Calculation of the electron mobility in III-V inversion layers with high-κ dielectrics

T. P. O'Regan' , M. V. Fischetti, B. Sorée, S. Jin' , W. Magnus, and M. Meuris

Citation: Journal of Applied Physics **108**, 103705 (2010); doi: 10.1063/1.3500553
View online: http://dx.doi.org/10.1063/1.3500553
View Table of Contents: http://aip.scitation.org/toc/jap/108/10
Published by the American Institute of Physics

Calculation of the electron mobility in III-V inversion layers with high-κ dielectrics

T. P. O'Regan,[1,a)] M. V. Fischetti,[1] B. Sorée,[2] S. Jin,[1,b)] W. Magnus,[2] and M. Meuris[2]
[1]*Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, Massachusetts 01003-9292, USA*
[2]*Interuniversity Microelectronics Center (IMEC vzw), Kapeldreef 75, B-3001 Leuven, Belgium*

(Received 19 July 2010; accepted 13 September 2010; published online 18 November 2010)

We calculate the electron mobility for a metal-oxide-semiconductor system with a metallic gate, high-κ dielectric layer, and III-V substrate, including scattering with longitudinal-optical (LO) polar-phonons of the III-V substrate and with the interfacial excitations resulting from the coupling of insulator and substrate optical modes among themselves and with substrate plasmons. In treating scattering with the substrate LO-modes, multisubband dynamic screening is included and compared to the dielectric screening in the static limit and with the commonly used screening model obtained by defining an effective screening wave vector. The electron mobility components limited by substrate LO phonons and interfacial modes are calculated for $In_{0.53}Ga_{0.47}As$ and GaAs substrates with SiO_2 and HfO_2 gate dielectrics. The mobility components limited by the LO-modes and interfacial phonons are also investigated as a function of temperature. Scattering with surface roughness, fixed interface charge, and nonpolar-phonons is also included to judge the relative impact of each scattering mechanism in the total mobility for $In_{0.53}Ga_{0.47}As$ with HfO_2 gate dielectric. We show that InGaAs is affected by interfacial-phonon scattering to an extent larger than Si, lowering the expected performance, but probably not enough to question the technological relevance of InGaAs. © *2010 American Institute of Physics*. [doi:10.1063/1.3500553]

I. INTRODUCTION

The continuation of Moore's law past the 22 nm node challenges the conventional scaling methods practiced over the last few decades.[1] Many options exist to continue scaling within Si technology: different device geometries [Tri-gates,[2] Fin-field-effect-transistors (FETs),[3] nanowires,[4] and multibridge FETs (Ref. 5)], fully depleted SOI,[6] strained substrates,[7] high-κ dielectrics, and the use of different wafer and device crystallographic orientations.[8] In addition, side-stepping Si technology in part or altogether, III-V compound semiconductors are considered as possible replacement for the conventional n-type Si metal-oxide-semiconductor FETs (nMOSFETs) because of their higher electron mobility (which suggests increased device performance). InGaAs single-gate (SG) MOSFETs with high-κ gate dielectrics have been fabricated[9] and show a peak mobility of 2200 $cm^2/V\,s$ compared to 500 $cm^2/V\,s$ for strained Si (Ref. 7) and 1100 $cm^2/V\,s$ for Ge.[10,11] The mobility enhancement may lead to overall device performance and faster switching speeds—making III-V's and especially InGaAs, a possible solution to continue Moore's law.[12]

While the use of electron mobility as a gauge of the performance of sub-22 nm devices is questionable,[13,14] its widespread use, its value as an indicator of the strength of the relevant scattering processes, and its importance in determining the behavior of the device in the linear regime render its theoretical study a worthwhile effort. We, therefore, focus on the calculation of the electron mobility in inversion layers of III-V substrates, while at the same time recognizing that the mobility is only one of the many indicators of device performance.

In this work, we account for scattering with the longitudinal-optical (LO) phonons of the substrate, and with the interfacial-optical modes (labeled "SO" for "surface optical," in the following to be consistent with the historical notation[15,16]) resulting from the mixing and coupling of the transverse-optical (TO) phonons of the polar insulating layer, of the substrate, and of the plasmons of the two-dimensional electron gas (2DED) constituting the inversion layer. Scattering with substrate LO phonons is treated by accounting for dynamic multisubband screening within the random-phase approximation—with the dynamic screening parameter accounting for the inelastic process with energy transfer determined by the energy of the bulk LO-mode. Static screening is compared to dynamic screening and the results deviate by less than 5%. Previous work has either ignored screening,[17,18] or included screening through the use of a reciprocal screening length.[19] An effective screening wave vector (reciprocal screening length) is defined in this work and compared to the dynamic and static screening approximations. Using an effective screening wave vector is found to only qualitatively reproduce the dynamic screening result.

The LO phonons associated with III-V substrates couple to interface plasmons and insulator-optical modes and thus, LO phonons themselves, induce coupled-optical interfacial-phonon modes at the dielectric/III-V interface. These modes, as well as those induced by the insulator LO-modes, generate

[a)]Current address: Tyndall National Institute, University College Cork, Ireland. Electronic mail: terrance.oregan@tyndall.ie.
[b)]Current address: Synopsys Inc., 700 East Middlefield Road, Mountain View, CA 94043.

a scattering potential in the substrate and reduce the electron mobility of the 2DEG. In this work we extend the existing model of Ref. 16 by accounting also for the substrate mode in obtaining the spectrum of the coupled interfacial TO-plasmons excitations and calculate the electron mobility for an ideal metallic gate with high-κ gate dielectric.

We shall consider GaAs and $In_{0.53}Ga_{0.47}As$ throughout this paper since the latter semiconductor, similarly to $In_{0.52}Al_{0.48}As$, is lattice matched to InP substrates, a combination of materials that have recently been considered as barrier layers in III-V nMOSFET design.[20] The Schrödinger and Poisson equations are solved self-consistently and the mobility is calculated using the Kubo–Greenwood formula with nonparabolic corrections. In addition to scattering with LO-phonons and SO-phonons, we also include scattering with fixed interface charge, bulk phonons (intra- and inter-valley phonons) and surface roughness (SR).

This paper is organized as follows. In Sec. II, we derive the LO-phonon momentum relaxation rate and describe dynamic screening, static screening, and a simpler effective screening method. Section III deals with SO-phonon scattering, the previous theory being extended so to include the additional SO-mode originating from the substrate TO-phonon. Section IV briefly considers Coulomb scattering with fixed interface charge. Nonparabolic corrections and the mobility calculation are described in Sec. V. Section VI presents results obtained for the LO-phonon-limited mobility as a function of electron sheet density, n_s, and temperature as well as a comparison of various approximations which may be employed to screen the potential of the LO phonons. The SO-phonon limited mobility is discussed as function of n_s and temperature. The various component-limited mobilities are shown and compared to experimental data. Finally, we draw conclusions in Sec. VII.

II. LO PHONONS

The LO-phonon scattering potential is given by Fröhlich's well-known expression[21]

$$V_q^{\mathrm{LO}}(r) = e\left[\frac{\hbar\omega_{\mathrm{LO}}}{2q^2}\left(\frac{1}{\epsilon_s^\infty}-\frac{1}{\epsilon_s^0}\right)\right]^{1/2} e^{i\mathbf{q}\cdot\mathbf{r}}, \tag{1}$$

where ϵ_s^∞ and ϵ_s^0 are the optical and static permittivity of the substrate, respectively, e the electron charge, and $\hbar\omega_{\mathrm{LO}}$ the energy of the substrate LO-phonon-mode ($\approx$35 meV for GaAs and $\approx$32.5 meV for $In_{0.53}Ga_{0.47}As$).[22] Here, $\mathbf{q}$ is the three-dimensional wave vector transfer $\mathbf{q}=\mathbf{k}'-\mathbf{k}$, where $\mathbf{k}$ and $\mathbf{k}'$ are the initial and final electron wave vectors, respectively. From here on, all capital letters denoting vectors (and magnitudes) represent 2D in-plane vectors. For example, aligning the quantization axis along the z-axis, we have $\mathbf{q}=(\mathbf{Q},q_z)$ with $Q^2=q_x^2+q_y^2$. Now we can write the wave function of an electron in subband μ with wave vector $\mathbf{K}$ as

$$\psi_{\mu\mathbf{K}}(\mathbf{r}) = \zeta_\mu(z)\frac{1}{\sqrt{2\pi}}e^{i\mathbf{K}\cdot\mathbf{R}}, \tag{2}$$

where $\zeta_\mu(z)$ is the solution of the Schrödinger equation (in this work we use the approximation that $\zeta_\mu(z)$ vanishes at the dielectric/substrate boundary). The squared magnitude of the matrix element between an initial state in subband μ and a final state in subband ν, is

$$|\langle\psi_{\mu\mathbf{K}}|V_q^{\mathrm{LO}}(r)|\psi_{\nu\mathbf{K}'}\rangle|^2 = \frac{e^2\hbar\omega_{\mathrm{LO}}}{2q^2}\left(\frac{1}{\epsilon_s^\infty}-\frac{1}{\epsilon_s^0}\right)\times\begin{Bmatrix} n_{\mathrm{LO}} \\ 1+n_{\mathrm{LO}}\end{Bmatrix} \times|F_{\nu\mu}(q_z)|^2\delta(\mathbf{K}-\mathbf{K}'-\mathbf{Q}), \tag{3}$$

where n_{LO} is the Bose occupation of the LO phonons. The upper value in the curly bracket should be considered in case of absorption processes, the lower value in case of emission processes. The form factor $F_{\nu\mu}(q_z)$ in Eq. (3) is given by

$$F_{\nu\mu}(q_z) = \int_0^\infty dz\zeta_\nu(z)e^{iq_z z}\zeta_\mu(z), \tag{4}$$

which accounts for the geometry of the device through the subband wave functions. Using Fermi's golden rule, the LO-phonon scattering rate from an initial subband μ to a final subband ν is

$$\frac{1}{\tau_{\mu\nu}(\mathbf{K})} = \frac{2\pi}{\hbar}\int\frac{d\mathbf{q}}{(2\pi)^3}\int\frac{d\mathbf{K}'}{(2\pi)^2}|\langle\psi_{\nu\mathbf{K}'}|V_q^{\mathrm{LO}}(r)|\psi_{\mu\mathbf{K}}\rangle|^2 \times\delta[E_\nu(\mathbf{K}')-E_\mu(\mathbf{K})\pm\hbar\omega_{\mathrm{LO}}]. \tag{5}$$

Now using polar coordinates with 'radial' variable E' instead of K', we obtain the general expression for the scattering rate

$$\frac{1}{\tau_{\mu\nu}(\mathbf{K})} = \frac{e^2\omega_{\mathrm{LO}}}{4\pi\hbar^2}\left(\frac{1}{\epsilon_s^\infty}-\frac{1}{\epsilon_s^0}\right)\begin{Bmatrix} n_{\mathrm{LO}} \\ 1+n_{\mathrm{LO}}\end{Bmatrix} \times\int_0^{2\pi}d\beta'\left[\frac{\cos^2\beta'}{m_x}+\frac{\sin^2\beta'}{m_y}\right]^{-1}\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}|\varphi_{Q,\mu\nu}^{(\mathrm{ex})}|^2, \tag{6}$$

having lumped the dependence on Q or q_z in the integrand into an "external" (unscreened) matrix element

$$\varphi_{Q,\mu\nu}^{(\mathrm{ex})} = \frac{F_{\nu\mu}(q_z)}{\sqrt{q_z^2+Q^2}}. \tag{7}$$

where $\beta(\beta')$ is the angle between $\mathbf{K}(\mathbf{K}')$ and the channel direction (aligned along the k_x direction), and m_x and m_y are the electron effective masses along the x and y axes, respectively. For the isotropic Γ-valley, the integration over q_z in Eq. (6) can be performed to obtain

$$\frac{1}{\tau_{\mu\nu}(\mathbf{K})} = \frac{e^2 m_d\omega_{\mathrm{LO}}}{8\pi\hbar^2}\left(\frac{1}{\epsilon_s^\infty}-\frac{1}{\epsilon_s^0}\right)\begin{Bmatrix} n_{\mathrm{LO}} \\ 1+n_{\mathrm{LO}}\end{Bmatrix} \times\int_0^{2\pi}d\beta'\frac{H_{\mu\nu}[Q(\beta')]}{Q(\beta')}, \tag{8}$$

with

$$Q^2(\beta') = K+K'^2-2KK'^2\cos\beta', \tag{9}$$

with the "form factor"[23]

$$H_{\mu\nu}[Q(\beta')] = \int_0^\infty dz\int_0^\infty dz'\zeta_\nu(z)\zeta_\nu(z') \times\zeta_\mu(z')\zeta_\mu(z)e^{-Q|z-z'|}. \tag{10}$$

The momentum relaxation rates needed to evaluate the electron mobility can be obtained in a similar way, inserting inside the β' integral the additional momentum-transfer factor

$$\frac{1-f(E')}{1-f(E)}\left(1-\frac{\tau^{\mathrm{tot}}_{x,\nu\mathbf{K}'}v_{x,\nu\mathbf{K}'}}{\tau^{\mathrm{tot}}_{x,\mu\mathbf{K}}v_{x,\nu\mathbf{K}}}\right), \tag{11}$$

where $\tau^{x,\mathrm{tot}}_{\mu\mathbf{K}}$ is the total momentum relaxation time including all scattering mechanisms, $v^{x}_{\nu\mathbf{K}}$ is the velocity along the x-direction, $f(E)$ is the Fermi–Dirac distribution function, and $E'=E\pm\hbar\omega_{\mathrm{LO}}$ is the final energy after absorption (+) or emission (−) of the bulk LO phonon, $\hbar\omega_{\mathrm{LO}}$. For elastic or isotropic scattering, the term in the square brackets of Eq. (11) simplifies to $1-[K'/K]\cos\phi$, where $\phi=\beta-\beta'$. In the case of LO-phonon scattering, this simplification is not strictly correct. In principle, a self-consistent iterative procedure can be employed to estimate the relaxation rates in the more general case of anisotropic and inelastic processes.[24,25] Here, however, we simplify the term inside the brackets of Eq. (11) $\rightarrow[1-\cos\phi]$, which corresponds to the zeroth order correction in the iteration scheme.

A. Multisubband screening

In the previous section we considered Fröhlich scattering with the unscreened potential of the LO phonons. Here, we follow Ref. 26, giving a summary of the general method to treat dielectric screening accounting for dynamic effects and including the response of electrons in all subbands. This will require the use of a complex dielectric function (or, equivalently, of a complex dynamic screening parameter).

Following Ref. 26, we ignore intersubband polarization—thus considering only the longitudinal dielectric response—so that the net potential for intrasubband scattering can be found by inverting the linear problem

$$\sum_{\lambda}\epsilon_{\mu\mu,\lambda\lambda}(Q,\omega)\varphi_{Q,\lambda\lambda}=\varphi^{(\mathrm{ex})}_{Q,\mu\mu}, \tag{12}$$

where $\varphi^{(\mathrm{ex})}_{Q,\mu\mu}$ is the matrix element of the unscreened or 'external' potential, $\varphi_{Q,\lambda\lambda}$ is the matrix element of the screened potential, and the dielectric matrix is

$$\epsilon_{\mu\nu,\lambda\lambda}(Q,\omega)=\delta_{\mu\nu}\delta_{\lambda\lambda}+\frac{\beta_{\lambda\lambda}(Q,\omega)}{Q}\Omega_{Q,\mu\nu,\lambda\lambda}, \tag{13}$$

where the real part, imaginary part, and magnitude of the intraground-state matrix element of the dielectric matrix, $\epsilon_{11,11}$, is shown in Fig. 1. The form factor expressing the multisubband matrix elements of the Green's function is

$$\Omega_{Q,\mu\nu,\lambda\lambda}=2\int dz\int dz' QG_Q(z,z')\zeta_{\mu(z)}\zeta_{\nu(z)}\zeta_{\lambda(z')}\zeta_{\lambda(z')}, \tag{14}$$

where $G_Q(z,z')$ is the Green's function. Using the intrasubband net matrix elements, the intersubband matrix elements are obtained directly[26]

$$\varphi_{Q,\mu\nu}=\varphi^{(\mathrm{ex})}_{Q,\mu\nu}-\sum_{\lambda}\epsilon_{\mu\nu,\lambda\lambda}\varphi_{Q,\lambda\lambda}. \tag{15}$$

This method speeds up the computation time and saves memory by ignoring the off-diagonal terms of the dielectric matrix which are usually small compared to the diagonal elements.

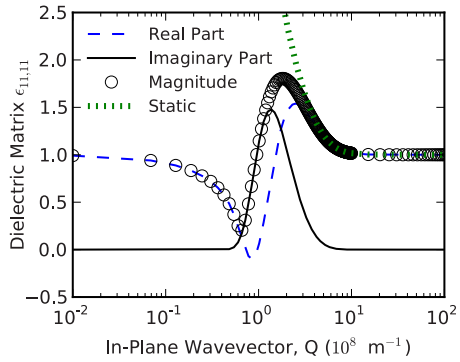

FIG. 1. (Color online) The real part, imaginary part, and magnitude of the intraground-state matrix element of the dielectric matrix, $\epsilon_{11,11}(Q,\omega_{\mathrm{LO}})$, for GaAs ($\hbar\omega=\hbar\omega_{\mathrm{LO}}\approx 35$ meV) and $n_s=2\times10^{12}$ cm^{-2}.

In Eq. (13), $\beta_{\mu\mu}(Q,\omega)$ is the 2D screening wave vector evaluated in the high temperature limit[26,27]

$$\mathrm{Re}[\beta_{\mu\mu}(Q,\omega)]=\beta_{\mathrm{DH}}\frac{\pi^{1/2}}{Ql_\mu}\left\{\Phi\left[\left(\frac{m_\mu}{2k_BT}\right)^{1/2}\left(\frac{\omega}{Q}+\frac{\hbar Q}{2m_\mu}\right)\right]-\Phi\left[\left(\frac{m_\mu}{2k_BT}\right)^{1/2}\left(\frac{\omega}{Q}-\frac{\hbar Q}{2m_\mu}\right)\right]\right\}, \tag{16}$$

and

$$\mathrm{Im}[\beta_{\mu\mu}(Q,\omega)]=\beta_{\mathrm{DH}}\frac{\pi\hbar\omega}{Ql_\mu k_BT}\exp\left(\frac{-m_\mu\omega^2}{2k_BTQ^2}-\frac{\hbar^2Q^2}{8m_\mu k_BT}\right)\times\frac{\sinh[\hbar\omega/(2k_BT)]}{\hbar\omega/(2k_BT)}, \tag{17}$$

where $\beta_{\mathrm{DH}}=e^2n_\mu/(2\epsilon_sk_BT)$ is the 2D Debye–Hückel limit of the static screening wave vector, $l_\mu=[2\pi\hbar^2/(m_\mu k_BT)]^{1/2}$ is the thermal wavelength of electrons in subband μ, and $\Phi(y)=2e^{-y^2}\int_o^y dte^{t^2}$ is the plasma dispersion function.[27] In principle, for highly degenerate systems, it is necessary to include degenerate statistics for the screening wave vector (as is done in Ref. 28 for static screening) starting from Maldegue's expression.[29] We find that the mobility changes less than 5% when comparing the nondegenerate dynamic screening approximation to the degenerate static screening case—so in this paper we use the nondegenerate limit for dynamic screening of the LO phonons, settting $\omega=\omega_{\mathrm{LO}}$ in Eqs. (16) and (17). When considering static perturbations (scattering with SR and fixed interface charge), we use the degenerate-screening wave vector summarized in Ref. 28.

The inclusion of dielectric screening adds the additional computational cost required to compute the screened potential (setting up and inverting the dielectric matrix) as well as introducing an additional integration step in the calculation of the scattering rate for the isotropic Γ-valley [i.e., Eq. (6) must be used instead of Eq. (8), and instead of tabulating $H_{\mu\nu}$ we must integrate over q_z "on the fly"]. A less computationally expensive, albeit approximate, method to account for dielectric screening is to define an effective screening wave vector

$$q_s = \sum_{\lambda} \mathrm{Re}(\beta_{Q,\omega_{LO};\lambda\lambda}) \tag{18}$$

and modify Eq. (7) as

$$\varphi^{(\mathrm{eff})}_{Q,\mu\nu} = \frac{F_{\nu\mu}(q_z)}{\sqrt{q_z^2 + Q^2 + q_s^2}}, \tag{19}$$

where q_s is also referred to as the reciprocal screening length.[19] The effective screening parameter, q_s, increases with sheet density through the β_{DH} term and the scattering potential is subsequently decreased—leading to stronger screening and an increased electron mobility at high sheet densities.

The general "rule of thumb" is that screening is strong and can be treated statically when the frequency of the perturbation (in our case, the frequency of the LO-modes) is much smaller than the plasma frequency, it is weak and can be ignored in the opposite case, while a full dynamic treatment is needed to treat the "in between" situations. When dealing with a 2DEG the situation is complicated by the fact that the plasma frequency is not constant, but 2D plasmons exhibit dispersion, so that one cannot give a simple rule and full calculations are required. When optical phonons and plasmons are coupled, the dispersion of the coupled modes must be computed for the particular geometry and material system of interest. In the simpler case of a single $Si-SiO_2$ interface the situation—together with plots showing how the effect of dielectric screening varies with frequency—has been discussed at length by in Refs. 26 and 48.

III. INTERFACE-OPTICAL PHONONS

Replacing SiO_2 with high-κ dielectrics, such as HfO_2, dramatically enhances the scattering with optical modes at the dielectric-substrate interface.[16,30,31] The electron mobility in Si and Ge with various high-κ dielectrics with polysilicon and metallic gates has been calculated before.[13,16,32–34] Here, we extend those calculations to two III-V substrates of interest: GaAs and InGaAs. Because of the polar nature of III-V materials, bulk optical modes are associated with a long-range dipole field. As it happens in the case of the insulator TO-modes, this field induces a polarization charge at the substrate dielectric interface. The resulting interfacial mode couples to other modes present (insulator SO modes, interfacial plasmons), thus increasing the complexity of the model required to extract the dispersion of these coupled modes.

In this work we consider the following physical system: infinitely thick metallic gate, insulator with a finite thickness, and a semiconductor substrate represented by a 2DEG. The SO-mode originating from the substrate is included in the dielectric function of the substrate as

$$\epsilon_s(Q,\omega) = \epsilon_s^{\infty}\left[1 - \frac{\omega_{p,s}^2(Q)}{\omega^2}\right] + (\epsilon_s^0 - \epsilon_s^{\infty})\frac{\omega_{TO3}^2}{\omega_{TO3}^2 - \omega^2}, \tag{20}$$

where the first term in Eq. (20) is the substrate plasmon contribution and the second term is due to the polar-optical phonon in the substrate. The plasmon frequency of the 2DEG is $\omega_{p,s}(Q) = (\Sigma_\nu e^2 n_\nu Q/2\epsilon_s^{\infty} m_\nu)^{1/2}$, where n_ν and m_ν are the electron density and conductivity mass of subband ν, respectively, and $\hbar\omega_{TO3}$ the energy of the substrate TO-phonon-mode. The notation "TO3" is used to distinguish the substrate mode from the two insulator modes. The dielectric function in the insulator is

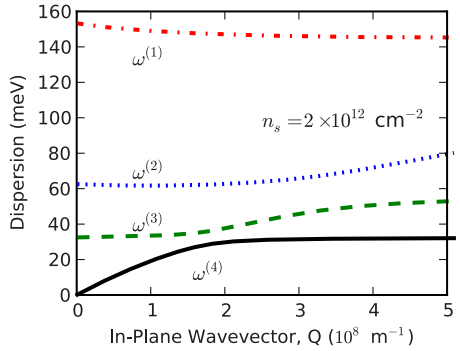

FIG. 2. (Color online) Calculated dispersion for the SO-phonon/substrate plasmon interface modes for the SiO_2/GaAs system. This is the solution of Eq. (23) yielding four mixed solutions (substrate plasmon and 3 TO-modes).

$$\epsilon_{ox}(\omega) = \epsilon_{ox}^{\infty}\frac{(\omega_{LO1}^2 - \omega^2)(\omega_{LO2}^2 - \omega^2)}{(\omega_{TO1}^2 - \omega^2)(\omega_{TO2}^2 - \omega^2)}, \tag{21}$$

where ϵ_{ox}^{∞} is the optical permittivity of insulator and the energies (taken from Ref. 16) of the two insulator TO phonons are $\hbar\omega_{TO1}$ and $\hbar\omega_{TO2}$.

For InGaAs, the second term in Eq. (20) could be replaced with

$$\epsilon_s^{\infty}\frac{(\omega_{LO3}^2 - \omega^2)(\omega_{LO4}^2 - \omega^2)}{(\omega_{TO3}^2 - \omega^2)(\omega_{TO4}^2 - \omega^2)}, \tag{22}$$

to account for two separate substrate modes (InAs- and GaAs-like modes) as measured experimentally.[35] However, because only one of these modes is active in scattering electrons,[36] we use Eq. (20) instead.

A. Dispersion

Following Ref. 16, we can obtain the dispersion of the coupled phonon-plasmons interfacial modes by solving the Laplace equation for the potential associated with these excitations subject to the proper boundary conditions at the interfaces (expressing the presence of the interfacial polarization charges). The homogeneous linear system expressing these boundary conditions admits nontrivial solutions only when the associated determinant vanishes, that is when the following secular equation is satisfied:

$$\epsilon_{ox}(\omega)\cosh(Qt_{ox}) + \epsilon_s(\omega)\sinh(Qt_{ox}) = 0, \tag{23}$$

where t_{ox} is the physical oxide thickness.

Equation (23) yields four solutions, each representing a mixture of the substrate plasmon, the SO-phonon originating in the substrate, and two SO modes from the insulator. An example of these four dispersions, a numerical solution of Eq. (23) for SiO_2 with GaAs at a sheet electron density $n_s = 2\times10^{12}$ cm^{-2}, is shown in Fig. 2.

Landau damping is approximately included by 'switching off' the substrate plasmon [i.e., setting $\epsilon_s(Q,\omega) = \epsilon_s^{\infty} + (\epsilon_s^0 - \epsilon_s^{\infty})\omega_{TO3}^2/(\omega_{TO3}^2 - \omega^2)$] when $\omega_{p,s} \leq \omega_{LDs}$, where

$\omega_{\mathrm{LDs}}=[\hbar Q/(2m_t)](Q+2K_f)$ is the boundary of the single-particle continuum of the 2DEG in the extreme quantum limit, $K_f=(\pi n_s)^{1/2}$, and m_t is the transverse mass.[16] When Landau damping is included in this way, Eq. (23) admits three solutions instead of four—basically three unscreened SO-phonons. We will examine the effect of Landau damping on the SO-limited mobility later in this paper.

B. Plasmon and phonon content

To define separate phonon contents of each mode, we consider three solutions $\omega_Q^{(-\mathrm{TO3},\alpha)}$ ($\alpha=1,2,3$) obtained from the secular equation, Eq. (23), by ignoring the response of phonon-3, that is $\epsilon_s(Q,\omega)\rightarrow\epsilon_s^{\infty}[1-\omega_{p,s}^2(Q)/\omega^2]$. Extending the theory of Ref. 16, the relative phonon-3 (the TO mode of the substrate) content of each branch i ($i=1,4$) is given by

$$R^{\mathrm{TO3}}(\omega_Q^{(i)})\approx\left|\frac{[\omega_Q^{(i)^2}-\omega_Q^{(-\mathrm{TO3},1)^2}][\omega_Q^{(i)^2}-\omega_Q^{(-\mathrm{TO3},2)^2}][\omega_Q^{(i)^2}-\omega_Q^{(-\mathrm{TO3},3)^2}]}{[\omega_Q^{(i)^2}-\omega_Q^{(j)^2}][\omega_Q^{(i)^2}-\omega_Q^{(k)^2}][\omega_Q^{(i)^2}-\omega_Q^{(l)^2}]}\right| \tag{24}$$

and the TO-phonon-3 content of mode "i" is

$$\Phi^{(\mathrm{TO3})}[\omega_Q^{(i)}]\approx\frac{R^{\mathrm{TO3}}[\omega_Q^{(i)}]}{R^{\mathrm{TO1}}[\omega_Q^{(i)}]+R^{\mathrm{TO2}}[\omega_Q^{(i)}]+R^{\mathrm{TO3}}[\omega_Q^{(i)}]}\times\{1-\Pi^{(s)}[\omega_Q^{(i)}]\}, \tag{25}$$

where the substrate plasmon content of mode i is

$$\Pi^{(s)}[\omega_Q^{(i)}]\approx\left|\frac{[\omega_Q^{(i)^2}-\omega_Q^{(-\mathrm{s},1)^2}][\omega_Q^{(i)^2}-\omega_Q^{(-\mathrm{s},2)^2}][\omega_Q^{(i)^2}-\omega_Q^{(-\mathrm{s},3)^2}]}{[\omega_Q^{(i)^2}-\omega_Q^{(j)^2}][\omega_Q^{(i)^2}-\omega_Q^{(k)^2}][\omega_Q^{(i)^2}-\omega_Q^{(l)^2}]}\right|, \tag{26}$$

and i,j,k,l are cyclical. The three solutions $\omega_Q^{(-s,\alpha)}$ ($\alpha=1,2,3$) are obtained from the secular equation, Eq. (23), by ignoring the substrate plasmon response, that is, with $\epsilon_s(Q,\omega)\rightarrow\epsilon_s^{\infty}$. Following the same logic, similar expressions are obtained for $\Phi^{(\mathrm{TO1})}[\omega_Q^{(i)}]$ and $\Phi^{(\mathrm{TO2})}[\omega_Q^{(i)}]$.

As an example, Fig. 3 shows the relative content of the phonon mode labeled "3" (the substrate optical mode) for the each of the dispersions shown in Fig. 2. For small Q, $\omega^{(3)}$ is essentially the TO3 mode, while for larger Q there is a crossover, with $\omega^{(4)}$ gradually becoming the TO3 mode. Figure 4 shows the substrate plasmon content for the dispersion in Fig. 2. The substrate plasmon is a mixture of all three SO-phonons across all four branches of the dispersion. For example, focusing on the $\omega^{(4)}$ curves in Figs. 2 and 4, as the substrate content decreases, the phonon-3-content increases in proportion. Including the substrate plasmon in the dielectric function of the substrate inherently incorporates the screening effects of the 2DEG. However, scattering with the substrate plasmon is ignored because there is no net effect on the momentum of the 2DEG (Ref. 16) (which is strictly true if collisional damping of the plasmons is weaker than Landau damping—we may slightly overestimate the mobility by ignoring this scattering process).

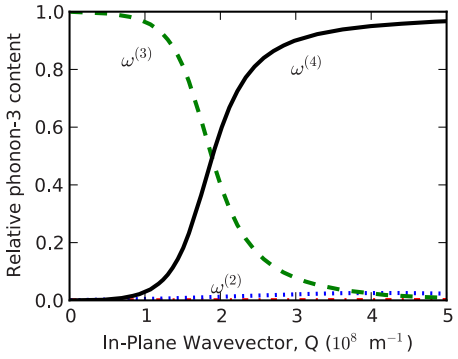

FIG. 3. (Color online) The relative phonon-3 content of each branch of the dispersion in Fig. 2. For small Q, $\omega^{(3)}$ is almost entirely composed of the substrate LO-mode. The substrate plasma frequency mixes the modes and $\omega^{(4)}$ becomes substrate modelike at large Q. The other modes, $\omega^{(2)}$ and $\omega^{(3)}$, are dominated by the dielectric LO-modes.

C. Scattering strength and momentum relaxation rate

Following again the arguments of Ref. 16, the "total" effective dielectric function for the metallic gate/high-κ stack is

$$\epsilon_{\mathrm{TOT}}(Q,\omega)=\epsilon_{\mathrm{substrate}}-\epsilon_{\mathrm{sub}}\frac{\epsilon_{\mathrm{insulator}}}{\epsilon_{\mathrm{ox}}}. \tag{27}$$

The scattering strengths for optical modes 1 and 2 are obtained with a semiclassical procedure outlined in Ref. 16. For

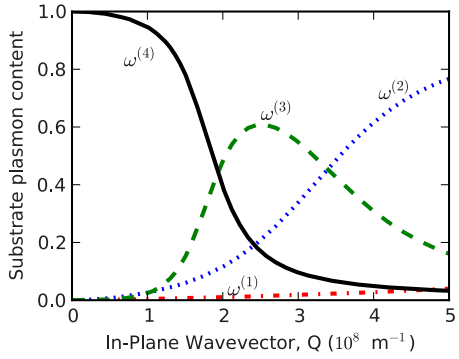

FIG. 4. (Color online) The substrate plasmon content of each branch of the dispersion in Fig. 2.

example, for mode 1 (and similarly for mode 2), setting $\epsilon_{\text{substrate}}(Q,\omega)=\epsilon_s(Q,\omega)$, and $\epsilon_{\text{insulator}}(\omega)=\epsilon_{\text{ox}}^{\infty}(\omega_{\text{LO2}}^2-\omega^2)/(\omega_{\text{TO2}}^2-\omega^2)$ (phonon-1 does not respond) results in the effective dielectric function $\epsilon_{\text{TOT,high}}^{(\text{TO1})}$. Again setting $\epsilon_{\text{substrate}}(Q,\omega)=\epsilon_s(Q,\omega)$, and now $\epsilon_{\text{insulator}}(\omega)=\epsilon_{\text{ox}}^{\infty}(\omega_{\text{LO2}}^2-\omega^2)/(\omega_{\text{TO2}}^2-\omega^2)(\omega_{\text{LO1}}/\omega_{\text{TO1}})^2$ (phonon-1 fully responds) results in the effective dielectric function $\epsilon_{\text{TOT,low}}^{(\text{TO1})}$. Then the scattering strength is

$$|A_Q|^2=\left|\frac{\hbar\omega_Q^{(i)}}{2Q}\left[\frac{1}{\epsilon_{\text{TOT,high}}^{(\text{TO1})}}-\frac{1}{\epsilon_{\text{TOT,low}}^{(\text{TO1})}}\right]\Phi^{(\text{TO1})}(\omega_Q^{(i)})\right|, \quad (28)$$

and a similar expression for mode 2.

For the TO mode originating in the bulk: setting $\epsilon_{\text{insulator}}(\omega)=\epsilon_{\text{ox}}(\omega)$, and $\epsilon_{\text{substrate}}(Q,\omega)=\epsilon_s^{\infty}(1-\omega_p^2/\omega^2)$ (phonon does not respond) results in the effective dielectric function $\epsilon_{\text{TOT,high}}^{(\text{TO3})}$. Again setting $\epsilon_{\text{insulator}}(\omega)=\epsilon_{\text{ox}}(\omega)$, and now $\epsilon_{\text{substrate}}(Q,\omega)=\epsilon_s^{\infty}(1-\omega_p^2/\omega^2)+(\epsilon_s^0-\epsilon_s^{\infty})$ (phonon fully responds) results in the effective dielectric function $\epsilon_{\text{TOT,low}}^{(\text{TO3})}$. Then the scattering strength is

$$|A_Q|^2=\left|\frac{\hbar\omega_Q^{(i)}}{2Q}\left[\frac{1}{\epsilon_{\text{TOT,high}}^{(\text{TO3})}}-\frac{1}{\epsilon_{\text{TOT,low}}^{(\text{TO3})}}\right]\Phi^{(\text{TO3})}(\omega_Q^{(i)})\right|. \quad (29)$$

Having obtained the scattering strength, the momentum relaxation rate can be expressed as[16]

$$\frac{1}{\tau_{\mu\nu}(K,\beta)}=\frac{e^2}{2\pi\hbar^3}\int_0^{2\pi}d\phi\begin{Bmatrix}n_Q\\1+n_Q\end{Bmatrix}\times\frac{1-f(E')}{1-f(E)}|A_Q|^2|F_{\mu\nu}|^2\times[1-\cos\phi]\theta(E'-E_\nu), \quad (30)$$

where

$$F_{\nu\mu}(q_z)=\int_0^{\infty}dz\zeta_\nu(z)e^{iq_zz}\zeta_\mu(z), \quad (31)$$

is the form factor, $f(E)$ is the Fermi–Dirac distribution function, $\phi-\beta-\beta'$ is the angle between the initial and final wave vector, $\theta(x)$ is the step function, and $E'=E\pm\hbar\omega_Q$ is the final energy after absorption (+) or emission (−) of an SO-phonon of energy $\hbar\omega_Q$. The approximation, $(1-\cos\phi)$ in Eq. (30), is the same as discussed for LO phonons after Eq. (11) above.

IV. FIXED INTERFACE CHARGE

In this section we briefly consider Coulomb scattering with fixed interface charge. The unscreened potential for scattering with fixed interface charge is[37–39]

$$V^{\text{it}}(z)=\frac{e}{(\epsilon_{\text{ox}}^0+\epsilon_{\text{sc}})Q}e^{-q_z|z|}, \quad (32)$$

where ϵ_{ox}^0 is the static dielectric constant of the insulator and ϵ_{sc} is the semiconductor dielectric constant. Dielectric screening has been handled using the degenerate static-multisubband-model described in Sec. II A above.[37,38] The scattering rate is written as follows:[26]

$$\frac{1}{\tau_{\mu\nu}(\mathbf{K})}=\frac{e^4N_{\text{it}}}{8\pi\hbar^3\epsilon_{\text{it}}^2}\times\int_0^{2\pi}d\beta'\left[\frac{\cos^2\beta'}{m_x}+\frac{\sin^2\beta'}{m_y}\right]^{-1}|\tilde{F}_{\nu\mu}^{\text{it}}|^2, \quad (33)$$

where $\tilde{F}_{\mu\nu}^{\text{it}}$ is the dielectrically screened form factor and the unscreened form factor is

$$F_{\nu\mu}^{\text{it}}=\int_0^{\infty}dz\zeta_\nu(z)e^{-q_zz}(z)\zeta_\mu(z), \quad (34)$$

where $\epsilon_{\text{it}}=(\epsilon_{\text{ox}}^0+\epsilon_{\text{sc}})/2$ and N_{it} is the areal density of interface charges. Below we will use N_{it} as a 'fitting parameter' to compare our calculations to experimental data.

V. MOBILITY CALCULATION

A. Nonparabolic corrections

The Γ valley of InGaAs is isotropic and strongly nonparabolic, making it necessary to include nonparabolicity to more accurately model the electron mobility. For the dispersion of a 2DEG, perturbation is used and the corrected energy minimum of subband μ is[26]

$$E_\mu\approx E_\mu^{(0)}+\alpha\int_0^{\infty}dz[E_\mu^{(0)}-U(z)]^2|\zeta_\nu^{(0)}(z)|^2, \quad (35)$$

where $U(z)$ is the potential energy. The problem arises when the parabolic subband energy, $E_\mu^{(0)}$, becomes larger than 0.5 eV. This problem is encountered when first-order perturbation fails—under strong quantization (as in very thin quantum wells in materials with a small mass along the confinement direction) and/or strong nonparabolicity parameter α (0.5 eV^{-1} for Si compared to 1 eV^{-1} for InGaAs[40] which we discuss further in Sec. V C below). This problem is overcome phenomenologically in Ref. 41 with the corrected subband energy of subband μ given as

$$E_\mu=U_\mu+\frac{-1+\sqrt{1+4\alpha[E_\mu^{(0)}-U_\mu]}}{2\alpha}, \quad (36)$$

with

$$U_\mu=\int dzU(z)|\zeta_\mu(z)|^2. \quad (37)$$

The phenomenological solution was chosen such that energy increases monotonically and gives the exact solution in the infinite square well potential.

To include nonparabolic corrections in calculation of the LO and SO momentum relaxation rates, the factor

$$[1-2\alpha(E-U_\mu)] \quad (38)$$

must be inserted outside the integral of Eqs. (6) and (8) and inside the integral of Eq. (30). Similarly, for an electron in subband μ, the relation between the magnitude of its wave vector K and its kinetic energy E becomes

$$K(E,\phi) = \frac{\{2[E - E_\mu^{(0)} - \alpha(E-U_\mu)^2]\}^{1/2}}{\hbar(\cos\ \phi^2/m_{\nu,1} + \sin\ \phi^2/m_{\nu,2})^{1/2}}, \quad (39)$$

where ϕ is the angle between the K-vector and the longitudinal axis of the ellipsoidal equienergy surface of the L and X valleys.

B. Schrödinger–Poisson

This section describes the calculation used to obtain the self-consistent potential energy profile, subband minima, wave functions, and quantum mechanical electron density.

The calculation[37] begins with the numerical solution to the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m_z}\frac{d^2}{dz^2} + V(z)\right]\zeta_\nu(z) = E_\nu\zeta_\nu(z), \quad (40)$$

(the separation of variables that leads to Eq. (40) is only generally valid[42] for the boundary condition ζ_ν go to zero for $z=0$ and $z\to\infty$) where m_z is the effective mass along the direction of quantization. The effective potential energy is

$$V(z) = V_{\rm H} + V_{\rm im}(z) + V_{\rm xc}(z), \quad (41)$$

where $V_{\rm H}$ is the Hartree energy obtained self-consistently with Poisson's equation

$$\frac{d^2V_{\rm H}}{dz^2} = \frac{e}{\epsilon_{\rm sc}}[p(z) - n_{\rm qm}(z) - N_{\rm a}], \quad (42)$$

where $p(z)$ is the classical hole density calculated with Fermi–Dirac statistics and $N_{\rm a}$ is the constant p-type doping. The image potential arising from the mismatch of the dielectric constants between the dielectric and the semiconductor is[43]

$$V_{\rm im}(z) = \frac{e^2\tilde{\epsilon}}{16\pi\epsilon_{\rm s}z} \quad (43)$$

in the limit of infinite semiconductor thickness and $\tilde{\epsilon}=(\epsilon_{\rm s}-\epsilon_{\rm ox})/(\epsilon_{\rm s}+\epsilon_{\rm ox})$. The exchange-correlation energy used here is given by the local-density approximation (LDA) (Refs. 28 and 44)

$$V_{\rm xc}(z) = -\frac{e^2}{4\pi^2\epsilon_{\rm s}a\alpha r_s}\left[1 + 0.7734x\ \ln\left(1+\frac{1}{x}\right)\right], \quad (44)$$

where $a=4\pi\epsilon_{\rm s}\hbar^2/(m^*e^2)$ is the effective Bohr radius, $\alpha=(4/9\pi)^{1/3}$, $r_s=a^{-1}[4\pi n_{\rm qm}(z)/3]^{-1/3}$, and $x=r_s/21$. The choice of m^* (Refs. 28 and 45) becomes even more ambiguous because of the occupation of the satellites valleys in III-V semiconductors. We choose to use the conductivity mass in the Γ valley as we only expect to capture the qualitative features of the exchange-correlation energy within the LDA.

Finally, the quantum mechanical electron density with nonparabolic corrections is calculated as[26,41]

$$n_{\rm qm}(z) = \frac{kT}{\pi\hbar^2}\sum_\mu g_\mu m_{d,\mu}\{[1+2\alpha(E_\mu - U_\mu)]F_0(\eta) + 2\alpha kTF_1(\eta)\}|\zeta_\mu(z)|^2, \quad (45)$$

where $F_0(\eta)$ and $F_1(\eta)$ are the Fermi–Dirac integrals of order 0 and 1, g_μ and $m_{d,\mu}$ are the degeneracy and DOS-effective mass in subband μ, respectively, T is the temperature and $E_{\rm F}=0$ is the Fermi-level.

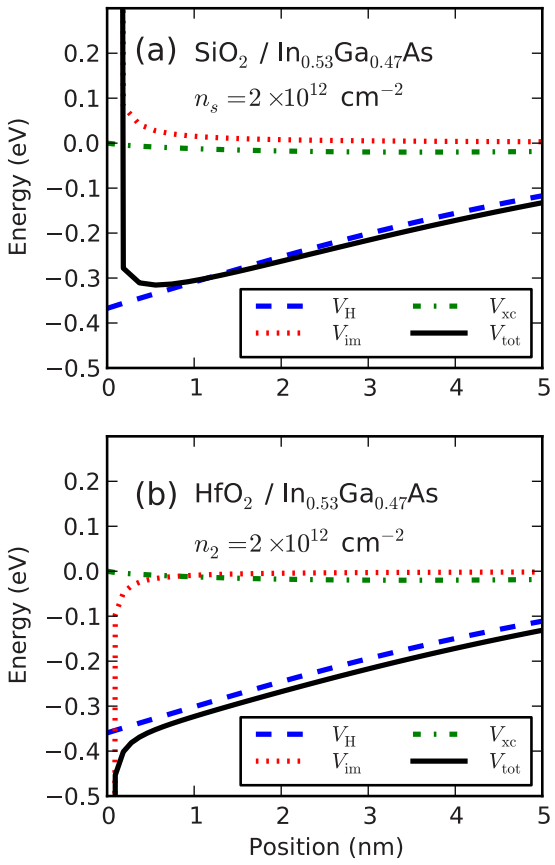

FIG. 5. (Color online) The effective potential energy and its components, Hartree energy $V_{\rm H}$, image potential $V_{\rm im}$, and exchange-correlation energy $V_{\rm xc}$, for an $In_{0.53}Ga_{0.47}As$ inversion layer with SiO_2 (a) and HfO_2 (b) gate dielectrics at $T=300$ K. The Fermi energy is the reference energy and is set to 0 eV.

Figure 5 shows the effective potential energy and its components for an $In_{0.53}Ga_{0.47}As$ inversion layer. We see that $V_{\rm im}>0$ for a SiO_2 gate dielectric, Fig. 5(a), while $V_{\rm im}<0$ for a HiO_2 gate dielectric, Fig. 5(b). The sign change in $V_{\rm im}$ is due to $\tilde{\epsilon}$ which accounts for the mismatch in the gate insulator and semiconductor dielectric constants. For SiO_2 the image potential and exchange-correlation energy have opposite signs—effectively canceling each other for $z>1$ nm. For HfO_2, the image potential and exchange-correlation energy are negative—lowering the effective potential energy throughout the inversion layer. The image term also has an effect on the SR-limited mobility, with high-κ dielectrics exhibiting a higher SR-limited mobility.[28]

C. Kubo–Greenwood formula

We can express the mobility tensor from the linearization of the 2D Boltzmann transport equation[46–48]

$$\mu_{ij} = -\frac{e}{\hbar}\left\langle \tau_{p,i}v_i\left(\frac{\partial f}{\partial K_j}\frac{1}{f}\right)\right\rangle_{th}, \quad (46)$$

where i and j run over the two in-plane Cartesian coordinates x and y, $\tau_{p,i}(\mathbf{K})$ is the momentum relaxation rate, $f(\mathbf{K})$ is the

TABLE I. GaAs parameters.

	E_{gap} (eV)	m_l (m_o)	m_t (m_o)	α (eV^{-1})
Γ	1.42	0.067	0.067	0.61
L	1.74	1.54	0.13	0.5
X	1.86	1.98	0.23	0.5

equilibrium distribution function, and $\langle\cdot\rangle_{th}$ is the thermal average

$$\langle A\rangle_{th} = \sum_{\nu} \frac{g_\nu}{n_\nu} \int \frac{d\mathbf{K}}{(2\pi)^2} A(\mathbf{K}) f_\nu(\mathbf{K}), \tag{47}$$

where n_ν is the electron density and g_ν the degeneracy of subband ν. The transport direction is arbitrarily aligned along the x-axis. For elliptical subbands, including nonparabolic corrections as discussed above, we calculate the xx mobility tensor component of subband ν as

$$\mu_{xx}^{(\nu)} = \frac{e g_\nu}{4\pi^2 m_{\nu,x}^2 k_B T n_\nu} \int_0^{2\pi} d\beta \int_{E_\nu}^{\infty} dE K^2 \cos^2\beta \tau_{p,x}(K,\beta) \times \frac{f(E)[1-f(E)]}{\cos^2\beta/m_{\nu,1} + \sin^2\beta/m_{\nu,2}} \frac{E - E_\mu^{(0)} + \alpha(E-U_\mu)}{1+2\alpha(E-U_\mu)}, \tag{48}$$

and the total mobility as

$$\mu_{xx} = \frac{1}{n_s} \sum_{\nu} n_\nu \mu_{xx}^{(\nu)}, \tag{49}$$

where n_s is the electron sheet density, n_ν is the electron sheet density in subband ν, $\mathbf{K}$ the initial wave vector, E the initial energy (integrating from each subband minimum up to a maximum energy above the last subband considered), and other notation defined in Ref. 48.

We have chosen the conventional [110] field direction on the (001) surface for calculations in this paper. We list, from Ref. 49, in Tables I and II, for GaAs and $In_{0.53}Ga_{0.47}As$, respectively, the values of the longitudinal, m_l, and transverse, m_t, effective masses for each valley. Also listed are the energy minima of the three valleys referenced to the top of the valence band and the nonparabolicity factors, α. The Γ valley of InGaAs is strongly nonparabolic with α=1 eV^{-1} compared to α=0.61 eV^{-1} for GaAs.[50] The InGaAs nonparabolicity parameter used in the Γ valley, is a compromise between the almost parabolic behavior seen for m_z along the (100) confinement direction, and the nonparabolic behavior seen for the DOS-effective-mass in 2D. Because of a lack of data, the silicon nonparabolicity parameter is used in the satellite valleys for GaAs and InGaAs, $\alpha_L = \alpha_X = 0.5$ eV^{-1}.

TABLE II. $In_{0.53}Ga_{0.47}As$ parameters.

	E_{gap} (eV)	m_l (m_o)	m_t (m_o)	α (eV^{-1})
Γ	0.73	0.048	0.048	1
L	1.49	1.57	0.23	0.5
X	1.98	2.26	0.25	0.5

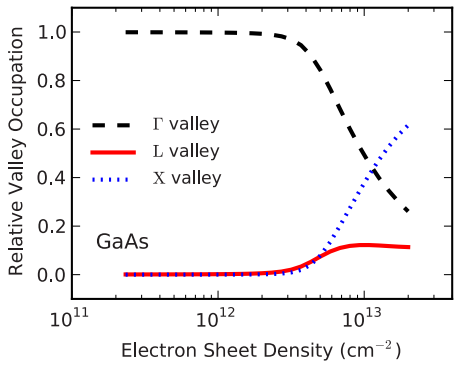

FIG. 6. (Color online) Relative valley-occupation as a function of electron sheet density for a GaAs SG structure.

VI. RESULTS AND DISCUSSION

From Fig. 6, we see, for GaAs, that the occupation of the X valley exceeds the occupation of the Γ valley at sheet densities above 10^{13} cm^{-2} with the L valley accounting for nearly 18% of the total electron population. This is the result of the strong quantization—pushing the energy levels associated with the Γ valley to higher energies and compensating for the higher energies of the satellite valleys, shown in Table I—and the larger density of states of the satellite valleys.

Figure 7 illustrates that for InGaAs the L valleys become significantly populated compared to the Γ valley at large electron sheet densities. Similarly to GaAs, this is due to stronger effect quantization has on the energies of the Γ-valley subbands because of its small quantization mass ($0.048m_0$ compared to $0.32m_0$ for the L valley). A similar result has been seen in the case of double-gate structures.[51] On the other hand, the X valleys do not become significantly occupied, mainly because of the large energy offset between the minima of the X and Γ valleys. We will see below that as the X and/or L valleys become significantly populated, the mobility degrades in InGaAs and GaAs as transport mainly occurs in the "slower" satellite valleys.

In Fig. 8, we compare the LO-limited mobility when V_{im} and V_{xc} are included and excluded for HfO_2 with InGaAs. The satellite valleys are ignored in this figure so that we may focus on the change in mobility due to V_{im} and V_{xc}. We see that including V_{im} and V_{xc} does not significantly change the calculated mobility, corresponding to the results in Fig. 5, where $V_{tot} - V_H < 0.1$ meV for $x > 0.5$ nm.

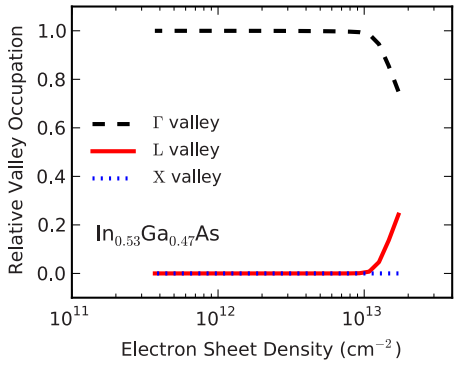

FIG. 7. (Color online) Relative valley-occupation as a function of electron sheet density for a $In_{0.53}Ga_{0.47}As$ SG structure.

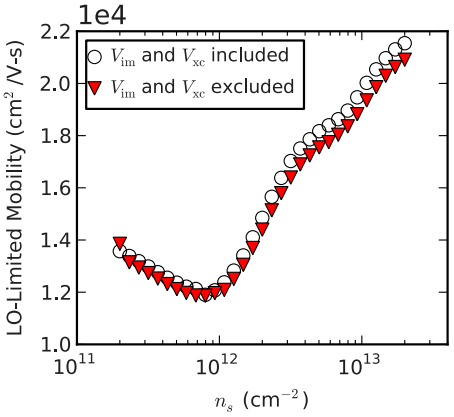

FIG. 8. (Color online) LO-phonon-limited mobility as a function of electron sheet density for the unscreened case for HfO_2 with InGaAs. Only the Γ valley is included and the satellite valleys are ignored.

In Fig. 9, we compare the LO-phonon-limited mobility for GaAs with an HfO_2 gate dielectric for the unscreened, static screened, dynamically screened, and effectively screened cases. For the unscreened case, we observe a mobility decrease in correlation with the increased occupation of the "heavier" X valleys. However, screening is stronger in the X valleys (as the heavier mass allows the slower electrons to screen more efficiently) resulting in a larger percentage mobility increase at large n_s. The use of a dynamic and of a static screening model yields similar results because of the large range of Q for which the dielectric matrix overlaps in Fig. 1. Using the effective screening wave vector tends to underestimate the screening correction at small n_s and overestimate at large n_s. Also shown in Fig. 9 is the dynamically screened LO-phonon-limited mobility when only the Γ valley is considered. We see, at large n_s, a significant increase in the LO-limited mobility when the satellite valleys are ignored.

Figure 10 compares the LO-phonon limited mobility for InGaAs with an HfO_2 gate dielectric for the unscreened, static screened, dynamically screened, and effectively screened cases. As expected, the LO-phonon-limited mobility for an InGaAs substrate is larger than for a GaAs substrate—corresponding to the lighter effective mass in the

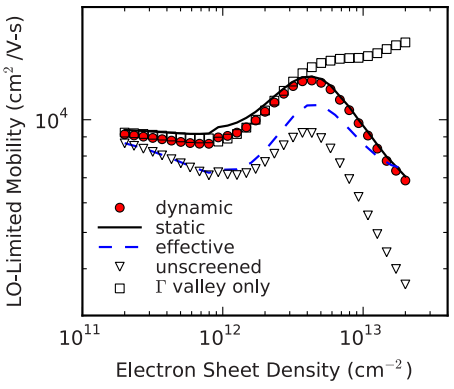

FIG. 9. (Color online) LO-phonon-limited mobility as a function of electron sheet density for the unscreened, static screened, dynamic screened, and effective screened cases for HfO_2 with GaAs. Also shown is the dynamically screened LO-phonon-limited mobility when only the Γ valley is considered.

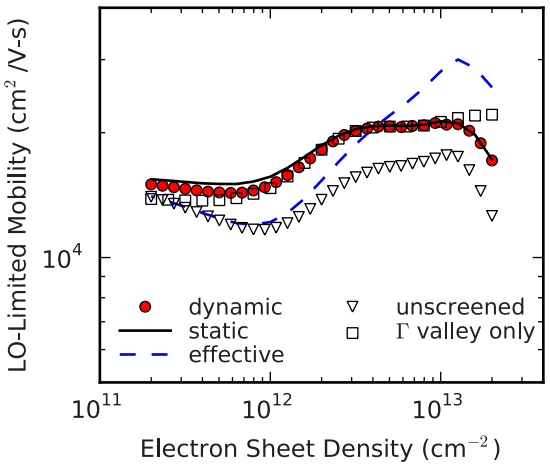

FIG. 10. (Color online) Same as in Fig. 9 but for InGaAs.

Γ valley of InGaAs. The LO-limited mobility peaks at $n_s = 10^{13}$ cm^{-2} and then drops in correlation with the L-valley-occupation seen in Fig. 7. As for GaAs, the effective screening wave vector tends to underestimate the screening correction at small n_s and overestimate at large n_s. We again see an increase in the LO-limited mobility when the heavier satellite valleys are ignored. The increase is not as significant because the Γ valley dominates over a large range of n_s.

Figure 11 shows the SO-phonon-limited mobility for a GaAs and InGaAs substrate with HfO_2 and SiO_2 gate dielectrics. The low-energy phonons associated with HfO_2 as compared to SiO_2 are 'easier' to excite, giving a larger Bose occupation and stronger scattering strength. Again, as in Figs. 9 and 10, the mobility decreases with increasing occupation of the L valley for InGaAs and the X and L valley for GaAs. We also see for HfO_2/InGaAs, above $n_s = 2 \times 10^{12}$ cm^{-2}, Landau damping (when compared to the no Landau damping mobility) decreases the mobility because of the screening effect (i.e., Landau damping turns off screening

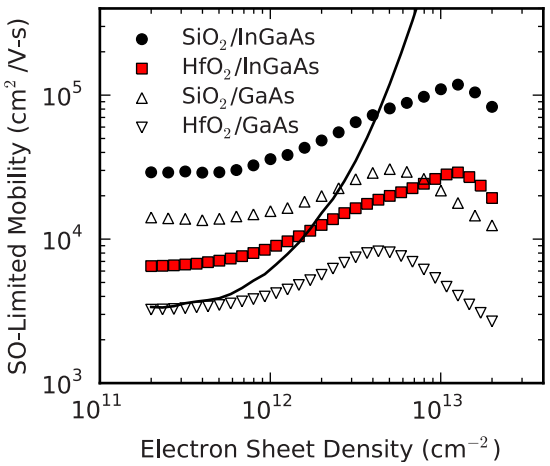

FIG. 11. (Color online) SO-phonon limited mobility as a function of electron sheet density for a metallic gate. The symbols include Landau damping and the solid line is for HfO_2/InGaAs without Landau damping. Comparing the solid squares to the solid line, above $n_s = 2 \times 10^{12}$ cm^{-2}, Landau damping (when compared to the no Landau damping mobility) decreases the mobility because of the screening effect. However, below $n_s = 2 \times 10^{12}$ cm^{-2}, mobility decreases because of the antiscreening effect (Ref. 16).

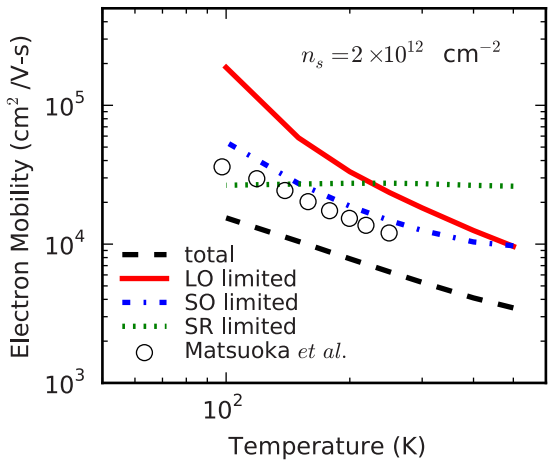

FIG. 12. (Color online) SO- and LO-phonon-limited, SR-limited, and total mobility as a function of temperature for a metallic gate with an InGaAs substrate and HfO_2 gate dielectric. The symbols are experimental data from Matsuoka *et al.* (Ref. 17) for an InGaAs/InAlAs heterostructure.

by 'switching off' the substrate plasmon). However, below $n_s = 2 \times 10^{12}$ cm^{-2}, mobility decreases because of the antiscreening effect[16] (i.e., when Landau damping is not included, the substrate plasmon is never "turned off" and we would expect a higher mobility unless antiscreening takes place).

Figure 12 shows the SO- and LO-phonon-limited, SR-limited, and total mobility (LO, SO, and SR) for an InGaAs substrate and HfO_2 gate dielectric as the temperature is varied from 100 K to 500 K for $n_s = 2 \times 10^{12}$ cm^{-2}. The symbols in Fig. 12 are experimental data from Matsuoka *et al.*[17] for an InGaAs/InAlAs heterostructure. Matsuoka *et al.* attribute the measured mobility primarily to LO and alloy scattering, while our calculations suggests SO-phonon scattering yields a similar temperature dependence—albeit for a HfO_2/InGaAs material system.

Figure 13 compares the SO-limited, LO-limited, SR-limited, nonpolar-phonon-limited, and total mobility for InGaAs ($N_a = 3 \times 17$ cm^{-3}) as a function of the n_s with HfO_2 gate dielectric. The nonpolar-phonon-limited mobility includes scattering with intravalley and intervalley phonons

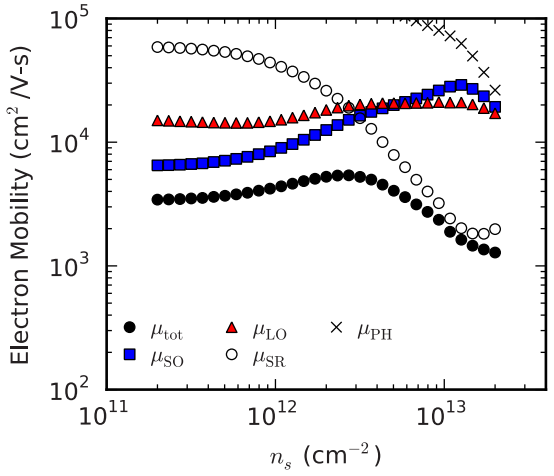

FIG. 13. (Color online) Scattering-limited mobilities as a function of electron sheet density for a metallic gate with InGaAs substrate and HfO_2 gate dielectric for $T = 300$ K. The scattering-limited mobilities shown are: total (μ_{tot}), interface-optical-phonon (μ_{SO}), LO-phonon (μ_{LO}), surface roughness (μ_{SR}), and nonpolar-phonon (μ_{PH}).

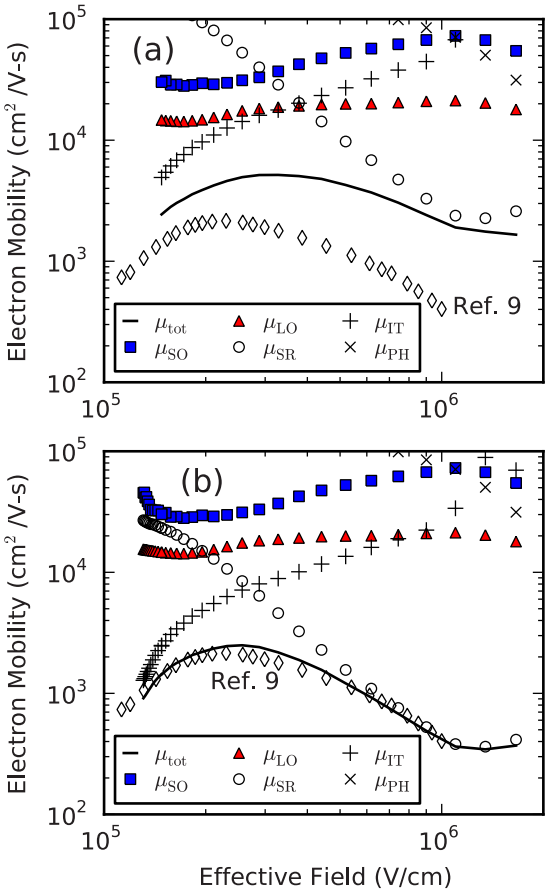

FIG. 14. (Color online) Comparison to experimental data from Ref. 9 is shown for an $Al_2O_3/In_{0.53}Ga_{0.47}As$ structure with p-type doping $N_a = 10^{17}$ cm^{-3}. Scattering with fixed interface charge, N_{it}, is also included. In (a), $N_{it} = 5 \times 12$ cm^{-2} and $\Delta = 0.48$ nm. In (b), $N_{it} = 1 \times 13$ cm^{-2} and $\Delta = 1.2$ nm. In both (a) and (b) the autocorrelation of the step distance in the SR model is $\Lambda = 1.2$ nm.

(see Ref. 26 for example). Ando's model,[26,42] with an exponential correlation spectrum, is used to calculate SR-scattering with degenerate-static-multisubband screening included. We see that SR-scattering dominates at large n_s, where confinement enhances the scattering with the rough interface. The increase in the SR-limited mobility at $n_s > 10^{13}$ cm^{-2} is attributed to the occupation of the "heavy" satellite-valleys and an increased mobility in these valleys because of stronger dielectric screening. At small n_s, SO-phonon scattering is dominant because of the low-energy phonons associated with HfO_2. LO-scattering is a significant contribution to the total mobility at lower to midrange values of n_s.

To facilitate comparison with experimental data from Ref. 9, we now consider the $Al_2O_3/In_{0.53}Ga_{0.47}As$ structure with $N_a = 1 \times 17$ cm^{-3}. Figure 14 compares the SO-limited, LO-limited, SR-limited, nonpolar-phonon-limited, and total mobility as a function of the effective normal field, $F_{eff} = (e/\epsilon_s)(n_s/2 + N_{dep})$, where N_{dep} is the depletion charge. In this figure we also include screened-Coulomb scattering with fixed interface charge. In Fig. 14(a) the experimental mobility is overestimated by the total-calculated mobility if we assume $N_{IT} = 5 \times 12$ cm^{-2} and the rms step height, $\Delta = 0.48$ nm, in the SR model. In Fig. 14(b), we see a much

better agreement by using $N_{IT}=1\times13$ cm^{-2} and $\Delta=1.2$ nm. These results suggest the fabricated device is dominated by scattering with interface charge at low effective-fields and SR at large effective-fields.

VII. CONCLUSIONS

The LO-limited electron mobility was calculated for GaAs and InGaAs substrates accounting for multisubband dynamic screening. We have shown that including screening is necessary to accurately model the LO-limited mobility. Dynamic and static screening yield nearly identical results with dynamic screening adding minimal computational cost over the static case. Defining an effective screening wave vector only captured the qualitative aspects of screening and significantly deviated from the full multisubband approach.

The SO-limited mobility was calculated for GaAs an InGaAs substrates with SiO_2 and HfO_2 gate dielectrics. The dispersion was calculated with the additional substrate LO-mode accounted for in the substrate dielectric function. The existing model was extended to account for scattering with the SO-mode originating from the substrate LO-mode. The SO-limited mobility was also calculated as a function of temperature showing good agreement with experimental data. The calculated total mobility is reduced by about 90% for InGaAs when switching from SiO_2 to HfO_2. While this is significantly more than the 38% predicted for Si,[13] we still expect a 3–4 fold mobility enhancement for HfO_2 on InGaAs compared to SiO_2 on Si.

The inclusion of the satellite valleys was shown to be important at large electron densities. Both the LO- and SO-limited electron mobilities are degraded as the Γ valley becomes deoccupied and the heavier satellite valleys begin to determine the transport properties in the inversion layer.

Finally, comparison to experimental data suggests that the electron mobility is dominated by scattering with interface charge at low effective-fields and SR at high effective-fields, suggesting the need for an improved high-κ/$In_{0.53}Ga_{0.47}As$ interface.

ACKNOWLEDGMENTS

The majority of this work was conducted at University of Massachusetts Amherst, supported in part by SRC (Research ID 1373) and by MARCO/MSD, and at IMEC. One of us (T.P.O.) acknowledges funding by the Marie Curie EDITH fellowship while interning at IMEC. Part of this work was completed by T.P.O. at Tyndall National Institute with support from the Science Foundation Ireland through the U.S.-Ireland Research Project (08/U.S./I1546). T.P.O. thanks P. K. Hurley for discussions on Coulomb scattering with fixed interface charge.

[1]W. Haensch, E. Nowak, R. Dennard, P. Solomon, A. Bryant, O. Dokumaci, A. Kumar, X. Wang, J. Johnson, and M. Fischetti, IBM J. Res. Dev. **50**, 339 (2006).
[2]B. Doyle, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalieros, T. Linton, A. Murthy, R. Rios, and R. Chau, IEEE Electron Device Lett. **24**, 263 (2003).
[3]E. Nowak, I. Aller, T. Ludwig, K. Kim, R. Joshi, C. Chuang, K. Bernstein, and R. Puri, IEEE Circuits Devices Mag. **20**, 20 (2004).
[4]C. Black, Appl. Phys. Lett. **87**, 163116 (2005).
[5]E.-J. Yoon, S.-Y. Lee, S.-M. Kim, M.-S. Kim, S. H. Kim, L. Ming, S. Suk, K. Yeo, C. W. Oh, J.-d.Choe, D. Choi, D.-W. Kim, D. Park, and B.-I. Ryu, Tech. Dig. - Int. Electron Devices Meet. **2004**, 627.
[6]Z. Krivokapic, V. Moroz, W. Maszara, and M. Lin, Tech. Dig. - Int. Electron Devices Meet. **2003**, 18.
[7]K. Rim, S. Koester, M. Hargrove, J. Chu, P. M. Mooney, J. Ott, T. Kanarsky, P. Ronsheim, M. Ieong, A. Grill, and H.-S. P. Wong, Tech. Dig. VLSI Symp. **2001**, 59.
[8]M. Yang, M. Ieong, L. Shi, K. Chan, V. Chan, A. Chou, E. Gusev, K. Jenkins, D. Boyd, Y. Ninomiya, D. Pendleton, Y. Surpris, D. Heenan, J. Ott, K. Guarini, C. D'Emic, M. Cobb, P. Mooney, B. To, N. Rovedo, J. Benedict, R. Mo, and H. Ng, Tech. Dig. - Int. Electron Devices Meet. **2003**, 18.7.1.
[9]Y. Xuan, Y. Wu, H. Lin, T. Shen, and P. Ye, IEEE Electron Device Lett. **28**, 935 (2007).
[10]C. Lee, T. Nishimura, N. Saido, K. Nagashio, K. Kita, and A. Toriumi, Tech. Dig. - Int. Electron Devices Meet. **2009**, 457.
[11]D. Kuzum, T. Krishnamohan, A. Nainani, Y. Sun, P. Pianetta, H.-S. Wong, and K. Saraswat, Tech. Dig. - Int. Electron Devices Meet. **2009**, 453.
[12]S. Laux, IEEE Trans. Electron Devices **54**, 2304 (2007).
[13]T. O'Regan and M. Fischetti, Jpn. J. Appl. Phys., Part 1 **46**, 3265 (2007).
[14]P. Solomon and S. Laux, Tech. Dig. - Int. Electron Devices Meet. **2001**, 5.
[15]B. G. Frederick, G. Apai, and T. N. Rhodin, Phys. Rev. B **44**, 1880 (1991).
[16]M. V. Fischetti, D. A. Neumayer, and E. A. Cartier, J. Appl. Phys. **90**, 4587 (2001).
[17]T. Matsuoka, E. Kobayashi, K. Taniguchi, C. Hamaguchi, and S. Sasa, Jpn. J. Appl. Phys., Part 1 **29**, 2017 (1990).
[18]D. Anderson, N. Zakhleniuk, M. Babiker, B. Ridley, and C. Bennett, Phys. Rev. B **63**, 245313 (2001).
[19]F. Riddoch and B. Ridley, J. Phys. C **16**, 6971 (1983).
[20]M. Fischetti, L. Wang, B. Yu, C. Sachs, P. Asbeck, Y. Taur, and M. Rodwell, Tech. Dig. - Int. Electron Devices Meet. **2007**, 109.
[21]H. Fröhlich, Proc. R. Soc. London, Ser. A **160**, 230 (1937).
[22]http://www.ioffe.ru/SVA/NSM
[23]K. Yokoyama and K. Hess, Phys. Rev. B **33**, 5595 (1986).
[24]D. Chattopadhyay and H. J. Queisser, Rev. Mod. Phys. **53**, 745 (1981).
[25]M. V. Fischetti, Phys. Rev. B **44**, 5527 (1991).
[26]M. V. Fischetti and S. E. Laux, Phys. Rev. B **48**, 2244 (1993).
[27]A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).
[28]S. Jin, M. V. Fischetti, and T.-W. Tang, IEEE Trans. Electron Devices **54**, 2191 (2007).
[29]P. F. Maldague, Surf. Sci. **73**, 296 (1978).
[30]K. Hess, Solid State Commun. **30**, 797 (1979).
[31]R. Chau, S. Datta, M. Doczy, B. Doyle, J. Kavalieros, and M. Metz, IEEE Electron Device Lett. **25**, 408 (2004).
[32]R. Kotlyar, M. Giles, P. Matagne, B. Obradovic, L. Shifren, M. Stettler, and E. Wang, Tech. Dig. - Int. Electron Devices Meet. **2004**, 391.
[33]B. Laikhtman and M. Solomon, J. Appl. Phys. **103**, 014501 (2008).
[34]R. Shah and M. De Souza, IEEE Trans. Electron Devices **54**, 2991 (2007).
[35]T. Mozume and J. Kasai, J. Cryst. Growth **278**, 178 (2005).
[36]T. P. Pearsall, R. Carles, and J. C. Portal, Appl. Phys. Lett. **42**, 436 (1983).
[37]F. Stern and W. E. Howard, Phys. Rev. **163**, 816 (1967).
[38]E. D. Siggia and P. C. Kwok, Phys. Rev. B **2**, 1024 (1970).
[39]T. H. Ning and C. T. Sah, Phys. Rev. B **6**, 4605 (1972).
[40]T. P. O'Regan, P. K. Hurley, B. Sorée, and M. V. Fischetti, Appl. Phys. Lett. **96**, 213514 (2010).
[41]S. Jin, M. V. Fischetti, and T.-w. Tang, J. Appl. Phys. **102**, 083715 (2007).
[42]T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[43]M. Kleefstra and G. C. Herman, J. Appl. Phys. **51**, 4923 (1980).
[44]L. Hedin and B. I. Lundqvist, J. Phys. C **4**, 2064 (1971).
[45]T. Ando, Phys. Rev. B **13**, 3468 (1976).
[46]R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).
[47]D. A. Greenwood, Proc. Phys. Soc. London **71**, 585 (1958).
[48]M. V. Fischetti, J. Appl. Phys. **89**, 1232 (2001).
[49]J. Kim and M. V. Fischetti, J. Appl. Phys. **108**, 013710 (2010).
[50]M. A. Littlejohn, J. R. Hauser, and T. H. Glisson, J. Appl. Phys. **48**, 4587 (1977).
[51]A. Pethe, T. Krishnamohan, D. Kim, S. Oh, H.-S. P. Wong, and K. Saraswat, in IEEE University/Government/Industry Microelectronics Symposium, 16th Biennial, San Jose, CA, 2006, pp. 47–50.