Understanding disorder in silicon quantum computing platforms:
Scattering mechanisms in Si/SiGe quantum wells

Yi Huang ⬤ and Sankar Das Sarma
Condensed Matter Theory Center and Joint Quantum Institute, Department of Physics,
University of Maryland, College Park, Maryland 20742, USA

Motivated by recent experiments on Si/SiGe quantum wells with a co-design of high electron mobility and large valley splitting [Paquelet Wuetz *et al.*, Nat. Commun. **14**, 1385 (2023); Esposti *et al.*, arXiv:2309.02832 [cond-mat.mes-hall]], suitable for a Si-based spin qubit quantum computing platform, we examine the role of disorder by theoretically calculating mobility and quantum mobility from various scattering mechanisms and their dependence on the electron density. At low electron densities $n_e < 4 \times 10^{11}$ cm^{-2}, we find that mobility is limited by remote Coulomb impurities in the capping layer, whereas interface roughness becomes the significant limiting factor at higher densities. We also find that alloy disorder scattering is not a limiting mechanism in the reported high-mobility structures. We estimate the critical density of the disorder-driven low-density metal-insulator transition using the Anderson-Ioffe-Regel localization criterion and qualitatively explain the breakdown of the Boltzmann-Born theory at low densities. We also estimate the critical density by considering inhomogeneous density fluctuations induced by long-range Coulomb disorder in the system, and find a larger critical density compared to the one obtained from the Anderson-Ioffe-Regel criterion. For quantum mobility, our calculation suggests that remote and distant background impurities are likely the limiting scattering sources across all densities. Future measurements of quantum mobility should provide more information on the distribution of background impurities inside the SiGe barriers. Moreover, we extend our theoretical analysis to the effect of quantum degeneracy on transport properties and predict the mobility and the critical density for the metal-insulator transition in spin-polarized high-mobility structures under an external parallel magnetic field.

I. INTRODUCTION

For spin qubits in gate-defined silicon quantum dots, it is important to minimize disorder potential fluctuations to better control charging energies and tunneling between quantum dots, which requires high mobility of the hosted two-dimensional electron gas (2DEG) [1–3]. Moreover, background charge noise, necessarily associated with unintentional random impurities in the environment, is the primary mechanism causing charge noise induced decoherence in semiconductor qubits, and therefore, a comprehensive understanding of the background impurity disorder is essential for good qubit performance [3–6]. Additionally, a large valley splitting energy is important to limit leakage from the computational Hilbert space in the lowest conduction band, which is crucial for maintaining high-fidelity qubit operations [3,7–12]. Recent advancements in isotopically purified ^{28}Si/SiGe quantum wells (QWs) at a low temperature of 1.7 K have shown promising improvements in channel static disorder, as indicated by an enhanced mobility 3×10^5 cm^2 V^{-1} s^{-1} and a smaller critical density of metal-insulator transition $n_c = 7 \times 10^{10}$ cm^{-2}, while maintaining a relatively high mean valley splitting 0.2 meV [13,14]. (Isotopic purification suppresses nuclear spin noise for semiconductor based spin qubits in the Si platform, providing long spin coherence times and enabling high-fidelity gate operations [15].) In comparison, previous experimental reports of large valley splitting

>0.2 meV in Si/SiGe QWs showed 5 times lower mobility <6×10^4 cm^2 V^{-1} s^{-1} [16–18]. On the other hand, high mobility of 6.5×10^5 cm^2 V^{-1} s^{-1} was reported in conventional Si/SiGe heterostructures but with low valley splitting 35–70 μeV [19]. Therefore the co-design of high electron mobility and large valley splitting calls for reinvestigating the scattering sources inside the Si/SiGe QWs reported in Refs. [13,14]. This is the goal of our theoretical work, particularly in the context of the importance of the Si/Ge quantum computing platform.

In this study, we employ the Boltzmann-Born transport theory [20] to theoretically investigate both transport and quantum mobility (μ and μ_q). Our focus is on quantifying the effects of various scattering mechanisms, namely remote impurities at the semiconductor-oxide interface (RI), background impurities (BI), interface roughness (IR), and alloy disorder (AD). By analyzing experimental mobility data from 5 nm and 7 nm Si/SiGe quantum wells reported in Refs. [13,14], we find that RI predominantly limits μ at electron densities below 4×10^{11} cm^{-2}, while IR becomes the limiting factor at higher densities ranging from $4 - 6 \times 10^{11}$ cm^{-2}. Despite limited experimental data on quantum mobility ($\mu_q = 3.0 \pm 0.5 \times 10^4$ cm^2 V^{-1} s^{-1} at the highest density 6×10^{11} cm^{-2} [21]), our theoretical models indicate that μ_q is likely constrained by RI and distant BI scattering across all electron densities. Since μ_q is much more

sensitive to RI and distant BI scattering compared to μ [22], we need more experimental data on μ_q across a wider range of densities to better understand the distribution for BI in SiGe barriers. We also show that AD is quantitatively irrelevant in all experimental density ranges for these high-mobility structures. We address the limitations of the Boltzmann-Born theory at low electron densities $n_e \lesssim 2 \times 10^{11}$ cm^{-2} with considerable disagreement between the theory and experimental data, although our theory explains the higher density data quantitatively. A possible explanation for this discrepancy at low densities is that the 2DEG is broken into inhomogeneous electron puddles separated by long-range Coulomb disorder potential barriers, so that both the perturbative Born approximation and linear screening are no longer applicable. The system eventually undergoes a metal-insulator transition (MIT) at a critical density $n_e = n_c$, where transport in the low-density $n_e \gtrsim n_c$ region should be described by percolation through charged puddles [23–31]. This MIT is, in fact, a crossover (happening sharply over a narrow density range around n_c because of the rapid failure of screening of the long-range Coulomb disorder) and not a quantum phase transition [26,27]. We estimate n_c in the experimental systems [13,14] by using the Anderson-Ioffe-Regel (AIR) criterion [32,33], although the failure of the Boltzmann-Born theory at low densities indicates that the critical density estimated from the AIR criterion should be smaller than the actual critical density (percolation threshold) observed in experiments and should be used as a lower bound estimate of the percolation threshold. We also estimate n_c by considering inhomogeneous density fluctuations induced by long-range Coulomb disorder in the system [34,35] and find a larger n_c compared to the one obtained from the AIR criterion. Moreover, we extend our theoretical analysis to the effect of quantum degeneracy on transport properties and predict the mobility and the critical density for the metal-insulator transition in spin-polarized high-mobility structures, which can be tested in future experiments by applying an external parallel magnetic field.

The Si/SiGe heterostructures reported in Refs. [13,14] consist of a strained Si QW of width w surrounded by a Si$_{0.69}$Ge$_{0.31}$ upper barrier of thickness $d = 30$ nm and a thick strained relaxed Si$_{0.69}$Ge$_{0.31}$ bottom barrier of several μm. The upper barrier is passivated with a 1 nm Si capping layer. A metallic top gate is separated from the capping layer by a $d_o = 10$ nm oxide dielectric spacer. The electron density n_e is tuned by the gate up to 6×10^{11} cm^{-2}. The bottom barrier is placed on top of a step-graded Si$_{1-x}$Ge$_x$ strain-relax buffer layer, which has a decreasing Ge concentration x until it reaches the Si substrate with $x = 0$. A schematic illustration of the Si/SiGe heterostructure is shown in Fig. 1. In our calculations, we model the possible disorder sources as follows. There are uniform background charged impurities outside (inside) the QW with 3D concentration N_1 (N_2). The remote charged impurities at the semiconductor-oxide interface have a 2D concentration n_r. Because the top gate is relatively close to the semiconductor-oxide interface, we take into account the gate screening when we calculate the remote and background charged impurity scattering. The interface roughness between the Si QW and the Si$_{0.69}$Ge$_{0.31}$ barrier is

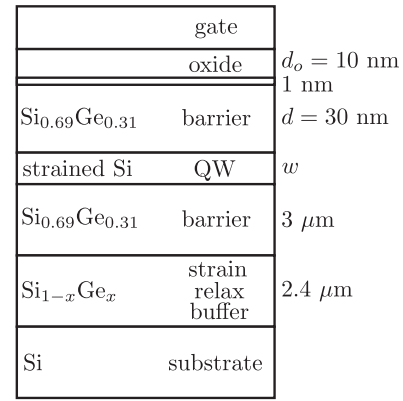| gate	
oxide	$d_o = 10$ nm
	1 nm
Si$_{0.69}$Ge$_{0.31}$ barrier	$d = 30$ nm
strained Si QW	w
Si$_{0.69}$Ge$_{0.31}$ barrier	3 μm
Si$_{1-x}$Ge$_x$ strain relax buffer	2.4 μm
Si substrate	

FIG. 1. A schematic illustration of the high-mobility Si/SiGe heterostructure reported in Refs. [13,14], where w is the thickness of the strained Si quantum well. The oxide and the top Si$_{0.69}$Ge$_{0.31}$ barrier is separated by a Si capping layer of thickness 1 nm.

characterized by a typical height Δ and a lateral size Λ. The conduction band offset from the twofold valley degenerate conduction band Δ_2 in the strained Si to the sixfold valley degenerate conduction band Δ_6 in the relaxed Si$_{0.69}$Ge$_{0.31}$ is taken to be $V_0 = 180$ meV [3,36]. The 2DEG wave function penetrates the barriers and scatters from the alloy disorder. Scattering mechanisms considered in our theory are the only low-temperature scattering mechanisms operational in high quality SiGe 2D structures, where the crystal quality is good enough to rule out short-range scattering by atomic point defects and vacancies [37], but random charged impurities are invariably present in the system along with interface roughness and alloy disorder. Since our interest is in understanding the low-temperature mobility, we neglect all phonon scattering and restrict our calculations to $T = 0$. We believe that our comprehensive transport theory is quantitatively accurate at high carrier densities, and poorer at lower densities although it should remain qualitatively (and semiquantitatively) valid at lower densities above the critical density for the 2D MIT.

The rest of this paper is structured as follows. Section II discusses the general formalism of the Boltzmann transport theory to calculate the mobility and quantum mobility. We evaluate the BI and RI scattering rates using the delta-layer 2DEG approximation in Sec. III. The infinite potential well approximation together with the local field correction, applied to RI, BI, and IR scattering, is discussed in Sec. IV, while Sec. V employs a finite potential well approximation to calculate the AD scattering. Gate-screened charged impurity scattering is explored in Sec. VI. In Sec. VII, we estimate the critical densities of MIT using the Anderson-Ioffe-Regel (AIR) criterion and qualitatively explain the breakdown of the Boltzmann-Born theory at low densities. We mention that the AIR criterion is often referred to as the Mott-Ioffe-Regel (MIR or IRM) or just the Ioffe-Regel criterion, but we believe that AIR is a more appropriate terminology here since our criterion is specific to disorder induced incoherent scattering associated with destructive interference as envisioned by Anderson rather than the Mott criterion involving the transport mean free path being equal to lattice spacing, which plays no role in doped semiconductors of interest to us. We also

estimate the critical density by considering inhomogeneous density fluctuations induced by long-range Coulomb disorder in the system and compare the result to the AIR critical density. Section VIII examines the effect of spin/valley degeneracy on the mobility and the AIR critical density. Finally, we conclude and summarize in Sec. IX.

II. MOBILITY AND QUANTUM MOBILITY

In this section, we apply the Boltzmann transport theory to calculate the mobility and quantum mobility at $T = 0$ [20,38–43]. The transport mobility is given by $\mu = e\tau/m^\star$, and the conductivity is $en_e\mu$, where n_e is the 2DEG concentration, $m^\star = 0.19\,m_0$ is the transverse effective mass of an electron in Si, m_0 is the free electron mass,

$$\frac{1}{\tau} = \frac{4m^\star}{\pi\hbar^3} \int_0^{2k_F} \frac{dq}{\sqrt{4k_F^2 - q^2}} \left(\frac{q}{2k_F}\right)^2 \langle|U(q)|^2\rangle, \quad (1)$$

is the transport scattering rate, and $U(q)$ is the screened potential of a given scattering source. $k_F = \sqrt{4\pi n_e/g}$ is the Fermi wave vector. $g = g_s g_v$ is the total degeneracy of electrons, where g_s is the spin degeneracy and g_v is the valley degeneracy. The quantum mobility is given by $\mu_q = e\tau_q/m^\star$, where the quantum (single particle) scattering rate reads

$$\frac{1}{\tau_q} = \frac{2m^\star}{\pi\hbar^3} \int_0^{2k_F} \frac{dq}{\sqrt{4k_F^2 - q^2}} \langle|U(q)|^2\rangle. \quad (2)$$

The quantum scattering rate can be indirectly measured through the Dingle temperature $T_D = \hbar/2\pi\tau_q$ which characterizes the damping of Shubnikov-de Haas (SdH) oscillations. In general, $\tau_q < \tau$ since the transport mobility ignores all forward scattering. (For strictly zero range s-wave scattering the two scattering times are trivially equal, but for long-range disorder scattering the two could vary by orders of magnitude.) For charged impurity scattering, the potential correlator averaged over impurity positions is given by

$$\langle|U(q)|^2\rangle = \int_{-\infty}^{+\infty} dz\, N(z)U_1^2(q,z), \quad (3)$$

where $N(z)$ is the 3D concentration of impurities at a distance z from the center of the 2DEG, and $U_1(q,z)$ is the screened Coulomb potential for an impurity located at z:

$$U_1(q,z) = \frac{2\pi e^2}{\kappa q \epsilon(q)} \int_{-\infty}^{+\infty} dz'\, |\psi(z')|^2 e^{-q|z-z'|}, \quad (4)$$

where $\psi(z)$ is the confinement wave function along z direction and the dielectric function is given by

$$\epsilon(q) = 1 + (q_{\mathrm{TF}}/q)F_c(qw)[1 - G(q)]. \quad (5)$$

Here, $q_{\mathrm{TF}} = g/a_B$ is the Thomas-Fermi screening wave vector, $a_B = \kappa\hbar^2/m^\star e^2$ is the effective Bohr radius, $\kappa = 11.9$ is the dielectric constant of Si, and w is the QW width. $G(q) = q/(g\sqrt{q^2 + k_F^2})$ is the local field correction using the Hubbard approximation [44,45] which reflects the suppressed screening at $q \gtrsim k_F$ (with $G = 0$ being the RPA, which overestimates short wavelength screening). The form factor

$F_c(qw)$ is given by

$$F_c(qw) = \iint_{-\infty}^{+\infty} dz dz' |\psi(z)|^2 |\psi(z')|^2 \exp(-q|z - z'|). \quad (6)$$

In Eq. (4), we ignore the screening from the top gate. To include the gate screening, we should add the Coulomb potential of the image charges and replace $e^{-q|z-z'|}$ by $(e^{-q|z-z'|} - e^{-q|2d_g-z-z'|})$, where $d_g = d + d_o + w/2$ is the distance from the gate to the center of the 2DEG.

III. DELTA LAYER 2DEG APPROXIMATION

For simplicity, assuming that the gate is far away $d_g \gg d$ and the wave function perpendicular to the quantum well has a delta function profile $|\psi(z)|^2 = \delta(z)$, the form factors reduce to identities and $U_1(q,z) = \frac{2\pi e^2}{\kappa q \epsilon(q)} e^{-q|z|}$. In this section, we disregard the local field correction, thus making the dielectric function the RPA form, $\epsilon(q) = 1 + q_{\mathrm{TF}}/q$. We will address the effect of form factors and local field correction in Sec. IV, and the effect of gate screening in Sec. VI. We use a two-impurity model to calculate the scattering rate following Refs. [42,46]. That is, we assume that there are two delta layers of impurities. One layer is located at a distance d from the 2DEG, which represents the remote charged impurities in the capping layer. The other layer is at $z = 0$, which represents oxygen-related background charges in the QW [47] or an effective description of other short-range scattering sources such as the interface roughness [48,49]. The impurity distribution is given by

$$N(z) = n_r \delta(z - d) + n_2 \delta(z). \quad (7)$$

The corresponding transport scattering rate $\tau^{-1} = \tau_{\mathrm{RI}}^{-1} + \tau_{\mathrm{S}}^{-1}$ reads [41,42,46]

$$\frac{1}{\tau_{\mathrm{RI}}} = n_r \frac{2\pi\hbar}{m^\star}\left(\frac{2}{g}\right)^2 f\left(2k_F d, \frac{q_{\mathrm{TF}}}{2k_F}\right), \quad (8)$$

$$\frac{1}{\tau_{\mathrm{S}}} = n_2 \frac{2\pi\hbar}{m^\star}\left(\frac{2}{g}\right)^2 f\left(0, \frac{q_{\mathrm{TF}}}{2k_F}\right). \quad (9)$$

The quantum scattering rate $\tau_q^{-1} = \tau_{q\mathrm{RI}}^{-1} + \tau_{q\mathrm{S}}^{-1}$ reads

$$\frac{1}{\tau_{q\mathrm{RI}}} = n_r \frac{2\pi\hbar}{m^\star}\left(\frac{2}{g}\right)^2 f_q\left(2k_F d, \frac{q_{\mathrm{TF}}}{2k_F}\right), \quad (10)$$

$$\frac{1}{\tau_{q\mathrm{S}}} = n_2 \frac{2\pi\hbar}{m^\star}\left(\frac{2}{g}\right)^2 f_q\left(0, \frac{q_{\mathrm{TF}}}{2k_F}\right), \quad (11)$$

and the dimensionless functions $f(a, s)$ and $f_q(a, s)$ are defined as

$$f(a, s) = \int_0^1 s^2 \frac{2x^2 e^{-2ax} dx}{\sqrt{1 - x^2}(x + s)^2}, \quad (12)$$

$$f_q(a, s) = \int_0^1 s^2 \frac{e^{-2ax} dx}{\sqrt{1 - x^2}(x + s)^2}. \quad (13)$$

In the density range

$$\frac{g}{16\pi d^2} \ll n_e \ll \frac{g^3}{16\pi a_B^2}, \quad (14)$$

the limits $2k_F d \gg 1$ and $q_{\mathrm{TF}}/2k_F \gg 1$ are satisfied and we have $f(a, s) = (2a^3)^{-1}$ and $f_q(a, s) = (2a)^{-1}$ [50] which lead to the well-known results for RI scattering [51–54]

$$\frac{1}{\tau_{\mathrm{RI}}} = n_r \left(\frac{2}{g}\right)^2 \frac{\pi\hbar}{8m^\star (k_F d)^3}, \tag{15}$$

$$\frac{1}{\tau_{q\mathrm{RI}}} = n_r \left(\frac{2}{g}\right)^2 \frac{\pi\hbar}{2m^\star k_F d}, \tag{16}$$

and the corresponding ratio is given by

$$\left(\frac{\tau}{\tau_q}\right)_{\mathrm{RI}} = (2k_F d)^2 \gg 1. \tag{17}$$

The large difference between τ and τ_q reflects the diffusive character of the electron dynamics on the Fermi surface scattered from remote impurities [22]. The physical meaning of τ/τ_q is the typical number of small-momentum forward scattering events needed to change the direction of momentum by an angle of order π [54]. Next, we calculate the short-range scattering rate τ_{S}^{-1} and $\tau_{q\mathrm{S}}^{-1}$ contributed from a delta layer of impurities at $z = 0$. The related integrals $f(0, s)$ and $f_q(0, s)$ at $s > 1$ can be calculated analytically:

$$f(0, s) = s^2\pi + \frac{2s^3}{1-s^2} + \frac{2s^3(2-s^2)\sec^{-1}(s)}{(s^2-1)^{3/2}}, \tag{18}$$

$$f_q(0, s) = \frac{s}{1-s^2} + \frac{s^3\sec^{-1}(s)}{(s^2-1)^{3/2}}. \tag{19}$$

In the density range $n_e \ll g^3/16\pi a_B^2$, the limit $s = q_{\mathrm{TF}}/2k_F \gg 1$ is satisfied and $f(0, s) = f_q(0, s) = \frac{\pi}{2}$, so that τ_{S}^{-1} and $\tau_{q\mathrm{S}}^{-1}$ are given by [41,46]

$$\frac{1}{\tau_{\mathrm{S}}} = \frac{1}{\tau_{q\mathrm{S}}} = n_2 \frac{\pi^2\hbar}{m^\star}\left(\frac{2}{g}\right)^2, \tag{20}$$

which is independent of the electron concentration. The ratio $\tau/\tau_q = 1$ reflects that short-range impurities are efficient in backscattering the electrons—this is the situation in regular 3D metals where τ and τ_q are basically the same.

In the other two-impurity model, we replace the delta-layer remote impurities by uniform background charged impurities outside the well with a 3D concentration N_1. The impurity distribution is given by

$$N(z) = N_1 \Theta(|z| - w/2) + n_2\delta(z). \tag{21}$$

The corresponding scattering rates due to uniform background charged impurities are given by

$$\frac{1}{\tau_{\mathrm{BI}}} = N_1 \frac{2\pi\hbar}{m^\star}\left(\frac{2}{g}\right)^2 \int_{|z|>w/2} f\left(2k_F z, \frac{q_{\mathrm{TF}}}{2k_F}\right)dz, \tag{22}$$

$$\frac{1}{\tau_{q\mathrm{BI}}} = N_1 \frac{2\pi\hbar}{m^\star}\left(\frac{2}{g}\right)^2 \int_{|z|>w/2} f_q\left(2k_F z, \frac{q_{\mathrm{TF}}}{2k_F}\right)dz. \tag{23}$$

In the electron density range

$$\frac{g}{16\pi w^2} \ll n_e \ll \frac{g^3}{16\pi a_B^2}, \tag{24}$$

the limits $2k_F w \gg 1$ and $q_{\mathrm{TF}}/2k_F \gg 1$ are satisfied and we get the analytical expressions of the BI scattering rates

$$\frac{1}{\tau_{\mathrm{BI}}} = N_1 w \left(\frac{2}{g}\right)^2 \frac{\pi\hbar}{2m^\star (k_F w)^3}, \tag{25}$$

$$\frac{1}{\tau_{q\mathrm{BI}}} = N_1 w \left(\frac{2}{g}\right)^2 \frac{\pi\hbar\ln(2d_s/w)}{m^\star k_F w}, \tag{26}$$

where d_s is the cutoff distance of the BI distribution so that impurities at $|z| > d_s$ can be ignored. Otherwise, $\tau_{q\mathrm{BI}}^{-1}$ diverges logarithmically due to the contribution of small-momentum scattering from the long distance. We choose $d_s = d_g/2$ as half the distance from the gate to the center of the 2DEG, considering that charged impurities close to the gate should be effectively screened [55]. We will discuss this choice of d_s and the gate screening effect in more detail in Sec. VI.

In the density range described by Eqs. (14) and (24), we see from Eqs. (15) and (25) that τ_{RI}^{-1} and τ_{BI}^{-1} are proportional to k_F^{-3}, which leads to a mobility proportional to $n_e^{3/2}$. On the other hand, τ_{S}^{-1} in Eq. (20) does not depend on k_F, so it results in a constant mobility that does not depend on n_e. As a result, the total mobility described by the two-impurity model $\tau^{-1} = \tau_{\mathrm{RI}}^{-1} + \tau_{\mathrm{S}}^{-1}$ or $\tau^{-1} = \tau_{\mathrm{BI}}^{-1} + \tau_{\mathrm{S}}^{-1}$ should increase as $n_e^{3/2}$ at low density and eventually saturate at high density. We can compare the density range $n_e = 1 - 6 \times 10^{11}$ cm^{-2} in the high mobility structures reported in Refs. [13,14] with the density ranges written in Eqs. (14) and (24). We use the spin/valley degeneracy $g_s = 2$ and $g_v = 2$ for the high mobility structures reported in Refs. [13,14]. In the density range $n_e = 1 - 6 \times 10^{11}$ cm^{-2}, the Fermi energy calculated using $g = 4$ is $E_F = 0.6$–3.8 meV which is much higher than the valley splitting 0.2 meV reported in Refs. [13,14], so our choice of $g = 4$ is correct. In Sec. VIII, we will discuss the effects of quantum degeneracy on the transport properties in more detail. Using $g = 4$, we find that the experimental density range is well within the range described in Eq. (14) where the upper bound is $g^3/16\pi a_B^2 \approx 10^{13}$ cm^{-2} and the lower bound is $g/16\pi d^2 \approx 10^{10}$ cm^{-2}. Therefore the asymptotic expressions for RI scattering Eqs. (15) and (16) should be a good approximation. On the other hand, the lower bound in Eq. (24) is $g/16\pi w^2 \approx 2.4 \times 10^{11}$ cm^{-2}, which means that the asymptotic expressions for BI scattering rates Eqs. (25) and (26) are good only when $n_e \gg 2.4 \times 10^{11}$ cm^{-2}. At low densities $n_e \ll 2.4 \times 10^{11}$ cm^{-2}, we have

$$\frac{1}{\tau_{\mathrm{BI}}} = \frac{1}{\tau_{q\mathrm{BI}}} = N_1(2d_s - w)\frac{\pi^2\hbar}{m^\star}\left(\frac{2}{g}\right)^2, \tag{27}$$

and the background impurity scattering behaves similarly to short-range impurity scattering with an effective 2D impurity concentration $N_1(2d_s - w)$. Hence, as the density is reduced, τ_{BI}^{-1} and $\tau_{q\mathrm{BI}}^{-1}$ transition from a long-range behavior, as expressed in Eqs. (25) and (26), to a short-range behavior, as expressed in Eq. (27). The crossover electron density is at $n_e = 2.4 \times 10^{11}$ cm^{-2}.

Figure 2 shows the results of the best-fit mobility curves (solid black) and the corresponding prediction of quantum mobility (dashed black) by performing the numerical integrals of Eqs. (8), (9), and (22). The experimental data from Refs. [13,14], represented by blue and red curves with shaded areas indicating error bars, are also included for comparison. The results of the combination of remote and short-range
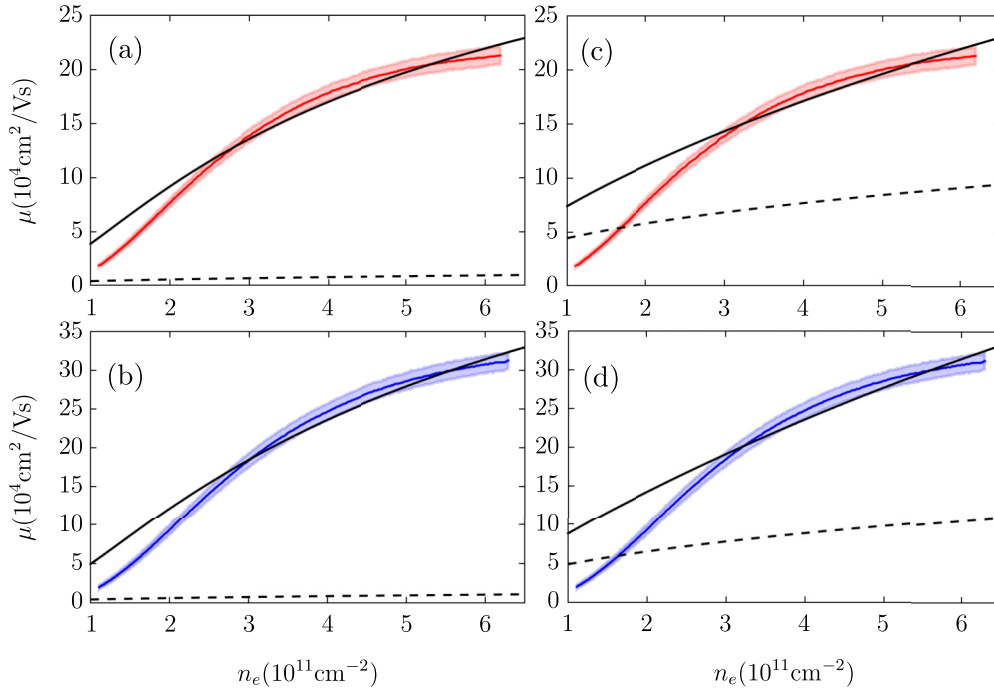
FIG. 2. Mobility (μ) versus electron density (n_e) for quantum wells with widths of 5 nm (red) and 7 nm (blue). The experimental data taken from Refs. [13,14] are depicted by red and blue curves, with corresponding error bars shown as shaded areas. The solid black curves represent the theoretical transport mobility, while the dashed black curves denote the quantum mobility, both calculated using the delta-layer approximation for the 2DEG without local field correction. Subfigures (a) and (b) incorporate remote and short-range impurity scattering [cf. Eqs. (8)–(10)], whereas (c) and (d) consider background and short-range impurity scattering [cf. Eqs. (9), (22), and (23)]. Fitting parameters: (a) $n_r = 1.9 \times 10^{12}$ cm^{-2} and $n_2 = 2.8 \times 10^{9}$ cm^{-2}; (b) $n_r = 1.7 \times 10^{12}$ cm^{-2} and $n_2 = 1.8 \times 10^{9}$ cm^{-2}; (c) $n_2 = 2.1 \times 10^{9}$ cm^{-2} and $N_1 = 11 \times 10^{15}$ cm^{-3}; (d) $n_2 = 1.4 \times 10^{9}$ cm^{-2} and $N_1 = 12 \times 10^{15}$ cm^{-3}.

impurity scattering are shown in Figs. 2(a) and 2(b), while the results of the combination of background and short-range impurity scattering are shown in Figs. 2(c) and 2(d). The fitting impurity densities are $n_r \approx 2 \times 10^{12}$ cm^{-2}, $n_2 \approx 1.4$–2.8×10^{9} cm^{-2} and $N_1 \approx 10 \times 10^{15}$ cm^{-3}, in reasonable agreement with the numbers reported in Refs. [36,41,47,56–59]. In both two-impurity models, short-range impurity scattering is the limiting factor for transport mobility at high densities $n_e = 4$–6×10^{11} cm^{-2}. The dominant scattering mechanism at lower densities is RI or BI, depending on the choice of the impurity model. We see that the first two-impurity model ($\tau^{-1} = \tau_{\mathrm{RI}}^{-1} + \tau_{\mathrm{S}}^{-1}$) fits the mobility data at intermediate to high electron density $n_e = 2.5$–6×10^{11} cm^{-2} within the experimental uncertainty. At the highest density $n_e = 6 \times 10^{11}$ cm^{-2} for both quantum wells, the predicted quantum mobility is around 1.0×10^{4} cm^{2} V^{-1} s^{-1}, while the measurements of SdH oscillations give a larger value $(3.0 \pm 0.5) \times 10^{4}$ cm^{2} V^{-1} s^{-1} [21]. On the other hand, the other two-impurity model ($\tau^{-1} = \tau_{\mathrm{BI}}^{-1} + \tau_{\mathrm{S}}^{-1}$) fits the experimental mobility data in a narrower density range $n_e = 3 - 6 \times 10^{11}$ cm^{-2}, with the predicted quantum mobility around 10×10^{4} cm^{2} V^{-1} s^{-1} at the highest density $n_e = 6 \times 10^{11}$ cm^{-2}, which is a factor of 3 larger than the experimental value [21]. Since the typical momentum transfer of RI scattering is much smaller than that of BI scattering, the ratio $(\mu/\mu_q)_{\mathrm{RI}} \gg (\mu/\mu_q)_{\mathrm{BI}}$ [22], which explains a factor of 10 difference in the predicted quantum mobility shown in Fig. 2. In Sec. VI, we will explore how including gate screening effects in RI scattering

calculations can improve the agreement with experimental quantum mobility data. However, it is important to note that the experimental accuracy for quantum mobility data is generally less reliable than for transport mobility, largely due to the indirect method of extracting quantum mobility through the Dingle temperature [22]. Given these considerations, a semi-quantitative agreement between our theoretical predictions and experimental data should be regarded as satisfactory. At low densities $n_e < 2 \times 10^{11}$ cm^{-2}, the theoretical mobility curves deviate and become higher than the experimental data. We should emphasize that this deviation at low densities does not arise from any new scattering mechanism ignored in the theory. The dominant scattering mechanism at low densities $n_e < 4 \times 10^{11}$ cm^{-2} is RI if we use the two-impurity model Eq. (7) [or BI if we use the two-impurity model Eq. (21)]. Rather, this deviation indicates the systematic failure of the Boltzmann-Born theory at low densities, a topic we will address in more detail in Sec. VII, particularly in the context of the metal-insulator transition.

IV. INFINITE POTENTIAL WELL

In this section, we use an infinite potential well approximation in order to address the quantum well confinement better, where the 2DEG wave function is given by $\psi(z) = \sqrt{2/w} \cos(\pi z/w) \Theta(w/2 - |z|)$ [46,60]. The justification of the infinite potential well approximation will be discussed at the end of this section. Substituting this wave function into
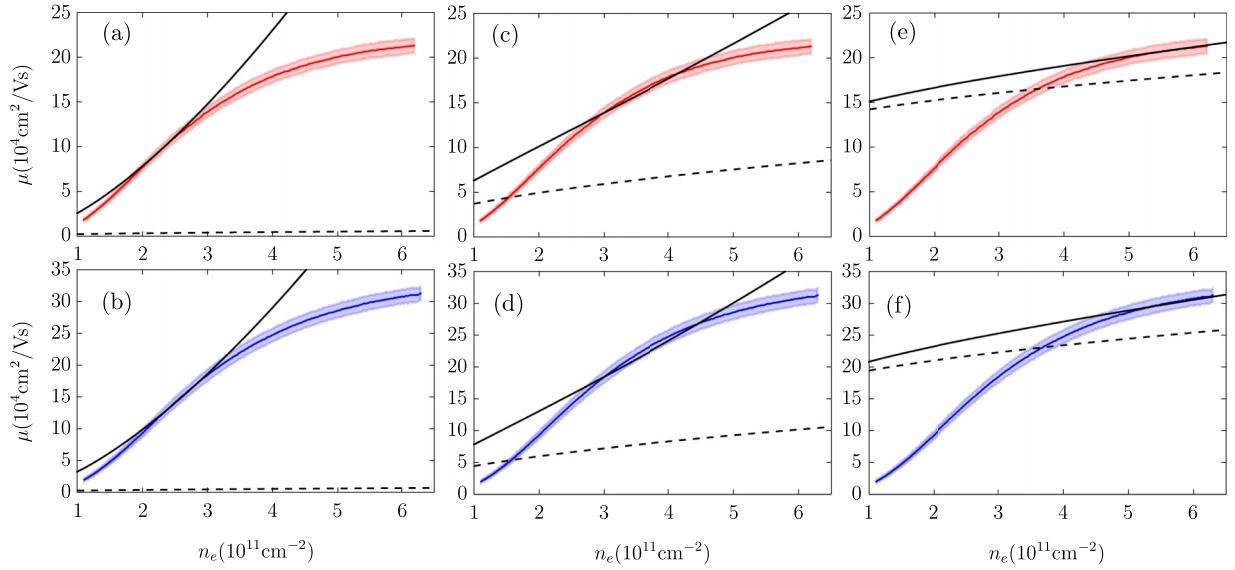
FIG. 3. Mobility (μ) vs electron density (n_e) for quantum wells with widths of 5 nm (red) and 7 nm (blue). The experimental data are depicted by red and blue curves, with corresponding error bars shown as shaded areas. The solid black curves represent the theoretical transport mobility, while the dashed black curves denote the quantum mobility, both calculated using the infinite potential well approximation without the local field correction. (a) and (b) assume only RI near the capping layer $n_r \neq 0$. (c) and (d) assume only BI outside the quantum well $N_1 \neq 0$ and no BI inside $N_2 = 0$, while (e) and (f) assume $N_1 = 0$ and $N_2 \neq 0$. Impurity concentrations and roughness parameters for fitting: (a) $n_r = 3.1 \times 10^{12}$ cm^{-3}; (b) $n_r = 2.6 \times 10^{12}$ cm^{-3}; (c) $N_1 = 14 \times 10^{15}$ cm^{-3} and $N_2 = 0$; (d) $N_1 = 12.6 \times 10^{15}$ cm^{-3} and $N_2 = 0$; (e) $N_1 = 0$ and $N_2 = 10.3 \times 10^{15}$ cm^{-3}; (f) $N_1 = 0$ and $N_2 = 5.5 \times 10^{15}$ cm^{-3}.

Eq. (4), we obtain the single-impurity Coulomb potential

$$U_1(q,z) = \frac{2\pi e^2}{\kappa q \epsilon(q)} \begin{cases} F_0(qw)e^{-q|z|}, & |z| > w/2, \\ G_0(q,z), & |z| < w/2, \end{cases} \quad (28)$$

where

$$F_0(x) = \frac{8\pi^2}{x(x^2 + 4\pi^2)} \sinh(x/2), \quad (29)$$

and

$$G_0(q,z) = \frac{8\pi^2(1 - e^{-qw/2}\cosh qz)}{qw(4\pi^2 + q^2 w^2)} + \frac{2qw[1 + \cos(2\pi z/w)]}{4\pi^2 + q^2 w^2}. \quad (30)$$

Both $F_0(qw)$ and $G_0(q,z)$ monotonically decrease as a function of q and weaken the scattering potential at large momentum q. The screening form factor $F_c(qw)$ in Eq. (6) reads

$$F_c(x) = \frac{20\pi^2 x^3 + 3x^5 - 32\pi^4(1 - e^{-x} - x)}{x^2(4\pi^2 + x^2)^2}. \quad (31)$$

This form factor enters the dielectric function $\epsilon(q)$ and weakens the screening at large momentum q.

First, we calculate the charged impurity scattering rates separately and see how well they can interpret the data individually. The 3D impurity concentration is given by

$$N(z) = n_r \delta(z - d - w/2)$$
$$+ N_1 \Theta(|z| - w/2) + N_2 \Theta(w/2 - |z|), \quad (32)$$

where n_r is the RI concentration in the capping layer, N_1 (N_2) is the BI concentration outside (inside) the quantum well. The theoretical results of mobility and quantum mobility are

shown in Fig. 3, based on the assumption that only one term in Eq. (32) is nonzero. Figures 3(a) and 3(b) show the results of $n_r \neq 0$ and $N_1 = N_2 = 0$. This RI-dominant mobility can fit the data at intermediate densities $n_e = 2$–3×10^{11} cm^{-2} within the margin of experimental uncertainty. The predicted quantum mobility is around 1.0×10^4 cm^2 V^{-1} s^{-1} at the highest density $n_e = 6 \times 10^{11}$ cm^{-2}. Figures 3(c) and 3(d) show the results of $N_1 \neq 0$ and $N_2 = n_r = 0$, which fit the data at intermediate densities $n_e = 2.5$–5×10^{11} cm^{-2}. The predicted quantum mobility is around 10×10^4 cm^2 V^{-1} s^{-1} at the highest density $n_e = 6 \times 10^{11}$ cm^{-2}. Figures 3(e) and 3(f) show the results of $N_2 \neq 0$ and $N_1 = n_r = 0$, which fit the data at high densities $n_e = 4$–6×10^{11} cm^{-2}. The corresponding ratio $\tau/\tau_q \approx 1$ is expected, as the background impurities inside the QW should behave similarly to the short-range impurities because of strong screening by the electrons. In general, switching the dominant impurity distribution from RI ($n_r \neq 0$) to BI ($N_1 \neq 0$) and finally to a different BI distribution ($N_2 \neq 0$) results in a progressive flattening of the mobility $\mu(n_e)$ and a decrease in the ratio μ/μ_q towards unity.

In order to better fit the mobility data, we use interface roughness as an additional scattering mechanism that limits mobility at high density [27,49,52,60–63]. Using the correlator of local well width variations $\langle \Delta(\mathbf{r})\Delta(\mathbf{r}') \rangle = \Delta^2 \exp(-|\mathbf{r} - \mathbf{r}'|^2/\Lambda^2)$, where Δ is the typical roughness height and Λ is the roughness lateral size, the scattering potential due to interface roughness is given by

$$\langle |U_{\text{IR}}(q)|^2 \rangle = \frac{\pi}{\epsilon^2(q)} \Delta^2 \Lambda^2 \left(\frac{\partial E}{\partial w}\right)^2 e^{-q^2 \Lambda^2/4}, \quad (33)$$

where E is the ground state energy of the potential well along z direction. For an infinite potential well, the ground state
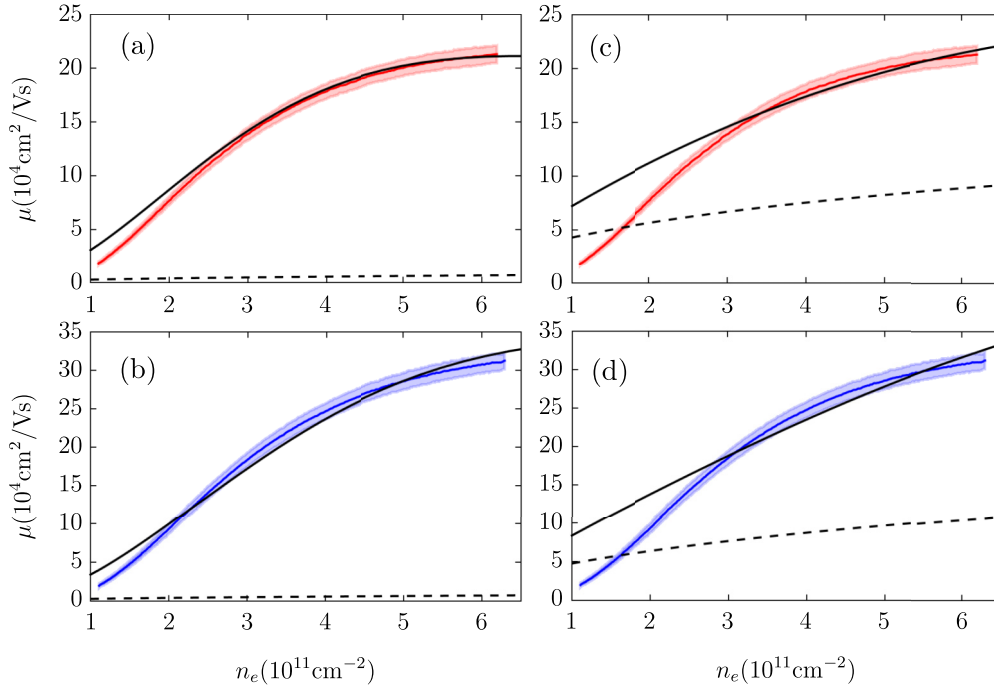
FIG. 4. Mobility (μ) vs electron density (n_e) for quantum wells with widths of 5 nm (red) and 7 nm (blue). The experimental data are depicted by red and blue curves, with corresponding error bars shown as shaded areas. The solid black curves represent the theoretical transport mobility, while the dashed black curves denote the quantum mobility, both calculated using the infinite potential well approximation without the local field correction. Subfigures (a) and (b) incorporate remote impurity and interface roughness scattering, whereas (c) and (d) consider background impurity and interface roughness scattering. Impurity concentrations and roughness parameters for fitting: (a) $n_r = 2.5 \times 10^{12}$ cm^{-2}, $\Delta = 4.5$ Å, and $\Lambda = 27$ Å; (b) $n_r = 2.5 \times 10^{12}$ cm^{-2}, $\Delta = 8$ Å, and $\Lambda = 28$ Å; (c) $N_1 = 12 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 5$ Å, and $\Lambda = 16$ Å; (d) $N_1 = 12 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 7$ Å, and $\Lambda = 20$ Å.

energy $E = E_0 = \hbar^2\pi^2/2m_z^\star w^2$ and $\partial E/\partial w = 2E_0/w$. Here $m_z^\star = 0.92m_0$ is the longitudinal effective mass of an electron in Si. Using Eqs. (33), (1), and (2), we obtain the transport relaxation rate τ_{IR}^{-1} and quantum rate $\tau_{q,IR}^{-1}$ due to interface roughness scattering. At low electron densities $n_e \ll g/4\pi\Lambda^2$, where $k_F\Lambda \ll 1$ and $q_{TF}/2k_F \gg 1$ are satisfied, we obtain

$$\tau_{IR}^{-1} = \frac{1}{\tau_0}\frac{3\pi}{8}\left(\frac{2k_F}{q_{TF}}\right)^2, \tag{34}$$

$$\tau_{qIR}^{-1} = \frac{1}{\tau_0}\frac{\pi}{4}\left(\frac{2k_F}{q_{TF}}\right)^2, \tag{35}$$

where τ_0^{-1} is given by

$$\tau_0^{-1} = \frac{2\pi^4\hbar}{m^\star}\left(\frac{m^\star}{m_z^\star}\right)^2\frac{\Delta^2\Lambda^2}{w^6}. \tag{36}$$

The corresponding mobility grows in thickness $\mu_{IR} \propto w^6$, which was experimentally demonstrated in modern GaAs electron quantum wells in Ref. [64] (see also Fig. 2(d) in Ref. [65]). We find that both μ_{IR} and μ_{qIR} are proportional to n_e^{-1} at low electron densities $n_e \ll g/4\pi\Lambda^2$. In the high-density range $g/4\pi\Lambda^2 \ll n_e \ll g^3/16\pi a_B^2$, we obtain

$$\tau_{IR}^{-1} = \frac{1}{\tau_0}\frac{3\sqrt{\pi}}{4}\left(\frac{2k_F}{q_{TF}}\right)^2(k_F\Lambda)^{-5}, \tag{37}$$

$$\tau_{qIR}^{-1} = \frac{1}{\tau_0}\frac{\sqrt{\pi}}{4}\left(\frac{2k_F}{q_{TF}}\right)^2(k_F\Lambda)^{-3}, \tag{38}$$

where we find $\mu_{IR} \propto n_e^{3/2}$ and $\mu_{qIR} \propto n_e^{1/2}$ [66], similar to the long-range remote impurity scattering results [cf. Eqs. (15) and (16)]. The crossover of the $\mu_{IR} \propto n_e^{-1}$ to the $\mu_{IR} \propto n_e^{3/2}$ behavior is at $n_e = g/4\pi\Lambda^2$, where μ_{IR} reaches its minimum

$$\mu_{IR,min} \approx \frac{e\tau_0}{m^\star}\frac{q_{TF}^2\Lambda^2}{4} = \frac{e}{8\pi^4\hbar}\left(\frac{m_z^\star}{m^\star}\right)^2\frac{q_{TF}^2w^6}{\Delta^2}. \tag{39}$$

For the high mobility structures reported in Refs. [14] with $w = 7$ nm, we equate the saturation mobility value $\approx 30 \times 10^4$ cm^2 V^{-1} s^{-1} to $\mu_{IR,min}$, and obtain a rough estimate of $\Delta \approx 5$ Å. This value is in reasonable agreement with the roughness height measured from scanning transmission electron microscopy (STEM) images reported in Refs. [13,14].

Figures 4(a) and 4(b) are calculated using the combination of RI and IR scattering, while (c) and (d) are calculated using the combination of BI and IR scattering. When we fit the experimental data, we restrict the RI (BI) concentration to be the same for both the 5 nm and 7 nm quantum wells. This restriction is based on the fact that the experimental uncertainty of mobility is relatively small $\sim 10\%$ for different devices with the same structure design [13,14], indicating that the sources of disorder should be similar from device to device. As a result, the best-fit charge impurity concentrations are $n_r = 2.5 \times 10^{12}$ cm^{-2} and $N_1 = 12 \times 10^{15}$ cm^{-3}. To explain the difference in the peak mobility between the 5 nm and 7 nm quantum wells, the roughness parameters are chosen differently within the range $\Delta = 4.5$–8.0 Å and $\Lambda = 16$–28 Å. The roughness parameters that we choose are of the same
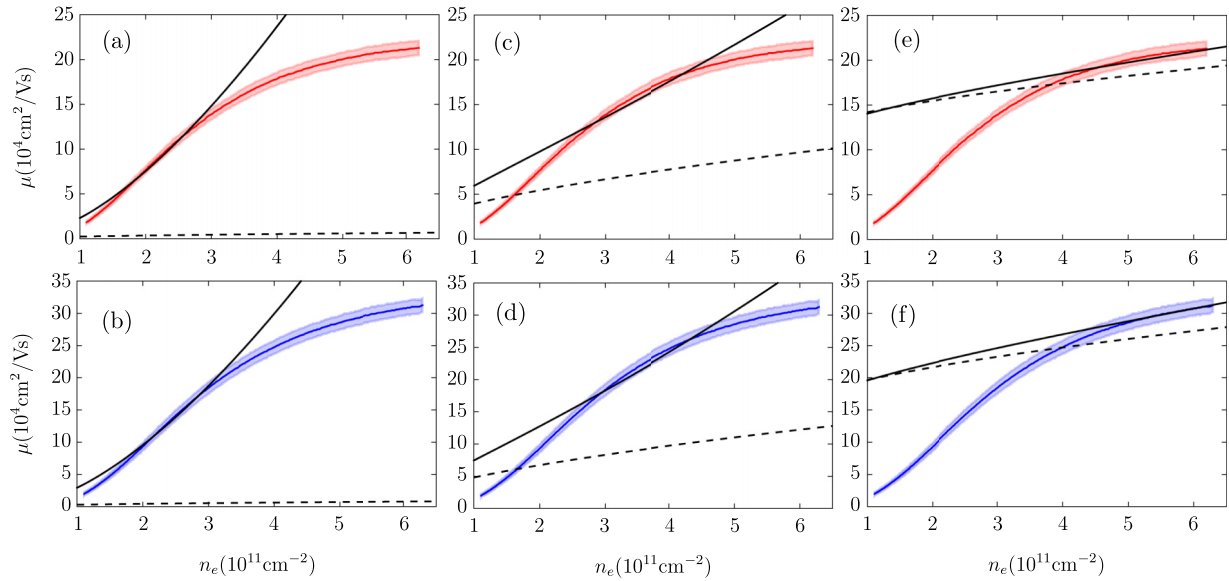
FIG. 5. Mobility (μ) vs electron density (n_e) for quantum wells with widths of 5 nm (red) and 7 nm (blue). The experimental data are depicted by red and blue curves, with corresponding error bars shown as shaded areas. The solid black curves represent the theoretical transport mobility, while the dashed black curves denote the quantum mobility, both calculated using the infinite potential well approximation with the local field correction. (a) and (b) assume only RI near the capping layer $n_r \neq 0$. (c) and (d) assume only BI outside the quantum well $N_1 \neq 0$ and no BI inside $N_2 = 0$, while (e) and (f) assume $N_1 = 0$ and $N_2 \neq 0$. To fit the data, the impurity concentrations and roughness parameters are chosen as (a) $n_r = 2.5 \times 10^{12}$ cm^{-2}; (b) $n_r = 2.1 \times 10^{12}$ cm^{-2}; (c) $N_1 = 9.7 \times 10^{15}$ cm^{-3} and $N_2 = 0$; (d) $N_1 = 8.7 \times 10^{15}$ cm^{-3} and $N_2 = 0$; (e) $N_1 = 0$ and $N_2 = 7.2 \times 10^{15}$ cm^{-3}; (f) $N_1 = 0$ and $N_2 = 3.8 \times 10^{15}$ cm^{-3}.

order of magnitude as those used in Refs. [27,48,49,63]. From Fig. 4, we see that the two-impurity model combining RI and IR scattering fits the mobility data in a wider density range $n_e = 1.8$–6×10^{12} cm^{-2}, while the other two-impurity model combining BI and IR fits the data in a narrower density range $n_e = 3$–6×10^{12} cm^{-2}. Therefore RI scattering is more likely the dominant scattering mechanism for transport mobility at low densities, whereas at high densities the transport mobility is limited by IR scattering. This behavior is similar to that in Si-SiO$_2$ MOSFET 2DEG systems, where the IR scattering is much larger leading consequently to lower maximum mobilities.

In the above calculations, we ignore the local field correction $G(q) = q/(g\sqrt{q^2 + k_F^2})$ in the dielectric function Eq. (5). Results that include the local field correction are presented in Figs. 5 and 6. Incorporating this correction leads to a reduction in the estimated impurity concentrations and the ratio of τ/τ_q. This can be understood because the local field correction weakens the screening at a large momentum $q \gtrsim k_F$, so the disorder is more effective in scattering the electrons. Given that τ is more sensitive to large-momentum scattering compared to τ_q, the local field correction consequently reduces τ to a greater extent than τ_q for the same disorder configuration, leading to a smaller τ/τ_q. To analytically incorporate the local field correction $G(q)$ into the scattering rate calculations, we evaluate the q integral in Eqs. (1) and (2) in an approximation where the square-root singularity at $q = 2k_F$ is integrated, but the form factors and $G(q)$ are taken at $q = 2k_F$. As a result, the expressions of charged impurity scattering rates Eqs. (15), (16), (20), and (25)–(27) should be multiplied by an additional factor $F_0(2k_F w)^2[(2k_F/q_{TF})\epsilon(2k_F)]^{-2}$, where the

local field correction $G(q)$ enters the dielectric function $\epsilon(q)$ defined in Eq. (5). At low densities $k_F \Lambda \ll 1$, the expressions of IR scattering rates Eqs. (34) and (35) should be multiplied by $[(2k_F/q_{TF})\epsilon(2k_F)]^{-2}$. While at high densities $k_F \Lambda \gg 1$, the typical momentum transfer is Λ^{-1} so Eqs. (37) and (38) should be multiplied by $[(2/q_{TF}\Lambda)\epsilon(2/\Lambda)]^{-2}$.

Now we address the justification of the infinite potential well approximation employed in our calculations. This approximation is justified under two conditions. The first condition is that the subband gap of the infinite potential well, $E_{sg} = 3\hbar^2\pi^2/2m_z^\star w^2$, should satisfy the inequality $E_F \ll E_{sg} \ll V_0$, where E_F is the Fermi energy of the 2DEG and V_0 is the conduction band offset of the quantum well. In the case of the 7 nm quantum well reported in Ref. [14], $E_{sg} \approx 25$ meV, $E_F < 4$ meV, and $V_0 \approx 180$ meV [3,36], so the first condition $E_F \ll E_{sg} \ll V_0$ is met. The second condition is that the asymmetric electric field from the top gate should not significantly change the effective well width. If the electric field is strong enough to tilt the bottom of the conduction band toward a triangular potential well, then the effective well width becomes smaller. Using the Fang-Howard wave function [67], we can roughly estimate the effective well width at the highest density where the effect of the electric field $E_{\text{gate}} = 4\pi e n_e/\kappa$ is the strongest. At the highest density $n_e = 6 \times 10^{11}$ cm^{-2}, the estimated smallest effective well width is $w_{\text{FH}} = 6/b \approx 5.5$ nm, where $b = (48\pi m_z^\star e^2 n_e/\kappa\hbar^2)^{1/3}$. Given that the estimated smallest effective well width from the Fang-Howard wave function, $w_{\text{FH}} \approx 5.5$ nm, does not significantly differ from the original well width $w = 5$–7 nm, it suggests that the bending of the conduction band bottom due to the electric field from the top gate should not substantially
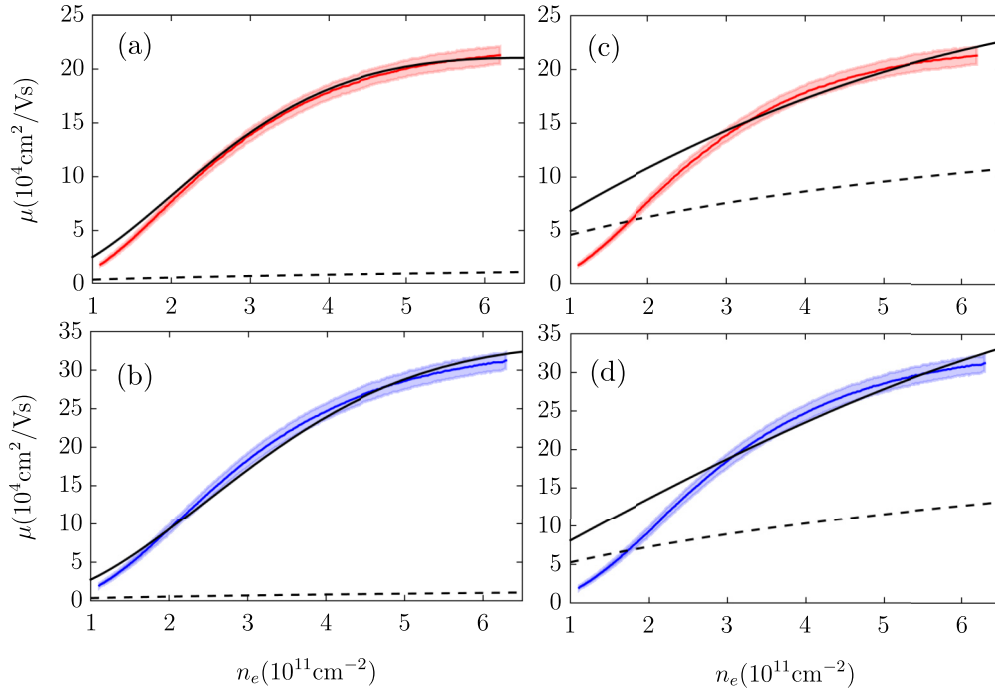
FIG. 6. Mobility (μ) vs electron density (n_e) for quantum wells with widths of 5 nm (red) and 7 nm (blue). The experimental data are depicted by red and blue curves, with corresponding error bars shown as shaded areas. The solid black curves represent the theoretical transport mobility, while the dashed black curves denote the quantum mobility, both calculated using the infinite potential well approximation with the local field correction. Subfigures (a) and (b) incorporate remote impurity and interface roughness scattering, whereas (c) and (d) consider background impurity and interface roughness scattering. Fitting parameters: (a) $n_r = 2.0 \times 10^{12}$ cm^{-2}, $\Delta = 3.5$ Å, and $\Lambda = 30$ Å; (b) $n_r = 2.0 \times 10^{12}$ cm^{-2}, $\Delta = 6.5$ Å, and $\Lambda = 30$ Å; (c) $N_1 = 8.1 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 3$ Å, and $\Lambda = 22$ Å; (d) $N_1 = 8.1 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 6$ Å, and $\Lambda = 22$ Å.

alter the results. If the density is further increased such that $w_{\text{FH}} \ll w$ (the wide quantum well limit), we should replace the well width w by w_{FH} in the expression of the subband gap E_{sg} and the scattering rate τ_{IR}^{-1} [cf. Eqs. (34)–(38)]. If the condition $E_F \ll E_{sg} \ll V_0$ is still satisfied, the density dependence of the corresponding mobility acquires an extra factor of $w_{\text{FH}}^6 \propto n_e^{-2}$. Additionally, the mobility increases by a factor of 2 because the wave function accumulates on the upper side of the quantum well in the presence of the top gate. This allows us to disregard the interface roughness on the bottom side of the quantum well. This reasoning justifies the use of the infinite potential well approximation in our calculations. One shortcoming of both the strict 2D and the infinite square well potential is that they cannot incorporate the short-range alloy disorder scattering arising from the SiGe alloy on the barrier outside the well. We address this shortcoming in the next section by using a finite well confinement model where the carrier wave function is allowed to tunnel into the barrier regime.

V. FINITE POTENTIAL WELL

In this section, to calculate the short-range alloy disorder scattering rate, we employ the finite potential well approximation to determine the wave function tail extending outside the quantum well. Furthermore, we recalculate the RI, BI, and IR scattering rates using the finite potential well wave function, which enters the expressions of the form factors Eqs. (5) and

(6) and the lowest subband energy in Eq. (33). The comprehensive results combining the scattering mechanisms of RI (BI), IR and AD are shown in Figs. 7 and 8. The finite potential well Schrödinger equation reads

$$-\frac{\hbar^2}{2m(z)}\frac{d^2\psi(z)}{dz^2} + V(z)\psi(z) = E\psi(z), \qquad (40)$$

where $m(z) = m_z^\star = 0.92\, m_0$ is the longitudinal effective mass for silicon at $|z| < w/2$ and $m(z) = m_B \simeq m_z^\star$ at $|z| > w/2$ inside the Si$_{1-x}$Ge$_x$ barrier, since at $x < 0.85$, Si$_{1-x}$Ge$_x$ alloys are considered as "Si-like" material with almost the same effective masses [68,69]. Note that Eq. (40) is an approximation for the situation with a variable effective mass, which applies very well in our system since $m(z)$ varies little with z. The effective finite potential well is described by

$$V(z) = \begin{cases} 0, & |z| < w/2, \\ V_0 - E_F, & |z| > w/2, \end{cases} \qquad (41)$$

where $V_0 = 180$ meV [3,36] and $E_F = \hbar^2 k_F^2/2m^\star$ is the Fermi energy of the 2DEG. In the following we use V to represent $V_0 - E_F$ for simplicity. The bound state solutions ($E < V$) for the lowest subband are

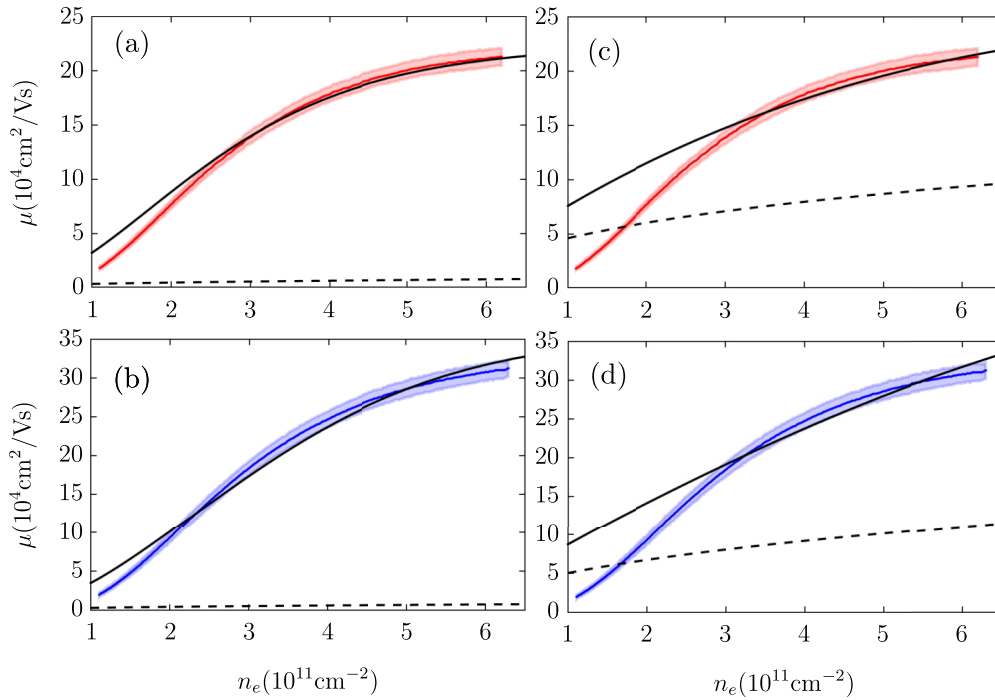$$\psi(z) = \begin{cases} Ce^{-\eta|z|}, & |z| > w/2, \\ D\cos(kz), & |z| < w/2, \end{cases} \qquad (42)$$

FIG. 7. Mobility (μ) vs electron density (n_e) for quantum wells with widths of 5 nm (red) and 7 nm (blue). The experimental data are depicted by red and blue curves, with corresponding error bars shown as shaded areas. The solid black curves represent the theoretical transport mobility, while the dashed black curves denote the quantum mobility, both calculated using the finite potential well approximation without the local field correction. Subfigures (a) and (b) incorporate RI, IR, and AD scattering, whereas (c) and (d) consider BI, IR, and AD scattering. Fitting parameters: (a) $n_r = 2.4 \times 10^{12}$ cm^{-2}, $\Delta = 4.5$ Å, and $\Lambda = 50$ Å; (b) $n_r = 2.4 \times 10^{12}$ cm^{-2}, $\Delta = 9$ Å, and $\Lambda = 37$ Å; (c) $N_1 = 11 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 5$ Å, and $\Lambda = 26$ Å; (d) $N_1 = 11 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 7$ Å, and $\Lambda = 30$ Å.
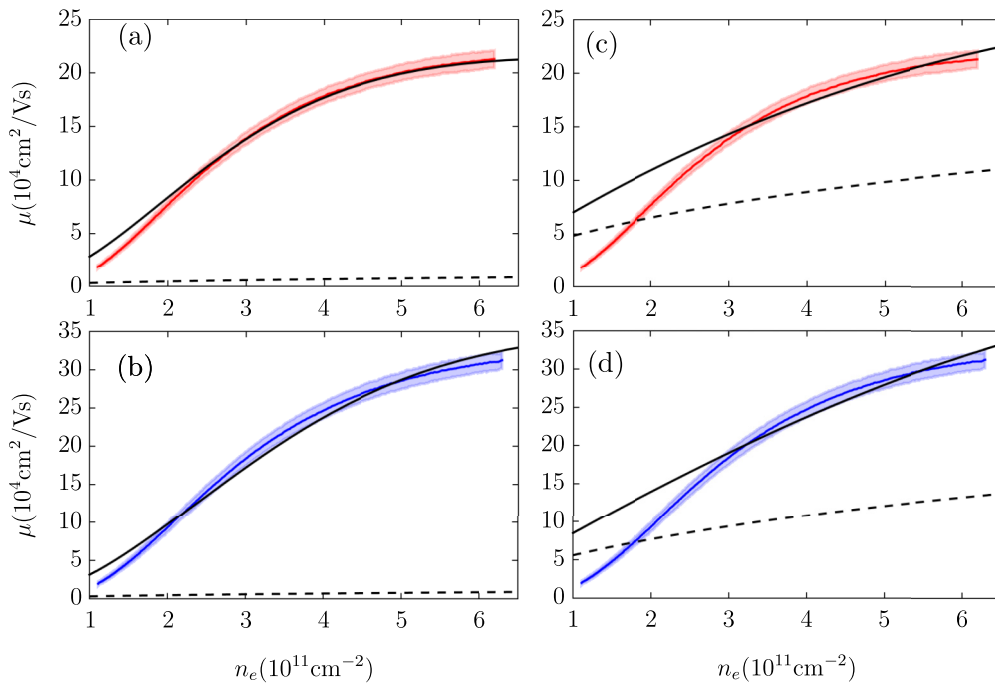

FIG. 8. Mobility (μ) vs electron density (n_e) for quantum wells with widths of 5 nm (red) and 7 nm (blue). The experimental data are depicted by red and blue curves, with corresponding error bars shown as shaded areas. The solid black curves represent the theoretical transport mobility, while the dashed black curves denote the quantum mobility, both calculated using the finite potential well approximation with the local field correction. Subfigures (a) and (b) incorporate RI, IR, and AD scattering, whereas (c) and (d) consider BI, IR, and AD scattering. Fitting parameters: (a) $n_r = 1.9 \times 10^{12}$ cm^{-2}, $\Delta = 4.5$ Å, and $\Lambda = 22$ Å; (b) $n_r = 1.9 \times 10^{12}$ cm^{-2}, $\Delta = 8$ Å, and $\Lambda = 35$ Å; (c) $N_1 = 7.5 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 5$ Å, and $\Lambda = 20$ Å; (d) $N_1 = 7.5 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 7.5$ Å, and $\Lambda = 24$ Å.

where $k = \sqrt{2m_z^\star E/\hbar^2}$ and $\eta = \sqrt{2m_B(V-E)/\hbar^2}$. The boundary condition reads

$$Ce^{-\eta w/2} = D\cos(kw/2), \tag{43}$$

$$\frac{C}{m_B}\eta e^{-\eta w/2} = \frac{D}{m_z^\star}k\sin(kw/2). \tag{44}$$

The normalization of $\psi(z)$ gives

$$C^2\int_{w/2}^\infty e^{-2\eta z}dz + D^2\int_0^{w/2}\cos^2(kz)dz = 1/2. \tag{45}$$

Combining with Eq. (43), one has

$$C = \left[\frac{1}{\eta}e^{-\eta w} + \frac{kw + \sin(kw)}{2k\cos^2(kw/2)}e^{-\eta w}\right]^{-1/2}. \tag{46}$$

Dividing Eq. (44) by Eq. (43), one obtains

$$\tan(kw/2) = \frac{m_z^\star}{m_B}\frac{\eta}{k}. \tag{47}$$

Introducing dimensionless quantities $\tilde{E} = E/E_0$, $\tilde{V} = V/E_0$, where $E_0 = \hbar^2\pi^2/2m_z^\star w^2$ is the ground state energy for infinite well, we arrive at [70]

$$\tilde{E} + \frac{m_B}{m_z^\star}\tilde{E}\tan^2(\sqrt{\tilde{E}}\pi/2) = \tilde{V}. \tag{48}$$

The local energy fluctuation due to interface roughness reads

$$\delta E(\mathbf{r}) = \frac{\partial E}{\partial w}\Delta(\mathbf{r}), \tag{49}$$

where $\Delta(\mathbf{r})$ is the local variation of the well width at a position $\mathbf{r}$. Taking the derivative with respect to w for both sides in Eq. (48), we obtain the expression for $\partial E/\partial w$

$$\frac{\partial E}{\partial w} = -\frac{2E}{w}\left\{1 + g\left(\frac{m_B}{m_z^\star}, \frac{E}{V}\right)\tilde{V}^{-1/2}\right\}^{-1}, \tag{50}$$

where

$$g(x,y) = \frac{2}{\pi}[x^{-1/2} + y(x^{1/2} - x^{-1/2})]^{-1}(1-y)^{-1/2}. \tag{51}$$

Substituting Eq. (50) into Eq. (33), we obtain the scattering potential for the interface roughness for a finite potential well.

We now estimate the correction introduced by the finite potential well model relative to the infinite potential well model. According to Eq. (48), we see that if $\tilde{V} \to \infty$ and $m_B = m_z^\star$, then $\tilde{E} = 1$. For finite values of $\tilde{V}$, we have $0 < \tilde{E} < 1$, implying that the effective well width is larger than w due to wave function penetration into the barriers. For example, at electron density $n_e = 6 \times 10^{11}$ cm^{-2}, we have $\tilde{V} \approx 10$ for a 5 nm QW and $\tilde{V} \approx 20$ for a 7 nm QW. In the limit $\tilde{V} \gg 1$, one can solve Eq. (48) perturbatively and obtain

$$\tilde{E} \approx 1 - \frac{4}{\pi\tilde{V}^{1/2}} + \frac{12}{\pi^2\tilde{V}} + \mathcal{O}(\tilde{V}^{-\frac{3}{2}}), \tag{52}$$

so that $\tilde{E} \approx 0.70$ for a 5 nm QW and $\tilde{E} \approx 0.77$ for a 7 nm QW. The reduced $\tilde{E}$ enters directly into the interface roughness scattering potential Eq. (33) and (50) and lowers the scattering rate by a factor of $\tilde{E}^2$. The parentheses in Eq. (50)

lead to a further reduction in the scattering rate. As a result, $\tau_{\text{IR,inf}}/\tau_{\text{IR,fin}} \approx 0.34$ for a 5 nm QW and $\tau_{\text{IR,inf}}/\tau_{\text{IR,fin}} \approx 0.45$ for a 7nm QW. To account for this reduction in the scattering rate due to the finite well potential, the product of $\Delta\Lambda$ must be increased to fit the mobility data, as evident from Eq. (33). Above, we see that the finite potential well approximation significantly alters the interface roughness scattering. The leading order correction to τ_{IR} is proportional to $\tilde{V}^{-1/2}$, which is substantial given that $\tilde{V}$ is around 10 for a 5 nm QW and 20 for a 7 nm QW. Next, we consider how the finite potential well model alters the background impurity scattering rates. The model alters the form factors as per Eqs. (4) and (5). Since the form factors are proportional to the electron density $|\psi(z)|^2$, the leading correction should be comparable to the leakage probability outside the quantum well

$$\text{Prob}_{|z|>w/2} = \int_{|z|>w/2}|\psi(z)|^2 dz = C^2\eta^{-1}e^{-\eta w}. \tag{53}$$

In the limit of $\tilde{V} \gg 1$, we have

$$\text{Prob}_{|z|>w/2} \approx \frac{2}{\pi\tilde{V}^{3/2}}. \tag{54}$$

Hence, $\text{Prob}_{|z|>w/2} \approx 0.02$ for a 5 nm QW and $\text{Prob}_{|z|>w/2} \approx 0.007$ for a 7 nm QW. This indicates that the finite potential well approximation only marginally modifies the background impurity scattering rates, affecting them by at most a few percent.

Alloy disorder within $\text{Si}_{1-x}\text{Ge}_x$ barriers is another short-range scattering mechanism that contributes equally to the transport and quantum scattering rate [71–73]

$$\frac{1}{\tau_{\text{AD}}} = \frac{m^\star}{\hbar^3}(\Delta E_c)^2\Omega x(1-x)\int_{|z|>w/2}|\psi(z)|^4 dz, \tag{55}$$

where $\Delta E_c \approx 0.8$ eV is the conduction band offset between Si and Ge [74], $\Omega = a^3/4$ is the scatter volume with $a = 5.4$ Å is the lattice constant. In the limit of $\tilde{V} \gg 1$, we have $\int_{|z|>w/2}|\psi(z)|^4 dz \approx 2\tilde{V}^{-5/2}/\pi w$, and the corresponding mobility is given by

$$\mu_{\text{AD}} \approx \frac{w^6 e(2m_z^\star V)^{5/2}}{2\pi^4\hbar^2 m^{\star 2}(\Delta E_c)^2\Omega x(1-x)}. \tag{56}$$

For a $w = 5$ nm quantum well, $\mu_{\text{AD}} \approx 1.4 \times 10^6$ cm^2 V^{-1} s^{-1}, while for a $w = 7$ nm quantum well, $\mu_{\text{AD}} \approx 1 \times 10^7$ cm^2 V^{-1} s^{-1}, which are much larger than the peak mobility reported in Refs. [13,14]. The complete numerical results for alloy disorder scattering using the full finite potential well approximation are shown in Fig. 9, and we see that the order of magnitude agrees with the rough analytical estimation of Eq. (56). Therefore AD scattering is negligible in the high-mobility quantum wells reported in Refs. [13,14], where there is no Ge present inside the well. (For other devices that incorporate Ge concentration within the quantum well [18], AD scattering is much stronger, resulting in an electron mobility of $2-3 \times 10^4$ cm^2 V^{-1} s^{-1}.) Regarding Eq. (55), we follow the approach of Refs. [71,75], omitting the dielectric function due to the short-range nature of AD around a lattice constant. Though several other studies have included the dielectric function in AD scattering
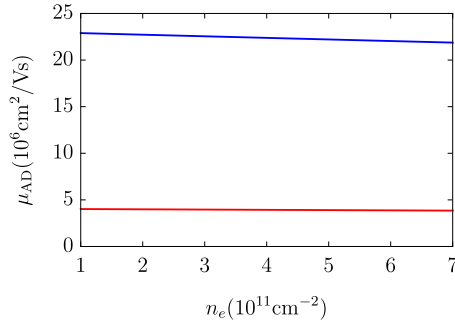
FIG. 9. Mobility and quantum mobility due to alloy disorder μ_{AD} vs electron density n_e for a $w = 5$ nm (red) and $w = 7$ nm (blue) quantum well, calculated using Eq. (55) with the finite potential well approximation.

calculations [49,52,76], the inclusion of dielectric screening would only increase mobility, further diminishing the role of AD. Therefore our conclusion that AD scattering is not a limiting factor in the high-mobility structures reported in Refs. [13,14] remains valid.

The results of mobility and quantum mobility calculated using the finite potential well approximation without the local field correction are shown in Fig. 7. Here, Figs. 7(a) and 7(b) are calculated using a combination of RI, IR, and AD scattering. (c) and (d) are calculated using a combination of BI, IR, and AD scattering. The corresponding results calculated with local field correction are shown in Fig. 8.

VI. GATE SCREENING

In our calculations above, the screening effect of the top gate is considered primarily through the cutoff distance d_s of the background impurity distribution outside the quantum well. In Sec. III, we argued that $d_s = d_g/2$ should be half the distance from the gate to the center of the quantum well because the charged impurities close to the gate are effectively screened by the image charges. In particular, we assume that the gate does not screen the remote impurities. In this section, we refine our approach by incorporating the Coulomb potential of the image charges induced by the gate and compare the mobility results of the RI and BI scattering. The best-fit mobility and the corresponding quantum mobility curves are depicted in Fig. 10, where the fitting concentrations of charge impurities are $n_r = 5.2 \times 10^{12}$ cm^{-2} and $N_1 = 7.5 \times 10^{15}$ cm^{-3}. To explain the difference in the peak mobility between the 5 nm and 7 nm quantum wells, the roughness parameters are chosen differently within the range $\Delta = 4.5$–8.0 Å and $\Lambda = 20$–39 Å. Our analysis confirms that $d_s = d_g/2$ serves as a reasonably accurate approximation for the cutoff distance of the background impurity distribution within the upper SiGe barrier. Furthermore, we observe that the gate-screened RI scattering results align more closely with the quantum mobility data compared to the RI results that disregard gate screening. Previous results shown in Fig. 8 indicate that at a high density of $n_e = 6 \times 10^{11}$ cm^{-2}, the quantum mobility for RI without gate screening is 1.0×10^4 cm^2 V^{-1} s^{-1}. This value is smaller than the experimental data point $(3.0 \pm 0.5) \times 10^4$ cm^2 V^{-1} s^{-1} [21]. On the
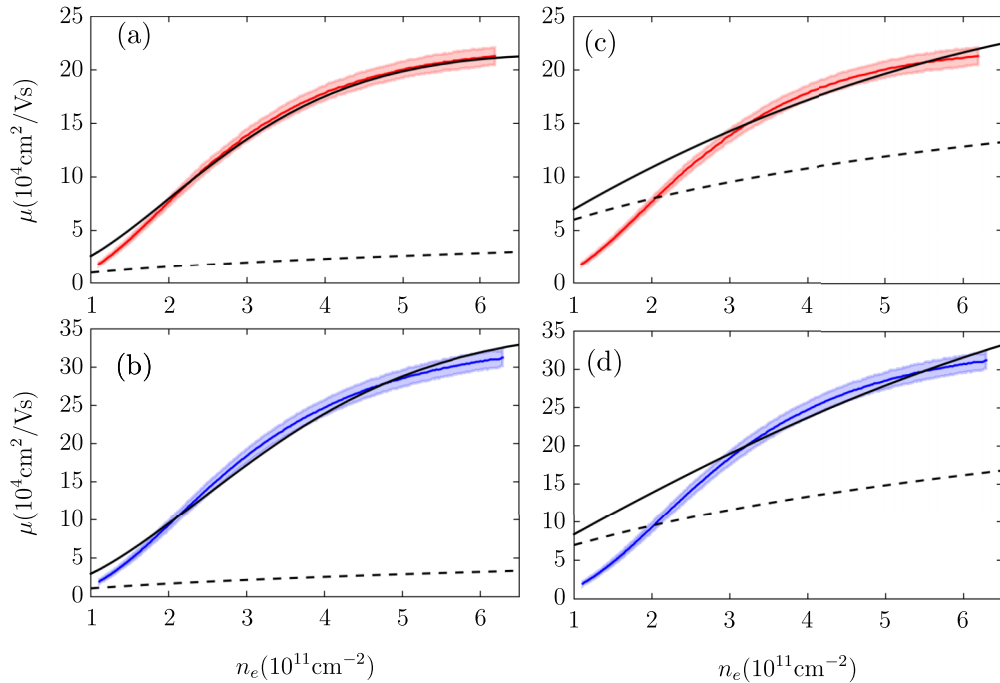

FIG. 10. Mobility (μ) vs electron density (n_e) for quantum wells with widths of 5 nm (red) and 7 nm (blue). The experimental data are depicted by red and blue curves, with corresponding error bars shown as shaded areas. The solid black curves represent the theoretical transport mobility, while the dashed black curves denote the quantum mobility, both calculated using the finite potential well approximation with gate screening and the local field correction. Subfigures (a) and (b) incorporate RI, IR, and AD scattering, whereas (c) and (d) consider BI, IR, and AD scattering. Fitting parameters: (a) $n_r = 5.2 \times 10^{12}$ cm^{-2}, $\Delta = 4.5$ Å, and $\Lambda = 39$ Å; (b) $n_r = 5.2 \times 10^{12}$ cm^{-2}, $\Delta = 8$ Å, and $\Lambda = 36$ Å; (c) $N_1 = 7.5 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 5$ Å, and $\Lambda = 20$ Å; (d) $N_1 = 7.5 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 7.5$ Å, and $\Lambda = 24$ Å.

other hand, as illustrated in Fig. 10, when gate screening is included, the quantum mobility for RI increases to 3.0×10^4 cm^2 V^{-1} s^{-1} at $n_e = 6 \times 10^{11}$ cm^{-2}, aligning well with the experimental data. This enhancement in quantum mobility due to the inclusion of gate screening can be attributed to the suppression of small momentum forward scattering by image charges, consequently reducing the ratio τ/τ_q. Therefore, for a given best-fit mobility curve, the corresponding quantum mobility is expected to be higher.

In the scattering rate calculations, the gate screening is taken care by including the image charge potential into the remote impurity potential Eq. (28) at $|z| > w/2$:

$$U_1(q, z) = \frac{2\pi e^2 F_0(qw)}{\kappa q \epsilon(q)}(e^{-q|z|} - e^{-q|2d_g - z|}), \quad (57)$$

where $2d_g = 2d + 2d_o + w$. As a result, the RI scattering rates Eqs. (15) and (16) are replaced by [77]

$$\frac{1}{\tau_{RI}} = \frac{\pi \hbar n_r}{8m^\star k_F^3 d^3}\left(1 + \frac{d^3}{(2d_g - d)^3} - \frac{2d^3}{d_g^3}\right)\left(\frac{2}{g}\right)^2, \quad (58)$$

$$\frac{1}{\tau_{qRI}} = \frac{\pi \hbar n_r}{2m^\star k_F d}\left(1 + \frac{d}{2d_g - d} - \frac{2d}{d_g}\right)\left(\frac{2}{g}\right)^2. \quad (59)$$

The ratio of the scattering rates is given by

$$\left(\frac{\tau}{\tau_q}\right)_{RI} = (2k_F d)^2 I_g(d, d_g), \quad (60)$$

where the dimensionless function $I_g(d, d_s)$ reads

$$I_g(d, d_g) = \frac{1 + \frac{d}{2d_g - d} - \frac{2d}{d_g}}{1 + \frac{d^3}{(2d_g - d)^3} - \frac{2d^3}{d_g^3}}. \quad (61)$$

A comparison between Eqs. (17) and (60) reveals that the inclusion of gate screening modifies the ratio by a factor of $I_g(d, d_s)$, which is independent of k_F. For the high mobility structures reported in Refs. [13,14], $d_g \approx 44$ nm and $d \approx 30$ nm, which gives $I_g(d, d_g) \approx 0.3$. Thus, incorporating gate screening leads to a decrease in the ratio $(\tau/\tau_q)_{RI}$, and consequently, the corresponding quantum mobility should increase by a factor of 3. This explains the difference in the quantum mobility values between Figs. 8, 10(a), and 10(b).

On the other hand, we see that the results in Figs. 8, 10(c), and 10(d) are very similar to each other, suggesting that our choice of $d_s = d_g/2$ describes the gate-screened BI scattering within the upper SiGe barrier reasonably well. We comment on the cutoff distance in the lower SiGe barrier in the context of the high-mobility structures reported in Refs. [13,14]. When we discuss the gate screening effects for BI in Sec. III, we assume that the cutoff distances are the same from both sides of the SiGe barriers, although in experiments the heterostructure is not symmetric. The bottom barrier (a few μm) is significantly thicker than the upper spacer (30 nm). Given that the top gate can effectively screen charge impurities only within the upper SiGe barrier, the cutoff distance for the lower SiGe barrier remains an open question. However, as indicated by Eq. (26), at high densities, the cutoff distance d_s only logarithmically changes the quantum mobility, without significantly affecting the transport mobility. Therefore the choice of the cutoff distance should not change the fit to the experimental mobility data. By considering the cutoff

distance as an additional fitting parameter, one can always adjust the model to match a specific quantum mobility data point. A more comprehensive future experimental investigation of high-mobility Si quantum wells would be instrumental in enhancing our understanding of the BI distribution within the SiGe barrieres.

In summary, the results of our model, which incorporates gate screening, suggest that RI scattering is likely the predominant scattering mechanism for transport mobility at low densities, whereas at high densities the transport mobility is restricted by IR scattering. The quantum mobility predicted for gate-screened RI scattering aligns closely with experimental data within the margin of error at the highest density examined. This observation suggests that RI scattering could be the limiting factor for quantum mobility across all densities. On the other hand, since the cutoff distance of the BI distribution inside the bottom SiGe barrier remains an open question, we cannot rule out the possibility that some distant background impurities also contribute to the quantum mobility. Future experiments focusing on the quantum mobility over a broader range of densities could yield more quantitative information on the concentration and the cutoff distance of background impurities inside the bottom SiGe barrier.

VII. METAL INSULATOR TRANSITION

In this section, we discuss the mobility data in the low-density regime, and estimate the critical density of the metal-insulator transition (MIT) using the Anderson-Ioffe-Regel (AIR) condition [32,33]. All our results in the previous sections are calculated using the Boltzmann-Born transport theory, which is valid at high densities such that $k_F l \gg 1$, where $l = v_F \tau$ is the transport mean free path. One might expect that the Boltzmann-Born theory fails at a critical density $n_e = n_c$ where $k_F l = 1$. This is the well-known AIR criterion of MIT. However, deviations in the metallic regime of our results, as shown in Fig. 10, suggest that the theoretical mobility curves, fitting well at high densities, start overestimating mobility at low densities compared to the experimental data. It is crucial to emphasize that these low-density deviations are not due to any neglected scattering mechanism in the theory. The dominant scattering mechanism at low densities is RI (or BI, depending on the choice of the impurity model). For RI-dominant cases, as depicted in Figs. 10(a) and 10(b), the deviation is observed at $n_e \lesssim 1.8 \times 10^{11}$ cm^{-2}, where $k_F l \lesssim 30$ or $k_F l_q \lesssim 6$ with $l_q = v_F \tau_q$ [59]. On the other hand, for BI dominant scenarios, shown in Figs. 10(c) and 10(d), the deviation occurs at a higher density $n_e \lesssim 2.8 \times 10^{11}$ cm^{-2}, around $k_F l \lesssim 100$ or $k_F l_q \lesssim 60$. The failure of the Boltzmann-Born theory at low densities is due to the following. Long-range Coulomb disorder leads to the fragmentation of the 2DEG into electron puddles, separated by disorder potential barriers, creating an inhomogeneous 2DEG landscape. Although $k_F l \gg 1$ is locally satisfied within each puddle, the uniform 2DEG assumption, fundamental to linear screening theory, is no longer valid. Instead, transport properties in the low-density region $n_e \geqslant n_c$ near MIT should be described by percolation through charged puddles, following $\sigma \propto (n_e - n_c)^p$ with $p = 1.31$ [23–31], and cannot be captured by the Boltzmann-Born theory even including all high-order scattering processes. The

failure of the Boltzmann-Born theory at low densities also indicates that the critical density estimated from the AIR criterion should be smaller than the actual critical density (percolation threshold) observed in experiments and should be used as a lower bound estimation of the percolation threshold.

Nevertheless, we apply the AIR criterion, characterized by $k_F l = 1$ and $k_F l_q = 1$, to estimate the critical densities n_c and n_{cq} [32,33]. Utilizing the delta-layer 2DEG approximation, as discussed in Secs. III and VI, we can analytically estimate the critical density of the metal-insulator transition (MIT). By substituting the gate-screened remote impurity scattering rates from Eqs. (58) and (59) into the AIR criterion, we derive {gather}

$$n_c^{(\mathrm{RI})} = \frac{g^{1/5}\left[1 + \frac{d^3}{(2d_g-d)^3} - \frac{2d^3}{d_g^3}\right]^{2/5}}{2^{12/5}\pi^{3/5}}\left(\frac{n_r}{d^3}\right)^{2/5}, \quad (62)$$

$$n_{cq}^{(\mathrm{RI})} = \frac{g^{-1/3}\left[1 + \frac{d}{2d_g-d} - \frac{2d}{d_g}\right]^{2/3}}{2^{4/3}\pi^{1/3}}\left(\frac{n_r}{d}\right)^{2/3}. \quad (63)$$

Equations (62) and (63) are applicable if $n_c^{(\mathrm{RI})}$ and $n_{cq}^{(\mathrm{RI})}$ exceed $g/16\pi d^2$. This condition is met in high-mobility quantum wells [13,14] where $g/16\pi d^2 \approx 10^{10}$ cm^{-2}. For BI scattering, substituting the rates from Eqs. (25) and (26) into the AIR criterion yields

$$n_c^{(\mathrm{BI})} = g^{1/5}2^{-8/5}\pi^{-3/5}\left(\frac{N_1}{w^2}\right)^{2/5}, \quad (64)$$

$$n_{cq}^{(\mathrm{BI})} = g^{-1/3}2^{-2/3}\pi^{-1/3}[\ln(2d_s/w)]^{2/3}N_1^{2/3}, \quad (65)$$

which are valid if $n_c^{(\mathrm{BI})}$ and $n_{cq}^{(\mathrm{BI})}$ are greater than $g/16\pi w^2$. Otherwise, we use Eq. (27) and obtain

$$n_c^{(\mathrm{BI})} = n_{cq}^{(\mathrm{BI})} = \frac{\pi}{g}N_1(2d_s - w), \quad (66)$$

which is applicable if $n_c^{(\mathrm{BI})} < g/16\pi w^2$. Given that the crossover density $g/16\pi w^2 \approx 2.4 \times 10^{11}$ cm^{-2} is significantly higher than the MIT critical density in high-mobility quantum wells [13,14], we should use Eq. (66) for estimating n_c in BI-dominated scenarios. Using $d = 30$ nm, $d_s = d_g/2 = 22$ nm, and the impurity concentration $n_r \approx 5.2 \times 10^{12}$ cm^{-2} and $N_1 \approx 7.5 \times 10^{15}$ cm^{-3} from Fig. 10, we obtain the critical densities $n_c^{(\mathrm{RI})} \approx 5 \times 10^{10}$ cm^{-2}, $n_{cq}^{(\mathrm{RI})} \approx 7 \times 10^{10}$ cm^{-2}, $n_c^{(\mathrm{BI})} = n_{cq}^{(\mathrm{BI})} \approx 2 \times 10^{10}$ cm^{-2}.

Next, we estimate the critical density of MIT by solving the AIR conditions $k_F l = (2/g)(hn_e\mu/e) = 1$ and $k_F l_q = (2/g)(hn_e\mu_q/e) = 1$ with μ and μ_q numerically calculated from the best fits shown in Fig. 10. The complete numerical results for n_c and n_{cq} are summarized in the first two rows of Table I. These values closely align with the analytical results from Eqs. (62), (63), and (66). Furthermore, Table I includes results for various quantum degeneracies $g = 4, 2,$ and 1, while keeping the same disorder parameters as used in Fig. 10. The changes in n_c and n_{cq} as functions of g are semi-quantitatively consistent with Eqs. (62), (63), and (66). We should emphasize that our estimate of n_c and n_{cq} using the AIR criterion shown in Table I should serve as a lower bound for the experimental critical density of MIT.

TABLE I. Critical density n_c and n_{cq} in units of 10^{11} cm^{-2} for different quantum degeneracy $g = 4, 2,$ and 1. The results are obtained from $k_F l = 1$ and $k_F l_q = 1$ using the same impurity parameters from Fig. 10 calculated in a finite potential well approximation with gate screening and local field correction. RI (BI) indicates the dominant scattering mechanism at low densities. 5 nm (7 nm) means the calculation for a 5 nm (7 nm) QW.

	RI (5 nm)	RI (7 nm)	BI (5 nm)	BI (7 nm)
n_c ($g = 4$)	0.52	0.49	0.15	0.14
n_{cq} ($g = 4$)	0.65	0.62	0.12	0.11
n_c ($g = 2$)	0.50	0.48	0.28	0.26
n_{cq} ($g = 2$)	0.83	0.79	0.28	0.25
n_c ($g = 1$)	0.49	0.47	0.42	0.38
n_{cq} ($g = 1$)	1.1	1.0	0.48	0.43

Considering that the nature of 2D MIT in high-mobility Si quantum wells likely resembles a density-inhomogeneity-driven percolation transition similar to Si MOSFETs [27], we performed a percolation fit of the experimental data to estimate the experimental critical density. References [13,14] reported the percolation density $n_c = 7 \times 10^{10}$ cm^{-2} ($n_c = 8 \times 10^{10}$ cm^{-2}) for their high-mobility 7 nm (5 nm) Si quantum wells. The percolation density can also be estimated approximately by considering inhomogeneous density fluctuations in the system [27,30,34,35,78,79], leading to $n_c \approx 0.1\sqrt{n_r}/d \approx 7.5 \times 10^{10}$ cm^{-2}, where we used $d = 30$ nm and $n_r \approx 5 \times 10^{12}$ cm^{-2}. This n_c is higher than the one obtained from the AIR criterion. This expression of $n_c \approx 0.1\sqrt{n_r}/d$ can be understood from simple theoretical argument [34], later confirmed by numerical simulations [80]. Namely, n_c is equal to the typical density variation (the charged puddle density) induced by random remote charged impurities within a square of $d \times d$, where the distance of the remote impurity d plays the role of the screening length. Furthermore, we individually performed the percolation fit $\sigma \propto (n_e - n_c)^p$ using the conductivity data $\sigma = en_e\mu$ at low densities $n_e = 1 - 2 \times 10^{11}$ cm^{-2} reported in Refs. [13,14]. The results of the best-fit n_c and p are presented in Table II. The percolation fit with fixed $p = 1.31$ yields $n_c \approx 0.85 \times 10^{11}$ cm^{-2}. Allowing p as a fitting parameter results in $p \approx 2.4$ and $n_c \approx 0.5 \times 10^{11}$ cm^{-2}. Our percolation fit results are in reasonable agreement with those reported in Refs. [13,14]. Comparing Tables I and II, we observe that the RI results provide a more stringent lower bound of the critical density, as $n_c^{(\mathrm{RI})}$

TABLE II. Critical density n_c in units of 10^{11} cm^{-2} obtained by the power law fit $\sigma \propto (n_e - n_c)^p$ to the conductivity data $\sigma = en_e\mu$ at low densities $n_e = 1-2 \times 10^{11}$ cm^{-2} reported in Refs. [13,14]. The parenthesis show the 95% confidence interval of the fitting parameters. The first column fixed $p = 1.31$ assuming the 2D percolation. The second and third columns are results treating p as a free parameter.

	n_c ($p = 1.31$)	n_c	p
5 nm	(0.837, 0.855)	(0.436, 0.514)	(2.20, 2.41)
7 nm	(0.869, 0.889)	(0.447, 0.515)	(2.34, 2.52)

and $n_{cq}^{(\mathrm{RI})}$ are closer to the percolation densities reported in Refs. [13,14]. This suggests that RI scattering, rather than BI, is likely the limiting mechanism at low densities in the studied high-mobility Si quantum wells. Given that $n_{cq}^{(\mathrm{RI})} > n_c^{(\mathrm{RI})}$, we find that the AIR criterion, which includes the single-particle scattering time τ_q [59], provides a more accurate estimate of the percolation density. In the future, a more systematic experimental study of the temperature and density dependence of the conductivity is needed to gain a more comprehensive understanding of the 2D MIT in high-quality Si quantum wells.

VIII. EFFECTS OF SPIN/VALLEY DEGENERACY

In earlier sections, our numerical calculations considered a total quantum degeneracy $g = g_s g_v = 4$. This assumption is valid under two conditions: first, the Fermi energy must substantially surpass the valley splitting, and second, the external magnetic field should be weak enough to avoid inducing spin polarization. In this scenario, both the valley and spin degrees of freedom maintain a twofold degeneracy, indicated by $g_v = g_s = 2$. This section theoretically addresses how the total quantum degeneracy g affects the electron transport properties in the high-mobility Si/SiGe structures for a given disorder configuration. For example, a decrease in degeneracy from $g = 4$ to $g = 2$ can be achieved by applying an in-plane magnetic field B parallel to the 2DEG, which is assumed to only affect the spin degeneracy of the 2DEG without orbital effect if the thickness of the 2DEG is much smaller than the cyclotron radius $w \ll c\hbar k_F/eB$ [41,81–83]. Such a reduced degeneracy is also possible without any spin polarization if the valley splitting happens to be larger than the Fermi energy so that $g_v = 1$ in the system.

In the strong screening limit $q_{\mathrm{TF}}/2k_F \gg 1$ (for the density range of interest to us $n_e \ll g^3/16\pi a_B^2$), mobility increases monotonically with increasing g for a given impurity configuration. This monotonic behavior of mobility versus g has been shown in a comprehensive analysis of valley-dependent 2D transport in (100), (110), and (111) Si inversion layers with the same bare Coulomb disorder [40,41]. A similar phenomenon was observed in an ambipolar gate-controlled Si(111)-vacuum field effect transistor, where by tuning the external gate voltage the 2D electron and hole transport can be studied within the same device against the same background Coulomb disorder [43]. For the ambipolar transistor discussed in Ref. [43], the effective valley degeneracy is 6 for a 2D electron system, whereas it is 1 for a 2D hole system, resulting in the peak electron mobility being approximately 20 times greater than the peak hole mobility at low temperatures. This can be qualitatively understood from the expression of the Thomas-Fermi screening wave vector $q_{\mathrm{TF}} = g/a_B$, implying that higher g leads to improved screening and consequently, higher mobility. However, the relationship between mobility and g is generally nontrivial, as g influences both $q_{\mathrm{TF}} \propto g$ and $k_F \propto g^{-1/2}$. For remote impurity scattering, it is evident from Eqs. (15), (16), (58), and (59) that $\mu_{\mathrm{RI}} \propto g^{1/2}$ and $\mu_{q\mathrm{RI}} \propto g^{3/2}$. On the other hand, for short-range impurity scattering, $\mu_S = \mu_{qS} \propto g^2$, as indicated by Eq. (20). As discussed in Sec. III, the g dependence of background impurity scattering crosses over from short-range to remote impurity

scattering behavior with increasing electron density, as detailed in Eqs. (25)–(27). The crossover density is at $n_e = g/16\pi w^2$. For interface roughness scattering, the mobilities μ_{IR} and $\mu_{q\mathrm{IR}}$ are proportional to g^3 at low densities $n_e < g/4\pi \Lambda^2$, transitioning to a long-range behavior where $\mu_{\mathrm{IR}} \propto g^{1/2}$ and $\mu_{q\mathrm{IR}} \propto g^{3/2}$ at higher densities, as shown in Eqs. (34)–(38). However, IR scattering at low densities $n_e < g/4\pi \Lambda^2$ is typically overshadowed by the dominant scattering mechanism at these densities, which is usually remote Coulomb disorder. At $n_e = g/4\pi \Lambda^2$, the minimum mobility due to IR scattering $\mu_{\mathrm{IR,min}} \propto g^2$, as per Eq. (39). Since the unscreened alloy disorder scattering rate, as given by Eq. (55), depends on k_F only through $V = V_0 - E_F$, the mobility μ_{AD} is not influenced by g provided that $V_0 \gg E_F$ is fulfilled. The results of the above analysis suggest that if a parallel magnetic field is used to lift the spin degeneracy without any orbital effect, the metallic resistance can at most be increased by a factor of 4 [39,82–86].

Figure 11 presents the results of mobility $\mu(n_e)$ for different quantum degeneracies $g = 4$, 2, and 1. For these calculations, the finite potential well approximation with gate screening and local field correction is utilized, employing the same disorder parameters as in Fig. 10. As expected, we see that the mobility decreases monotonically with the reduction in g. The scattering mechanism that limits the transport mobility is IR at high densities, while the limiting factor of mobility at low densities is RI (or BI, depending on the choice of the impurity model). The dashed curves in Fig. 11 replicate the $g = 4$ mobility results from Fig. 10. The solid curves represent our prediction for the mobility in the presence of a parallel magnetic field without orbital effect. In this scenario, the spin is fully polarized ($g_s = 1$), resulting in a total quantum degeneracy of $g = g_v = 2$. The observed ratio of $\mu_{g=4}/\mu_{g=2}$ falls within the range from $\sqrt{2}$ to 4, which aligns with the predicted $g^{1/2}$ to g^2 scaling of μ across various scattering mechanisms. Our choice of the valley degeneracy $g_v = 2$ is appropriate for the high mobility structures reported in Refs. [13,14], where the Fermi energy in the density range $n_e = 1 - 6 \times 10^{11}$ cm^{-2} is $E_F \in (0.6, 3.8) \times (4/g)$ meV. This Fermi energy is significantly greater than the valley splitting of 0.2 meV. Should the valleys split further such that only one valley is occupied, combined with a strong parallel magnetic field, one could achieve a quantum degeneracy of $g = 1$. The dot-dashed curves in Fig. 11 are our predictions in this scenario, indicating a further reduction in mobility.

The quantum degeneracy g also has an effect on the critical density n_c estimated from the AIR criterion. The specific relationship between n_c and g depends on the dominant scattering mechanism responsible for MIT and the dimensionless parameters such as $q_{\mathrm{TF}}/2k_F$, $k_F d$ for RI, $k_F w$ for BI, and $k_F \Lambda$ for IR. These parameters determine the effective screened potential that scatters electrons, thereby altering the MIT critical density. In the strong screening limit $q_{\mathrm{TF}}/2k_F \gg 1$ or $n_e \ll g^3/16\pi a_B^2$, which is of interest to us for the high-mobility structures reported in Refs. [13,14], the critical densities estimated from the AIR criterion for the RI and BI scatterings are summarized in Eqs. (62)–(66). We see that for long-range Coulomb impurity scattering, $n_c \propto g^{1/5}$ while $n_{cq} \propto g^{-1/3}$. For MIT induced by short-range impurity scattering, the
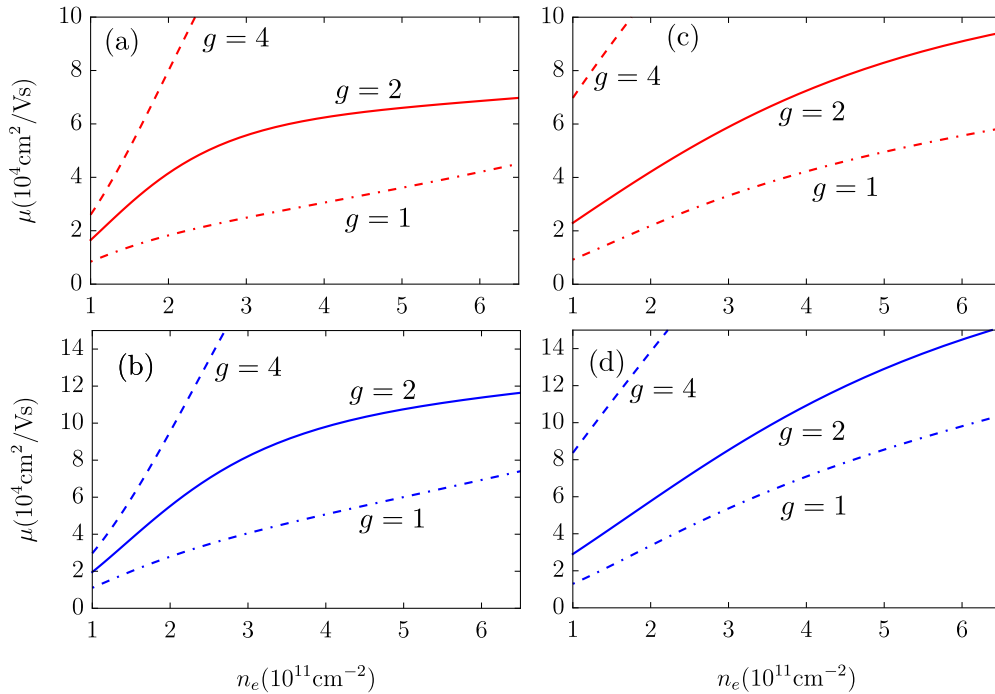
FIG. 11. Mobility (μ) vs electron density (n_e) for quantum wells with widths of 5 nm (red) and 7 nm (blue). The dashed, solid, and dot-dashed curves represent the theoretical mobility assuming the total quantum degeneracy $g = 4$, 2, and 1 respectively. The calculations are done using the finite potential well approximation with gate screening and local field correction. Subfigures (a) and (b) incorporate RI, IR, and AD scattering, whereas (c) and (d) consider BI, IR, and AD scattering. The corresponding disorder parameters are the same as in Fig. 10.

critical density is given by

$$n_c^{(S)} = n_{cq}^{(S)} = \frac{\pi}{g} n_2, \tag{67}$$

which is the same as Eq. (66) with $N_1(2d_s - w)$ replaced by the effective 2D short-range impurity concentration n_2 nearby the 2DEG. For example, Eq. (67) describes the case in Si MOSFETs where n_c increases as g (and μ) decreases [59,78,87]. Table I shows the value of n_c and n_{cq} for the same disorder parameters taken from Fig. 10 but with different g. We should emphasize that the disorder parameters in Fig. 10 are obtained by fitting the experimental data with the theoretical mobility curves using $g = 4$, which is justified since E_F is much larger than the valley splitting in the high-mobility Si quantum wells. For completeness, Appendix presents the best-fit results of the mobility curves and the corresponding critical densities assuming $g = 2$. In this scenario, the fitting disorder parameters are lower than those in Fig. 10 due to weaker screening.

IX. CONCLUSION

In summary, as the main goal of our theoretical work in the context of the importance of the Si/Ge quantum computing platform, we have systematically examined the roles of different scattering sources on the transport and quantum mobilities in the Si/Si$_{0.69}$Ge$_{0.31}$ quantum wells with a co-design of high electron mobility and large valley splitting reported in Refs. [13,14]. We identify remote charged impurities and interface roughness as primary limitations for

transport mobility at low and high electron densities, respectively. Quantum mobility, on the other hand, influenced by both remote and distant background impurities, requires further experimental investigation to understand the charged impurity distribution in SiGe barriers. Our analysis finds alloy disorder scattering to be negligible in high-mobility structures. Other scattering sources, such as inhomogeneous strain distributions and threading dislocations [37] that originated from the bottom metamorphic substrate of several μm, are also quantitatively irrelevant for the reported mobility using similar estimations from Ref. [53,88]. We estimate the metal-insulator transition critical density using the Anderson-Ioffe-Regel criterion and qualitatively explain the breakdown of the Boltzmann-Born theory at low densities due to the long-range Coulomb disorder. We also estimate the critical density by considering inhomogeneous density fluctuations induced by long-range Coulomb disorder in the system and find a larger critical density compared to the one obtained from the Anderson-Ioffe-Regel criterion. To gain a more comprehensive understanding of the metal-insulator transition, more extensive future experimental studies of the temperature and density dependence of the conductivity should be conducted. A brief theoretical prediction for the finite-temperature resistivity is as follows. At low temperature $T < E_F$, it is possible to observe the linear-in-T resistivity due to screened Coulomb impurities (see, for example, Ref. [89] and the references there). On the other hand, the $\ln T$ insulating behavior due to quantum weak localization is difficult to observe because of the linear-in-T metallic Boltzmann resistivity, which may only become apparent at extremely low temperatures outside
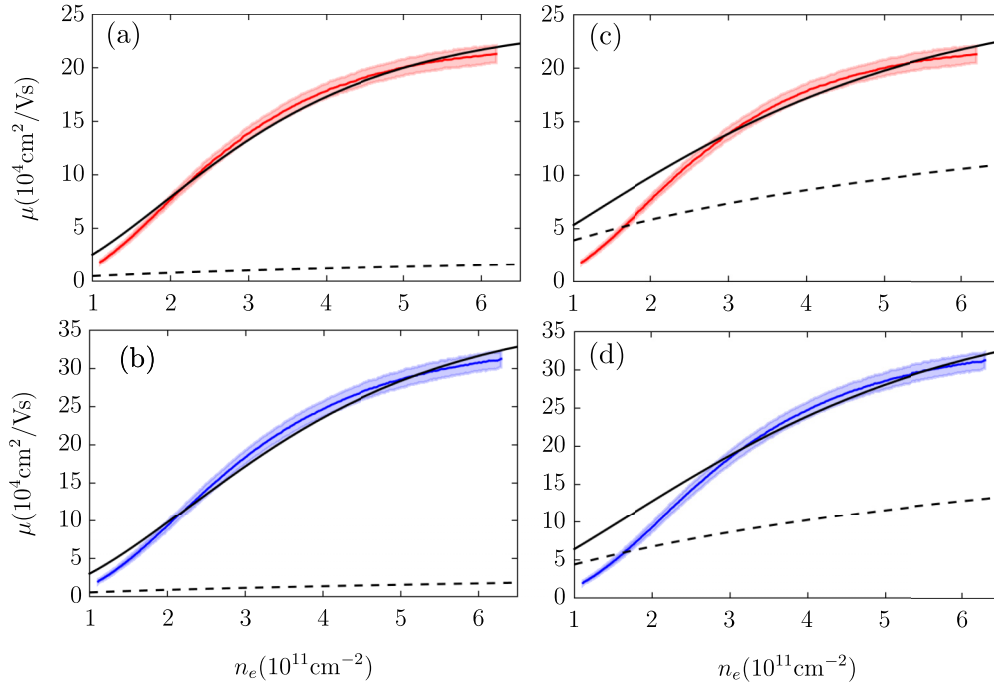
FIG. 12. Mobility (μ) vs electron density (n_e) for quantum wells with widths of 5 nm (red) and 7 nm (blue). The experimental data are depicted by red and blue curves, with corresponding error bars shown as shaded areas. The solid black curves represent the theoretical transport mobility, while the dashed black curves denote the quantum mobility, both calculated using the finite potential well approximation with gate screening and the local field correction, assuming the total quantum degeneracy $g = 2$. Subfigures (a) and (b) incorporate RI, IR, and AD scattering, whereas (c) and (d) consider BI, IR, and AD scattering. Fitting parameters: (a) $n_r = 3.6 \times 10^{12}$ cm^{-2}, $\Delta = 3.5$ Å, and $\Lambda = 20$ Å; (b) $n_r = 3.6 \times 10^{12}$ cm^{-2}, $\Delta = 8$ Å, and $\Lambda = 15$ Å; (c) $N_1 = 3.3 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 4$ Å, and $\Lambda = 15$ Å; (d) $N_1 = 3.3 \times 10^{15}$ cm^{-3}, $N_2 = 0$, $\Delta = 6.5$ Å, and $\Lambda = 18$ Å.

the experimental temperature range [31]. At high temperture $T > E_F$, the resistivity due to screened Coulomb scattering decreases in T, while the acoustic phonon scattering leads to a resistivity increasing linearly in T at $T > T_{\text{BG}} = 2\hbar v_s k_F$ where v_s is the sound velocity. This interplay of mechanisms may lead to a complex, possibly nonmonotonic, resistivity behavior on the metallic side of the 2D MIT at low carrier densities [90]. We also predict the electron mobility with reduced quantum degeneracy, where the spin is fully polarized by an external parallel magnetic field, for future experimental validation.

the total quantum degeneracy $g = 2$. Figure 12 shows the mobility and quantum mobility as a function of the electron density. By comparing Figs. 10 and 12, it is evident that the disorder parameters for $g = 4$ are generally higher than those for $g = 2$, due to the fact that screening increases monotonically as g increases. Table III shows the corresponding results of MIT critical densities. We see that n_c and n_{cq} are in general smaller than the $g = 2$ results shown in Table I, because the values of best-fit mobility and quantum mobility are higher.

ACKNOWLEDGMENTS

APPENDIX: BEST-FIT RESULTS ASSUMING $g = 2$

In this Appendix, we present the best-fit results of mobility curves and the corresponding MIT critical densities assuming

TABLE III. Critical density n_c and n_{cq} in units of 10^{11} cm^{-2}. The results are obtained from $k_F l = 1$ and $k_F l_q = 1$ using the same impurity parameters as in Fig. 12 calculated in the finite potential well approximation with gate screening and local field correction. RI (BI) indicates the dominant scattering mechanism at low densities. 5 nm (7 nm) means the calculation for a 5 nm (7 nm) QW.

	RI (5 nm)	RI (7 nm)	BI (5 nm)	BI (7 nm)
n_c ($g = 2$)	0.44	0.41	0.16	0.15
n_{cq} ($g = 2$)	0.68	0.63	0.15	0.14

[1] L. M. K. Vandersypen, H. Bluhm, J. S. Clarke, A. S. Dzurak, R. Ishihara, A. Morello, D. J. Reilly, L. R. Schreiber, and M. Veldhorst, npj Quantum Inf. **3**, 34 (2017).

[2] G. Scappucci, P. J. Taylor, J. R. Williams, T. Ginley, and S. Law, MRS Bull. **46**, 596 (2021).

[3] G. Burkard, T. D. Ladd, A. Pan, J. M. Nichol, and J. R. Petta, Rev. Mod. Phys. **95**, 025003 (2023).

[4] X. Hu and S. Das Sarma, Phys. Rev. Lett. **96**, 100501 (2006).

[5] D. Culcer, X. Hu, and S. Das Sarma, Appl. Phys. Lett. **95**, 073102 (2009).

[6] S. Ahn, S. Das Sarma, and J. P. Kestner, Phys. Rev. B **103**, L041304 (2021).

[7] D. Culcer, L. Cywiński, Q. Li, X. Hu, and S. Das Sarma, Phys. Rev. B **82**, 155312 (2010).

[8] D. Buterakos and S. Das Sarma, PRX Quantum **2**, 040358 (2021).

[9] F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. L. Hollenberg, G. Klimeck, S. Rogge, S. N. Coppersmith, and M. A. Eriksson, Rev. Mod. Phys. **85**, 961 (2013).

[10] P. Huang and X. Hu, Phys. Rev. B **90**, 235315 (2014).

[11] M. L. V. Tagliaferri, P. L. Bavdaz, W. Huang, A. S. Dzurak, D. Culcer, and M. Veldhorst, Phys. Rev. B **97**, 245412 (2018).

[12] A. M. J. Zwerver, S. V. Amitonov, S. L. de Snoo, M. T. Madzik, M. Rimbach-Russ, A. Sammak, G. Scappucci, and L. M. K. Vandersypen, PRX Quantum **4**, 030303 (2023).

[13] B. Paquelet Wuetz, D. Degli Esposti, A.-M. J. Zwerver, S. V. Amitonov, M. Botifoll, J. Arbiol, A. Sammak, L. M. K. Vandersypen, M. Russ, and G. Scappucci, Nat. Commun. **14**, 1385 (2023).

[14] D. D. Esposti, L. E. A. Stehouwer, Önder Gül, N. Samkharadze, C. Déprez, M. Meyer, I. N. Meijer, L. Tryputen, S. Karwal, M. Botifoll, J. Arbiol, S. V. Amitonov, L. M. K. Vandersypen, A. Sammak, M. Veldhorst, and G. Scappucci, arXiv:2309.02832 [cond-mat.mes-hall].

[15] W. M. Witzel, M. S. Carroll, A. Morello, L. Cywiński, and S. Das Sarma, Phys. Rev. Lett. **105**, 187602 (2010).

[16] M. G. Borselli, R. S. Ross, A. A. Kiselev, E. T. Croke, K. S. Holabird, P. W. Deelman, L. D. Warren, I. Alvarado-Rodriguez, I. Milosavljevic, F. C. Ku, W. S. Wong, A. E. Schmitz, M. Sokolich, M. F. Gyure, and A. T. Hunter, Appl. Phys. Lett. **98**, 123118 (2011).

[17] A. Hollmann, T. Struck, V. Langrock, A. Schmidbauer, F. Schauer, T. Leonhardt, K. Sawano, H. Riemann, N. V. Abrosimov, D. Bougeard, and L. R. Schreiber, Phys. Rev. Appl. **13**, 034068 (2020).

[18] T. McJunkin, B. Harpt, Y. Feng, M. P. Losert, R. Rahman, J. P. Dodson, M. A. Wolfe, D. E. Savage, M. G. Lagally, S. N. Coppersmith, M. Friesen, R. Joynt, and M. A. Eriksson, Nat. Commun. **13**, 7777 (2022).

[19] X. Mi, J. V. Cady, D. M. Zajac, J. Stehlik, L. F. Edge, and J. R. Petta, Appl. Phys. Lett. **110**, 043502 (2017).

[20] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).

[21] G. Scappucci and D. D. Esposti (private communication).

[22] S. Das Sarma and F. Stern, Phys. Rev. B **32**, 8442 (1985).

[23] B. J. Last and D. J. Thouless, Phys. Rev. Lett. **27**, 1719 (1971).

[24] S. Kirkpatrick, Rev. Mod. Phys. **45**, 574 (1973).

[25] B. I. Shklovskii and A. L. Efros, Sov. Phys. Usp. **18**, 845 (1975).

[26] S. D. Sarma and E. Hwang, Solid State Commun. **135**, 579 (2005).

[27] L. A. Tracy, E. H. Hwang, K. Eng, G. A. Ten Eyck, E. P. Nordberg, K. Childs, M. S. Carroll, M. P. Lilly, and S. Das Sarma, Phys. Rev. B **79**, 235307 (2009).

[28] S. Das Sarma, M. P. Lilly, E. H. Hwang, L. N. Pfeiffer, K. W. West, and J. L. Reno, Phys. Rev. Lett. **94**, 136401 (2005).

[29] M. J. Manfra, E. H. Hwang, S. Das Sarma, L. N. Pfeiffer, K. W. West, and A. M. Sergent, Phys. Rev. Lett. **99**, 236402 (2007).

[30] S. Das Sarma, E. H. Hwang, and Q. Li, Phys. Rev. B **88**, 155310 (2013).

[31] S. Das Sarma, E. H. Hwang, K. Kechedzhi, and L. A. Tracy, Phys. Rev. B **90**, 125410 (2014).

[32] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).

[33] A. F. Ioffe and A. R. Regel, in *Progress in Semiconductors*, edited by A. F. Gibson (Wiley, New York, 1960), Vol. 4, pp. 237–291.

[34] A. Efros, Solid State Commun. **67**, 1019 (1988).

[35] F. G. Pikus and A. L. Efros, Zh. Eksp. Teor. Fiz. **96**, 985 (1989) [Sov. Phys. - JETP **69**, 558 (1989)].

[36] F. Stern and S. E. Laux, Appl. Phys. Lett. **61**, 1110 (1992).

[37] The STEM images of the step-graded SiGe buffer layer beneath the quantum well, as reported in Refs. [13,14], demonstrate that strain-release defects and dislocations are limited to the multiple and sharp interfaces within the compositionally graded buffer layer, which emphasizes the overall crystalline quality of the SiGe strain-released barrier below the quantum well.

[38] S. Das Sarma and E. H. Hwang, Phys. Rev. Lett. **83**, 164 (1999).

[39] E. H. Hwang and S. Das Sarma, Phys. Rev. B **72**, 085455 (2005).

[40] E. H. Hwang and S. Das Sarma, Phys. Rev. B **75**, 073301 (2007).

[41] E. H. Hwang and S. Das Sarma, Phys. Rev. B **87**, 075306 (2013).

[42] S. Das Sarma and E. H. Hwang, Phys. Rev. B **90**, 035425 (2014).

[43] B. Hu, M. M. Yazdanpanah, B. E. Kane, E. H. Hwang, and S. Das Sarma, Phys. Rev. Lett. **115**, 036801 (2015).

[44] M. Jonson, J. Phys. C: Solid State Phys. **9**, 3055 (1976).

[45] A. Gold, Phys. Rev. B **50**, 4297 (1994).

[46] E. H. Hwang and S. Das Sarma, Phys. Rev. B **87**, 125411 (2013).

[47] X. Mi, T. M. Hazard, C. Payette, K. Wang, D. M. Zajac, J. V. Cady, and J. R. Petta, Phys. Rev. B **92**, 035304 (2015).

[48] S.-H. Huang, T.-M. Lu, S.-C. Lu, C.-H. Lee, C. W. Liu, and D. C. Tsui, Appl. Phys. Lett. **101**, 042111 (2012).

[49] S. Das Sarma and E. H. Hwang, Phys. Rev. B **89**, 121413(R) (2014).

[50] These results are also true even in the weak screening limit $q_{TF}/2k_F \ll 1$ where the density is high $g/16\pi d^2 \ll g^3/16\pi a_B^2 \ll n_e$, as long as $q_{TF}d \gg 1$ is satisfied.

[51] P. J. Price, J. Vac. Sci. Technol. **19**, 599 (1981).

[52] A. Gold, Phys. Rev. B **38**, 10798 (1988).

[53] D. Monroe, Y. H. Xie, E. A. Fitzgerald, P. J. Silverman, and G. P. Watson, J. Vac. Sci. Technol. **11**, 1731 (1993).

[54] I. A. Dmitriev, A. D. Mirlin, D. G. Polyakov, and M. A. Zudov, Rev. Mod. Phys. **84**, 1709 (2012).

[55] M. Sammon, M. A. Zudov, and B. I. Shklovskii, Phys. Rev. Mater. **2**, 064604 (2018).

[56] C. M. Tanner, Y.-C. Perng, C. Frewin, S. E. Saddow, and J. P. Chang, Appl. Phys. Lett. **91**, 203510 (2007).

[57] A. Gold, Europhys. Lett. **92**, 67002 (2010).

[58] D. Laroche, S.-H. Huang, E. Nielsen, Y. Chuang, J.-Y. Li, C. W. Liu, and T. M. Lu, AIP Adv. **5**, 107106 (2015).

[59] S. Ahn and S. Das Sarma, Phys. Rev. Mater. **6**, L091001 (2022).

[60] A. Gold, J. Appl. Phys. **108**, 063710 (2010).

[61] R. M. Feenstra, M. A. Lutz, F. Stern, K. Ismail, P. M. Mooney, F. K. LeGoues, C. Stanis, J. O. Chu, and B. S. Meyerson, J. Vac. Sci. Technol. **13**, 1608 (1995).

[62] U. Penner, H. Rücker, and I. N. Yassievich, Semicond. Sci. Technol. **13**, 709 (1998).

[63] S. Das Sarma and E. H. Hwang, Phys. Rev. B **69**, 195305 (2004).

[64] Y. J. Chung, K. A. Villegas Rosales, K. W. Baldwin, P. T. Madathil, K. W. West, M. Shayegan, and L. N. Pfeiffer, Nat. Mater. **20**, 632 (2021).

[65] J. Pack, Y. Guo, Z. Liu, B. S. Jessen, L. Holtzman, S. Liu, M. Cothrine, K. Watanabe, T. Taniguchi, D. G. Mandrus, K. Barmak, J. Hone, and C. R. Dean, arXiv:2310.19782 [cond-mat.mes-hall].

[66] These results are also true even in the weak screening limit $q_{TF}/2k_F \ll 1$ where the density is high $g/4\pi\Lambda^2 \ll g^3/16\pi a_B^2 \ll n_e$, as long as $q_{TF}\Lambda \gg 1$ is satisfied.

[67] F. F. Fang and W. E. Howard, Phys. Rev. Lett. **16**, 797 (1966).

[68] F. Schäffler, Semicond. Sci. Technol. **12**, 1515 (1997).

[69] *Properties of Advanced Semiconductor Materials: GaN, AlN, InN, BN, SiC, SiGe*, edited by M. E. Levinshtein, S. L. Rumyantsev, and M. S. Shur (John Wiley & Sons, Inc., New York, 2001), Chap. 6, pp. 149–188.

[70] See for example, Eq. (5c) in Ref. [91] or Sec. 4.9 in Ref. [92].

[71] G. Bastard, Appl. Phys. Lett. **43**, 591 (1983).

[72] B. Laikhtman and R. A. Kiehl, Phys. Rev. B **47**, 10515 (1993).

[73] Y. Huang, B. I. Shklovskii, and M. A. Zudov, Phys. Rev. Mater. **6**, L061001 (2022).

[74] V. Venkataraman, C. W. Liu, and J. C. Sturm, Appl. Phys. Lett. **63**, 2795 (1993).

[75] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Halsted, New York, 1988), Chap. 6.

[76] T. Ando, J. Phys. Soc. Jpn. **51**, 3900 (1982).

[77] We assume the delta-layer approximation $w = 0$ to get the analytical expressions.

[78] S. Das Sarma and E. H. Hwang, Phys. Rev. B **89**, 235423 (2014).

[79] S. Ahn and S. Das Sarma, Phys. Rev. B **107**, 195435 (2023).

[80] A. L. Efros, F. G. Pikus, and V. G. Burnett, Phys. Rev. B **47**, 2233 (1993).

[81] S. Das Sarma and E. H. Hwang, Phys. Rev. Lett. **84**, 5596 (2000).

[82] S. Das Sarma and E. H. Hwang, Phys. Rev. B **72**, 035311 (2005).

[83] S. Das Sarma and E. H. Hwang, Phys. Rev. B **72**, 205303 (2005).

[84] T. Okamoto, K. Hosoya, S. Kawaji, and A. Yagi, Phys. Rev. Lett. **82**, 3875 (1999).

[85] V. T. Dolgopolov and A. Gold, J. Exp. Theor. Phys. Lett. **71**, 27 (2000).

[86] I. F. Herbut, Phys. Rev. B **63**, 113102 (2001).

[87] M. P. Sarachik, Europhys. Lett. **57**, 546 (2002).

[88] K. Ismail, F. K. LeGoues, K. L. Saenger, M. Arafa, J. O. Chu, P. M. Mooney, and B. S. Meyerson, Phys. Rev. Lett. **73**, 3447 (1994).

[89] S. Das Sarma and E. H. Hwang, Sci. Rep. **5**, 16655 (2015).

[90] S. Das Sarma and E. H. Hwang, Phys. Rev. B **61**, R7838 (2000).

[91] A. Gold, Z. Phys. B **74**, 53 (1989).

[92] J. H. Davies, *The Physics of Low-Dimensional Semiconductors: An Introduction* (Cambridge University Press, New York, 1997).