2012-10-12

Theory of Interfacial Plasmon-Phonon Scattering in Supported Graphene

Zhun-Yong Ong and Massimo V. Fischetti

Theory of interfacial plasmon-phonon scattering in supported graphene

Zhun-Yong Ong* and Massimo V. Fischetti†
Department of Materials Science and Engineering, University of Texas at Dallas RL10, 800 W Campbell Rd RL10, Richardson, Texas 75080, USA
(Received 24 July 2012; published 12 October 2012)

One of the factors limiting electron mobility in supported graphene is remote phonon scattering. We formulate the theory of the coupling between graphene plasmon and substrate surface polar phonon (SPP) modes and find that it leads to the formation of interfacial plasmon-phonon (IPP) modes, from which the phenomena of dynamic antiscreening and screening of remote phonons emerge. The remote phonon-limited mobilities for SiO_2, HfO_2, h-BN, and Al_2O_3 substrates are computed using our theory. We find that hexagonal boron nitride (h-BN) yields the highest peak mobility, but in the practically useful high-density range, the mobility in HfO_2-supported graphene is high, despite the fact that HfO_2 is a high-κ dielectric with low-frequency modes. Our theory predicts that the strong temperature dependence of the total mobility effectively vanishes at very high carrier concentrations. The effects of polycrystallinity on IPP scattering are also discussed.

DOI: 10.1103/PhysRevB.86.165422 PACS number(s): 79.60.Jv, 73.50.Dn, 74.78.Fk, 72.80.Vp

I. INTRODUCTION

Graphene, a single-layer of hexagonally arranged carbon atoms,[1] has been long considered a promising candidate material for post-Si CMOS technology and other nanoelectronic applications on account of its excellent electrical[2] and thermal transport[3] properties. In suspended single-layer graphene (SLG), the electron mobility has been demonstrated to be as high as 200 000 $cm^2V^{-1}s^{-1}$.[4] However, in real applications such as a graphene field-effect transistor (GFET), the graphene is physically supported by an insulating dielectric substrate such as SiO_2, and the carrier mobility in such supported-graphene structures is about one order of magnitude lower.[4] This reduction in carrier mobility is further exacerbated in top-gated structures[5] in which a thin layer of a high-κ dielectric, such as HfO_2 or Al_2O_3, is deposited or grown on the graphene sheet.[6–8] The degradation of the electrical transport properties is a result of exposure to environmental perturbations such as scattering by charge traps, surface roughness, and remote optical phonons, which are a kind of surface excitation. Such environmental effects are encountered in metal-oxide-semiconductor (MOS) structures.[9] Hess and Vogl first suggested that remote phonons [sometimes also known as Fuchs-Kliewer (FK)[10] surface optical (SO) phonons] can have a substantial effect on the mobility of Si inversion layer carriers.[11] Fischetti and coworkers later studied the effects of remote phonon scattering in MOS structures and found that high-κ oxide layers have a significant effect on carrier mobility in Si[9] and Ge.[12] This method was later applied by Xiu to study remote phonon scattering in Si nanowires.[13] In Refs. 9 and 13, it was found that the plasmons in the channel material (Si) hybridized with the surface polar phonons (SPP) in the nearby dielectric material to form interfacial plasmon-phonon (IPP) modes. This hybridization occurrence naturally leads to the screening and antiscreening of the SPP from the dielectric material. Scattering with these IPP modes results in a further reduced channel electron mobility in 2D Si and Si nanowires.

Likewise, remote phonon scattering is one of the mechanisms believed to reduce the mobility of supported graphene, with the form of the scattering mechanism varying with the material properties of the dielectric substrate. Experimentally, hexagonal boron nitride (h-BN) has been found to be a promising dielectric material for graphene, and it is commonly believed that this is at least partially due to the fact that remote phonon scattering is weak with a h-BN substrate.[14] On other substrates such as SiO_2[15,16] and SiC,[17,18] the mobility of supported graphene is lower. Thus it is important to develop an accurate understanding of remote phonon scattering in order to find an optimal choice of substrate that will minimize the degradation of carrier mobility in supported graphene.

Although the subject of remote phonon scattering in graphene[19–23] and carbon nanotubes[20,21] has been broached in the recent past, the basic approach used in the aforementioned works does not deal adequately with the *dynamic* screening of the SPP modes. In graphene, dynamic screening of SPP modes has its origin in SPP-plasmon coupling, and the two time-dependent phenomena have to be treated within the same framework. Typically, the coupling phenomenon is ignored, and screening of the SPP modes is approximated with a Thomas-Fermi (TF) type of static screening,[19,22,24] which is adequate for the case of impurity scattering[25,26] but can lead to a miscalculation of the scattering rates since the use of static screening underestimates the electron-phonon coupling strength,[19] especially for higher-frequency modes. The failure to incorporate correctly SPP-plasmon coupling into the approach means that the dispersion relation of the SPP (or, more accurately, of the IPP) modes is incorrect and that the dynamic screening of the remote phonons is not accounted for in a natural manner.

To understand the screening phenomenon in our situation, let us first give a bird's eye view of the physical picture. This picture is somewhat different from what is found in the more familiar semiconductor-inversion-layer, high-κ-dielectric geometry, since the absence of a gap in bulk graphene renders its dielectric response stronger and qualitatively different—almost metal-like, as testified by the presence of Kohn anomalies in the phonon spectra[27,28]—than the response of a two-dimensional electron gas. Graphene plasmons interact with the SPP modes through the time-dependent electric field generated by the latter, and the former are forced to oscillate at

the frequency of the latter (ω). When ω is less than the natural frequency of the plasmon (ω_p), i.e., $\omega < \omega_p$, the electrons can respond to the SPP mode and screen its electric field. On the other hand, when $\omega > \omega_p$, the motion of the plasmons lags that of the SPP mode, resulting in poor or no screening, or even in antiscreening, which can actually augment the scattering field.[29] In bulk SiO_2, the main TO-phonon frequencies are around 56 and 138 meV. At long wavelengths ($\lambda > 10^{-8}$ m), the plasmon frequencies for a carrier density of 10^{12} cm^{-2} are comparable or smaller than the TO-phonon frequencies. Thus a TF-type approximation is inadequate especially for describing the screening (or more accurately, the *antiscreening*) of the 138 meV TO phonon modes. Our calculations suggest that, contrary to what is found in the semiconductor/high-κ case[9] and to the claims made in Ref. 19, the higher-frequency SPP modes cannot be ignored despite their reduced Bose-Einstein occupation factors at room temperature.

It is our intention in this paper to provide a systematic description of the coupling between the substrate SPP and the graphene plasmon modes, and relate this coupling to the dynamic screening phenomenon. Our theory can be generalized to graphene heterostructure such as double-gated graphene although this falls outside the scope of our paper and will be the subject of a future work. We begin by deriving our model of the IPP system. Its dispersion is then calculated from the model. The pure SO phonon and graphene plasmon branches are compared with the IPP branches. Also, we compute the electron-IPP and the electron-SPP coupling coefficients for different substrates (SiO_2, h-BN, HfO_2, and Al_2O_3). We show that the IPP modes can be interpreted as dynamically screened SPP modes. Scattering rates are then calculated and used to compute the remote phonon-limited mobility μ_{RP} for different substrates at room temperature (300 K) with varying carrier density. The temperature dependence of μ_{RP} at low and high carrier densities is compared. Using the μ_{RP} results, we analyze the suitability of the various dielectric materials for use as substrates or gate insulators in nanoelectronics applications. We also discuss the effects of polycrystallinity on remote phonon scattering.

II. MODEL

A. Coupling between substrate polar phonons and graphene plasmons

Our approach to constructing the theoretical model of the coupled plasmon-phonon systems follows closely that of Fischetti, Neumayer, and Cartier[9] although some modifications are needed to describe the plasmon-phonon coupling. One of the primary difficulties in describing the coupled system is the anisotropy in the dielectric response of graphene: graphene is polarizable in the plane but its out-of-plane response is presumably negligible. If the graphene sheet is modeled as a slab of finite thickness with a dynamic dielectric response in the in-plane direction [$\epsilon_{gr}^{\parallel}(\omega) = \epsilon_{gr}(1 - \omega_p^2/\omega^2)$, where ω_p is the plasma frequency] and none in the out-of-plane direction [$\epsilon_{gr}^{\perp}(\omega) =$ constant], the dispersion of the SPPs remains unchanged, indicating that the SPP and plasmon modes are uncoupled. This absence of coupling is implicitly assumed in much of the current literature on SPP scattering in graphene[19,22–24,30] although it has already been shown to be untrue in 2D Si[9] and Si nanowires.[13] Furthermore, there is considerable experimental support for the coupling of graphene plasmons to the SPPs.[31–34] As we will show later, accounting for this coupling results in the formation of IPP modes, which are screened (antiscreened) and scatter charge carriers in graphene more weakly (strongly) than the unhybridized SPP modes. This "uncoupling problem" persists even when one inserts a vacuum region between the graphene slab and the substrate. Ultimately, this alleged lack of coupling can be traced back to the continuity of the electric displacement field **D** at the interface between the graphene slab and the substrate or vacuum. Given that the dynamic response of the graphene is only in the in-plane directions and that the coupling should be with the *p*-polarized waves of the substrate, the slab approach is not likely to be correct. To overcome this difficulty, we find it is necessary to treat the graphene as a polarizable charge sheet (as shown in Fig. 1) rather than as a finite slab with a particular in-plane dielectric function. This polarization charge then generates a discontinuity in the electric displacement along the surface of the graphene. It is this discontinuity that couples the dielectric response of the substrate to that of the graphene sheet. The basic setup is shown in Fig. 1. The graphene is an infinitely thin sheet coplanar with the (x,y) plane and floating at a height d above the substrate that occupies the semi-infinite region $z < 0$. Notation wise, we try to follow Ref. 9. In the direction perpendicular to the interface, the (ionic) dielectric response of the substrate is assumed to be due to two optical phonon modes, an approximation used in Ref. 9, that is

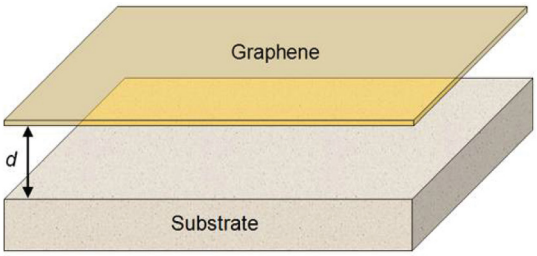

FIG. 1. (Color online) Schematic of set up of graphene-substrate system. The SLG is modeled as an infinitely thin (in the z direction) layer of polarizable charge. A gap of d separates the graphene charge sheet and the substrate surface.

$$\epsilon_{ox}(\omega) = \epsilon_{ox}^{\infty} + \left(\epsilon_{ox}^{i} - \epsilon_{ox}^{\infty}\right)\frac{\omega_{\mathrm{TO2}}^2}{\omega_{\mathrm{TO2}}^2 - \omega^2} + \left(\epsilon_{ox}^{0} - \epsilon_{ox}^{i}\right)\frac{\omega_{\mathrm{TO1}}^2}{\omega_{\mathrm{TO1}}^2 - \omega^2}, \quad (1)$$

where ω_{TO1} and ω_{TO2} are the first and second transverse optical (TO) angular frequencies (with $\omega_{\mathrm{TO1}} < \omega_{\mathrm{TO2}}$), and ϵ_{ox}^{∞}, ϵ_{ox}^{i}, and ϵ_{ox}^{0} are the optical, intermediate, and static permittivities. We can also express $\epsilon_{ox}(\omega)$ in the generalized Lyddane-Sachs-Teller form:

$$\epsilon_{ox}(\omega) = \epsilon_{ox}^{\infty}\frac{\left(\omega_{\mathrm{LO2}}^2 - \omega^2\right)\left(\omega_{\mathrm{LO1}}^2 - \omega^2\right)}{\left(\omega_{\mathrm{TO2}}^2 - \omega^2\right)\left(\omega_{\mathrm{TO1}}^2 - \omega^2\right)},$$

where ω_{LO1} and ω_{LO2} are the first and second longitudinal optical angular frequencies. The variables **Q** and **R** represent

the two-dimensional wave and coordinate vector in the (x,y) plane of the interface, respectively.

As in Ref. 9, we try to derive the longitudinal electric eigenmodes of the system since the transverse modes (given by poles of the total electric response) correspond to a vanishing electric field and so to a vanishing coupling with the graphene carriers. In effect, the longitudinal modes are the transverse-magnetic (TM) solutions of Maxwell's equations. It was also shown in Ref. 9 that one may ignore the effects of retardation. Therefore we need only to employ simpler electrostatics instead of the full set of Maxwell's equations.

We begin our derivation by writing down the Poisson equation for the *bare* scalar potential Φ,

$$-\nabla^2\Phi(\mathbf{R},z) = \frac{1}{\epsilon_0}\rho_{ox}(\mathbf{R},z,t), \tag{2}$$

where ρ_{ox} is the (periodic) polarization charge distribution at the surface of the substrate that is the source of scattering, and ϵ_0 is the permittivity of vacuum. Equation (2) describes the electrostatic potential within the graphene. However, the *effective* scalar potential felt by the graphene carriers is different and should include the collective screening effect of the induced electrons/holes, which changes the right-hand side (RHS) of Eq. (2). Hence we modify Eq. (2) by adding a screening charge term on its RHS, and we obtain the Poisson equation for the *screened* scalar potential $\Phi_{\rm scr}$,

$$-\nabla^2\Phi_{\rm scr}(\mathbf{R},z,t) = \frac{1}{\epsilon_0}\left[\rho_{ox}(\mathbf{R},z,t) + \rho_{\rm scr}(\mathbf{R},z,t)\right], \tag{3}$$

where $\rho_{\rm scr}$ is the screening charge term. The integral form of Eq. (3) is

$$\Phi_{\rm scr}(\mathbf{R},z,t) = \Phi(\mathbf{R},z,t) + \int d\mathbf{R}'dz'G(\mathbf{R}z,\mathbf{R}'z')\rho_{\rm scr}(\mathbf{R}'z',t), \tag{4}$$

where $G(\mathbf{R}z,\mathbf{R}'z')$ is the Green function that satisfies the boundary conditions [see Eq. (14)], and the equation

$$-\nabla^2[\epsilon(\mathbf{R},z)G(\mathbf{R}z,\mathbf{R}'z')] = \delta(\mathbf{R}-\mathbf{R}',z-z')\,. \tag{5}$$

The bare potential $\Phi(\mathbf{R},z,t)$ is defined as $\Phi(\mathbf{R},z,t) = \int d\mathbf{R}'dz'G(\mathbf{R}z,\mathbf{R}'z')\rho_{ox}(\mathbf{R}'z',t)$. The second term on the RHS of Eq. (4) represents the screening charge distribution. The bare and screened potentials can be written as sums of their Fourier components:

$$\Phi(\mathbf{R},z,t) = \sum_{\mathbf{Q}}\phi_{Q,\omega}(z)e^{-i\mathbf{Q}\cdot\mathbf{R}}e^{i\omega t}\,, \tag{6a}$$

$$\Phi_{\rm scr}(\mathbf{R},z,t) = \sum_{\mathbf{Q}}\phi^{\rm scr}_{Q,\omega}(z)e^{-i\mathbf{Q}\cdot\mathbf{R}}e^{i\omega t}, \tag{6b}$$

where it must be understood that only the real part of Eq. (6) is to be taken here and in the following sections. Given the cylindrical symmetry of the problem, the Fourier components $\phi_{Q,\omega}$ and $\phi^{\rm scr}_{Q,\omega}$ depend only on the magnitude of the wave vector $\mathbf{Q}$.

From Eq. (4), we obtain the following expression for the z-dependent part of the Fourier-transformed screened potential:

$$\phi^{\rm scr}_{Q,\omega}(z)e^{i\omega t} = \phi_{Q,\omega}(z)e^{i\omega t} + \int dz'G_Q(z,z')\rho^{\rm scr}_{Q,\omega}(z',t). \tag{7}$$

Equation (7) is solvable if the polarization charge $\rho^{\rm scr}_{Q,\omega}$ is expressed as a function of the screened scalar potential. Here, we assume that $\rho^{\rm scr}_{Q,\omega}$ responds linearly to $\phi^{\rm scr}_{Q,\omega}$ and write the screening charge term as

$$\rho^{\rm scr}_{Q,\omega}(z,t) = e^2\Pi(Q,\omega)f(z)\phi^{\rm scr}_{Q,\omega}(z)e^{i\omega t}, \tag{8}$$

where $\Pi(Q,\omega)$ is the in-plane 2D polarization charge term and $f(z)$ governs the polarization charge distribution in the out-of-plane direction. For convenience, we model the graphene as an infinitely thin sheet of polarized charge and set $f(z) = \delta(z-d)$. Combining Eqs. (7) and (8), we obtain the expression

$$\phi^{\rm scr}_{Q,\omega}(z) = \phi_{Q,\omega}(z) + e^2\int dz'G_Q(z,z')\Pi(Q,\omega)f(z')\phi^{\rm scr}_{Q,\omega}(z'). \tag{9}$$

The expression in Eq. (9) becomes

$$\begin{aligned}\phi^{\rm scr}_{Q,\omega}(z) &= \phi_{Q,\omega}(z) + e^2G_Q(z,d)\Pi(Q,\omega)\phi^{\rm scr}_{Q,\omega}(d)\\ &= \phi_{Q,\omega}(z) + e^2G_Q(z,d)\Pi(Q,\omega)\phi_{Q,\omega}(d)\\ &\quad + e^4G_Q(z,d)\Pi(Q,\omega)G_Q(d,d)\Pi(Q,\omega)\phi^{\rm scr}_{Q,\omega}(d)\\ &= \cdots\\ &= \phi_{Q,\omega}(z) + \frac{e^2G_Q(z,d)\Pi(Q,\omega)}{1-e^2G_Q(d,d)\Pi(Q,\omega)}\phi_{Q,\omega}(d),\end{aligned}$$

and the corresponding component of the electric field perpendicular to the interface at $z=0$ is

$$\hat{\mathbf{z}}\cdot\mathbf{E}_{Q,\omega}|_{z=0} = -\frac{\partial}{\partial z}\phi^{\rm scr}_{Q,\omega}(z)\bigg|_{z=0} = -\frac{\partial\phi_{Q,\omega}(z,t)}{\partial z} - \frac{\partial G_Q(z,d)}{\partial z} \times \frac{e^2\Pi(Q,\omega)}{1-e^2G_Q(d,d)\Pi(Q,\omega)}\phi_{Q,\omega}(d)\bigg|_{z=0}. \tag{10}$$

For notational simplicity, we write

$$\phi^{\rm scr}_{Q,\omega}(z) = \phi_{Q,\omega}(z) + G_Q(z,d)\mathcal{P}_{Q,\omega}\phi_{Q,\omega}(d), \tag{11a}$$

$$\hat{\mathbf{z}}\cdot\mathbf{E}_{Q,\omega} = -\frac{\partial\phi_{Q,\omega}(z,t)}{\partial z} - \frac{\partial G_Q(z,d)}{\partial z}\mathcal{P}_{Q,\omega}\phi_{Q,\omega}(d), \tag{11b}$$

where

$$\mathcal{P}_{Q,\omega} = \frac{e^2\Pi(Q,\omega)}{1-e^2G_Q(d,d)\Pi(Q,\omega)}\,. \tag{12}$$

Here, we emphasize that Eq. (11a) is the key to determining the dispersion relation as we shall show later.

The Green function $G_Q(z,z')$ in Eq. (7) obeys the relation

$$-\left(\frac{\partial^2}{\partial z^2} - Q^2\right)G_Q(z,z') = \frac{1}{\epsilon_0}\delta(z-z')\,. \tag{13}$$

We require the Green function to satisfy the following conditions at and away from the interface ($z=0$):

$$\epsilon_0\frac{dG_Q(z=0^+,z')}{dz} = \epsilon^\infty_{ox}\frac{dG_Q(z=0^-,z')}{dz}, \tag{14a}$$

$$G_Q(z=0^+,z') = G_Q(z=0^-,z'), \tag{14b}$$

$$G_Q(z<0,d) = G_Q(0,d)e^{+Qz}, \tag{14c}$$

$$G_Q(z>d,d) = G_Q(d,d)e^{-Q(z-d)}. \tag{14d}$$

The solution to Eq. (13) is[35]

$$G_Q(z,z') = \begin{cases} \frac{1}{2\epsilon_0 Q}(e^{-Q|z-z'|} - \lambda e^{-Q|z+z'|}), & z > 0, \\ \frac{1}{2\epsilon_0 Q}(1-\lambda)\, e^{-Q|z-z'|}, & z \leqslant 0, \end{cases} \tag{15}$$

where

$$\lambda = \frac{\epsilon_{ox}^{\infty} - \epsilon_0}{\epsilon_{ox}^{\infty} + \epsilon_0}\,.$$

The bare potential in Eq. (11a) can be written as

$$\phi_{Q,\omega}(z) = \begin{cases} A_1 e^{-Qz}, & z > 0, \\ A_2 e^{+Qz}, & z \leqslant 0, \end{cases}$$

where A_1 and A_2 are the amplitudes of the bare potential for $z > 0$ and $z \leqslant 0$, respectively. Thus the expression for the screened potential in Eq. (11a) is

$$\phi_{Q,\omega}^{\rm scr}(z) = \begin{cases} A_1 e^{-Qz} + G_Q(z,d)\mathcal{P}_{Q,\omega} A_1 e^{-Qd}, & z > 0, \\ A_2 e^{+Qz} + G_Q(z,d)\mathcal{P}_{Q,\omega} A_1 e^{-Qd}, & z \leqslant 0. \end{cases} \tag{16}$$

At the interface $z = 0$, the continuity of the component of the electric field parallel to the interface requires the continuity of $\phi_{Q,\omega}^{\rm scr}$, i.e., $\phi_{Q,\omega}^{\rm scr}(z=0^+) = \phi_{Q,\omega}^{\rm scr}(z=0^-)$, giving us

$$\begin{aligned} &A_1 + A_1 G_Q(z=0^+,d)\mathcal{P}_{Q,\omega} e^{-Qd} \\ &\quad = A_2 + A_1 G_Q(z=0^-,d)\mathcal{P}_{Q,\omega} e^{-Qd}\,. \end{aligned} \tag{17}$$

Similarly, the continuity of the perpendicular component of the electric displacement, i.e., $\epsilon_0 \frac{d\phi_{Q,\omega}^{\rm scr}(z=0^+)}{dz} = \epsilon_{ox}(\omega)\frac{d\phi_{Q,\omega}^{\rm scr}(z=0^-)}{dz}$ leads to

$$\begin{aligned} &\epsilon_0\left[A_1 - A_1\frac{1}{Q}\frac{dG_Q(z=0^+,d)}{dz}\mathcal{P}_{Q,\omega}e^{-Qd}\right] \\ &= \epsilon_{ox}(\omega)\left[-A_2 - A_1\frac{1}{Q}\frac{dG_Q(z=0^-,d)}{dz}\mathcal{P}_{Q,\omega}e^{-Qd}\right]. \end{aligned} \tag{18}$$

Substituting Eqs. (14a) and (14b) into Eqs. (17) and (18), we obtain the following relations:

$$A_1 = A_2, \tag{19a}$$

$$\begin{aligned} &\epsilon_0\left[1 - \frac{1}{Q}\frac{\partial G_Q(z=0^+,d)}{\partial z}\mathcal{P}_{Q,\omega}e^{-Qd}\right] \\ &= \epsilon_{ox}(\omega)\left[-1 - \frac{1}{Q}\frac{\partial G_Q(z=0^-,d)}{\partial z}\mathcal{P}_{Q,\omega}e^{-Qd}\right]. \end{aligned} \tag{19b}$$

Rearranging the terms in Eq. (19b), we obtain

$$\epsilon_0 + \epsilon_{ox}(\omega) + \left[\epsilon_{ox}(\omega) - \epsilon_{ox}^{\infty}\right] G_Q(0,d)\mathcal{P}_{Q,\omega}e^{-Qd} = 0, \tag{20}$$

which can be rewritten as

$$\begin{aligned} &[\epsilon_0 + \epsilon_{ox}(\omega)]\{1 - [G_Q(d,d) - G_Q(0,d)e^{-Qd}]e^2\Pi(Q,\omega)\} \\ &\quad - \left(\epsilon_0 + \epsilon_{ox}^{\infty}\right) G_Q(0,d)e^2\Pi(Q,\omega)e^{-Qd} = 0, \end{aligned}$$

or more explicitly,

$$\begin{aligned} &[\epsilon_{ox}(\omega) + \epsilon_0]\left[1 - (1 - e^{-2Qd})\frac{e^2\Pi(Q,\omega)}{2\epsilon_0 Q}\right] \\ &\quad - \frac{e^2\Pi(Q,\omega)}{Q}e^{-2Qd} = 0\,. \end{aligned} \tag{21}$$

Equation (20) gives us the dispersion of the coupled plasmon-phonon modes and is sometimes called the secular equation.[9] Physically, we expect three branches (two phonon and one plasmon). We write the coupled plasmon-phonon modes as $\omega_Q^{(\rm I)}$, $\omega_Q^{(\rm II)}$, and $\omega_Q^{(\rm III)}$ for each $\mathbf{Q}$ point. In the limit $d \to \infty$, Eq. (21) becomes

$$[\epsilon_{ox}(\omega) + \epsilon_0]\left[1 - \frac{e^2\Pi(Q,\omega)}{2\epsilon_0 Q}\right] = 0,$$

which gives us as expected the dispersion for the two uncoupled SPP branches and the single plasmon branch in isolated graphene.

B. Plasmon and phonon content

The solutions of Eq. (21) represent excitations of the IPP modes. However, the effective scattering amplitude of a particular mode may not be substantial if it is plasmonlike. Scattering with a plasmonlike excitation does not necessarily lead to loss of momentum since the momentum is simply transferred to the constituent carriers of the plasmon excitation and there is no change in the total momentum of all the carriers. On the other hand, scattering with a phononlike excitation does lead to a loss of momentum since phonons belong to a different set of degrees of freedom. Therefore, as in Ref. 9, it is necessary to define the *phonon content*[36] of each IPP mode. The phonon content quantifies the modal fraction that is phononlike and modulates its scattering strength. Likewise, we can also define the plasmon content of the mode. To find the plasmon content, we first consider the two solutions $\omega_Q^{(-g,\alpha)}$ ($\alpha = 1,2$) obtained from the secular equation Eq. (21) by ignoring the polarization response [setting $\Pi(Q,\omega) = 0$]. Following Ref. 9, the plasmon content of the IPP mode $\omega_Q^{(i)}$ is defined here as

$$\Phi^{(g)}\left(\omega_Q^{(i)}\right) = \left|\frac{\left[\omega_Q^{(i)2} - \omega_Q^{(-g,1)2}\right]\left[\omega_Q^{(i)2} - \omega_Q^{(-g,2)2}\right]}{\left[\omega_Q^{(i)2} - \omega_Q^{(j)2}\right]\left[\omega_Q^{(i)2} - \omega_Q^{(k)2}\right]}\right|, \tag{22}$$

where the indices (i,j,k) are cyclical. Note that the expected "sum rule,"[9]

$$\sum_{i=1}^{3}\Phi^{(g)}\left[\omega_Q^{(i)}\right] = 1\,, \tag{23}$$

holds. Equation (23) implies that the total plasmon weight of the three solutions is equal to one (as it would be without hybridization). The (nonplasmon) phonon content is then defined as $1 - \Phi^{(g)}[\omega_Q^{(i)}]$. In order to distinguish the TO1- and TO2-phonon parts of the nonplasmon content, we need to define the relative individual phonon content. For TO1 phonon, the content is computed by ignoring its response and replacing $\epsilon_{ox}(\omega)$ in Eq. (21) with $\epsilon_{ox}^{\infty}(\omega_{\rm LO2}^2 - \omega^2)/(\omega_{\rm TO2}^2 - \omega^2)$. From the solutions of the modified secular equation [$\omega_Q^{(-{\rm TO1},1)}$ and

$\omega_Q^{(-\mathrm{TO1},2)}]$, the relative TO1-phonon content of mode i will be

$$R^{(\mathrm{TO1})}\big[\omega_Q^{(i)}\big] = \left|\frac{\big[\omega_Q^{(i)2} - \omega_Q^{(-\mathrm{TO1},1)2}\big]\big[\omega_Q^{(i)2} - \omega_Q^{(-\mathrm{TO1},2)2}\big]}{\big[\omega_Q^{(i)2} - \omega_Q^{(j)2}\big]\big[\omega_Q^{(i)2} - \omega_Q^{(k)2}\big]}\right|, \tag{24}$$

where, as before, i, j, and k are cyclical. The relative TO2-phonon content can be similarly defined by replacing the superscript $(-\mathrm{TO1},\alpha)$ with $(-\mathrm{TO2},\alpha)$. Hence, the TO1-phonon content will be

$$\Phi^{(\mathrm{TO1})}\big[\omega_Q^{(i)}\big] = \frac{R^{(\mathrm{TO1})}\big[\omega_Q^{(i)}\big]}{R^{(\mathrm{TO1})}\big[\omega_Q^{(i)}\big] + R^{(\mathrm{TO2})}\big[\omega_Q^{(i)}\big]}\big\{1 - \Phi^{(g)}\big[\omega_Q^{(i)}\big]\big\}. \tag{25}$$

The TO2-phonon content $\Phi^{(\mathrm{TO2})}[\omega_Q^{(i)}]$ can be similarly defined. Given Eqs. (22) and (25), the following sum rules have been numerically verified:

$$\sum_{i=1}^{3}\Phi^{(\mathrm{TO1})}\big[\omega_Q^{(i)}\big] = \sum_{i=1}^{3}\Phi^{(\mathrm{TO2})}\big[\omega_Q^{(i)}\big] = 1, \tag{26a}$$

$$\Phi^{(g)}\big[\omega_Q^{(i)}\big] + \Phi^{(\mathrm{TO1})}\big[\omega_Q^{(i)}\big] + \Phi^{(\mathrm{TO2})}\big[\omega_Q^{(i)}\big] = 1, \tag{26b}$$

for each mode $\omega_Q^{(i)}$.

C. Scattering strength

As we have seen earlier, the IPP modes that result from the SPP-plasmon coupling have a different dispersion from that of the uncoupled SPP and plasmon modes. The electric field generated by the IPP modes is also different from that of the uncoupled SPP and plasmon modes. Since the remote phonon-electron coupling is derived from the quantization of the energy density of the electric field,[9] we expect this difference in the electric field to be reflected in the scattering strength of the IPP modes.

To find the scattering strength of an IPP mode, we have to determine the amplitude of its electric field. In Eq. (16), there are three unknowns (A_1, A_2, and ω), two of which (A_1 and ω) can only be eliminated through Eqs. (17) and (18). To find A_1, we follow the semiclassical approach in Ref. 9, where the time-averaged total energy of the scattering field is set equal to the zero-point energy. In the following discussion, we set $A_1 = A_Q$. We first compute the time-averaged electrostatic energy $\langle \mathcal{U}^{\mathrm{scr}}_{Q,\omega}\rangle$ associated with the screened field:

$$\big\langle \mathcal{U}^{\mathrm{scr}}_{Q,\omega}\big\rangle = \left\langle \frac{1}{2}\int dz d\mathbf{R}\,\epsilon\big[\omega_Q^{(i)}\big]\big|\nabla\big[\phi^{\mathrm{scr}}_{Q,\omega}(z)e^{i\mathbf{Q}\cdot\mathbf{R}-i\omega_Q^{(i)}t}\big]\big|^2\right\rangle. \tag{27}$$

The angle brackets $\langle\ldots\rangle$ denote time average. The volume integral in Eq. (27) is the result of three contributions: one from the substrate ($z \leqslant 0$), one from the graphene-substrate gap ($0 < z \leqslant d$), and one from the region above the graphene ($z > d$). Each term can be converted into a surface integral. Adopting a "piecewise approach" to compute the integral in Eq. (27), we must evaluate three surface integrals. To do so, we need the explicit expressions for $G_Q(z,d)$:

$$G_Q(z,d) = \begin{cases} \frac{1}{2\epsilon_0 Q}(1-\lambda)\,e^{+Q(z-d)}, & z \leqslant 0, \\ \frac{1}{2\epsilon_0 Q}[e^{+Q(z-d)} - \lambda e^{-Q(z+d)}], & 0 < z \leqslant d, \\ \frac{1}{2\epsilon_0 Q}[e^{-Q(z-d)} - \lambda e^{-Q(z+d)}], & z > d, \end{cases} \tag{28}$$

and for $-\frac{\partial}{\partial z}G_Q(z,d)$:

$$-\frac{\partial}{\partial z}G_Q(z,d) = \begin{cases} -\frac{1}{2\epsilon_0}(1-\lambda)\,e^{+Q(z-d)}, & z \leqslant 0, \\ \frac{1}{2\epsilon_0}[-e^{+Q(z-d)} - \lambda e^{-Q(z+d)}], & 0 < z \leqslant d, \\ \frac{1}{2\epsilon_0}[e^{-Q(z-d)} - \lambda e^{-Q(z+d)}], & z > d. \end{cases} \tag{29}$$

We can now evaluate the electrostatic energy in the regions $z \leqslant 0$, $0 < z \leqslant d$, and $z > d$:

$$\big\langle \mathcal{U}^{\mathrm{scr}}_{Q,\omega}\big\rangle = \big\langle \mathcal{U}^{\mathrm{scr}}_{Q,\omega}(z \leqslant 0)\big\rangle + \big\langle \mathcal{U}^{\mathrm{scr}}_{Q,\omega}(0 < z \leqslant d)\big\rangle + \big\langle \mathcal{U}^{\mathrm{scr}}_{Q,\omega}(z > d)\big\rangle. \tag{30}$$

As mentioned earlier, the volume integrals in Eq. (27) can be recast as surface integrals. Thus

$$\big\langle \mathcal{U}^{\mathrm{scr}}_{Q,\omega}(z \leqslant 0)\big\rangle = \frac{\epsilon_0\Omega A_Q^2 Q}{2}\left[1 - \frac{1}{Q}\frac{\partial G_Q(z=0^+,d)}{\partial z}\mathcal{P}_{Q,\omega}e^{-Qd}\right][1 + G_Q(z=0^+,d)\mathcal{P}_{Q,\omega}e^{-Qd}], \tag{31a}$$

$$\begin{aligned}\big\langle \mathcal{U}^{\mathrm{scr}}_{Q,\omega}(0 < z \leqslant d)\big\rangle = {}& \frac{\bar{\epsilon}_{ox}\big[\omega_Q^{(i)}\big]\Omega A_Q^2 Q}{2}\left[1 + \frac{1}{Q}\frac{\partial G_Q(z=0^-,d)}{\partial z}\mathcal{P}_{Q,\omega}e^{-Qd}\right][1 + G_Q(z=0^-,d)\mathcal{P}_{Q,\omega}e^{-Qd}] \\ & + \frac{\epsilon_0\Omega A_Q^2 Q}{2}\left[e^{-Qd} - \frac{1}{Q}\frac{\partial G_Q(z=d^+,d)}{\partial z}\mathcal{P}_{Q,\omega}e^{-Qd}\right][e^{-Qd} + G_Q(z=d^+,d)\mathcal{P}_{Q,\omega}e^{-Qd}],\end{aligned} \tag{31b}$$

$$\big\langle \mathcal{U}^{\mathrm{scr}}_{Q,\omega}(z > d)\big\rangle = -\frac{\epsilon_0\Omega A_Q^2 Q}{2}\left[e^{-Qd} - \frac{1}{Q}\frac{\partial G_Q(z=d^-,d)}{\partial z}\mathcal{P}_{Q,\omega}e^{-Qd}\right][e^{-Qd} + G_Q(z=d^-,d)\mathcal{P}_{Q,\omega}e^{-Qd}]. \tag{31c}$$

In Eq. (31b), $\bar{\epsilon}_{ox}(\omega)$ is the dielectric function of the substrate, which we distinguish with the overhead bar, and distinct from $\epsilon_{ox}(\omega)$. As we shall see later, as in Ref. 9, the function $\bar{\epsilon}_{ox}(\omega)$ is chosen in a way consistent with the particular excitation that we want. Let us regroup the terms in Eqs. (31) into those on the substrate surface at $z = 0$ and those on the graphene at $z = d$. At

$z = 0$, we have

$$\langle \mathcal{U}^{\rm scr}_{Q,\omega}(z=0)\rangle = \frac{\epsilon_0 \Omega A_Q^2 Q}{2}\left[1 - \frac{1}{Q}\frac{\partial G_Q(z=0^+,d)}{\partial z}\mathcal{P}_{Q,\omega}e^{-Qd}\right][1 + G_Q(z=0^+,d)\mathcal{P}_{Q,\omega}e^{-Qd}] + \frac{\bar{\epsilon}_{ox}[\omega_Q^{(i)}]\Omega A_Q^2 Q}{2}\left[1 + \frac{1}{Q}\frac{\partial G_Q(z=0^-,d)}{\partial z}\mathcal{P}_{Q,\omega}e^{-Qd}\right][1 + G_Q(z=0^-,d)\mathcal{P}_{Q,\omega}e^{-Qd}], \quad (32)$$

while at $z = d$, we have

$$\langle \mathcal{U}^{\rm scr}_{Q,\omega}(z=d)\rangle = \frac{\epsilon_0 \Omega A_Q^2 Q}{2}\left[e^{-Qd} - \frac{1}{Q}\frac{\partial G_Q(z=d^+,d)}{\partial z}\mathcal{P}_{Q,\omega}e^{-Qd}\right][e^{-Qd} + G_Q(z=d^+,d)\mathcal{P}_{Q,\omega}e^{-Qd}] - \frac{\epsilon_0 \Omega A_Q^2 Q}{2}\left[e^{-Qd} + \frac{1}{Q}\frac{\partial G_Q(z=d^-,d)}{\partial z}\mathcal{P}_{Q,\omega}e^{-Qd}\right][e^{-Qd} + G_Q(z=d^-,d)\mathcal{P}_{Q,\omega}e^{-Qd}]. \quad (33)$$

We have to be careful in computing $\langle \mathcal{U}^{\rm scr}_{Q,\omega}\rangle$. Mathematically, it may seem that we ought to set $\langle \mathcal{U}^{\rm scr}_{Q,\omega}\rangle = \langle \mathcal{U}^{\rm scr}_{Q,\omega}(z=0)\rangle + \langle \mathcal{U}^{\rm scr}_{Q,\omega}(z=d)\rangle$. However, note that the term $\langle \mathcal{U}^{\rm scr}_{Q,\omega}\rangle$ accounts for the various excitation effects, ionic and electronic, but the term $\langle \mathcal{U}^{\rm scr}_{Q,\omega}(z=d)\rangle$ corresponds to the charge singularity in the zero-thickness graphene sheet. It is a "self-interaction" of the charge distribution in the graphene which has no dependence on ω, unlike $\langle \mathcal{U}^{\rm scr}_{Q,\omega}(z=0)\rangle$, so that we should not expect it to contribute physically to the scattering of the graphene carriers. To see this more clearly, we rewrite Eq. (27) as

$$\langle \mathcal{U}^{\rm scr}_{Q,\omega}\rangle = \left\langle \frac{1}{2}\int dz d\mathbf{R}\, \mathbf{D}\cdot\mathbf{E}\right\rangle = \left\langle \frac{1}{2}\int dz d\mathbf{R}\, (\epsilon_0\mathbf{E} + \mathbf{P}_L + \mathbf{P}_e)\cdot\mathbf{E}\right\rangle,$$

where $\mathbf{P}_L$ and $\mathbf{P}_e$ are the polarization fields of the lattice (substrate) and the graphene electronic excitation, respectively. The following identification can be made:

$$\langle \mathcal{U}^{\rm scr}_{Q,\omega}(z=0)\rangle = \left\langle \frac{1}{2}\int dz d\mathbf{R}\, (\epsilon_0\mathbf{E} + \mathbf{P}_L)\cdot\mathbf{E}\right\rangle, \quad \langle \mathcal{U}^{\rm scr}_{Q,\omega}(z=d)\rangle = \left\langle \frac{1}{2}\int dz d\mathbf{R}\, \mathbf{P}_e\cdot\mathbf{E}\right\rangle,$$

and since we are only interested in the interaction of the *lattice* polarization field with the graphene carriers, we set $\langle \mathcal{U}^{\rm scr}_{Q,\omega}\rangle = \langle \mathcal{U}^{\rm scr}_{Q,\omega}(z=0)\rangle$, i.e.,

$$\langle \mathcal{U}^{\rm scr}_{Q,\omega}\rangle = \frac{\Omega A_Q^2 Q}{2}\{\epsilon_0 + \bar{\epsilon}_{ox}[\omega_Q^{(i)}] + [\bar{\epsilon}_{ox}[\omega_Q^{(i)}] - \epsilon^{\infty}_{ox}]G_Q(z=0,d)\mathcal{P}_{Q,\omega}e^{-Qd}\} : [1 + G_Q(z=0,d)\mathcal{P}_{Q,\omega}e^{-Qd}], \quad (34)$$

where we have used the relationship $\frac{1}{Q}\frac{\partial}{\partial z}G_Q(z=0^-,d) = G_Q(z=0,d)$. In Eq. (34), the first factor $[\epsilon_0 + \bar{\epsilon}_{ox}(\omega) - \cdots]$ resembles the secular equation (20). Indeed, if we replace $\bar{\epsilon}_{ox}(\omega)$ with $\epsilon_{ox}(\omega)$, then $\langle \mathcal{U}^{\rm scr}_{Q,\omega}\rangle = 0$ as expected. This is no coincidence since $\mathcal{U}^{\rm scr}_{Q,\omega}$ represents the energy of the charge distribution present at the substrate-vacuum interface and the secular equation in Eq. (20) is a statement about the *absence* of charges at the substrate-vacuum interface. This also confirms our earlier choice of excluding the contribution from the surface charges at $z = d$, since that contribution does not disappear when we replace $\bar{\epsilon}_{ox}(\omega)$ with $\epsilon_{ox}(\omega)$. By regrouping the terms according to the position of their charge distribution in Eq. (30), we make the relationship to the secular equation Eq. (20) manifest. Using Eq. (12), the expression for the screened electrostatic energy can be rewritten as

$$\langle \mathcal{U}^{\rm scr}_{Q,\omega}\rangle = \frac{\Omega A_Q^2 Q}{2}\{\bar{\epsilon}_{ox}[\omega_Q^{(i)}] - \epsilon_{ox}[\omega_Q^{(i)}]\}\left\{\frac{1 - G_Q(d,d)e^2\Pi[Q,\omega_Q^{(i)}] + G_Q(0,d)e^2\Pi[Q,\omega_Q^{(i)}]e^{-Qd}}{1 - G_Q(d,d)e^2\Pi[Q,\omega_Q^{(i)}]}\right\}^2. \quad (35)$$

We use the relationship $\langle \mathcal{W}^{\rm scr}_{Q,\omega}\rangle = 2\langle \mathcal{U}^{\rm scr}_{Q,\omega}\rangle$ to obtain the time-averaged total energy, and set it equal to the zero-point energy, i.e., $\frac{1}{2}\hbar\omega_Q^{(i)} = \langle \mathcal{W}^{\rm scr}_{Q,\omega}\rangle$, so that

$$\frac{1}{2}\hbar\omega_Q^{(i)} = \Omega A_Q^2 Q\{\bar{\epsilon}_{ox}[\omega_Q^{(i)}] - \epsilon_{ox}[\omega_Q^{(i)}]\}\left\{\frac{1 - G_Q(d,d)e^2\Pi[Q,\omega_Q^{(i)}] + G_Q(0,d)e^2\Pi[Q,\omega_Q^{(i)}]e^{-Qd}}{1 - G_Q(d,d)e^2\Pi[Q,\omega_Q^{(i)}]}\right\}^2.$$

Therefore the squared amplitude of the field is

$$A_Q^2 = \frac{\hbar\omega_Q^{(i)}}{2\Omega Q\{\bar{\epsilon}_{ox}[\omega_Q^{(i)}] - \epsilon_{ox}[\omega_Q^{(i)}]\}}\left\{\frac{1 - G_Q(d,d)e^2\Pi[Q,\omega_Q^{(i)}] + G_Q(0,d)e^2\Pi[Q,\omega_Q^{(i)}]e^{-Qd}}{1 - G_Q(d,d)e^2\Pi[Q,\omega_Q^{(i)}]}\right\}^{-2}.$$

To determine the strength of the scattering field, say for the TO1 phonon, we take the difference between (i) the squared amplitude of the field with the TO1 mode frozen and (ii) that of the field with the mode in full response. In (i), we

set

$$\bar{\epsilon}_{ox}^{\mathrm{TO1},\infty}(\omega)=\epsilon_{ox}^{\infty}\left(\frac{\omega_{\mathrm{LO2}}^2-\omega^2}{\omega_{\mathrm{TO2}}^2-\omega^2}\right)$$

and

$$\bar{\epsilon}_{ox}^{\mathrm{TO1},0}(\omega)=\epsilon_{ox}^{\infty}\left(\frac{\omega_{\mathrm{LO2}}^2-\omega^2}{\omega_{\mathrm{TO2}}^2-\omega^2}\right)\frac{\omega_{\mathrm{LO1}}^2}{\omega_{\mathrm{TO1}}^2}\,.$$

The squared amplitude of the TO1 scattering field for $\omega=\omega_Q^{(i)}$ is

$$A_{\mathrm{TO1}}\big[Q,\omega_Q^{(i)}\big]^2=\frac{\hbar\omega_Q^{(i)}}{2\Omega Q}\left\{\frac{1}{\bar{\epsilon}_{ox}^{\mathrm{TO1},\infty}\big[\omega_Q^{(i)}\big]-\epsilon_{ox}\big[\omega_Q^{(i)}\big]}-\frac{1}{\bar{\epsilon}_{ox}^{\mathrm{TO1},0}\big[\omega_Q^{(i)}\big]-\epsilon_{ox}\big[\omega_Q^{(i)}\big]}\right\}\Phi^{(\mathrm{TO1})}\big[\omega_Q^{(i)}\big]$$
$$\times\left\{\frac{1-G_Q(d,d)e^2\Pi\big[Q,\omega_Q^{(i)}\big]+G_Q(0,d)e^2\Pi\big[Q,\omega_Q^{(i)}\big]e^{-Qd}}{1-G_Q(d,d)e^2\Pi\big[Q,\omega_Q^{(i)}\big]}\right\}^{-2}. \tag{36}$$

The expression for the TO2 scattering field can be similarly obtained. Therefore the TO1 *effective* scattering field can be written as

$$\phi_{Q,\omega}^{\mathrm{scr}}(d)=A_{\mathrm{TO1}}\big[Q,\omega_Q^{(i)}\big]\left\{e^{-Qd}+G_Q(z,d)\frac{e^2\Pi\big[Q,\omega_Q^{(i)}\big]}{1-e^2G_Q(d,d)\Pi\big[Q,\omega_Q^{(i)}\big]}e^{-Qd}\right\}. \tag{37}$$

The scattering potential is

$$V(\mathbf{R},z)=\sum_{l=1}^{3}\sum_{\mu=1}^{2}eA_{\mathrm{TO1}}\big[Q,\omega_Q^{(l)}\big]\big[e^{-Qz}+G_Q(z,d)\mathcal{P}_{Q,\omega_Q^{(l)}}e^{-Qd}\big]e^{i\mathbf{Q}\cdot\mathbf{R}-i\omega_Q^{(l)}t}\big[a_{\mathbf{Q}}^{(l)}+a_{-\mathbf{Q}}^{(l)\dagger}\big], \tag{38}$$

where $a_{\mathbf{Q}}^{(l)}$ ($a_{\mathbf{Q}}^{(l)\dagger}$) is the annihilation (creation) operator for the mode corresponding to $\mathbf{Q}$ and $\omega_Q^{(l)}$. Generally, the graphene field operator can be written in the spinorial form as

$$\Psi(\mathbf{R},z)=\frac{1}{\sqrt{2\Omega}}\sum_{s=\pm1}\sum_{\mathbf{K}}\left[\begin{pmatrix}1\\ se^{i\theta_{\mathbf{K}}}\end{pmatrix}c_{\mathbf{K}}^{s\mathcal{K}}+\begin{pmatrix}e^{i\theta_{\mathbf{K}}}\\ s\end{pmatrix}c_{\mathbf{K}}^{s\mathcal{K}'}\right]e^{i\mathbf{K}\cdot\mathbf{R}}\sqrt{\delta(z-d)}, \tag{39}$$

where $\mathcal{K}(\mathcal{K}')$ denotes the $\mathcal{K}(\mathcal{K}')$ valley, and the $+(-)$ sign corresponds to the $\pi(\pi^*)$ band; $c_{\mathbf{K}}^{s\mathcal{K}}(c_{\mathbf{K}}^{s\mathcal{K}\dagger})$ is the annihilation (creation) operator of the s-band $\mathbf{K}$ electron state at the $\mathcal{K}$ valley. Therefore the interaction term is

$$H_{\mathrm{int}}=\int dz\int d\mathbf{R}\Psi^{\dagger}(\mathbf{R},z)V(\mathbf{R},z)\Psi(\mathbf{R},z)$$

and, if we neglect the intervalley terms, simplifies to

$$H_{\mathrm{int}}\approx\sum_{l=1}^{3}\sum_{s_1,s_2}\sum_{\mathbf{K},\mathbf{Q}}M_Q^{(l)}\alpha_{s_1\mathbf{K}+\mathbf{Q},s_2\mathbf{K}}\big(c_{\mathbf{K}+\mathbf{Q}}^{s_1\mathcal{K}\dagger}c_{\mathbf{K}}^{s_2\mathcal{K}}+s_1s_2c_{\mathbf{K}+\mathbf{Q}}^{s_1\mathcal{K}'\dagger}c_{\mathbf{K}}^{s_2\mathcal{K}'}\big)\big[a_{\mathbf{Q}}^{(l)}+a_{-\mathbf{Q}}^{(l)\dagger}\big], \tag{40}$$

where

$$\alpha_{s_1\mathbf{K_1},s_2\mathbf{K_2}}=\frac{1+s_1s_2e^{-i(\theta_{\mathbf{K}_1}-\theta_{\mathbf{K}_2})}}{2}$$

is the overlap integral that comes from the inner product of the spinors, and

$$M_Q^{(l)}=\left(\sum_{\mu=1}^{2}\frac{e^2\hbar\omega_Q^{(l)}}{2\Omega Q}\left\{\frac{1}{\bar{\epsilon}_{ox}^{\mathrm{TO}\mu,\infty}\big[\omega_Q^{(l)}\big]-\epsilon_{ox}\big[\omega_Q^{(l)}\big]}-\frac{1}{\bar{\epsilon}_{ox}^{\mathrm{TO}\mu,0}\big[\omega_Q^{(l)}\big]-\epsilon_{ox}\big[\omega_Q^{(l)}\big]}\right\}\Phi^{(\mathrm{TO}\mu)}\big[\omega_Q^{(l)}\big]\right)^{1/2}$$
$$\times\big|1-G_Q(d,d)e^2\Pi\big[Q,\omega_Q^{(l)}\big]+G_Q(0,d)e^2\Pi\big[Q,\omega_Q^{(l)}\big]e^{-Qd}\big|^{-1} \tag{41}$$

is the electron-phonon coupling coefficient corresponding to the $\omega_Q^{(l)}$ mode.

D. Landau damping

At sufficiently short wavelengths, plasmons cease to be proper quasiparticle excitations because of Landau damping.[29] To model this phenomenon, albeit approximately, we take that to be the case when the pure graphene plasmon excitation, of which the dispersion $\omega=\omega_p(Q)$ is determined by the

expression $1 - e^2 G_Q(d,d)\Pi(Q,\omega) = 0$, enters the intraband single-particle excitation (SPE) continuum.[37] This happens when the plasmon branch crosses the electron dispersion curve, i.e., when $\omega_p = \hbar v_F Q$, and the wave vector at which this happens is Q_c. A common cutoff Q_c is used for all three IPP branches because we are constrained by the necessity to maintain the sum rule in Eq. (23). Thus, although the top IPP branch may undergo Landau damping from interband transitions, we assume that the IPP branch III modes are still well defined and do not undergo significant broadening. Selective omission of the top IPP branch while still retaining the lower two coupled plasmon-phonon branches would violate the sum rules $\Phi^{(g)}[\omega_Q^{(\mathrm{IPP}_1)}] + \Phi^{(g)}[\omega_Q^{(\mathrm{IPP}_2)}] + \Phi^{(g)}[\omega_Q^{(\mathrm{IPP}_3)}] = \Phi^{(\mathrm{TO}\mu)}[\omega_Q^{(\mathrm{IPP}_1)}] + \Phi^{(\mathrm{TO}\mu)}[\omega_Q^{(\mathrm{IPP}_2)}] + \Phi^{(\mathrm{TO}\mu)}[\omega_Q^{(\mathrm{IPP}_3)}] = 1$. If we set Q_c to be the point where the plasmonlike branch III undergoes interband SPE Landau damping, then the lower branches would have to be replaced by SPP branches. This approximation is unreasonable because it assumes that the modes are broadened when they are not. On the other hand, it is less severe to assume that interband SPE Landau damping does not lead to significant broadening of the plasmonlike IPP modes. Admittedly, this is not a perfect way of handling the issue of Landau damping and is a limitation of our theory, which assumes the spectral weight of the modes to be sharp delta function in frequency space.

When $Q < Q_c$, the electron-phonon coupling coefficient in Eq. (40) is that of Eq. (41). Although the lower-frequency IPP branches may undergo Landau damping from intraband SPE as $\omega_Q^{(l)} < v_F Q$, we still retain them because the sum rules in Eqs. (23) and (26) require us to maintain charge conservation.[29] On the other hand, when $Q > Q_c$, Landau damping is assumed to dominate all the IPP modes and the coupling between the substrate SPP modes and the graphene plasmons can be ignored. Instead of scattering with three IPP modes for each given wave vector, we revert to using only two SPP modes. This allows us to satisfy the sum rules in Eq. (26). In this case, the electron-phonon coupling coefficient in Eq. (41) can be rewritten as

$$M_Q^{(l)} = \left(\sum_{\mu=1}^{2} \frac{e^2\hbar\omega_Q^{(l)}}{2\Omega Q} \left\{ \frac{1}{\bar{\epsilon}_{ox}^{\mathrm{TO}\mu,\infty}[\omega_Q^{(l)}] + \epsilon_0} - \frac{1}{\bar{\epsilon}_{ox}^{\mathrm{TO}\mu,0}[\omega_Q^{(l)}] + \epsilon_0} \right\} \Phi^{(\mathrm{TO}\mu)}[\omega_Q^{(l)}] \right)^{1/2}, \quad (42)$$

where $l = \mathrm{SO}_1, \mathrm{SO}_2$ indexes the SPP branch. The expression in Eq. (42) can be derived from Eq. (41) obtained in the limit of $\Pi \to 0$. If there is no polarization charge, the secular equation is $\epsilon_{ox}[\omega_Q^{(l)}] + \epsilon_0 = 0$, which gives us $\epsilon_{ox}[\omega_Q^{(l)}] = -\epsilon_0$. If we substitute the latter back into Eq. (41), we obtain Eq. (42).

III. RESULTS AND DISCUSSION

A. Numerical evaluation

Having set up the theoretical framework for electron-IPP interaction, we compute the dispersion of the coupled interfacial plasmon-phonon modes and study the electrical transport properties.

1. Interfacial plasmon-phonon dispersion

In this section we compute the scattering rates from the remote phonons by employing the dispersion relation $[\omega_Q^{(l)}]$ and the electron-phonon coupling coefficient $[M_Q^{(l)}]$, which can be determined by solving Eqs. (20) and (40), respectively. For simplicity, to solve the latter equations, we use the zero-temperature, long-wavelength approximation for $\Pi(Q,\omega)$:[37,38]

$$\Pi(Q,\omega) = \frac{Q^2 E_F}{\pi\hbar^2\omega^2}, \quad (43)$$

where E_F is the Fermi level which can be determined from the carrier density n via the relation $n = E_F^2/(\pi\hbar^2 v_F^2)$.

In Fig. 2, we show the dispersion relation for an SiO_2 substrate with $n = 10^{12}$ cm^{-2}. The three coupled IPP branches are drawn with solid lines and labeled I, II, and III, while the dispersion of the uncoupled modes is drawn in dashed lines in the figure. The branches labeled "SO_1" (61 meV) and "SO_2" (149 meV) have a flat dispersion and are determined from the equation

$$\epsilon_0 + \epsilon_{ox}(\omega) = 0,$$

while the branch labeled "Pure plasmon" is determined from the zeros of the equation

$$1 - G_Q(d,d)e^2\Pi(Q,\omega), \quad (44)$$

which gives the dispersion of the pure graphene plasmons when the frequency dependence of the substrate dielectric function is neglected and only the effect of the substrate image charges is taken into account. We observe that in the long wavelength limit ($Q \to 0$), branches I, II, and III converge asymptotically to the pure plasmon, SO_1, and SO_2 branches respectively. On the other end, as $Q \to \infty$, branches I, II, and III converge asymptotically to the pure SO_1, SO_2, and plasmon branches respectively. At intermediate values of Q, the IPP branches are a mixture of the pure branches. The coupling between pure SO phonons and graphene plasmons has often been ignored in transport studies based on the dispersionless unscreened, uncoupled SO modes[19,20,22–24,30,39] On the other

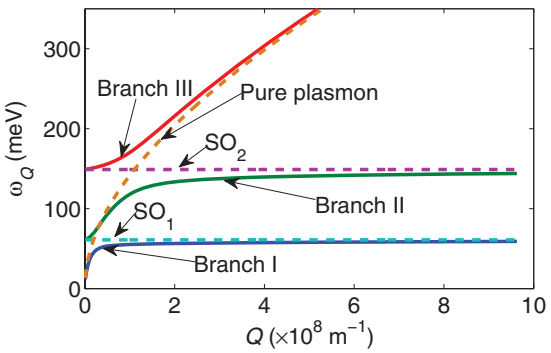

FIG. 2. (Color online) Dispersion relation of coupled interfacial plasmon-phonon system with $n = 10^{12}$ cm^{-2} for the SiO_2 substrate. The three hybrid IPP branches are labeled Branch I, II, and III. The uncoupled pure SO phonon (1 and 2) and plasmon branches are drawn with dashed lines. In the limits $Q \to 0$ and $Q \to \infty$, the IPP branches converge to the pure phonon and plasmon branches. In between, they are a mix of the pure branches.

hand, using many-body techniques, Hwang, Sensarma, and Das Sarma[40] studied the remote phonon-plasmon coupling in supported graphene and were able to reproduce the coupled plasmon-phonon dispersion observed by Liu and Willis[31,32] in their angle-resolved electron-energy-loss spectroscopy experiments on epitaxial graphene grown on SiC. Similar results of strongly coupled plasmon-phonon modes were reported by Koch, Seyller, and Schaefer.[34] Fei and coworkers also found evidence of this plasmon-phonon coupling in the graphene-SiO_2 system in their infrared nanoscopy experiments.[33] Given the increasing experimental support for the hybridization of the SPPs with the graphene plasmons, it is interesting to investigate the effect of these coupled modes on carrier transport in graphene.

2. *Electron-phonon coupling*

Here, the electron-phonon coupling coefficients $M_Q^{(l)}$ of the IPP and the SPP modes are compared at $n = 10^{12}$ cm^{-2} for the SiO_2 substrate. Recall that the IPP modes are formed through the hybridization of the SPP and graphene plasmon modes, and their coupling to the graphene electrons are different to that of the SPP modes. It is sometimes assumed[19,22] that the SPP modes are screened by the plasmons, and the IPP-electron coupling is weaker than the SPP-electron coupling. As we have discussed above, this assumption does not hold when the frequency of the IPP mode is higher than the plasmon frequency. We plot the $M_Q^{(l)}(\Omega/A)^{1/2}$, where A is the area of the primitive unit cell in SLG, for the SPP and IPP modes in Fig. 3(a) . We first notice that at small Q, the coupling terms for branches I and II are actually larger than those for SO_1 and SO_2. This is because at long wavelengths, $\omega_p < \omega_Q^{(l)}$ for $l =$ I and II, resulting in antiscreening, which enhances the IPP electric field. However, at longer wavelengths in the limit $Q \to 0$ [see Fig. 3(b)], the strength of the antiscreening diminishes, and the electron-IPP coupling coefficients converge to those of the electron-SPP coupling coefficients, i.e.,

$$\lim_{Q\to 0}\begin{pmatrix} M_Q^{(\mathrm{II})} \\ M_Q^{(\mathrm{III})} \end{pmatrix} = \begin{pmatrix} M_Q^{(\mathrm{SO1})} \\ M_Q^{(\mathrm{SO2})} \end{pmatrix} .$$

The electron-IPP coupling coefficients $M_Q^{(\mathrm{II})}$ and $M_Q^{(\mathrm{III})}$ scale as $\sim 1/Q^{1/2}$. $M_Q^{(\mathrm{I})}$ scales as $\sim 1/Q^{1/4}$ because $M_Q^{(\mathrm{I})} \propto [\omega_Q^{(\mathrm{I})}/Q]^{1/2}$ and $\lim_{Q\to 0}\omega_Q^{(\mathrm{I})} \sim Q^{1/2}$ since branch I is plasmonlike. In the short wavelength limit ($Q \to \infty$), the phononlike branch I and II electron-IPP coupling coefficients appear to scale as $\sim 1/Q^2$. This can be interpreted as a signature of dynamic screening. $M_Q^{(\mathrm{III})}$ scales as $\sim 1/Q^{1/4}$ since branch III is plasmonlike in the $Q \to \infty$ limit. Also, $M_Q^{(\mathrm{III})}$ is actually much larger than $M_Q^{(\mathrm{SO1})}$ and $M_Q^{(\mathrm{SO2})}$ over the entire range of Q values. When we take Landau damping into account, we use the coupling coefficients [shaded in Fig. 3(a)] of branches I, II, and II for $Q < Q_c$ and of SO_1 and SO_2 for $Q \geqslant Q_c$.

To understand the relationship between the coupling strength and the phonon content of the IPP modes, we take the expression for the polarizability in Eq. (43), and plot the phonon and plasmon contents [see Eqs. (22) and (25)] of the IPP branches in Figs. 3(a)–3(c). We also plot the dimensionless Q-dependent quantity

$$\Lambda^{(l)}(Q) = M_Q^{(l)}\left[\frac{e^2\hbar\omega_Q^{(l)}}{2\Omega Q\epsilon_0}\right]^{-1/2} , \qquad (45)$$

which serves as a measure of the dipole coupling strength of the corresponding $\omega_Q^{(l)}$ mode, for $l =$ I, II, and III in Figs. 3(d)–3(f), respectively.

Figure 4(a) shows the change in the TO1 phonon content [$\Phi^{(\mathrm{TO1})}$] of the branch I modes as Q increases from 0 to $2Q_F$

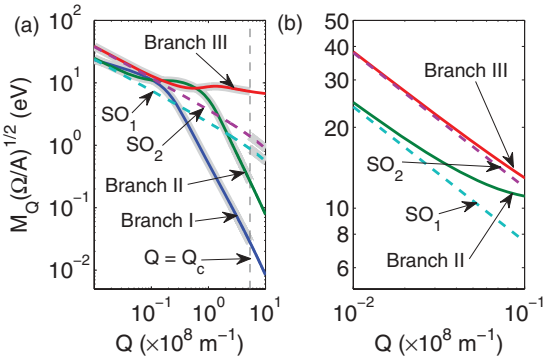

FIG. 3. (Color online) (a) Plot of $M_Q^{(l)}(\Omega/A)^{1/2}$ for $n = 10^{12}$ cm^{-2} in SiO_2-supported graphene. The IPP branches are labeled Branch I, II, and III, and the SPP branches are labeled SO_1 and SO_2. The cutoff wave vector Q_c is drawn in gray dashed lines. When $Q < Q_c$, we use the part of the IPP branches shaded in gray, and when $Q \geqslant Q_c$, we use the part of the SPP branches shaded in gray. (b) Plot of $M_Q^{(l)}(\Omega/A)^{1/2}$ at small Q for $l =$ II, III, SO_1, and SO_2. Note that $\lim_{Q\to 0} M_Q^{(\mathrm{II,III})} = M_Q^{(\mathrm{SO1,SO2})}$ as expected, since $\lim_{Q\to 0}\omega_Q^{(\mathrm{II,III})} = \omega_Q^{(\mathrm{SO1,SO2})}$.

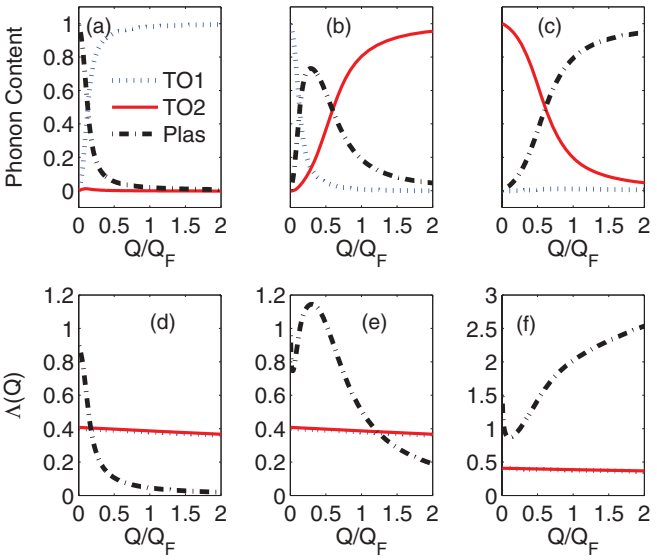

FIG. 4. (Color online) We plot $\Phi^{(\mathrm{TO1})}[\omega_Q^{(l)}]$, $\Phi^{(\mathrm{TO2})}[\omega_Q^{(l)}]$ and $\Phi^{(g)}[\omega_Q^{(l)}]$, which we label as "TO1" (dotted line), "TO2" (solid line), and "Plas" (dash-dot line), for $l =$ (a) IPP_1, (b) IPP_2, and (c) IPP_3. The corresponding dimensionless dipole coupling strength $\Lambda^{(l)}(Q)$ are drawn in dash-dot lines for $l =$ (d) IPP_1, (e) IPP_2 and (f) IPP_3. For comparison, $\Lambda^{(\mathrm{SO1})}(Q)$ and $\Lambda^{(\mathrm{SO2})}(Q)$ are also drawn in solid and dotted lines, respectively, in each of the subdiagrams (d) to (f).

where $Q_F = E_F/(2\hbar v_F)$. At small Q, the branch I modes are plasmonlike. Likewise, in Fig. 2, the branch I dispersion is close to that of the pure graphene plasmons. Since the branch I modes are plasmonlike, the corresponding $\Lambda^{(\mathrm{I})}(Q)$ values in Fig. 4(d) are also larger than $\Lambda^{(\mathrm{SO1})}(Q)$ and $\Lambda^{(\mathrm{SO2})}(Q)$. This is because the plasmon part of the branch I modes is antiscreened. As Q increases, the branch I dispersion moves closer to the SO_1 branch, and this is reflected in the increase in the TO1 phonon content in Fig. 4(a). The modes become more TO1 phononlike. Therefore the corresponding $\Lambda^{(\mathrm{I})}(Q)$ values in Fig. 4(d) undergo a reduction and are smaller than those of the SO_1 and SO_2 branches because of increased screening.

A similar trend can be seen in Fig. 4(b). Branch II is TO1 phononlike initially and becomes more plasmonlike, and finally TO2 phononlike as Q increases. In Fig. 2, the branch II dispersion is initially close to the SO_1 phonon branch. As Q increases, it approaches the plasmon branch and also becomes more plasmonlike. In Fig. 4(e), the associated $\Lambda^{(\mathrm{II})}(Q)$ values also undergo an increase initially. Then, as it converges to the SO_2 phonon branch, it becomes more TO2 phononlike, and the $\Lambda^{(\mathrm{II})}(Q)$ values then start to fall below $\Lambda^{(\mathrm{SO1})}(Q)$ and $\Lambda^{(\mathrm{SO2})}(Q)$ because of screening. A similar analysis can be performed for branch III. The reason for the enhanced coupling strength of the branch III modes is clear: being more plasmonlike as $Q \to \infty$, the modes are weakly screened. Hence, in Fig. 4(f), $\Lambda^{(\mathrm{III})}(Q)$ is always greater than $\Lambda^{(\mathrm{SO1})}(Q)$ and $\Lambda^{(\mathrm{SO2})}(Q)$.

B. Substrate-limited mobility

The momentum relaxation rate for an electron in band s with wave vector $\mathbf{K}$ can be written as

$$\Gamma_{\mathrm{RP}}(s,\mathbf{K}) = \frac{2\pi}{\hbar}\sum_{l}\sum_{s'}\sum_{\mathbf{Q}}\left|M_Q^{(l)}\alpha_{s\mathbf{K}+\mathbf{Q},s'\mathbf{K}}\right|^2[1-ss'\cos(\theta_{\mathbf{K}+\mathbf{Q}}-\theta_{\mathbf{K}})]\left(\left\{1+N_B\left[\omega_Q^{(l)}\right]\right\}[1-f(E_{s'\mathbf{K}+\mathbf{Q}})]\right.$$
$$\left.\times\,\delta\left(E_{s\mathbf{K}}-E_{s'\mathbf{K}+\mathbf{Q}}-\hbar\omega_Q^{(l)}\right)+N_B\left[\omega_Q^{(l)}\right][1-f(E_{s'\mathbf{K}+\mathbf{Q}})]\delta\left[E_{s\mathbf{K}}-E_{s'\mathbf{K}+\mathbf{Q}}+\hbar\omega_Q^{(l)}\right]\right), \tag{46}$$

where $N_B(\omega) = (e^{\hbar\omega/k_BT}-1)^{-1}$, $f(E) = [e^{(E-E_F)/k_BT}+1]^{-1}$, and $E_{s\mathbf{K}} = s\hbar|\mathbf{K}|$. In assuming the latter expression, we use the Dirac-conical approximation. Equation (46) automatically includes the Fermi-Dirac distribution of the final states and remains applicable when the doping level is high. The individual scattering rates for the screened (I, II, and III) and unscreened (SO_1 and SO_2) branches at the carrier concentration of $n = 10^{12}$ cm^{-2} in SiO_2 and HfO_2 are plotted in Fig. 5. Landau damping is taken into account by setting the coupling coefficient of the IPP (SPP) modes to zero when $Q < Q_c$ ($Q \geqslant Q_c$). We observe that at low energies, the IPP scattering rates are much higher than the SPP ones. At higher energies, the SPP scattering rates increase rapidly. The dominant scattering mechanism around the Fermi level appears to be due to the plasmonlike branch III in SiO_2 and HfO_2. In addition, at the Fermi level in HfO_2, the SPP branches have scattering rates comparable to those of branch III. This explains why the low density mobility of HfO_2 is less than that of SiO_2.

The expression for the IPP/SPP-limited part of the electrical conductivity is

$$\sigma_{\mathrm{RP}} = \frac{g_s g_v e^2}{4\pi\hbar^2 k_B T}\int_0^\infty f(E-E_F)\times[1-f(E-E_F)]\Gamma_{\mathrm{tr}}(E)^{-1}EdE\,, \tag{47}$$

where $g_s = 2$ and $g_v = 2$ are the spin and valley degeneracies, respectively. Only the contribution from the conduction band is included in Eq. (47). We use Eqs. (46) and (47) to compute the IPP or SPP-limited electrical conductivity by setting

$$\Gamma_{\mathrm{tr}}(E) = \Gamma_{\mathrm{RP}}(s,\mathbf{K})\,. \tag{48}$$

In making this approximation, we ignore the other effects (ripples, charged impurities, acoustic phonons, optical phonons, etc.). The scattering rates from the acoustic and optical phonons tend to be significantly smaller and are not the limiting factor in electrical transport in supported graphene.[41] Impurity scattering tends to be the dominant limiting factor, but its effects can be reduced by varying fabrication conditions. Thus the conductivity using Eq. (48) gives us its upper bound. We calculate the remote phonon-limited mobility as

$$\mu_{\mathrm{RP}} = \frac{\sigma_{\mathrm{RP}}}{en}, \tag{49}$$

where $n = \frac{g_s g_v}{2\pi\hbar^2 v_F^2}\int_0^\infty f(E-E_F)EdE$ is the carrier density. For $n = 10^{12}$ cm^{-2} in SiO_2, we obtain $\mu_{\mathrm{RP}} \approx 40\,000$ cm^{2}V^{-1}s^{-1}. This is more than the corresponding values reported in the literature (∼1000–20 000 cm^{2}V^{-1}s^{-1})[1,42] although we have to bear in mind that it is an upper limit. Nonetheless, it suggests that IPP and SPP scattering imposes a bound on the electron mobility.

C. Mobility results

Although suspended graphene has an intrinsic mobility limit of 200 000 cm^{2}V^{-1}s^{-1} at room temperature,[4] typical numbers for graphene on SiO_2 tend to fall in the range 1000–20 000 cm^{2}V^{-1}s^{-1}.[15] One significant reason for this drastic reduction in mobility is believed to be the presence of charged impurities in the substrate which causes long-range Coulombic scattering[25,26,43] and much effort has been directed towards the amelioration of the effects of these charged impurities. For example, it has been suggested that modifying the dielectric environment of the graphene, either through immersion in a high-κ liquid or an overlayer of high-κ dielectric material, can lead to a weakening of the Coulombic interaction and an increase in electron mobility.[43] On the other hand, actual experimental evidence in favor of this

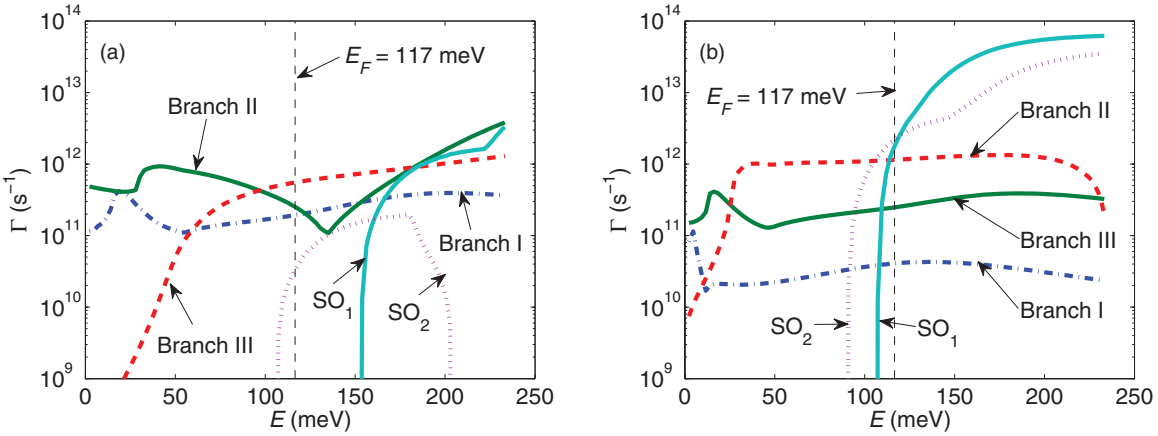

FIG. 5. (Color online) Plot of scattering rates at $n = 10^{12}$ cm^{-2} for different substrates: (a) SiO_2 and (b) HfO_2. In SiO_2, the plasmonlike branch III dominates the scattering rate at $E = E_F$. In HfO_2, branches III, SO_1 and SO_2 dominate the scattering rate at $E = E_F = 117$ meV. The SPP branches (SO_1 and SO_2) do not contribute much to the Fermi-level scattering rate in SiO_2 because of their higher frequencies and smaller occupation factors.

theory is ambiguous. Electrical conductivity data from Jang and coworkers[43] as well as Ponomarenko and coworkers[44] indicate a smaller-than-expected increase in mobility when a liquid overlayer is used. This suggests that mechanisms other than long and short-range impurity scattering are at play here. Here, we turn to the problem of scattering by IPP modes.

1. *Comparing different substrates*

Having set up the theoretical framework in the earlier sections, we now apply it to the study of the remote phonon-limited mobility of four commonly-used substrates: SiO_2, HfO_2, h-BN, and Al_2O_3. Their parameters are given in Table I. Silicon dioxide is the most common substrate material while HfO_2 and Al_2O_3 are high-κ dielectrics commonly used as top gate oxides.[39,45] Hexagonal boron nitride shows much promise as both a substrate and a top gate dielectric material.[14] The study of the remote phonon-limited mobility in these substrates allows us to understand how electronic transport in supported graphene depends on the frequencies and relative permittivities of the substrate phonons.

From Eq. (47) with the effects of Landau damping taken into account, we compute the remote phonon-limited mobility numerically, using the well-known Gilat-Raubenheimer method[46] to discretize the sum in Eq. (47). We plot μ_{RP} as a function of carrier density ($n = 0.3 \times 10^{12}$ to 5.2×10^{12} cm^{-2}) at 300 K in Fig. 6. Note that the mobility values of the high-κ substrates (HfO_2 and Al_2O_3) are substantially lower compared to SiO_2 and h-BN in the carrier density range $n < 2.0 \times 10^{12}$ cm^{-2}. Similar results have been found in MOS systems.[9] Hexagonal BN has the highest mobility at low carrier densities because of its high phonon frequencies, which corresponds to low Bose-Einstein occupancy, as well as its weak dipole coupling to graphene. In general, μ_{RP} for all four substrates increases with n because the dynamic screening effect becomes stronger at higher carrier densities. At low carrier densities, the mobility is low for all the substrates because there is a large proportion of plasmons modes whose frequencies are lower than the SPP mode frequencies. Thus, their coupling to the SPP modes results in the formation of *antiscreened* IPP modes that couple *more* strongly to the carriers, a phenomenon that has been studied for polar semiconductors.[29] However, as n increases, the mobility for

TABLE I. Parameters [see Eq. (1)] used in computing dispersion relation and scattering rates for SiO_2, h-BN, HfO_2, and Al_2O_3. They are taken from Refs. 9 and 30.

	SiO_2	h-BN	HfO_2	Al_2O_3
ϵ^0_{ox} (ϵ_0)	3.90	5.09	22.00	12.53
ϵ^i_{ox} (ϵ_0)	3.05	4.57	6.58	7.27
ϵ^∞_{ox} (ϵ_0)	2.50	4.10	5.03	3.20
ω_{TO1} (meV)	55.60	97.40	12.40	48.18
ω_{TO2} (meV)	138.10	187.90	48.35	71.41

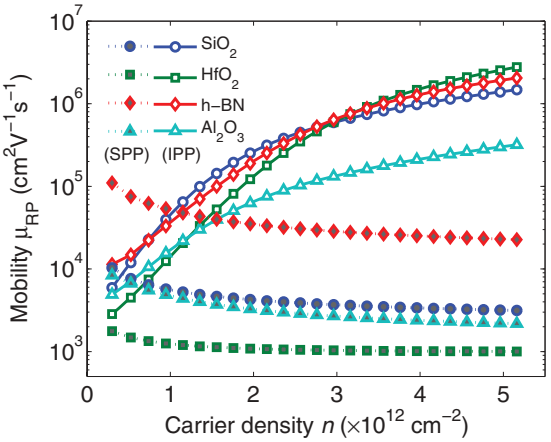

FIG. 6. (Color online) Calculated conductivity remote phonon-limited mobility for different values of carrier density and different substrates (SiO_2, HfO_2, h-BN, and Al_2O_3) at room temperature (300 K). The IPP-limited mobility values are plotted using solid lines with unfilled symbols while the SPP-limited mobility values are plotted using dotted lines with solid symbols.

all four substrates rises because the plasmon frequency scales as $\omega_p \propto n^{1/4}$, resulting in higher-frequency plasmon modes. Thus, the plasmon-phonon coupling forms screened IPP modes that are weakly coupled to the carriers. Furthermore, at higher carrier densities, Landau damping becomes less important as a result of the increasing magnitude of the plasmon wave vector Q_c. Contrary to expectation, we find that the mobility for HfO_2 exceeds those of other substrates at larger densities ($n = 5 \times 10^{12}$ cm^{-2}). At $n = 5 \times 10^{12}$ cm^{-2}, HfO_2 has the highest remote-phonon mobility followed by h-BN, SiO_2, and Al_2O_3. This is because the proportion of screened IPP modes increases with increasing carrier density. Given the small values of ω_{TO1} and ω_{TO2} for HfO_2, its coupling coefficients are smaller as a result of stronger dynamic screening. This weaker coupling compensates in part the higher occupation factors. In contrast, the larger values of ω_{TO1} and ω_{TO2} for h-BN imply that screening does not play a significant role at low carrier densities. Hence, its coupling to the graphene carriers does not diminish as rapidly as carrier density increases. The computed μ_{RP} values for HfO_2 and h-BN highlight the role of low-frequency excitations in carrier scattering. The low-frequency modes are highly occupied at room temperature and induce carrier significant scattering at low n. At higher n when dynamic screening becomes important, the low-frequency modes are more strongly screened and their coupling to the carriers becomes diminished more rapidly than that of high-frequency modes.

2. *Dynamic screening effects*

To compute the mobility for the case without any screening or antiscreening effects, the Landau damping cutoff wave vector is decreased, i.e., $Q_c \to 0$, resulting in the replacement of all the IPP modes with SPP modes. We plot the SPP-limited mobility as a function of carrier density in Fig. 6 (solid symbols), and compare these results for the IPP-limited mobility. The SPP-limited mobility for different substrates spans a range of values varying over nearly two orders of magnitude. In the absence of dynamic screening or antiscreening, the SPP-limited mobility for HfO_2 is only around 1000 cm^{2}V^{-1}s^{-1} at $n = 10^{12}$ cm^{-2}, more than an order of magnitude smaller than the corresponding IPP-limited mobility, because of its low phonon frequencies. This result is also clearly inconsistent with experimental observations, since significantly higher mobility values have been reported for HfO_2-covered graphene.[39,47] The drastic reduction of the computed mobility suggests that screening is very important for the determination of scattering rates in a coupled plasmon-phonon system with low frequency modes. In contrast, h-BN gives an SPP-limited mobility of ~110 000 cm^{2}V^{-1}s^{-1} at $n = 0.3 \times 10^{12}$ cm^{-2}, which is still close to the IPP-limited mobility, indicating that its high frequency modes are relatively unaffected by screening. The maximum SPP-limited mobility for Al_2O_3 is around 8400 cm^{2}V^{-1}s^{-1} at $n = 0.3 \times 10^{12}$ cm^{-2}, which is much smaller than the 19 000 cm^{2}V^{-1}s^{-1} extracted by Jandhyala and coworkers[48] who used Al_2O_3 for their top gate dielectric. This disagreement reinforces the necessity of including dynamic screening effects. Furthermore, the carrier density dependence of SPP-limited mobility is different from that of IPP-limited mobility. The IPP-limited μ_{RP} increases rapidly with carrier density because dynamic screening becomes stronger at higher n, an effect that is not found in SPP-limited mobility. In contrast, SPP-limited μ_{RP} decreases monotonically with increasing n.

Our results suggest that HfO_2 remains a promising candidate material for integration with graphene since its high static permittivity can reduce the effect of charged impurities[22] while its IPP scattering rates are relatively low when the carrier density is high. Although its surface excitations are low-frequency, which results in high Bose-Einstein occupation, this is offset by its relatively strong dynamic screening at higher carrier densities. Thus IPP scattering does not represent a problem for its integration with graphene field-effect transistors. As expected, h-BN is also a good dielectric material since its high phonon frequencies imply a low Bose-Einstein occupation factor. Furthermore, its smooth interface results in a smaller interface charge density and is less likely to induce mobility-limiting ripples in graphene.

3. *Temperature dependence*

Remote phonon scattering exhibits a strong temperature dependence—stronger than for ionized impurity scattering—because the Bose-Einstein occupation of the remote phonons decreases with lower temperature. This change in the distribution of the remote phonons (IPP or SPP) necessarily implies that the electronic transport character of the SLG must change with temperature. At lower temperatures, scattering with the remote phonons decreases, resulting in a higher remote phonon-limited carrier mobility. The dependence of the change in mobility with temperature is related to the dispersion of the remote phonons and their coupling to the graphene electrons. By measuring the dependence of the mobility or conductivity with respect to temperature, it is possible to determine the dominant scattering mechanisms in the supported graphene. Given that our model of electron-IPP scattering differs from the more common electron-SPP scattering model, comparing the temperature dependence of the substrate-limited mobility can enable us to distinguish between the two models.

The mobility of supported graphene over the temperature range of 100 to 500 K for the different substrates is computed at carrier densities of $n = 10^{12}$ and 10^{13} cm^{-2}. For the purpose of comparison, we perform the calculation for the case with screening (IPP) and without screening (SPP). The results ($1/\mu_{RP}$ versus T) are shown in Figs. 7(a) and 7(b). In Fig. 7(a), we plot the IPP- and SPP-limited inverse mobility at $n = 10^{12}$ cm^{-2}. As expected, the substrate-limited mobility decreases with rising temperatures for both the screened and unscreened cases. From the plots, we observe that there exists an "activation" temperature for each substrate at which the inverse mobility increases precipitously. For SiO_2, that temperature is around 200 K in the screened case and around 120 K in the unscreened case. This difference is striking and may be used to distinguish the IPP model from the SPP model at low carrier densities. In all four substrates, the slope of $1/\mu_{RP}$ with respect to T is also steeper in the IPP-limited case than in the SPP-limited case. In Fig. 7(b), we plot again the IPP- and SPP-limited inverse mobility but at a much higher carrier density of $n = 10^{13}$ cm^{-2}. The IPP-limited $1/\mu_{RP}$ is about

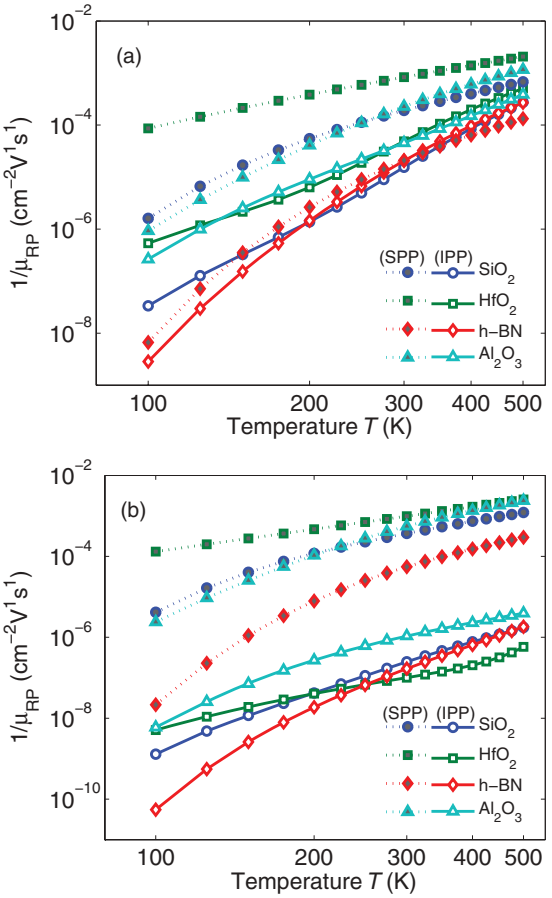

FIG. 7. (Color online) Inverse SPP and IPP-limited mobility versus temperature at (a) $n = 10^{12}$ cm^{-2} and (b) $n = 10^{13}$ cm^{-2} for SiO_2, HfO_2, h-BN, and Al_2O_3. $1/\mu_{RP}$ is strongly temperature dependent at low carrier densities only.

three orders of magnitude smaller than the SPP-limited $1/\mu_{RP}$ from 100 to 500 K. At high carrier densities, IPP scattering is insignificant and any changes in total mobility with respect to temperature cannot be attributed to IPP scattering.

The results in Fig. 7 suggest that if IPP modes are the surface excitations that limit carrier transport in SiO_2-supported graphene at room temperature, then the mobility would have a significant increase at around 200 K for $n = 10^{12}$ cm^{-2}. However, this IPP temperature dependence disappears at much higher carrier densities (approximately $n = 10^{13}$ cm^{-2}) because the IPP coupling to electrons becomes so weak that it no longer contributes significantly to carrier scattering. The results in Fig. 7 also shows that the μ_{RP} increases monotonically with n. This should be contrasted with the result of Fratini and Guinea[19] who found that μ_{RP} *decreases* as $\sim 1/\sqrt{n}$ at room temperature. This is because the coupling coefficient $\lim_{Q\to\infty} M_{\mathbf{Q}}$, which is proportional to the matrix element, scales as $1/\sqrt{Q}$ in the SPP model with static screening. In Fig. 3(a), $\lim_{Q\to\infty} M_{\mathbf{Q}}$ of branches I and II scales as Q^{α}, where $\alpha \sim -2$. In other words, the coupling coefficient vanishes more rapidly with Q in the IPP model than the SPP model. Our μ_{RP} results parallel those in Ref. 49 in which the remote phonon-limited mobility increases with the carrier density in a two-dimensional electron gas system in the Si inversion layer with high-κ insulators.

This strong density dependence is the result of dielectric screening and we should emphasize once more the strong effect of screening in graphene as compared, for example, to the case of semiconductors. As we have already mentioned, this is due to the absence of a gap and to the fact that free carrier have no inertia—they do not have to be accelerated to reach high velocities—so that they can respond efficiently to high-frequency perturbations and so the effect of screening is much stronger. The presence of Kohn anomalies in the phonon dispersion[27,28] and the breakdown of the Born-Oppenheimer approximation[28] are spectacular manifestations of this fact, as we have mentioned before. For these reasons, dynamic screening, known to be less effective than static screening in semiconductors,[50] is significantly more effective in graphene. A simple argument can explain, for example, the dramatic difference shown in Fig. 6 between the vastly different values of the screened and unscreened mobility, 3 000 versus 10^6 cm^{2}V^{-1}s^{-1}, in SiO_2-supported graphene at a sheet density of 5×10^{12} cm^{-2}. The ratio r between the screened and unscreened values can be estimated from Ref. 9 as $r = (\epsilon_{gr} + \epsilon_0)(\epsilon_{gr} + \epsilon_{ox})/[1 + \epsilon_0(1 + \epsilon_{ox})]$, where $\epsilon_{gr} = 1 + e^2 G_Q(d,d)\Pi(Q,\omega)$ is the dielectric "constant" of graphene evaluated at the frequency ω of the dominant mode and at the magnitude Q of the wave vector transfer for those collisions involving carriers which contribute mostly to the mobility (that is, those near the Fermi surface). Estimating Q as the Fermi wave vector $Q_F = E_F/(\hbar v_F)$, evaluating the Fermi energy E_F and the polarizability $\Pi(Q,\omega)$ from Eq. (43) at $n = 5 \times 10^{12}$ cm^{-2} for the low-frequency "bare" optical mode SO_1, $\omega = 59$ meV, we obtain $\epsilon_{gr} \approx 80$ and so $r \approx 680$. Thus an unscreened mobility of 3000 cm^{2}V^{-1}s^{-1} is boosted by screening to a value of about of 2×10^6 cm^{2}V^{-1}s^{-1}. The calculated screened value shown in Fig. 6, $\mu_{RP} \approx 10^6$ cm^{2}V^{-1}s^{-1}, is indeed very close to this simple estimate which ignores the effect of the high-energy mode SO_2.

In supported SLG, the carrier mobility is limited by three scattering mechanisms: long-range charged impurity, short-range defect, and remote phonon scattering.[51] The intrinsic phonon scattering processes in graphene can be effectively neglected. Of the three scattering mechanisms, only remote phonon scattering is strongly temperature dependent. The IPP model suggests that remote phonon scattering diminishes with increasing carrier density. Thus the experimental consequence is that the temperature dependence of the mobility in supported-SLG should weaken at higher carrier densities. On the other hand, the SPP model predicts that the temperature dependence of the mobility should increase at higher carrier densities.[19] This difference in the temperature dependence of the total mobility between the two models should be easily discriminable in experiments.

4. *Disordered graphene*

We discuss qualitatively the interfacial plasmon-phonon phenomenon in disordered graphene. It is well known that graphene grown by chemical vapor deposition (CVD)[52] is generally polycrystalline and contains a high density of defects. In supported graphene, charged impurities from the substrate

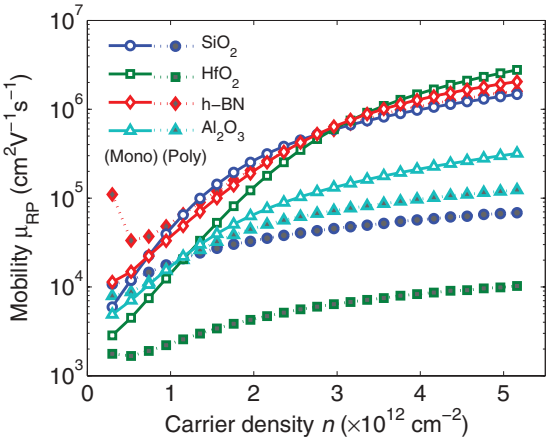

FIG. 8. (Color online) Remote phonon-limited mobility in perfect monocrystalline (clear symbols) and defective polycrystalline graphene (solid symbols) for SiO_2, HfO_2, h-BN, and Al_2O_3. As carrier density increases, μ_{RP} also increases. The use of long-wavelength SPP modes leads to a significant decrease in μ_{RP} in polycrystalline graphene.

and other defects scatter graphene carriers. These defects can affect the dynamics of plasmons in graphene, which may in turn affect the hybridization between the plasmon and the SPP modes. At short wavelengths, the plasmon lifetime rapidly decreases as a result of Landau damping which results in the decay of the plasmons into single-particle excitations. At long wavelengths, the plasmon lifetime can be affected by defects in the graphene. As far as we know, there is no theory of graphene plasmon damping from defects. However, it has been pointed out that long-wavelength plasmons in polycrystalline metal undergo anomalously large damping due to scattering with structural defects.[53] If this is also true in polycrystalline or defective graphene, then it implies that the long-wavelength surface excitation in supported graphene are SPP, not IPP, modes.

To model phenomenologically this damping of long-wavelength plasmon modes in polycrystalline graphene with defects, we set another cutoff wave vector Q_d below which the surface excitations are SPP and not IPP modes. Q_d is possibly related to the length scale λ of the inhomogeneities or defects in graphene. As a guess, we choose $\lambda = 6$ nm, which is a typical autocorrelation length of "puddles" in neutral supported graphene,[54] and set $Q_d = 1/\lambda$. Hence, in our model, for $Q \geqslant Q_c$ and $Q \leqslant Q_d$, the surface excitations are SPP modes, while for $Q_d < Q < Q_c$, they are IPP modes. We compute the remote phonon-limited mobility at 300 K and plot the results in Fig. 8. We find that the long-wavelength SPP dramatically alters the carrier dependence of μ_{RP} in SiO_2, HfO_2, and Al_2O_3. In perfect monocrystalline graphene, μ_{RP} reaches $\sim 2 \times 10^6$ $cm^2V^{-1}s^{-1}$ in HfO_2 and SiO_2 at $n = 5 \times 10^{12}$ cm^{-2}. On the other hand, in polycrystalline graphene with defects, it drops to the range of 10^4 to 10^5 $cm^2V^{-1}s^{-1}$. For h-BN, μ_{RP} is quite relatively unaffected by the long-wavelength SPP modes except at low carrier densities ($n < 0.5 \times 10^6$ cm^{-2}).

This change in remote phonon-limited mobility highlights the possible role of defects in the surface excitations of supported graphene. We emphasize that our treatment is purely phenomenological and a more rigorous treatment of plasmon damping is needed in order to obtain a more quantitatively accurate model. Nevertheless, it highlights the relationship between dynamic screening and plasmons. In highly defective graphene, the surface excitations may be unscreened SPPs rather than IPPs because of plasmon damping. This should be taken into account when interpreting electronic transport experimental data of exfoliated and CVD-grown graphene. Our results also suggest that the h-BN-supported graphene is significantly less sensitive than HfO_2 and SiO_2-supported graphene to the increase in remote phonon scattering induced by defects in the graphene.

IV. CONCLUSION

We have studied coupled interfacial plasmon-phonon excitations in supported graphene. The coupling between the pure graphene plasmon and the surface polar phonon modes of the substrates results in the formation of the IPP modes, and this coupling is responsible for the screening and antiscreening of the IPP modes. Accounting for these modes, we calculate the room temperature scattering rates and substate-limited mobility for SiO_2, HfO_2, h-BN, and Al_2O_3 at different carrier densities. The results suggest that, despite being a high-κ oxide with low frequency modes, HfO_2 exhibits a substrate-limited mobility comparable to that of h-BN at high carrier densities. We attribute this to the dynamic screening of the HfO_2 low-frequency modes. The disadvantage of the higher Bose-Einstein occupation of these low-frequency modes is offset by the stronger dynamic screening which suppresses the electron-IPP coupling. Our study also indicates that the contribution to scattering by high-frequency substrate phonon modes cannot be neglected because of they are less weakly screened by the graphene plasmons. The temperature dependence of the remote phonon-limited mobility is also calculated within out theory. Its change with temperature is different at low and high carrier densities. We find that in the IPP model, the temperature dependence of the mobility diminishes with increasing carrier density only, in direct contrast to the predictions of the more commonly used SPP models. The implications of the damping of long-wavelength plasmons have also been studied. We find that the it leads to a substantial reduction in the remote phonon-limited mobility in SiO_2, HfO_2, and Al_2O_3. Our results also suggest that the h-BN-supported graphene is less sensitive than HfO_2 and SiO_2-supported graphene to the quality of the graphene.

ACKNOWLEDGMENTS

We gratefully acknowledge the support provided by Texas Instruments, the Semiconductor Research Corporation (SRC), the Microelectronics Advanced Research Corporation (MARCO), the Focus Center Research Project (FCRP) for Materials, Structures and Devices (MSD), and Samsung Electronics Ltd. We also like to thank David K. Ferry (Arizona State University), Eric Pop (University of Illinois), and Andrey Serov (University of Illinois) for engaging in valuable technical discussions.

*zhunyong.ong@utdallas.edu
†max.fischetti@utdallas.edu

[1] K. Novoselov, A. Geim, S. Morozov, D. Jiang, M. Grigorieva, S. Dubonos, and A. Firsov, Nature (London) **438**, 197 (2005).
[2] E. H. Hwang, S. Adam, and S. Das Sarma, Phys. Rev. Lett. **98**, 186806 (2007).
[3] A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, and C. Lau, Nano Lett. **8**, 902 (2008).
[4] K. Bolotin, K. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H. Stormer, Solid State Commun. **146**, 351 (2008).
[5] Zhun-Yong Ong and Massimo V. Fischetti, Phys. Rev. B **86**, 121409(R) (2012).
[6] M. Lemme, T. Echtermeyer, M. Baus, B. Szafranek, J. Bolten, M. Schmidt, T. Wahlbrink, and H. Kurz, Solid-State Electron. **52**, 514 (2008).
[7] J. Moon *et al.*, IEEE Electron Device Lett. **31**, 260 (2010).
[8] J. Pezoldt, C. Hummel, A. Hanisch, I. Hotovy, M. Kadlecikova, and F. Schwierz, Phys. Status Solidi C **7**, 390 (2010).
[9] M. V. Fischetti, D. A. Neumayer, and E. A. Cartier, J. Appl. Phys. **90**, 4587 (2001).
[10] R. Fuchs and K. L. Kliewer, Phys. Rev. **140**, A2076 (1965).
[11] K. Hess and P. Vogl, Solid State Commun. **30**, 797 (1979).
[12] T. O'Regan and M. Fischetti, J. Comput. Electron. **6**, 81 (2007).
[13] K. Xiu, in *Simulation of Semiconductor Processes and Devices (SISPAD), 2011 International Conference* (IEEE, New York, 2011), pp. 35–38.
[14] C. Dean *et al.*, Nat. Nanotechnology **5**, 722 (2010).
[15] J. Chen, C. Jang, S. Xiao, M. Ishigami, and M. Fuhrer, Nat. Nanotechnology **3**, 206 (2008).
[16] V. E. Dorgan, M.-H. Bae, and E. Pop, Appl. Phys. Lett. **97**, 082112 (2010).
[17] J. Robinson *et al.*, Nano Lett. **9**, 2873 (2009).
[18] P. Sutter, Nat. Mater. **8**, 171 (2009).
[19] S. Fratini and F. Guinea, Phys. Rev. B **77**, 195415 (2008).
[20] S. Rotkin, V. Perebeinos, A. Petrov, and P. Avouris, Nano Lett. **9**, 1850 (2009).
[21] V. Perebeinos, S. Rotkin, A. Petrov, and P. Avouris, Nano Lett. **9**, 312 (2008).
[22] A. Konar, T. Fang, and D. Jena, Phys. Rev. B **82**, 115452 (2010).
[23] J. K. Viljas and T. T. Heikkilä, Phys. Rev. B **81**, 245404 (2010).
[24] X. Li, E. Barry, J. Zavada, M. Nardelli, and K. Kim, Appl. Phys. Lett. **97**, 232105 (2010).
[25] S. Adam, E. Hwang, E. Rossi, and S. Das Sarma, Solid State Commun. **149**, 1072 (2009).
[26] S. Adam, E. H. Hwang, V. M. Galitski, and S. Das Sarma, Proc. Natl. Acad. Sci. USA **104**, 18392 (2007).
[27] M. Lazzeri and F. Mauri, Phys. Rev. Lett. **97**, 266407 (2006).
[28] S. Pisana, M. Lazzeri, C. Casiraghi, K. Novoselov, A. Geim, A. Ferrari, and F. Mauri, Nat. Mater. **6**, 198 (2007).
[29] B. K. Ridley, *Quantum Processes in Semiconductors* (Oxford University Press, Oxford, 1999).
[30] V. Perebeinos and P. Avouris, Phys. Rev. B **81**, 195442 (2010).
[31] Y. Liu, R. F. Willis, K. V. Emtsev, and T. Seyller, Phys. Rev. B **78**, 201403 (2008).
[32] Y. Liu and R. F. Willis, Phys. Rev. B **81**, 081406 (2010).
[33] Z. Fei *et al.*, Nano Lett. **11**, 4701 (2011).
[34] R. J. Koch, T. Seyller, and J. A. Schaefer, Phys. Rev. B **82**, 201413 (2010).
[35] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1999).
[36] M. E. Kim, A. Das, and S. D. Senturia, Phys. Rev. B **18**, 6890 (1978).
[37] E. H. Hwang and S. Das Sarma, Phys. Rev. B **75**, 205418 (2007).
[38] B. Wunsch, T. Stauber, F. Sols, and F. Guinea, New J. Phys. **8**, 318 (2006).
[39] K. Zou, X. Hong, D. Keefer, and J. Zhu, Phys. Rev. Lett. **105**, 126601 (2010).
[40] E. H. Hwang, R. Sensarma, and S. Das Sarma, Phys. Rev. B **82**, 195406 (2010).
[41] R. S. Shishir and D. K. Ferry, J. Phys.: Condens. Matter **21**, 232204 (2009).
[42] Y. W. Tan, Y. Zhang, K. Bolotin, Y. Zhao, S. Adam, E. H. Hwang, S. Das Sarma, H. L. Stormer, and P. Kim, Phys. Rev. Lett. **99**, 246803 (2007).
[43] C. Jang, S. Adam, J. H. Chen, E. D. Williams, S. Das Sarma, and M. S. Fuhrer, Phys. Rev. Lett. **101**, 146805 (2008).
[44] L. A. Ponomarenko, R. Yang, T. M. Mohiuddin, M. I. Katsnelson, K. S. Novoselov, S. V. Morozov, A. A. Zhukov, F. Schedin, E. W. Hill, and A. K. Geim, Phys. Rev. Lett. **102**, 206603 (2009).
[45] N. Garces, V. Wheeler, J. Hite, G. Jernigan, J. Tedesco, N. Nepal, C. Eddy, and D. Gaskill, J. Appl. Phys. **109**, 124304 (2011).
[46] G. Gilat and L. Raubenheimer, Phys. Rev. **144**, 390 (1966).
[47] B. Fallahazad, S. Kim, L. Colombo, and E. Tutuc, Appl. Phys. Lett. **97**, 123105 (2010).
[48] S. Jandhyala, G. Mordi, B. Lee, G. Lee, C. Floresca, P.-R. Cha, J. Ahn, R. M. Wallace, Y. J. Chabal, M. J. Kim, L. Colombo, K. Cho, and J. Kim, ACS Nano **6**, 2722 (2012).
[49] Z. Ren, M. V. Fischetti, E. P. Gusev, E. A. Cartier, and M. Chudzik, in *Electron Devices Meeting, 2003. IEDM'03 Technical Digest* (IEEE, New York, 2003), pp. 33.2.1–33.2.4.
[50] S. Das Sarma, J. K. Jain, and R. Jalabert, Phys. Rev. B **37**, 6290 (1988).
[51] W. Zhu, V. Perebeinos, M. Freitag, and P. Avouris, Phys. Rev. B **80**, 235402 (2009).
[52] X. Li *et al.*, Science **324**, 1312 (2009).
[53] V. Krishan and R. H. Ritchie, Phys. Rev. Lett. **24**, 1117 (1970).
[54] S. Adam, S. Jung, N. N. Klimov, N. B. Zhitenev, J. A. Stroscio, and M. D. Stiles, Phys. Rev. B **84**, 235421 (2011).