Engineering Plasmons in Graphene Nanostructures in THz Frequencies: Compact Modeling and Performance Analysis for On-chip Interconnects

Shaloo Rakheja, *Member* IEEE

Department of Electrical and Computer Engineering, New York University, NY, USA 11201

Abstract— In this paper, transverse magnetic (TM) propagation modes of surface plasmon polaritons (SPPs) in graphene micro/nano ribbons are exhaustively characterized by accounting for the finite lateral dimensions of graphene, screening of Fermi level in multilayer graphene stack, and the impact of dielectric permittivity and the associated charge impurities at the dielectric-graphene interface. Fermi level screening leads to a non-uniform carrier density across multiple layers, which changes the electron relaxation rate and considerably alters the complex dynamical conductivity of multilayer GNRs. It is shown that ignoring the screening effects in multilayer GNRs overestimates both the SPP propagation length and its propagation velocity. Graphene plasmonic interconnects are envisaged as low energy, high frequency on-chip interconnects for future technology nodes. Simulations are performed over a broad frequency spectrum to identify the merits of future graphene plasmonic interconnects over the conventional electrical Cu/low-κ at a minimum feature size of 10 nm. Using energy-per-bit as a figure-of-merit, a range of SPP propagation lengths is identified for graphene plasmonic interconnects to outperform Cu interconnects.

Keywords— **Graphene nanoribbons, surface plasmon polaritons, plasmonic interconnects,**

I. INTRODUCTION

II. PLASMON DISPERSION RELATIONSHIP IN GRAPHENE

In this work, we consider a graphene sheet embedded in a dielectric environment as shown in Fig. 1. To obtain the dispersion relation for transverse magnetic (TM) surface plasmon polaritons (SPPs) supported by graphene, we solve Maxwell's equations with appropriate boundary conditions. For TM SPPs propagating along the x- direction as indicated in Fig. 1, electric field has components in both x- and z- directions, while the magnetic field lies in the y- direction. Due to its finite non-zero conductivity, graphene imposes an impedance boundary condition in Maxwell's equations. For TM plasmons, the dispersion relation is given as [1]

$$\frac{\varepsilon_1}{\sqrt{\beta^2 - \varepsilon_1 \left(\frac{\omega}{c}\right)^2}} + \frac{\varepsilon_2}{\sqrt{\beta^2 - \varepsilon_2 \left(\frac{\omega}{c}\right)^2}} = -i\frac{\sigma(\omega)}{\omega\varepsilon_0} \quad (1)$$

where β denotes the plasmon propagation constant, ε_1 and ε_2 are the dielectric constants of the surrounding media, ω is the operation frequency, c is the speed of light, ε_0 is the

permittivity of free space, $\sigma(\omega)$ is the complex AC electrical conductivity of graphene, where non-local effects are not considered. Note that (1) does not make any approximations of the non-retarded regime of operation for SPPs as done in previous works 13]. The plasmon dispersion relation can be used to obtain the propagation velocity as $v_g = \partial \omega / \partial[\text{Re}(\beta)]$ and propagation length $L_{prop} = 1/\text{Im}(\beta)$ of plasmons. The $\text{Im}(\beta)$ is related with Ohmic losses in graphene and mathematically described using a phenomenological parameter Γ in the conductivity description as given in the next sub-section (See Eqs. (2)-(5).)
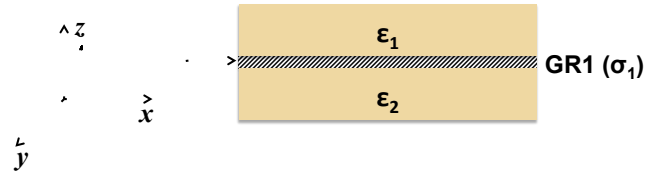

Figure 1: Graphene sheet embedded in dielectric with dielectric constants ε_1 and ε_2. The electrical conductivity of graphene is denoted with σ_1. The SPPs in graphene propagate in x-direction.

A. Complex dynamical conductivity in graphene and electron scattering rate

The net complex dynamical conductivity of graphene is given by the sum of its intra-band and inter-band components. In the presence of inter-layer screening, the conductivity of multilayer graphene is given as [2]

$$\sigma_{intra}(\omega) = \sum_{j=1}^{N_{layer}} \frac{e^2}{\pi\hbar^2} \frac{k_B T}{\Gamma_j - i\omega} \left[\frac{E_{f,j}}{k_B T} + 2\ln\left(1 + \exp\left(-\frac{E_{f,j}}{k_B T}\right)\right) \right]$$

$$\sigma_{inter}(\omega) = \sum_{j=1}^{N_{layer}} \frac{e^2}{4\hbar} \left[\frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{\hbar\omega - 2E_{f,j}}{k_B T}\right) - \frac{i}{2\pi}\ln\left(\frac{\left(\hbar\omega + 2E_{f,j}\right)^2}{\left(\hbar\omega - 2E_{f,j}\right)^2 + k_B T}\right) \right] \quad (2)$$

$$E_{f,j} = E_{f0} \exp\left(-\frac{0.35nm(2j-1)}{\lambda_s}\right)$$

where e is the elementary charge, $E_{f,j}$ is the Fermi level in the jth layer, $k_B T$ is the thermal energy and Γ_j is a phenomenological parameter that characterizes the electron disorder scattering processes in graphene in the jth layer, and N_{layer} is the total number of layers in the multilayer graphene. An estimate of Γ is obtained from the D.C. relaxation time,

which arises mainly from scatterings due to intrinsic phonons and charges impurities of the substrate. For narrow ribbons, electron scatterings due to edge roughness are also significant and must be accounted for in the description of Γ.

Phonon scattering in graphene fundamentally limits the mobility of carriers. In its simplest form, the scattering rate due to acoustic phonons, Γ_{ac}, is given as [3]

$$\Gamma_{ac} = \frac{1}{\tau_{ac}} = \frac{D_{ac}}{4\hbar^3} \frac{k_B T}{\rho_m v_f^2 v_{ph}^2} |E_f| \tag{3}$$

where D_{ac} is the acoustic deformation potential (≈ 9 eV), $v_f = 8 \times 10^5$ m/s is the Fermi velocity of Dirac fermions in graphene, and v_{ph} is the speed of acoustic phonons in graphene.

The scattering rate due to charged impurities at the interface between graphene and the substrate is given as [3]

$$\Gamma_{imp} = \frac{1}{\tau_{imp}} = \frac{\hbar(4\varepsilon_0 \varepsilon(1+\gamma))^2}{N_{imp} Z^2 e^4} |E_f|$$

$$\gamma = \frac{\rho(E_f)}{2\varepsilon_0 \varepsilon k_F} \tag{4}$$

where N_{imp} is the impurity concentration, Ze is the net charge of the impurity atom, $\rho(E_f)$ is the 2D density-of-states of graphene at Fermi level, and k_F is the Fermi wave-vector. The 2D density-of-states in graphene is given as $\rho(E_f) = \frac{2|E_f|}{\pi(\hbar v_f)^2}$.

Finally, scatterings due to edge roughness in narrow ribbons are modeled using an edge scattering coefficient, P_{GNR} that lies between zero and unity. The electron relaxation rate due to edge roughness scatterings is given as [4]

$$\Gamma_{edge} = \frac{1}{\tau_{edge}} = \frac{P_{GNR}}{W} \frac{1}{\sqrt{\left(\frac{E}{E_{sub,m}}\right)^2 - 1}}$$

$$E_{sub,m} = \frac{\hbar v_f}{2W} \left| m + \frac{1}{3} \right| \tag{5}$$

Where $E_{sub,m}$ is the energy of the bottom of the mth sub-band in the graphene ribbon. The net electron relaxation rate is given by using the Mattheissen's sum rule. That is, $\Gamma_{net} = \Gamma_{ac} + \Gamma_{imp} + \Gamma_{edge}$.

The Fermi level in the jth layer in (2) depends on the screening length, λ_s. For a large screening length ($\lambda_s \to \infty$), all layers are perfectly coupled such that $\sigma_{multi} = N_{layer}\sigma_{mono}$, where σ_{multi} corresponds to the conductivity of multilayer graphene and σ_{mono} denotes the conductivity of monolayer graphene. However, a small screening length leads to degradation in carrier concentration in upper layers in the multilayer stack for which the Fermi level drops significantly. Note that 0.35 nm in the formulation of $E_{f,j}$ denotes the inter-layer screening between the multiple layers in graphene.

The relationship between the Fermi level and the carrier concentration in graphene is obtained using the Fermi Dirac integrals given as

$$n_s = \sum_{j=1}^{j=N_{layer}} \left| n_{elec,j} - n_{hole,j} \right|$$

$$n_{elec,j} = \int_0^\infty DOS(E) \frac{1}{1+\exp\left(\frac{E - E_{f,j}}{k_B T}\right)} dE;$$

$$n_{elec,j} = \int_0^\infty DOS(E) \frac{1}{1+\exp\left(\frac{E + E_{f,j}}{k_B T}\right)} dE \tag{6}$$

$$DOS(E) = \frac{2}{(\pi \hbar v_f)^2} \left[\frac{2\sigma_{dis}}{\sqrt{2\pi}} \exp\left(-\frac{E^2}{2\sigma_{dis}^2}\right) + E \times erf\left(\frac{E}{\sigma_{dis}\sqrt{2}}\right) \right]$$

Here, σ_{dis} is the broadening in eV of the DOS around the Dirac point [5].

Figure 2 shows the both the real and the imaginary components of σ_{intra} and σ_{inter} as functions of frequency, ω, and the Fermi level, E_{f0}. Inter-layer screening is not considered in this plot. It is evident from this figure that inter-band conductivity plays a role only when $\omega > 2\omega_f$, where ω_f is the frequency corresponding to the Fermi level in graphene. Hence, in the frequency region of interest (the THz band) in this work, we ignore the contributions from inter-band scatterings in graphene.
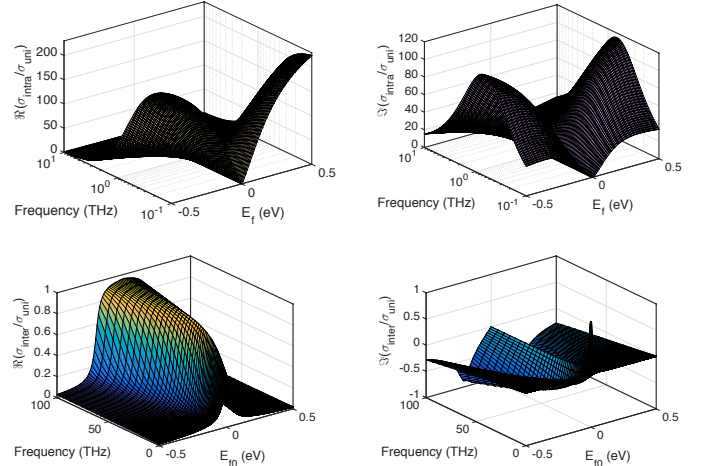

Figure 2: Real (left panel) and imaginary (right panel) components of intra-band (top row) and inter-band (bottom row) conductivity in graphene.

In Fig. 3, the impact of E_f-screening on $\sigma_{intra}(\omega)$ is analyzed. For screening length, $\lambda_s = 0.6$ nm, only up to two layers contribute significantly to the overall conductivity; however, for relatively small screening ($\lambda_s = 6$ nm), up to five layers contribute to the total conductivity and beyond that the conductivity saturates as a function of N_{layer}. For multilayer graphene to be profitable compared to monolayer graphene, it is important to have a large screening length, which can be obtained by chemically doping all the layers in a multilayer stack.
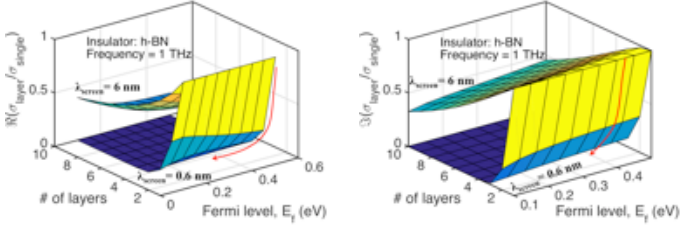
Figure 3: Real part (left) and imaginary part (right) of conductivity contribution from the different layers in multilayer GNR for different screening lengths.

The impact of dielectric permittivity and the interface charged impurities on the SPP resonant frequency for TM modes is shown in Fig. 4. It is emphasized that using an approximate NR regime expression (shown in circles) of TM modes leads to inaccurate resonant frequencies below a few THz. Hence, in this work we use the complete description of the dispersion relationship of TM modes as given in (1).
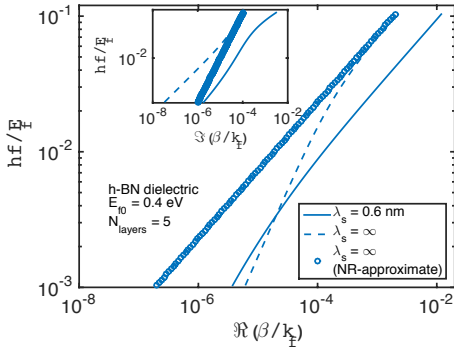
Figure 4: Resonant frequency of SPP plasmons as a function of Re(β). NR-approximation becomes highly inaccurate for lower frequencies and wave-vectors. The inset shows the relation between resonant frequency and Im(β).

Propagation length is an important figure of merit for plasmonic interconnects. Propagation length is defined as the length at which the initial energy of the surface plasmons degrades to 37% of their initial energy. A large plasmon propagation length is desirable as it signifies a lower interconnect energy dissipation. In Fig. 5, we plot the propagation length, L_{prop}, of SPPs in multilayer graphene for different values of the screening length λ_s. It is seen that L_{prop} saturates with N_{layer} for small λ_s and is 60× lower than the propagation length for a large λ_s. One way to overcome the degradation in carrier concentration n_s due to finite λ_s is to explore chemical doping of all the layers in graphene multilayer stack. Further, scatterings due to edge roughness can severely degrade L_{prop} for a given λ_s as also demonstrated in Fig. 5. The degradation in L_{prop} has a significant impact on the performance and energy-per-bit of on-chip plasmonic interconnects as discussed in Section III.

III. PERFORMANCE ANALYSIS OF GRAPHENE PLASMONIC ON-CHIP INTERCONNECTS

To assess the performance of on-chip local interconnects, the metrics we use are interconnect delay (τ_D), signal drive
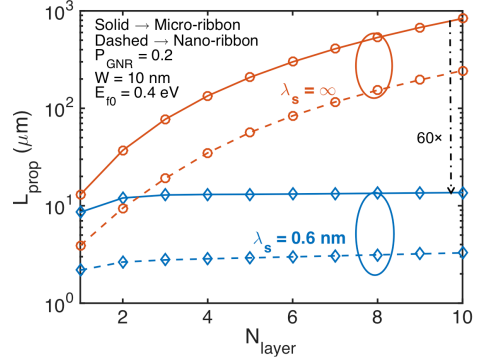

Figure 5: Fig. 5: Propagation length of TM SPPs in graphene @ 1 THz in h-BN dielectric as a function of the number of layers. P_{GNR} denotes the edge-roughness coefficient. P_{GNR} = 0 (specular scattering). P_{GNR} =1 (diffusive scatterings).

distance (SDD), and energy-per-bit (E_{bit}) for local communication. SDD is a more powerful, scalable metric to understand the increasing wire delays and is defined as the interconnect length for which the delay through the interconnect segment matches the intrinsic gate delay [6].

The propagation delay for Cu/low-k interconnects is modeled using the circuit shown in Fig. 6, while for plasmonic interconnects, the interconnect delay is given as $\tau_{SPP} = L/v_g$, where v_g is the group velocity of TM surface plasmons. For plasmonic interconnects, we only account for the delay associated with the interconnect, while switches are assumed to be ideal, allowing us to obtain an upper bound on their performance.
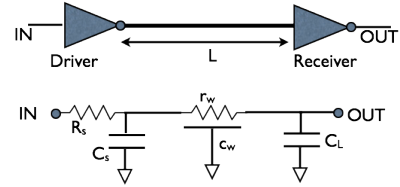

Figure 6: The schematic of the CMOS system with a CMOS driver, an interconnect, and a CMOS load (top). The equivalent circuit representation of the CMOS system (bottom).

Figure 7 shows a comparison of the propagation delay for both Cu/low-k and SPP interconnects at a width of 10 nm. An increase in the frequency of operation leads to a degradation in the plasmon propagation velocity. Hence, the delay of plasmonic interconnects increases with an increase in the operating frequency. However, for a signal at 1 THz, the best-case performance of plasmonic interconnects is more than two orders of magnitude superior than Cu interconnects.

In Fig. 8, the SDD for plasmonic interconnects is obtained as a function of the SPP propagation velocity and the ratio of delays of CMOS and plasmonic switch. As the interconnect delay increases, SDD decreases necessitating the insertion of buffers. Conversely, a larger SDD means that there is less disparity in the delays of devices and interconnects.
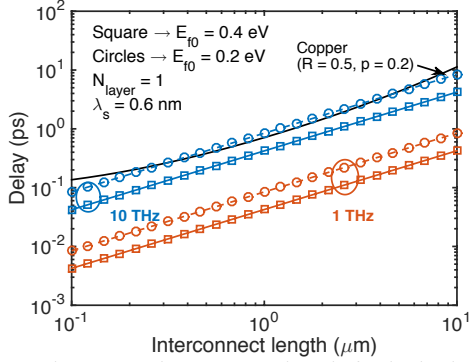
Figure 7: Delay versus interconnect length for both electrical and plasmonic interconnects. Due to a reduction in the propagation velocity of SPPs in graphene at higher frequency and lower E_f, the delay of plasmonic interconnects increases for higher frequency and/or lower E_f.
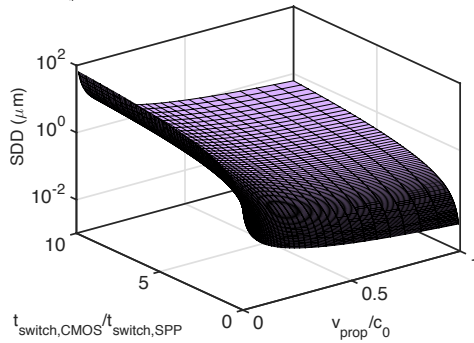

Figure 8: SDD versus propagation velocity of SPPs and the ratio of the delay of CMOS and plasmonic switch.

The energy-per-bit of both interconnects is given as

$$E_{Cu} = \frac{1}{2}(C_S + C_L + c_w L_{int})V_{dd}^2$$

$$E_{spp} = \frac{1}{2}\hbar\omega <m> \exp\left(\frac{L_{int}}{L_{prop}}\right)$$

(7)

where $<m>$ denotes the average number of plasmons needed to represent a logic bit. Shot noise limited detection of an ideal coherent noise source with a mean number of $<m>$ plasmons for state "1" and "0" results in a bit error rate of $0.5exp(-<m>)$. If the bit error rate is equal to 10^{-30}, then mean number of plasmons $<m>$ is equal to 68 [7].

As seen from (7), the energy-per-bit of plasmonic interconnects increases exponentially with interconnect length, while that of electrical interconnects increases linearly with interconnect length. For a fixed SPP propagation length, the ratio of E_{bit} of electrical and plasmonic interconnects degrades with interconnect length, L_{int} as shown in **Fig. 9**. Plotting constant contours of $E_{bit}(electrical)/E_{bit}(plasmonic)$ in **Fig. 10** shows that for a given frequency of operation of plasmonic interconnects, there necessarily exists a maximum ratio of L_{int}/L_{prop}. The shaded region between different contours indicates the energy budget that can be allocated toward plasmonic switches and other energy transducers in a heterogeneous electrical-photonic logic framework.
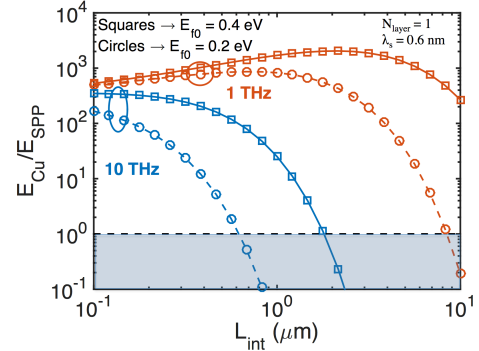

Figure 9: Ratio of energy-per-bit of Cu and plasmonic interconnects as a function of interconnect length for an interconnect width of 10 nm. In the shaded region, Cu interconnects outperform plasmonic interconnects.
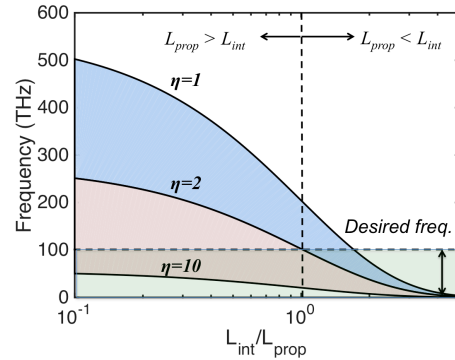

Figure 10: Constant $\eta = E_{bit}(electrical)/E_{bit}(plasmonic)$ contours. The shaded regions between contours denote the energy budget available to be allocated toward computation in plasmonic domain and transducers for energy conversion. The desired frequency band of $f < 100$ THz to avoid inter-band losses.

IV. CONCLUSIONS

In this paper, we have provided comprehensive modeling of TM SPP plasmon modes in graphene nanostructures and evaluated their performance for future high-speed, low-power on-chip interconnects. The physics-based modeling approach considered in this work accurately accounts for Fermi-level screening in multilayer graphene stack, edge roughness in narrow ribbons, and impact of dielectric permittivity on plasmon response.

REFERENCES

[1] M. Jablan et al., *Physical Review B* 80, 245435, 2009.
[2] L. A. Falkovsky, *Journal of Physics: Conference Series*, vol. 129, no. 1, p. 012004, 2008.
[3] T. Stauber et al., *Physical Review B*, 77, no. 11, p. 085418, 2008.
[4] S. Rakheja et al., *Proceedings of the IEEE*, vol. 101, no. 7, 2013.
[5] S. Rakheja et al., *IEEE Transactions on Nanotechnology*, vol. 13, no. 5, 2014.
[6] D. Matzke, *Computer*, vol. 30, no. 9, pp. 37-39, 2007.
[7] S. Rakheja and V. Kumar, *International Symposium on Quality Electronic Design (ISQED)*, March 18-20, 2012.