and K. S. Novoselov, "Graphene plasmonics," Nat. Photonics **6**, 749–758 (2012).
34. R. Maiti, T. K. Sinha, S. Mukherjee, B. Adhikari, and S. K. Ray, "Enhanced and selective photodetection using graphene-stabilized hybrid plasmonic silver nanoparticles," Plasmonics **11**, 1305–1306 (2016).
35. N. Youngblood, Y. Anugrah, R. Ma, S. J. Koester, and M. Li, "Multifunctional graphene optical modulator and photodetector integrated on silicon waveguides," Nano Lett. **14**, 2741–2746 (2014).

36. C. Qiu, W. Gao, R. Vajtai, P. M. Ajayan, J. Kono, and Q. Xu, "Efficient modulation of 1.55 μm radiation with gated graphene on a silicon micro-ring resonator," Nano Lett. **14**, 6811–6815 (2014).

37. Y. Gao, R.-J. Shiue, X. Gan, L. Li, C. Peng, I. Meric, L. Wang, A. Szep, D. Walker, Jr., and J. Hone, "High-speed electro-optic modulator integrated with graphene-boron nitride heterostructure and photonic crystal nanocavity," Nano Lett. **15**, 2001–2005 (2015).

38. M. Liu, X. Yin, and X. Zhang, "Double-layer graphene optical modulator," Nano Lett. **12**, 1482–1485 (2012).

39. V. J. Sorger, N. D. Lanzillotti-Kimura, R. M. Ma, and X. Zhang, "Ultra-compact silicon nanophotonic modulator with broadband response," Nanophotonics **1**, 17–22 (2012).

40. G. A. Keeler, S. Campione, M. G. Wood, D. K. Serkland, S. Parameswaran, J. Ihlefeld, T. S. Luk, J. R. Wendt, and K. M. Geib, "Reducing optical confinement losses for fast, efficient nanophotonic modulators," in *IEEE Photonics Society Summer Topical Meeting Series (SUM)* (2017), pp. 201–202.

41. C. Haffner, W. Heni, Y. Fedoryshyn, J. Niegemann, A. Melikyan, D. L. Elder, B. Baeuerle, Y. Salamin, A. Josten, U. Koch, C. Hoessbacher, F. Ducry, L. Juchli, A. Emboras, D. Hillerkuss, M. Kohl, L. R. Dalton, C. Hafner, and J. Leuthold, "All-plasmonic Mach-Zehnder modulator enabling optical high-speed communication at the microscale," Nat. Photonics **9**, 525–528 (2015).

42. S. Khan, Z. Ma, J. Jeon, C. Lee, and V. J. Sorger, "Sub-voltage graphene-plasmon based electro-absorption modulator," in *Advanced Photonics 2017 (IPR, NOMA, Sensors, Networks, SPPCom, PS)*, OSA Technical Digest (online) (Optical Society of America, 2017), paper ITu3A.4.

43. Z. Ma, M. H. Tahersima, S. Khan, and V. J. Sorger, "Two-dimensional material-based mode confinement engineering in electro-optic modulators," IEEE J. Sel. Top. Quantum Electron. **23**, 81–88 (2017).

44. A. Kumar, T. Low, K. H. Fung, P. Avouris, and N. X. Fang, "Tunable light-matter interaction and the role of hyperbolicity in graphene-hBN system," Nano Lett. **15**, 3172–3180 (2015).

45. G. Hanson, "Dyadic Green's functions and guided surface waves for a surface conductivity model of graphene," J. Appl. Phys. **103**, 064302 (2008).

46. W. Gao, J. Shu, C. Qiu, and Q. Xu, "Excitation of plasmonic waves in graphene by guided-mode resonances," ACS Nano **6**, 7806–7813 (2012).

47. R. Amin, C. Suer, Z. Ma, I. Sarpkaya, J. B. Khurgin, R. Agarwal, and V. J. Sorger, "A deterministic guide for material and mode dependence of on-chip electro-optic modulator performance," Solid State Electron. **136**, 92–101 (2017).

48. V. P. Gusynin, S. G. Sharapov, and J. P. Carbotte, "Magneto-optical conductivity in graphene," J. Phys. Condens. Matter **19**, 026222 (2007).

49. T. Stauber, N. M. R. Peres, and A. K. Geim, "Optical conductivity of graphene in the visible region of the spectrum," Phys. Rev. B **78**, 085432 (2008).

50. K. J. A. Ooi, L. K. Ang, and D. T. H. Tan, "Waveguide engineering of graphene's nonlinearity," Appl. Phys. Lett. **105**, 111110 (2014).

51. J. Gosciniak and D. T. H. Tan, "Graphene-based waveguide integrated dielectric-loaded plasmonic electro-absorption modulators," Nanotechnology **24**, 185202 (2013).

52. N. M. R. Peres, F. Guinea, and A. H. Castro Neto, "Electronic properties of disordered two-dimensional carbon," Phys. Rev. B **73**, 125411 (2006).

53. B. Wunsch, T. Stauber, F. Sols, and F. Guinea, "Dynamical polarization of graphene at finite doping," New J. Phys. **8**, 318 (2006).

54. E. H. Hwang and S. Das Sarma, "Dielectric function, screening, and plasmons in two-dimensional graphene," Phys. Rev. B **75**, 205418 (2007).

55. E. Simsek, "A closed-form approximate expression for the optical conductivity of graphene," Opt. Lett. **38**, 1437–1439 (2013).

56. L. A. Falkovsky and A. A. Varlamov, "Space-time dispersion of graphene conductivity," Eur. Phys. J. B **56**, 281–284 (2007).

57. H. Yan, T. Low, W. Zhu, Y. Wu, M. Freitag, X. Li, F. Guinea, P. Avouris, and F. Xia, "Damping pathways of mid-infrared plasmons in graphene nanostructures," Nat. Photonics **7**, 394–399 (2013).

58. A. Woessner, M. B. Lundeberg, Y. Gao, A. Principi, P. Alonso-Gonzalez, M. Carrega, K. Watanabe, T. Taniguchi, G. Vignale, M. Polini, J. Hone, R. Hillenbrand, and F. H. L. Koppens, "Highly confined low-loss plasmons in graphene-boron nitride heterostructures," Nat. Mater. **14**, 421–425 (2015).

59. M. Jablan, M. Soljacic, and H. Buljan, "Plasmons in graphene: fundamental properties and potential applications," Proc. IEEE **101**, 1689–1704 (2013).

60. C. R. Dean, A. F. Young, I. Meric, C. Lee, L. Wang, S. Sorgenfrei, K. Watanabe, T. Taniguchi, P. Kim, K. L. Shepard, and J. Hone, "Boron nitride substrates for high-quality graphene electronics," Nat. Nanotechnol. **5**, 722–726 (2010).

61. K. Bolotin, K. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H. Stormer, "Ultrahigh electron mobility in suspended graphene," Solid State Commun. **146**, 351–355 (2008).

62. M. Jablan, H. Buljan, and M. Soljačić, "Plasmonics in graphene at infrared frequencies," Phys. Rev. B **80**, 245435 (2009).

63. J. Dionne, K. Diest, L. Sweatlock, and H. A. Atwater, "PlasMOStor: a metal-oxide-Si field effect plasmonic modulator," Nano Lett. **9**, 897–902 (2009).

64. 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 Lett. **14**, 6463–6468 (2014).

65. L. Lafone, T. P. H. Sidiropoulos, and R. F. Oulton, "Silicon-based metal-loaded plasmonic waveguides for low-loss nanofocusing," Opt. Lett. **39**, 4356–4359 (2014).

66. R. F. Oulton, V. J. Sorger, D. A. Genov, D. F. P. Pile, and X. Zhang, "A hybrid plasmonic waveguide for subwavelength confinement and long-range propagation," Nat. Photonics **2**, 496–500 (2008).

67. V. J. Sorger, Z. Ye, R. F. Oulton, Y. Wang, G. Bartal, X. Yin, and X. Zhang, "Experimental demonstration of low-loss optical waveguiding at deep sub-wavelength scales," Nat. Commun. **2**, 331 (2011).

68. C. Ye, S. Khan, Z. R. Li, E. Simsek, and V. J. Sorger, "λ-Size ITO and graphene-based electro-optic modulators on SOI," IEEE J. Sel. Top. Quantum Electron. **20**, 40–49 (2014).

69. P. R. West, S. Ishii, G. V. Naik, N. K. Emani, V. M. Shalaev, and A. Boltasseva, "Searching for better plasmonic materials," Laser Photon. Rev. **4**, 795–808 (2010).

70. G. V. Naik, V. M. Shalaev, and A. Boltasseva, "Alternative plasmonic materials: beyond gold and silver," Adv. Mater. **25**, 3264–3294 (2013).

71. W. S. Leong, H. Gong, and J. T. L. Thong, "Low-contact-resistance graphene devices with nickel-etched-graphene contacts," ACS Nano **8**, 994–1001 (2014).

72. W. Li, Y. Liang, D. Yu, L. Peng, K. P. Pernstich, T. Shen, A. R. H. Walker, G. Cheng, C. A. Hacker, C. A. Richter, Q. Li, D. J. Gundlach, and X. Liang, "Ultraviolet/ozone treatment to reduce metal-graphene contact resistance," Appl. Phys. Lett. **102**, 183110 (2013).

73. D. A. B. Miller, "Attojoule optoelectronics for low-energy information processing and communications," J. Lightwave Technol. **35**, 346–396 (2017).

74. R. Amin, J. B. Khurgin, and V. J. Sorger, "Waveguide-based electro-absorption modulator performance," arXiv:1712.02837 (2017).

75. C. Ye, K. Liu, R. Soref, and V. J. Sorger, "3-waveguide 2 × 2 plasmonic electro-optic switch," Nanophotonics **4**, 261–268 (2015).

76. S. Sun, V. Narayana, I. Sarpkaya, J. Crandall, R. A. Soref, T. El-Ghazawi, and V. J. Sorger, "Hybrid photonic-plasmonic non-blocking broadband 5 × 5 router for optical networks," IEEE Photon. J. **PP**, 1 (2017).

77. K. Liu and V. J. Sorger, "Electrically-driven carbon nanotube-based plasmonic laser on silicon," Opt. Mater. Express **5**, 1910–1919 (2015).