Graphene Modulators and Switches Integrated on Silicon and Silicon Nitride Waveguide

Leili. A. Shiramin, Dries. Van Thourhout

Abstract— The optoelectronic properties of graphene attracted a lot of interest in recent years. Several demonstrations of integrated graphene based modulators, switches, detectors and non-linear devices have been reported. We present here a comprehensive study investigating the different design trade-offs involved in realizing in particular graphene based modulators and switches. We compare 4 representative hybrid graphene-waveguide configurations, focusing on optimizing their dimensions, the gate-oxide thickness, the polarization, the operating wavelength and contact definition. We study both static and dynamic behavior, defining a relevant figure of merit. We find that a 20 µm device based on silicon waveguides should allow for 25 GBit/s modulation rate and an extinction ratio of 5 dB. A 200 µm long SiN-device on the other hand should allow for 23dB extinction ratio and switching speeds down to 0.4 ns.

Index Terms— Double layer, Graphene, Silicon, Silicon nitride, waveguide, interconnects.

I. INTRODUCTION

GRAPHENE has been demonstrated to exhibit unique optoelectronics properties [1]. It has been exploited to demonstrate several photonics devices including photodetectors [2,3,4], modulators [5,6,7,8], transistors [9,10], polarizers [11], saturable absorbers [12,13] and heaters [14].

A promising route to enhance the interaction between light and the graphene layer is to integrate it with an optical waveguide. Integration with different waveguide types, including silicon and silicon nitride waveguides has been demonstrated in literature, but thus far the trade-offs involved in the selection of the optimal waveguide configuration have only be discussed to a limited extent. E.g. crystalline silicon waveguides provide more confined modes, possibly leading to more compact devices, while silicon nitride, as a deposited material, is possibly more cost effective and provides more freedom in the positioning of the graphene sheet. The discussed trade-offs are relevant for all devices mentioned above but in this paper we will focus on modulators and switches. For modulators the objective is reaching the highest modulation rate, while maintaining a reasonable extinction ratio and insertion loss. Both graphene silicon and graphene silicon nitride based modulators have been proposed in literature. For switches the focus is on obtaining a high extinction ratio and low insertion loss, for a switching speed of ~1-10ns (or slower). Switch architectures proposed in literature so far mostly rely on semiconductor optical amplifiers (SOA) [15,16], liquid crystals [17,18] or the thermo-optic effect [19]. The only graphene-based switches demonstrated thus far are a graphene plasmonic mode switch operating at THz frequencies [20] and a device switching the reflectivity for mid-infrared light using back-gated single layer graphene [21]. In this work to the contrary, we focus on the on-off switching of a guided mode in the telecom bands at wavelengths of 1310nm and 1550 nm.

The waveguide configurations studied in this work are shown in Fig. 1. The parameters to be optimized include the material platform (Si or Si3N4), the polarization of the guided mode, the waveguide dimensions (width and height), the device length and the contact spacing. We focus on non-resonant devices so that we can fully exploit the broadband response of graphene.

II. DEVICE CONFIGURATIONS

The first two configurations are based on a silicon waveguide and were originally proposed in [5,6]. The main difference is that the waveguides are assumed to be planarized by CMP (chemical mechanical polishing) to ease the integration process with the graphene layer (see also [7]). The single layer graphene on silicon configuration (SLG-Si) shown in Fig. 1(a) has a single layer of graphene transferred on the silicon waveguide, with a thin oxide layer in between. The silicon waveguide here serves as the back gate. The double layer graphene on silicon configuration (DLG-Si) has two layers of graphene separated by a thin dielectric, on top [Fig. 1(b)]. Typically a layer of Aluminum Oxide deposited by atomic layer deposition is used as the dielectric in this capacitive stack. The voltage to drive the device is now applied between top and bottom graphene layer. The double layer graphene on silicon nitride device (DLG-SiN, Fig. 1(c)) is conceptually identical to the DLG-Si, except for the fact that the waveguide is now formed by a lower index Si3N4 strip (nSi3N4=2.0 vs. nSi=3.5). Finally, given that Si3N4 is a deposited material, it is possible to embed the double layer graphene stack within the waveguide core, leading to the double layer graphene embedded in SiN device (DLG-E-SiN) shown in Fig. 1(d).
Our simulations were performed based on the surface conductivity method using the finite-element solver COMSOL. The conductivity of graphene was extracted from the Kubo formula [22] taking into account both interband and intraband transitions and depends on the wavelength of the light, the chemical potential $\mu$ of the graphene layer (controlled through the applied voltage), the carrier relaxation time ($\tau_r$) of the graphene layer (depending on the quality of the graphene after full fabrication) and the temperature of the device. In all simulations we assumed a temperature $T=300 K$.

Fig. 2 plots the absorption for the different proposed waveguide configurations, as function of the waveguide dimensions, at the neutrality point ($\mu=0$ eV). The absorption is a measure for the interaction between the optical mode and the graphene layer, hence a higher value is preferred. Fig. 2(a) shows that for all structures the absorption increases with increasing waveguide width (assuming waveguide height of 220nm for the Si structures and 300nm for the SiN structures). However there are important differences between different configurations. Assuming quasi TM polarization, the DLG-Si devices exhibit almost twice the absorption of the SLG-Si device. For TE-polarization, the absorption in the Si-based devices is considerably lower (up to 0.12 dB/μm and 0.08 dB/μm for DLG-Si and SLG-Si devices respectively, e.g. see [7], not shown in figure). This is related to the interaction with the strong longitudinal electric field component at the top interface of the silicon waveguides (note that the graphene sheet does not interact with the out-of-plane field components [22]). In the SiN-devices on the other hand, the quasi TE-polarized modes exhibit higher interaction with the graphene film than the quasi TM modes. The DLG-SiN has lower absorption than the DLG-Si device but by embedding the graphene stack within the SiN waveguide, the absorption level of DLG-E-SiN can be increased again, almost to the level of the DLG-Si, albeit at larger waveguide dimensions. The simulations shown in Fig. 2(a) were carried out for a gate oxide $d_{Al2O3} = 5$ nm. Varying the gate oxide from 1nm to 15nm has very limited impact on the absorption (<0.03 dB/μm).

Finally, it can be seen that, e.g. for the DLG-E-SiN with optimized layer thickness, the absorption at a wavelength of 1310nm is significantly higher than for 1550 nm. This is related to a higher overlap of the optical mode with the graphene layer at 1310 nm, rather than to a change in the intrinsic absorption properties of graphene, which are fairly constant over the wavelength interval considered.

Fig. 2(b) depicts the effect of the SiN waveguide thickness on the absorption of the DLG-E-SiN device both at a wavelength of 1310 nm and 1550 nm, assuming the graphene layers are embedded symmetrically in the center of the waveguide. At 1550 nm, the absorption peaks for a total waveguide thickness of 300 nm (150 nm bottom layer, 150 nm top layer), while at 1310 nm it reaches its maximum for a 250 nm thick waveguide. In the remainder of the paper we assume a 300 nm thick waveguide for both wavelengths. Simulations further also showed that the 220 nm thick silicon waveguide widely used in the industry is close to optimal in terms of height.
while remaining sufficiently far from the multimode regime.

### Table I

<table>
<thead>
<tr>
<th>Structure</th>
<th>Polarization</th>
<th>Width (nm)</th>
<th>Waveguide thickness (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLG-E-SiN</td>
<td>TE</td>
<td>1200</td>
<td>300</td>
</tr>
<tr>
<td>DLG-SiN</td>
<td>TE</td>
<td>1200</td>
<td>300</td>
</tr>
<tr>
<td>DLG-Si</td>
<td>TM</td>
<td>750</td>
<td>220</td>
</tr>
<tr>
<td>SLG-Si</td>
<td>TM</td>
<td>750</td>
<td>220</td>
</tr>
</tbody>
</table>

Applying a voltage over the graphene layer changes its chemical potential $\mu_c$, equivalent to shifting its Fermi level. This implies higher interband transition energy and therefore the absorption of the graphene layer is suppressed. Given the high carrier mobility of graphene ($\sim 10^3$ cm$^2$/V·s$^{-1}$ at room temperature, taking into account a realistic fabrication process [23,24]) this effect is quasi instantaneous. The change in the absorption as function of voltage is illustrated in Fig. 3 for an Aluminum Oxide layer as a gate oxide with relative permittivity of 5 and thickness of 1nm, 3nm and 5 nm. In calculating these curves we took into account the geometrical capacitance of the stack, using the approach described e.g. in [7, 22] and $V_0$ which is the offset induced by the background doping $n_b$ of the graphene with $n_b = n_{ox e} \cdot V_0$ [25], whereby we assumed that both layers are equally but oppositely doped. If this is not the case, e.g. if both layers have the same type of doping, both the maximum achievable absorption and the modulation efficiency will decrease, in worst case to half the level of Fig. 3. The curves are plotted for two values of $\tau$, 65fs and 13fs respectively, whereby the higher $\tau$ is equivalent with less scattering and hence higher quality graphene (similar values were used in [26]). The change in $\tau$ has little impact on the maximum absorption but strongly affects the remaining absorption at higher voltages, with strong consequences on the insertion loss and overall figure of merit of the devices as will be discussed further in the paper. The shape of the curves is only determined by $\mu_c$. However, the voltage needed to reach a given $\mu_c$ in the graphene layer strongly depends on the gate oxide thickness $d_{Al2O3}$ and hence the voltage needed to reach transparency increases for the thicker gate oxides.

From Fig. 3, we can now calculate the modulation efficiency defined as $(\alpha_{2203} - \alpha_{2203})/IL(\alpha_{2203})$, the insertion loss (IL) and extinction ratio (ER) for the different devices. The IL is determined at high forward voltage, where the absorption is quasi-minimal while the ER is calculated as the ratio of the maximum and minimum absorption.

![Figure 3](image)

The results are shown in Fig. 4, as function of length of the device for $\tau =$ 65fs. The modulation efficiency strongly depends on $d_{Al2O3}$ and a 5 nm oxide reduces the modulation efficiency considerably in comparison to a 1 nm oxide [Fig. 4(a)]. IL and ER on the other hand, are nearly independent of $d_{Al2O3}$ [Fig. 4(b)]. Further the modulation efficiency improves for devices, which interact stronger with the graphene film (DLG-Si > DLG-E-SiN > SLG-Si > DLG-SiN). This is also true for the ER but the IL shows an opposite behavior: stronger interacting structures exhibit higher excess loss for a given length.

To get more insight in this trade-off, we calculate a figure of merit often used for electro-absorption based modulators, $FOM= ER(\alpha_{2203} - \alpha_{2203})/IL(\alpha_{2203})$ as function of the applied drive voltage $\Delta V$ (with $\Delta V = V_{f}$...
V_2 and V_{bias}=(V_1+V_2)/2 chosen such that the FOM is maximized. The results are shown in Fig. 5. The FOM strongly depends on d_{Al2O3}, the applied drive voltage and t_r. Introducing lower quality graphene with shorter scattering time (t_r = 13 fs vs t_r = 65 fs) reduces the best achievable FOM by almost a factor of five. But, interestingly, the FOM is independent of the chosen device configuration. A similar conclusion was reached in a more general context in [27]. Hence, only considering the static performance of the device one should select the thinnest oxide layer technologically feasible while the detailed device configuration can be decided on other grounds, e.g. compatibility with other devices or cost. This conclusion will no longer be correct however when taking also the dynamic response into account as discussed in the next section.

Thus far we only considered the insertion loss related to the graphene layer itself. However, we also have to take into account the losses caused by the metal contacts and by the resistive part of the graphene outside of the capacitive stack (see Fig. 6(a)). Here we have assumed Palladium as the contact because of its low contact resistance on graphene [28,29,30]. Fig. 6(b) shows the impact of the metal contacts as function of its spacing (s) with the waveguide, assuming, for technological reasons, that the graphene capacitive stack itself extends 300nm besides the waveguide (d_o=300nm). Extending the spacing beyond 0.75 \mu m (DLG-Si) and 1 \mu m (others) is sufficient to suppress the loss of the metal contacts almost completely. The remaining loss then stems from the overlap with the non-transparent graphene (absorber graphene, d_A) outside the capacitive stack. Fig. 6(c) shows the IL as function of the parameter d_o, which indicates how far the two overlapping graphene layers extent besides the
waveguide. Increasing $d_0$ reduces the loss due to the non-transparent graphene. However, note that this also has a strong impact on the overall capacitance of the device, negatively impacting its modulation speed. Similarly increasing the metal contact spacing $s$ [Fig. 6(b)] increases the series resistance of the device again resulting in a reduced modulation speed. Therefore, as will be shown in the next section, in particular for high speed modulators it might be beneficial to tolerate some added IL and reduce both $d_0$ and $d_A$ to maximize the speed.

All the devices discussed above operate over a wide wavelength range, inheriting the broadband modulation properties of graphene. As an example Fig. 7 shows the modulation curves for a DLG-SiN device ($d_{AlOx}$=3nm, $\tau_e$ = 65fs). Nearly wavelength independent operation is obtained from 1500nm to 1600nm and from 1250 nm to 1350 nm. Operating at 1310 nm results in a higher absorption as discussed above but also requires a higher drive voltage.

IV. Dynamic Performance

The frequency response of the devices can be calculated using the electrical circuit model shown in Fig. 8(a), similar to the one used in ref [26]. $R_{sub}$ and $C_{box}$ present the substrate resistance and buried oxide capacitance, $C_{g,sh}$ is the capacitance between the electrodes through the air and $C_{GIG}$ is the capacitance of the graphene-insulator-graphene stack. The series resistance, $R_s$, is the sum of the contact resistance ($R_{g,c}$) and sheet resistance ($R_{g,sh}$) of the first and second graphene layer.

$C_{GIG}$ comprises the oxide capacitance ($C_{ox}$) and quantum capacitance ($C_q$) of top and bottom graphene layers [31] where $C_{ox}$ was calculated using a simple parallel plate approximation and $C_q$ is dependent on the carrier density in the graphene sheets. This means $C_q$ and hence the modulation bandwidth $f_{3db}$ are voltage dependent. Therefore we calculated $f_{3db}$ at the bias voltage exhibiting the highest modulation efficiency (1.6 V, 4.27 V and 6.9 V for 1 nm, 3 nm and 5 nm oxide thickness respectively), $C_q$ was calculated assuming a background doping of $n_b = 9 \times 10^{12}$ cm$^{-2}$ and the doping from bias voltage. The relative permittivity of Aluminum Oxide varies for different fabrication processes, hence in our calculations it is assumed to be 5 or 9. The graphene contact resistance is assumed to be $R_{g,c} = 150 \Omega \mu m$, a value achievable e.g. through edge contacting of graphene [23,24,27]. The graphene sheet resistance $R_{g,sh}$ is varied between 100 $\Omega$/sq and 300 $\Omega$/sq, equivalent with a mobility of 6900 cm$^2$/V.s and 2300 cm$^2$/V.s, respectively, assuming the same back ground doping as used for the $C_{g,c}$ calculation ($n_b = 9 \times 10^{12}$ cm$^{-2}$) and the spacing $s$ between the waveguide edge and the metal contact edge was varied between 1 $\mu$m and 2$\mu$m. The effect of $s$ is included in the capacitance width ($d_c$), which comprises the waveguide width and twice the overlap region $d_o$. $C_{air}$, $C_{box}$ and $R_{sub}$ were taken from ref [32].

The electrical bandwidth was calculated assuming a 50 $\Omega$ source resistance [Fig. 8(b)-(c)]. Fig. 8(b) shows how the modulation bandwidth changes as function of device length for varying oxide thickness and capacitor width ($d_c$). The fact that the 50 $\Omega$ source resistance is of the same order as the intrinsic device resistance comprising $R_{g,c}$ and $R_{g,sh}$ and varies roughly between 1.85 $\Omega$ and 90 $\Omega$ for the parameters considered in Fig. 8(b)-(c), makes that the overall modulation bandwidth strongly decreases with increasing length. Further one notes that a smaller $d_c$ and a thicker oxide lead to higher bandwidth. However, as discussed in the previous paragraph, these also lead to lower modulation efficiency and higher IL. A 200 $\mu$m long device with $d_c=1.8 \mu$m, $d_{AlOx}=3$ nm has a bandwidth 2.5 GHz, compatible with 0.4 nanosecond switching time. Assuming a DLG-SiN configuration, and a 3V drive voltage this device also exhibits a low IL (1dB) and a high ER (23dB). Alternatively, a 20 $\mu$m long DLG-Si device with $d_c = 0.8 \mu$m and $s = 0.35$ $\mu$m, $d_{AlOx}=5$nm is compatible with 16 GHz modulation (shown in Fig. 8(c)) and exhibits reduced IL (0.25dB) but also a reduced ER (5.2dB).
length for different relative permittivity, spacing and graphene sheet resistance all for $d_{c} = 0.8 \mu m$ and oxide thickness of 5 nm. Fig. 8(c) also shows the effect of the sheet resistance $R_{\text{she}}$ and the relative permittivity of the oxide layer. Obviously a lower sheet resistance and reduced spacing s are beneficial for the performance of the device. Increasing the relative permittivity of the oxide layer on the other hand is equivalent to decreasing its speed.

V. Conclusion

In summary, we investigated the performance of four representative configurations, relying either on Silicon or on Silicon Nitride waveguides, for realizing graphene based switches and modulators operating in the wavelength regions around 1310 nm and 1550 nm. We first studied their properties at static operating conditions, which might actually also be relevant for other types of devices such as voltage controlled non-linear devices and detectors. Interestingly we found that the figure-of-merit of these devices, defined as the ratio of the extinction ratio over the insertion loss, is independent of the precise waveguide configuration and only depends on the quality of the used graphene and the gate-oxide thickness. Hence the waveguide configuration can be selected based on other criteria. For ultimate speed, the high confinement offered by higher index contrast silicon waveguides offers more compact and hence faster devices. For switches or non-linear devices with voltage controlled doping, which have less stringent requirements on operating speed, deposited silicon nitride waveguides can provide more versatility and do not suffer from two photon absorption. In all applications accurately controlling the gate-oxide thickness is extremely important as it directly impacts modulation efficiency, power consumption and operation bandwidth.

References


Leili A. Shiramin received the B.Sc. and M.Sc. degree in physics from University of Tabriz and K.N.Tooosi University of technology, Tehran, Iran, in 2008 and 2011, respectively. From 2011 to 2013 she joined to the Photonics Group at University of Tabriz as a researcher. Since 2013, she works as a PhD researcher at Photonics Research Group, Ghent University-Imec, Ghent, Belgium. Her current interest includes 2D material based optical devices fabrication and characterization. She is a student member of IEEE Photonics Society.

Dries Van Thorhout received the Master degree in physical engineering and the Ph.D. degree from Ghent University, Ghent, Belgium in 1995 and 2000 respectively. From Oct. 2000 to Sep. 2002 he was with Lucent Technologies, Bell Laboratories, New Jersey, USA, working on the design,
processing and characterization of InP/InGaAsP monolithically integrated devices. In Oct. 2002 he joined the Department of Information Technology (INTEC), Ghent University, Belgium, where he is now a full professor. His research focuses on the design, fabrication and characterization of integrated photonic devices. Main topics involve Silicon nanophotonic devices and the integration of novel materials (graphene, III-V, ferro-electrics, quantum dots, ...) on the waveguides to expand their functionality. He is working on applications for telecom, diatom, optical interconnect and sensing.