

Roughness scattering induced insulator-metal-insulator transition in a quantum wire

Han Fu, M. Sammon,* and B. I. Shklovskii

Fine Theoretical Physics Institute, University of Minnesota, Minneapolis, Minnesota 55455, USA



(Received 26 May 2017; revised manuscript received 17 November 2017; published 10 January 2018)

We investigated theoretically the influence of interface roughness scattering on the low-temperature mobility of electrons in quantum wires when electrons fill one or many subbands. We find that the Drude conductance of the wire with length \mathcal{L} first increases with increasing linear concentration of electrons η and then decreases at larger concentrations. For small radius R of the wire with length \mathcal{L} the peak of the conductance G_{\max} is below e^2/h so that electrons are localized. The height of this peak grows as a large power of R , so that at large R the conductance G_{\max} exceeds e^2/h and a window of concentrations with delocalized states (which we call the metallic window) opens around the peak. Thus, we predict an insulator-metal-insulator transition with increasing concentration for large enough R . Furthermore, we show that the metallic domain can be subdivided into three smaller domains: (1) single-subband ballistic conductor, (2) many-subband ballistic conductor, and (3) diffusive metal, and use our results to estimate the conductance in these domains. Finally, we estimate the critical value of $R_c(\mathcal{L})$ at which the metallic window opens for a given length \mathcal{L} and find it to be in reasonable agreement with experiment.

DOI: [10.1103/PhysRevB.97.035304](https://doi.org/10.1103/PhysRevB.97.035304)

I. INTRODUCTION

Semiconductor nanowires have attracted lots of attention due to their potential applications, such as field-effect transistors, elementary logic circuits, resonant tunneling diodes, light-emitting diodes, lasers, and biochemical sensors [1,2]. Advances in the nanowire growth have also led to the development of novel quantum devices [3–7], which allow the exploration of mesoscopic transport in highly confined systems. Recently, hybrid superconductor-semiconductor nanowire devices have been identified [8,9] as a platform to study Majorana end modes [10], which exhibit topological properties [11–13]. To further improve this topological system, a reduction of the disorder in the nanowire is essential [14,15]. The performance of some of these devices is limited by scattering of electrons on surface roughness [16–21]. A theory of roughness scattering limited mobility of nanowires as a function of their radius R and linear electron concentration η controlled by a back gate would be helpful. In spite of some attempts to create such a theory [22–24] the full picture of roughness limited transport in nanowires currently is not available. This is not surprising because, as we show below, even in the case of quantum wells there are big gaps in the roughness limited mobility theory; namely, for wells with many subbands filled. In this paper, we fill the gaps in the theoretical description of roughness limited mobility both for quantum wells and quantum wires.

Much of the focus in nanowire technology is in creating ballistic nanowires that can support the Majorana zero edge modes for quantum computation [14,15]. We show below that the possibility to achieve ballistic transport depends strongly on the radius R and the length \mathcal{L} of the wire. Namely, we show that, for a fixed \mathcal{L} , there exists a critical value $R_c(\mathcal{L})$ such

that electrons in wires with $R < R_c(\mathcal{L})$ are localized, while for $R > R_c(\mathcal{L})$ there is a window of concentrations where a metallic phase exists.

Before addressing why such a window exists, let us describe the conventional models of roughness developed for quantum wells. In a quantum well confined by interfaces at $z = 0$ and $z = L$, the surface roughness is a random shift of the interface position $\Delta(\vec{r})$ from the average level so that $\langle \Delta(\vec{r}) \rangle = 0$, where $\vec{r} = (x, y)$ is the coordinate in the $z = 0$ (or $z = L$) interface plane. The roughness is described by the height correlator and its Fourier transform

$$\begin{aligned} \langle \Delta(\vec{r}) \Delta(\vec{r}') \rangle &= W(\vec{r} - \vec{r}'), \\ \langle |\Delta(q)|^2 \rangle &= W(q). \end{aligned} \quad (1)$$

The first theories of surface roughness scattering assumed the correlator to be Gaussian [25–29]:

$$\begin{aligned} W(\vec{r} - \vec{r}') &= \Delta^2 e^{-\sqrt{2}|\vec{r} - \vec{r}'|^2/d^2}, \\ W(q) &= \pi \Delta^2 d^2 e^{-q^2 d^2/4}. \end{aligned} \quad (2)$$

However, experimental observations using transmission electron microscopy (TEM) and scanning tunneling microscope (STM) measurements of Si/SiO₂ interfaces and InAs/GaSb interfaces found that the spatial correlations follow an exponential behavior [30,31]:

$$\begin{aligned} W(\vec{r} - \vec{r}') &= \Delta^2 e^{-\sqrt{2}|\vec{r} - \vec{r}'|/d}, \\ W(q) &= \pi \Delta^2 d^2 (1 + q^2 d^2/2)^{-3/2}. \end{aligned} \quad (3)$$

This correlator describes randomly distributed flat islands of typical thickness Δ and diameter d on the top of the last complete layer of the crystal [32]. On the other hand, Gaussian roughness can be visualized as randomly positioned stacks of total height Δ and diameter d made of progressively smaller islands of flat atomic layers on the top of bigger ones [32],

*sammo017@umn.edu

TABLE I. Mobility μ in units of $(e/\hbar)(d^4/\Delta^2)$ as a function of the linear electron concentration η at $d < a_B$ for different regions.

| A | B | C |
|--------------------------|--|--------------------------------------|
| $a_B^2 R^2 / \eta^2 d^6$ | $a_B^{8/5} R^{11/5} / d^6 \eta^{11/5}$ | $a_B R / \eta d^3$ |
| D | E | F |
| $a_B^2 R^3 / \eta d^6$ | $a_B^{1/2} R^{11/3} / \eta^{11/6} d^6$ | $a_B^{1/2} R^{5/3} / \eta^{5/6} d^3$ |
| G | H | I |
| $R^7 \eta / d^6$ | $R^{14/3} / \eta^{4/3} d^6$ | $R^{8/3} / \eta^{1/3} d^3$ |

similar to the ancient Mayan pyramids. As we show below, in many cases the two correlators give the same expression for the mobility in terms of Δ and d , and so the difference in parameter values can have serious implications. Only at very large electron densities when $k_F d \gg 1$ (k_F is the Fermi wave number), do the two correlators give different expressions for the mobility. This difference is relatively unimportant for this work, so we give results only for the exponential correlator.

While the above isotropic roughness models were designed for quantum wells with flat interfaces, they are valid for quantum wires of characteristic size $R > d$.¹ In most of this paper we deal with such roughness. However, TEM images of InAs wires [21] suggest that in quantum wires another model of roughness in which the radius of the wire varies along its axis may be more realistic. We discuss this “variable-radius model” (VRM) and its implications in Sec. VI.

In this paper we consider wires with linear electron concentration η doped by a relatively distant back gate (we assume that there are no chemical donors in the wire). Then the interplay between the concentration η , the radius of the wire R , and the semiconductor Bohr radius a_B determines the number of filled subbands of radial quantization, the Fermi wave number k_F of electrons, and whether the confinement is electrostatic or by the surface barriers (referred to as geometric confinement). Here the effective Bohr radius $a_B = \kappa \hbar^2 / m^* e^2$, where κ is the effective dielectric constant, \hbar is the reduced Planck constant, and m^* is the effective electron mass. This means that, for quantum wires, there are five lengths Δ , d , η^{-1} , R , and a_B , or four dimensionless lengths when all are scaled by a_B , that determine the Drude mobility.

Below we use the scaling theory to calculate the low-temperature roughness limited Drude mobility μ in units $(e/\hbar)(d^4/\Delta^2)$ as a function of the dimensionless variables R/a_B and ηa_B . Here the use of Drude’s name signifies that we ignore interference effects and electron-electron correlations. We summarize our results for different regions in Fig. 1 as a “phase diagram” in the plane of R/a_B and ηa_B , the details of which are elaborated in Sec. IV. For the most interesting case $\Delta \ll d \ll a_B$ we find a total of nine regions A–I whose

¹For the case of cylindrical wires, the characteristic size R would be the radius of the wire. However, the results presented are applicable to any cross section that can be described with a single characteristic length, such as a square wire with side length $2R$ or a regular hexagonal wire in which R is the distance from the center of the wire to each vertex.

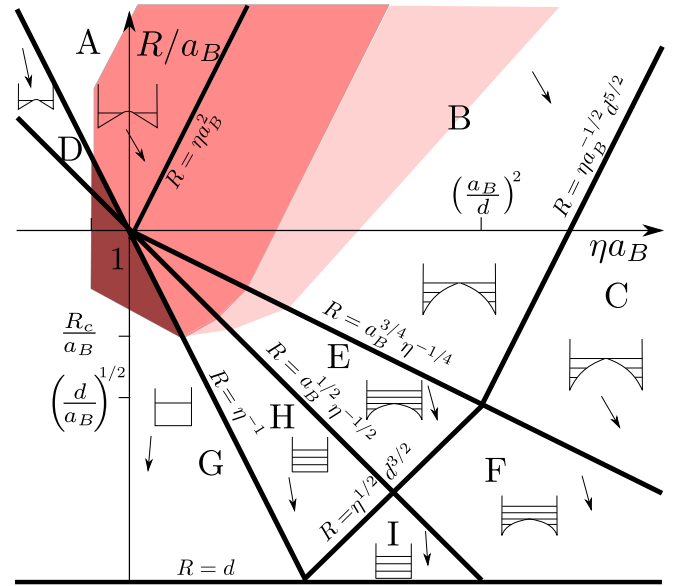


FIG. 1. The scaling phase diagram of roughness limited electron Drude mobility of a long quantum wire plotted as a function of radius R and linear electron concentration η for $d < a_B$ on a log-log scale. Different “phases” or regions are denoted by capital letters. Drude mobility expressions corresponding to these regions are given in Table I. Region boundaries are given by the equations next to them. The schematic self-consistent electron potential-energy profile along the wire diameter and subbands occupied by electrons are shown for each region. Small arrows show the direction of mobility decrease in each region. The colored areas illustrate where the wire of length \mathcal{L} is metallic. The dark-red, light-red, and pink regions correspond to the single-subband ballistic conductor, many-subband ballistic conductor, and diffusive metal regions for $R_c(\mathcal{L}) < a_B$, respectively. Electrons are localized in all the colorless regions. The border between them and colored regions is determined by the length of the wire \mathcal{L} . We assumed that $\mathcal{L} \sim 1 \mu\text{m}$ as is used in quantum devices. For shorter wires $R_c(\mathcal{L})$ decreases, and the colored metallic regions expand to cover most of the area of the phase diagram.

mobilities are listed in Table I. Note that, due to the limitations of the scaling theory, the mobility expressions for the different regions of the phase diagram are valid only away from the borders between different regions. In the vicinity of the border between two regions, there is a smooth crossover between the two mobilities, the details of which are beyond the scope of this paper. While the scaling approach only gives the dependence of mobility on the different parameters without numerical precision, its simplicity allows for a clear picture of the different physical domains and the approximate limits under which they occur.

Now we are ready to address the origin of the Drude conductance peak which leads to a metallic window for large R , illustrated by the colored regions in Fig. 1. Schematic plots of the Drude conductance (in units e^2/h) $G = \eta \mu h / \mathcal{L} e$ of the wire with length \mathcal{L} are shown in Fig. 2 for two representative values of R by full lines. They are obtained from cross sections of Fig. 1 and the mobilities in Table I. At low concentrations, we see that the Drude conductance increases with increasing concentration. This corresponds to region G of Fig. 1, where there is a single radial subband occupied and the electrons are

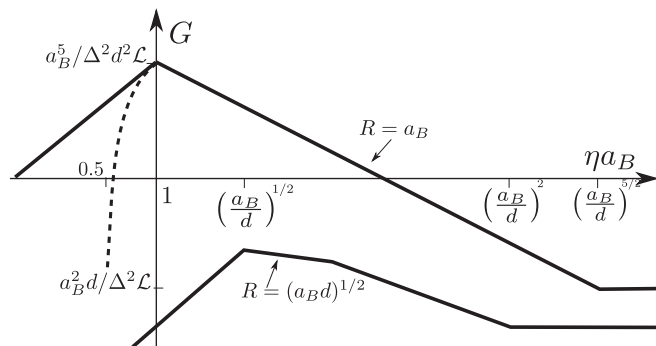


FIG. 2. The scaling behavior of the dimensionless Drude conductance of a quantum wire with length \mathcal{L} and radius R as a function of the linear electron concentration η at different wire radii for $d \ll a_B$ on a log-log scale (full lines). The upper curve corresponds to $R = a_B$ and the lower one is $R = (a_B d)^{1/2}$. They are obtained from cross sections of the phase diagram in Fig. 1 and the mobilities in Table I. The dashed line on the upper curve shows the metal-insulator crossover near $\eta a_B = 0.5$ induced by electron-electron interactions. We see that for $R = a_B$ the metallic window is open, while for $R = (a_B d)^{1/2}$ the window is closed. $\mathcal{L} = a_B^{7/2} \Delta^{-2} d^{-1/2}$ was chosen.

confined geometrically. We know from Fermi's golden rule that the relaxation time τ is inversely proportional to the density of states at the Fermi energy, which in the one-dimensional (1D) case goes like $1/k_F \sim 1/\eta$. The scattering potential, however, is independent of concentration in this regime. Therefore, the relaxation time τ , the mobility μ , and conductance G increase with concentration due to the decrease in the density of states. This trend continues until the concentration is large enough that multiple subbands become occupied. Now electrons have more states to scatter into, and the relaxation time quickly decreases with increasing concentration. Thus the conductance peaks at the border concentration η_c when electrons begin to populate multiple subbands. The peak of the Drude conductance for the most interesting cases of $R \leq a_B$ is given by

$$G_{\max} = \frac{R^5}{\mathcal{L} \Delta^2 d^2}. \quad (4)$$

So far we have ignored electron-electron interactions and quantum interference effects. They dramatically change the conductivity of one-dimensional systems at low temperatures. For single subband wires (regions D and G), electron-electron interactions result in Wigner-crystal-like correlations and pinning of the electron gas leading to the metal-insulator crossover near $\eta a_B = 0.5$ [33]. In Fig. 2 the corresponding collapse of conductance at $\eta a_B < 0.5$ is shown by the dashed lines. According to Luttinger liquid theory [34,35] similar effects persist at very low temperatures in very long wires. We are interested here in relatively short wires with $\mathcal{L} \sim 1 \mu\text{m}$, where plasmon quantization does not allow such effects to develop [35]. Therefore, for $\eta a_B > 1$ we can ignore electron-electron interactions. However, in this case we should still take into account quantum interference effects. They lead to one-electron localization when Drude $G < 1$. This means that, when $G_{\max} < 1$ (see lower curve in Fig. 2), the wire is an insulator at any concentration η . On the other hand, for $G_{\max} > 1$ (see upper curve in Fig. 2) the wire has a

TABLE II. Metal-insulator border $R_{MI}(\eta)$, ballistic-diffusive border $R_{BD}(\eta)$, and the total number of subbands K_{\max} for regions G, H, E, and B of Fig. 1.

| Region | $R_{MI}(\eta)$ | $R_{BD}(\eta)$ | K_{\max} |
|--------|--|---|--------------------------------|
| G | $\eta^{-2/7} R_c(\mathcal{L})^{5/7}$ | | 1 |
| H | $\eta^{1/14} R_c(\mathcal{L})^{15/14}$ | $\eta^{2/13} R_c(\mathcal{L})^{15/13}$ | $(\eta R)^{2/3}$ |
| E | $a_B^{-3/22} \eta^{5/22} R_c(\mathcal{L})^{15/11}$ | $a_B^{-3/7} \eta^{5/7} R_c(\mathcal{L})^{15/7}$ | $(\eta R)^{2/3}$ |
| B | $a_B^{-8/11} \eta^{6/11} R_c(\mathcal{L})^{25/11}$ | $a_B^{-7/9} \eta R_c(\mathcal{L})^{25/9}$ | $\eta^{3/5} R^{2/5} a_B^{1/5}$ |

concentration window of metallic behavior. The critical radius $R_c(\mathcal{L})$ in which the metallic window opens is then determined by the condition that $G_{\max} = 1$. For G_{\max} defined by Eq. (4) we find

$$R_c(\mathcal{L}) = (\Delta^2 d^2 \mathcal{L})^{1/5}. \quad (5)$$

Note that the restriction that $\eta a_B > 0.5$ necessary for the single subband wires to be metallic requires $R_c < 2a_B$.

Thus, we predict a zero-temperature reentrant insulator-metal-insulator transition with increasing η in quantum wires with $R > R_c(\mathcal{L})$. Such a transition was first predicted for a two-dimensional electron gas (2DEG) in silicon MOSFETs [36]. However it was later shown [32] that there is no second reentrant metal-insulator transition at large concentrations of a 2DEG because the dimensionless conductance saturates at a value larger than unity. As our paper shows, the idea of Ref. [36] is realized in quantum wires. For more details see Sec. II below.

The metallic regimes for a wire with $R_c(\mathcal{L}) < a_B$ are shown in different colors in Fig. 1, while regions where the electrons are localized are left blank. The dark-red, light-red, and pink regions of the metallic regime specify a single-subband ballistic conductor, a many-subband ballistic conductor, and a diffusive metal, respectively. It should be emphasized that the metal-insulator and ballistic-diffusive borders depend on the wire length. With decreasing \mathcal{L} and $R_c(\mathcal{L})$ the colored regions expand dramatically and for short wires eventually cover most of the phase diagram. In Fig. 1 we used $\mathcal{L} \sim 1 \mu\text{m}$ as in Fig. 2, which is typically used in quantum devices (see details in Sec. VII).

The detailed derivation of all the metallic border equations are given in Sec. V and in Table II. Here we give a brief summary of the derivation. Let us begin with the metal-insulator border. For $\eta a_B > 0.5$ this border comes from the condition that the Drude conductance $G_D = e^2/h$ and gives rise to the sequence of border lines between the colored and uncolored regions with minimum at $R_c(\mathcal{L})$ in Fig. 1. For $\eta a_B < 0.5$ there is no metallic regime for the single subband regions (G and H), as illustrated by the vertical line that cuts off the dark-red region of Fig. 1 at low concentrations. This line continues vertically to the asymptotic line $\eta a_B \sim C(R/a_B)$, which can be understood as the Wigner crystallization of the 2DEG at $na_B^2 = C \ll 1$, where $n = \eta/2\pi R$. Finally, we address the ballistic-diffusive border which only exists in the regions with many subbands occupied. Typically, a diffusive metal becomes ballistic when the mean-free path $l = \mathcal{L}$. However, for the many subband regions there is an ambiguity, because we can have different l for different subbands. Fortunately, the conductance in these cases is determined by a small subset of subbands which have

identical l and we define the ballistic-diffusive border by the line where $l = \mathcal{L}$ for these subbands.

Let us discuss the conductance in the different-colored regions of Fig. 1. We begin with the ballistic regimes (red regions of Fig. 1). Here the dimensionless conductance $G \approx 2K$, where K is the number of ballistic channels of a wire with finite length \mathcal{L} , and the factor of two comes from the spin degeneracy. Estimates of K can be found in Sec. V. Within the diffusive regime (pink regions of Fig. 1) $G = (h/e)\eta\mu/\mathcal{L}$, where the mobility is given in Table I. Finally, in the insulating regions electrons are localized at temperature $T = 0$. At finite T wires conduct via phonon-assisted hopping. Calculations of the hopping conductivity are relatively straightforward but are beyond the scope of this paper.

The plan of this paper is as follows: In Secs. II and III we study the roughness limited mobility of quantum wells as a function of their width L and two-dimensional concentration of electrons n arriving at the phase diagram for $\mu(L, n)$ with nine different regions. In Sec. IV we use the quantum well phase diagram to construct the phase diagram $\mu(R, \eta)$ for quantum wells with surface roughness described by Eq. (3). In Sec. V we use our results for the Drude mobility to estimate the wire conductance in the ballistic and diffusive regions. In Sec. VI we discuss quantum wires within the variable-radius model (VRM). In Sec. VII we dwell upon some experimental implications; namely, the peak mobility and the value of radius $R_c(\mathcal{L})$ in which the metallic window opens. We conclude with Sec. VIII.

II. ROUGHNESS LIMITED MOBILITY RESULTS FOR QUANTUM WELLS

To understand the roughness limited mobility of quantum wells, it is convenient to first make clear of that in quantum wells. We start from a quantum well confined by two high potential barriers at $z = 0, L$. It has the two-dimensional (2D) electron concentration n created either by two positive donor layers located symmetrically on both sides of the well or by two symmetric metallic gates. In both cases, at $z = 0, L$ there is an electric field pointing into the well with $|E| = 2\pi ne$, where e is the electron charge. Interplay of effects of the electric field and barrier confinement creates five different types of wells shown in Fig. 3 in regions I–IX. In a narrow well the electric field E plays a minor role in level quantization compared with confining barriers so that we assume that all subbands are geometrically confined in the small- L regions VI, VII, VIII, and IX in Fig. 3. When the concentration is relatively small, electrons occupy only the first subband. At larger n , electrons populate many subbands (see the level schematics in regions VI and VII in Fig. 3). In the wider wells shown in regions I, II, III, IV, and V the electric field becomes important compared with the surface barriers. In turn this leads to the splitting of the electron density in two peaks. With growing L , in the beginning (regions IV and V) this splitting is moderate and affects only the lowest subbands. In regions II and III the splitting results in two separate accumulation layers in response to the electric field each side of the well. Finally, at large L and small n , we again reach the single-subband limit; however, the confinement is electrostatic rather than geometric (region I in Fig. 3).

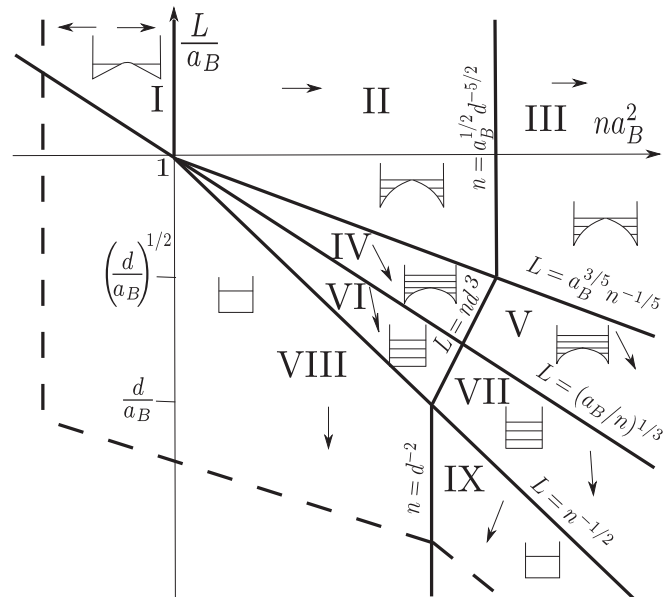


FIG. 3. The scaling phase diagram of roughness limited electron mobility of quantum well at different well width L and 2D electron concentration n for $d \ll a_B$ on a log-log scale. Different “phases” or regions are denoted by Roman numerals. Mobility expressions corresponding to these regions are given in Table III. Region boundaries are given by the equations next to them. The schematic self-consistent electron potential-energy profile along the z axis of wells and levels (subbands) occupied by electrons are shown for each region. Small arrows show the direction of mobility decrease in each region. Apparently the maximum mobility is achieved in region I. The dashed line indicates schematically the border of the metal-insulator transition (MIT) at small enough n . At large n there is no reentrant MIT in spite of the decreasing mobility.

The roughness limited mobility of a single-subband electron gas of a quantum well (regions I, VIII, and IX) was thoroughly studied in Refs. [25–28] more than 30 years ago. On the other hand, the roughness limited mobility of accumulation layers was calculated recently in Ref. [32], the results of which are directly applicable to regions II and III. However, no work has been done in the intermediate regions where many subbands are occupied but the electric field is weak so that some or all of the subbands are confined geometrically (regions IV, V, VI, and VII). In this paper we fill this gap. Below, because of the complexity of the problem, we first present the final results in this section and then give their derivations in next section.

The complete results at $d \ll a_B$ are shown in Fig. 3 and Table III. The single subband results I, VIII, and IX are taken from Refs. [25–28] and the accumulation layer results II and III are from Ref. [32]. For the intermediate regions IV, V, VI, and VII, we obtain their results in this paper.

Let us first look at the physical meaning and corresponding equations of boundary lines in Fig. 3. Across the line between region I and region II, the concentration becomes so large that electrons have to occupy multiple subbands (see level schematics in Fig. 3). With n further increased, $k_F d$ becomes larger than unity in region III where k_F is the three-dimensional (3D) electron Fermi wave number here. Instead of averaging over different islands, the electron hits only a single island

TABLE III. Mobility μ in units of $(e/\hbar)(d^4/\Delta^2)$ as a function of the 2D electron concentration n at $d < a_B$ for different regions.

| I | II | III |
|---------------------------------|-------------------------------|-----------------------|
| a_B^2/n^2d^6 | $a_B^{8/5}/n^{11/5}d^6$ | a_B/nd^3 |
| IV | V | VI |
| $a_B^{1/2}L^{11/6}/n^{11/6}d^6$ | $a_B^{1/2}L^{5/6}/n^{5/6}d^3$ | $L^{10/3}/n^{4/3}d^6$ |
| VII | VIII | IX |
| $L^{7/3}/n^{1/3}d^3$ | L^6/d^6 | $L^6n^{3/2}/d^3$ |

now. This leads to the change of the mobility result at the II-III border.

For regions I, II, and III, all subbands are electrostatically confined. For moderately smaller well width L , some of the subbands become geometrically confined. This happens when the well width L becomes smaller than the characteristic thickness D of the accumulation layer, where [37,38]

$$D \simeq \frac{a_B}{(na_B^2)^{1/5}}. \quad (6)$$

The criterion $L = D$ then gives the line between II, III and IV, V. At the line between IV and V, $k_F d = 1$.

With further reduction of L , all subbands would be geometrically confined (see the level schematic in Fig. 3). This happens when the electrostatically confined distance of the lowest-subband electrons from the surface is equal to the well width L . This distance is $D_0 \simeq a_B^{1/3}/n^{1/3}$ (see Refs. [25,28]), which is the smallest among all subbands since the lowest subband has the smallest kinetic energy in the z direction. The condition $L = D_0$ gives the line between IV, V and VI, VII. The border between VI and VII corresponding to the critical point of $k_F d = 1$ is a continuation of the line between regions IV and V. Moving to even smaller L from regions VI and VII, we cross over to the single subband (see level schematic in Fig. 3). This corresponds to the line $k_F L = 1$ between VI, VII and VIII, IX. The border of the VIII and IX regions is the line of $k_F d = 1$ where k_F is the 2D electron Fermi wave number here. In Fig. 3, one can see that there is another border line between I and VIII, which both correspond to a single subband gas. However, region I corresponds to two electrostatically split electron subbands near the two well interfaces, while region VIII represents the case that the electron subband is spatially restricted by the well width L (see the level schematic in Fig. 3). Their crossover happens at the point that both electrostatic and geometric confinements give the same thickness of the electron gas. Remember that the electrostatically confined thickness of the first subband is D_0 . Then the condition $L = D_0$ gives the border. So this line between I and VIII is just an extension of the line between IV, V and VI, VII.

One should note that, in Table III, all results are shown without numerical coefficients, i.e., we present only the scaling behavior. Previous works have already found the exact coefficients in the single-subband regions I, VIII, and IX [25,26]. Results of many subband regions II, IV, and VI with $k_F d \ll 1$ can also be obtained with the approximate coefficients, as seen later in Sec. III. We cannot get coefficients analytically in

remaining regions III, V, and VII. Thus we focus only on the scaling behavior in all tables and derivations.

In Fig. 3 the metal-insulator transition (MIT) is shown schematically by the dashed lines. Let us dwell on the meaning of these lines. The lower line is related to the localization physics of a noninteracting electron gas. Strictly speaking all states are localized in 2D infinite samples; however, at $k_F l \gg 1$ the localization length grows exponentially as $\zeta = l \exp(k_F l)$, where l is the mean-free path. In finite square samples of area A we have in mind that ζ quickly becomes larger than the sample size $A^{1/2}$. This allows one to discuss the metallic conductivity and expect the insulator-metal transition near $\sigma = (e^2/\hbar) \ln(A^{1/2}/l)$. Ignoring the logarithm and using the expressions of mobility μ for VIII and IX in Table III as well as $\sigma = ne\mu$, one gets that the low- L MIT border of region VIII obeys $L = \Delta^{1/3}d^{1/3}n^{-1/6}$. We also find the MIT border of region IX is $L = \Delta^{1/3}d^{-1/6}n^{-5/12}$. We have used $\Delta/d = (d/a_B)^{8/5}$ in order to draw these lines. The vertical line $na_B^2 = C \ll 1$ reflects the role of the Coulomb interaction between electrons in a degenerate electron gas. At $na_B^2 \ll 1$ strong Coulomb repulsion leads to Wigner crystallization. The Wigner crystal is pinned by relatively small disorder and electrons become localized.

III. ROUGHNESS LIMITED MOBILITY DERIVATIONS FOR QUANTUM WELLS

In the previous section, we have presented the physical picture of all nine regions and their border lines and summarized the mobility results. In this section, we derive the new expressions of mobility for regions IV, V, VI, and VII. First, let us derive μ for region VI. According to Fermi's golden rule and the Boltzmann equation, the relaxation time τ_N of a particular state with the wave function $\xi(z, \vec{r})$ and with in-plane velocity \vec{v}_k in the N th (counted from the bottom lowest subband) subband is

$$\frac{1}{\tau_N} = \frac{2\pi}{\hbar} \sum_{N'} \int \frac{d^2k'}{(2\pi)^2} \frac{|V(q)|^2}{\epsilon(q)^2} \delta(\epsilon - \epsilon_F) \left(1 - \frac{\vec{v}_k \cdot \vec{E} \tau'_N}{\vec{v}_k \cdot \vec{E} \tau_N} \right), \quad (7)$$

where τ_N, τ'_N denote the relaxation time for N, N' th subbands, \vec{v}_k is the in-plane velocity for the final state with the wave function ξ' in the N' subband with in-plane momentum k' , ϵ is the energy of the final state ξ' and ϵ_F is the Fermi energy, q is the transferred momentum in the x - y plane between ξ and ξ' . Here V is the scattering matrix element arising from the scattering potential. Due to the electronic screening, the Fourier transform of the scattering potential $V(q)$ is reduced by the dielectric function $\epsilon(q)$ [25]. One should note that here the last term inside the parentheses of Eq. (7) does not reduce to $\cos \theta$, where θ is the angle between the initial and final total momenta. This is because, due to the 2D nature of the surface roughness and thus of the scattering potential, the multisubband electron gas experiences anisotropic scattering, i.e., different subbands have different relaxation times.² As

²One should note here that the definition of relaxation time τ is still valid according to Ref. [39], which might be broken in more complicated cases.

a result, $\cos\theta$ in Eq. (7) is replaced by the ratio of the out-of-equilibrium part of the distribution function of the states ξ' and ξ represented by $(\vec{v}_{k'} \cdot \vec{E}/\vec{v}_k \cdot \vec{E})(\tau'_N/\tau_N)$ (see Ref. [39]). For brevity, we refer to this term as the distribution function ratio (DFR) from now on.

It is known that the roughness-caused scattering potential $V(\vec{r})$ and corresponding scattering matrix element $V(q)$ satisfy the equation [25,32]

$$V(r) = \frac{\hbar^2}{m^*} \Delta(\vec{r}) \frac{\partial \xi}{\partial z} \frac{\partial \xi'}{\partial z} \Big|_{z=0,L},$$

$$\langle |V(q)|^2 \rangle \simeq \left(\frac{\hbar^2}{m^*} \right)^2 \frac{k_z^2}{Z} \frac{k_z'^2}{Z'} W(q), \quad (8)$$

where $k_z \simeq N/Z$, $k_z' \simeq N'/Z'$ are the z -direction momenta of ξ and ξ' , Z and Z' are the z -direction widths of the N th and N' th subbands, which are determined by the confinement. For example, when the subband N is electrostatically confined, $Z = \varepsilon_z/eE \simeq \hbar^2 k_z^2/m^* e^2 n \simeq a_B k_z^2/n$ (ε_z is the kinetic energy in the z direction), while when geometrically confined, $Z = L$. For region VI, all subbands are geometrically confined. So

$$\langle |V(q)|^2 \rangle \simeq \left(\frac{\hbar^2}{m^*} \right)^2 \frac{N^2}{L^3} \frac{N'^2}{L^3} W(q). \quad (9)$$

Since in region VI $k_F d \ll 1$, $W(q) \simeq \Delta^2 d^2$ is independent of q according to Eq. (3). The scattering is isotropic for a given subband N' with respect to different directions of $\vec{v}_{k'}$. The scattering rate is then reduced to

$$\frac{1}{\tau_N} = \frac{2\pi}{\hbar} \sum_{N'} \int \frac{d^2 k'}{(2\pi)^2} \left(\frac{\hbar^2}{m^*} \right)^2 \frac{N^2 N'^2 \Delta^2 d^2}{L^6 \varepsilon(q)^2} \delta(\varepsilon - \varepsilon_F). \quad (10)$$

The (2D) screening radius is $a_B/k_F L$ where $k_F L$ is the total number of subbands in region VI. Since $L \ll a_B$ in this region, this screening radius is much larger than the Fermi wavelength $1/k_F$. So $\varepsilon(q) \approx 1$ and the screening can be ignored for the scattering between the N th subband and the typical subbands with $k_z' \simeq k_F$ and thus $q \sim k_F$. Equation (10) then yields

$$\frac{1}{\tau_N} \simeq \frac{\hbar}{m^*} \frac{N^2 \Delta^2 d^2}{L^6} \sum_{N'} N'^2 \simeq \frac{\hbar}{m^*} \frac{N^2 \Delta^2 d^2}{L^6} (k_F L)^3$$

$$\simeq \frac{\hbar}{m^*} \frac{N^2 \Delta^2 d^2 k_F^3}{L^3}, \quad (11)$$

where the 3D wave number $k_F = (n/L)^{1/3}$, and the scattering rate is mainly determined by scattering between the N th subband and typical subbands with large N' . The absence of screening in the scattering-rate calculation is then self-consistently justified. Also, from Eq. (11), one can easily see that $\tau_N \propto 1/N^2$ so the lowest subband with $N = 1$ has the largest relaxation time while for typical subbands with $k_z \simeq k_F$ and, thus, $N \simeq k_F L$, the corresponding relaxation time is $(k_F L)^2$ times smaller. Since there are $\sim k_F L$ subbands in total with each subband having a 2D concentration $n/k_F L$ and the number of typical subbands is close to the total number $k_F L$, the final conductivity is dominated by the lowest

subband as

$$\sigma = \frac{n}{k_F L} \frac{e^2}{\hbar} \frac{L^3}{\Delta^2 d^2 k_F^3} = n e \frac{e}{\hbar} \frac{L^2}{\Delta^2 d^2 k_F^4}, \quad (12)$$

and the effective mobility is

$$\mu = \frac{\sigma}{ne} = \frac{e}{\hbar} \frac{L^2}{\Delta^2 d^2 k_F^4} = \frac{e}{\hbar} \left(\frac{d^4}{\Delta^2} \right) \frac{L^{10/3}}{d^6 n^{4/3}}. \quad (13)$$

This is the result shown in Table III in Sec. II.

Now let us move to region IV. This region is a crossover between completely geometrically confined region VI to completely electrostatically confined region II. The lowest M subbands are electrostatically confined due to their relatively small distances to the surface while the $k_F L - M$ higher subbands are geometrically confined. So for the lowest M subbands, $k_z^2/Z \sim n/a_B$ is a constant independent of the subband index N determined only by the surface electric field E or the 2D electron concentration n . As a result, the lowest M subbands have comparable relaxation times. The rest $k_F L - M$ subbands are geometrically confined and their contribution to the conductivity is dominated by the lowest subband of the group, i.e., by the $(M+1)$ st subband. Here the index M is obtained by the condition that its electrostatically confined width is equal to the well width L :

$$\frac{a_B k_z^2}{n} = L, \quad k_z \simeq \frac{M}{L}. \quad (14)$$

As a result, $M = (nL^3/a_B)^{1/2}$. Now Eq. (11) is modified for subbands from 1 to M as

$$\frac{1}{\tau_{1-M}} \simeq \frac{\hbar}{m^*} \frac{n \Delta^2 d^2}{a_B} \left(\sum_{N'=1, \dots, M} \frac{n}{a_B} + \sum_{N'=M+1, \dots, k_F L} \frac{N'^2}{L^3} \right)$$

$$\simeq \frac{\hbar}{m^*} \frac{n \Delta^2 d^2}{a_B} \sum_{N'=M+1, \dots, k_F L} \frac{N'^2}{L^3} \simeq \frac{\hbar}{m^*} \frac{n \Delta^2 d^2 k_F^3}{a_B}, \quad (15)$$

where $N^2/L^3 \gg n/a_B$ for $N' > M$ and the total number of subbands is still $k_F L \gg M$ in region IV. Therefore the scattering rate of each subband is always determined by its scattering into the typical subbands which are geometrically confined to a width L and have the momentum $k_z = k_F$ in the z direction. One can easily check that in region IV, i.e., at $L < n^{-1/5} a_B^{3/5}$, the conductivity is determined by the lowest M subbands and the effective mobility

$$\mu = \frac{\sigma}{ne} = \left(M \frac{n}{k_F L} \frac{e^2}{m^*} \frac{m^*}{\hbar} \frac{a_B}{n \Delta^2 d^2 k_F^3} \right) \frac{1}{ne}$$

$$= \frac{e}{\hbar} \left(\frac{d^4}{\Delta^2} \right) \frac{a_B^{1/2} L^{11/6}}{n^{11/6} d^6} \quad (16)$$

is obtained in a way similar to that of region VI discussed before. This is the result given in Table III.

Now let us talk about the $k_F d \gg 1$ case for regions V and VII. In this case, $W(q)$ is no longer a constant but can be much smaller than $\Delta^2 d^2$ for some values of q . The scattering is no longer isotropic in the x - y plane and one cannot ignore the DFR term $\vec{v}_{k'} \cdot \vec{E} \tau'_N / \vec{v}_k \cdot \vec{E} \tau_N$ in Eq. (7). As we show in Appendix B, the scattering is dominated by events with $q \simeq k_F$ instead

of small $q \lesssim 1/d$. It can be easily seen quasiclassically that only when an electron hits the sharp edge of an island can the nonspecular reflection happen. This is an event on a length scale $k_F^{-1} \ll d$ so that the scattering is dominated by $q \simeq k_F$.

For the dominant large-angle scattering, although the term $(1 - \vec{v}_k' \cdot \vec{E} \tau_N' / \vec{v}_k \cdot \vec{E} \tau_N)$ after averaging over different ϕ is not exactly unity as in the $k_F d \ll 1$ case, it is still of order unity. Thus in the scaling sense, the difference brought by $k_F d \gg 1$ is only in the $(k_F d)^3$ times reduction of $W(q)$. Note that, for the large-angle scattering, the rate is dominated by scattering into typical subbands of $k_z' \simeq k_F$ similarly to the $k_F d \ll 1$ case discussed before. The screening here is again ignored since the large-angle scattering has $q \simeq k_F$ and the screening radius $a_B/k_F L$ is much larger than the electron Fermi wavelength k_F^{-1} , similarly to the case in regions IV and VI. As a result, from region IV to V, the scattering rate decreases by $(k_F d)^3$ for each subband and the effective mobility increases by $(k_F d)^3$. A similar increase by a factor $(k_F d)^3$ happens across the border from region VI to region VII. So far we have derived all the new results in Fig. 3 and Table III.

One can see from Fig. 3 and Table III that the results of mobility in different regions match each other at all borders between the regions. Actually, by using the derived result Eq. (13) for region VI together with the results for regions II, III, and IX, which are already known, one can uniquely identify the mobility expressions in regions IV, V, and VII by matching them with the neighboring mobilities on the borders.

So far, we have focused on the $d \ll a_B$ case, which is generic for large- a_B semiconductors such as InAs and InSb. Now we would like to briefly discuss the $d \gg a_B$ case, which may take place, say, in silicon. Let us start from the case when $d = a_B$. In this case, the phase diagram Fig. 3 is dramatically simplified because the middle regions II, IV, and VI vanish and the border line $k_F d = 1$ merges with the vertical axis $na_B^2 = 1$. Let us now move to the case $d \gg a_B$. Since at $na_B^2 \ll 1$, the electron gas is two dimensional for all values of L , there is only one line $nd^2 = 1$ for the critical border $k_F d = 1$. We assume that this line is located already in the insulator regime, so that in the whole metallic region $k_F d \gg 1$. This leads to an additional factor $(k_F d)^3$ for the mobility result in region I and gives $\mu = (e/\hbar)(d^4/\Delta^2)(a_B^2/n^{1/2}d^3)$ (see Ref. [32]). Mobility results for the extended regions III, V, VII, and IX remain the same as in Table III.

IV. ROUGHNESS LIMITED MOBILITY IN QUANTUM WIRES

In the previous sections, we described the roughness limited mobility in a quantum well as a function of the 2D electron concentration n and the well width L . Here we would like to generalize these results to that of a nanowire with linear electron concentration η and radius R . We assume that an electric field $E = 2e\eta/R$ applied radially inward at the surface of the wire. Such a system can be realized by a metallic gate surrounding the nanowire, or by a planar gate located a distance greater than the wire radius R .

Our results are summarized in Fig. 1 as a phase diagram in the plane (η, R) , where each “phase” or region marked by a capital letter denotes a different dependence of the mobility on R and η , as shown in Table I. Just as for quantum wells,

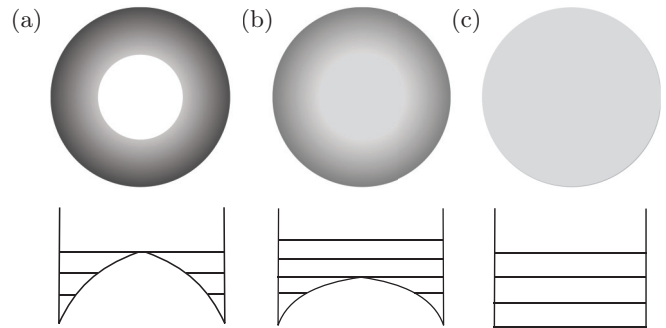


FIG. 4. Electron concentration within the nanowire (top) and the corresponding level schematic along the diameter (bottom). Regions of higher concentration correspond to darker shading. (a) Regions B and C of the scaling phase diagram Fig. 1, where all subbands are confined electrostatically forming an accumulation layer of thickness D near the surface. (b) Regions E and F, where the lowest subbands are confined electrostatically, while the top subbands are confined geometrically. (c) Regions H and I, where all the subbands are confined geometrically.

many different regions appear due to the interplay between the electrostatic and geometric confinements. The electronic structure of each region is illustrated with a radial level (subband) schematic similar to those in Fig. 3. One can divide all regions into three groups. In regions D and G the electron gas is strictly one dimensional (1DEG), i.e., it occupies a single subband in the wire cross section. In region A electrons occupy a single radial subband and many azimuthal subbands (2DEG). Finally, in regions B, C, E, F, H, and I, electrons occupy many subbands in both the radial and azimuthal directions and the gas is three dimensional (3DEG). To clarify the meaning of the level schematics, Fig. 4 provides an illustration of the electronic structure in the 3DEG regions. Each top image shows the electron density (shaded regions) in a cross section of the wire while its bottom image shows the corresponding level schematic along the wire diameter.

Let us first concentrate on the 2DEG and 3DEG regions, where the circumference $2\pi R$ is much larger than the typical electron wavelength k_F^{-1} . This means that we can generalize our results of the quantum well by treating the wire along the x axis as a stripe-like quantum well whose y -direction size is $2\pi R$ and 2D concentration $n = \eta/2\pi R$. As a result each of the regions I–VII of Fig. 3 has an analogous region in Fig. 1 in which the electronic structure near the surface and the mobility are the same upon substituting $n = \eta/2\pi R$ and $L \simeq R$ everywhere. For example, in region B electrons are confined electrostatically near the wire surface and form an accumulation layer [see Fig. 4(a)] whose thickness is given by Eq. (6) with $n = \eta/2\pi R$, similar to region II for the quantum well. By using the correspondence between regions A, C, E, F, H, and I of Fig. 1 with regions I, III, IV, V, VI, and VII of Fig. 3 we find the wire mobility values for each of these regions as listed in Table I.

So far we have shown that, in the 2DEG and 3DEG limits of the nanowire, there is a corresponding region in Fig. 3 from which the mobility of the wire may be obtained upon substituting $n = \eta/R$. In regions D and G however, the electron gas in the wire forms a 1DEG for which there is

no corresponding region in the quantum well. Let us first concentrate on region G, where the gas is geometrically confined to a single subband in the plane of its cross section (y, z) with energy $E_R = \hbar^2/2m^*R^2$ and its wavelength along the wire axis is $k_F^{-1} = \eta^{-1}$. Here y is the azimuthal direction along the wire circumference and z is the radial direction. Due to the roughness, the radius of the wire varies along the wire surface in the x and azimuthal directions by an amount $\delta R = \Delta(k_F^{-1}R/d^2)^{-1/2}$, where $k_F^{-1}R/d^2$ is the typical number of islands over which the electron averages the roughness. These variations lead to a change in the confinement energy that acts as a random scattering potential given by $V = E_R(\delta R/R)$. Using $\hbar/\tau \approx V^2/(\hbar^2k_F^2/2m^*)$ to estimate the scattering rate, we find the mobility in region G to be

$$\mu = \frac{e}{\hbar} \frac{\eta R^7}{\Delta^2 d^2}. \quad (17)$$

If we increase R so that we enter region D, the electron gas will instead be confined electrostatically to a single subband of width $D_0 = (a_B R/\eta)^{1/3}$. This change amounts to replacing R by D_0 in the confinement energy E_R . The mobility can thus be obtained by replacing the R^6 factor in Eq. (17) by $D_0^6 = a_B^2 R^2/\eta^2$ and so the mobility in region D is given by

$$\mu = \frac{e}{\hbar} \frac{a_B^2 R^2}{\eta^2 \Delta^2 d^2} (\eta R) = \frac{e}{\hbar} \frac{a_B^2 R^3}{\eta \Delta^2 d^2}. \quad (18)$$

The factor ηR is unchanged because this came from averaging over an area $k_F^{-1}R$ on the surface and was independent of the confinement in the radial direction. The mobility values given in Eqs. (17) and (18) are shown in Table I.

We can make the previous discussion more rigorous by considering the scattering rate by using Fermi's golden rule. In the IDEG limit there is only one radial or azimuthal subband occupied so that the scattering rate given by Eq. (7) then simplifies to

$$\begin{aligned} \frac{1}{\tau} &= \frac{2\pi}{\hbar} \frac{1}{R} \sum_{k'_y} \int \frac{dk'_x}{2\pi} \langle |V(q)|^2 \rangle \delta(\varepsilon_F - \varepsilon') \\ &= \frac{2\pi}{\hbar} \frac{1}{R} \int \frac{dk'_x}{2\pi} \langle |V(q)|^2 \rangle \delta(\varepsilon_F - \varepsilon'). \end{aligned} \quad (19)$$

Here the marginal one-dimensional screening is ignored and $\langle |V(q)|^2 \rangle$ is defined to be

$$\langle |V(q)|^2 \rangle \simeq \left(\frac{\hbar^2}{m^* Z^3} \right)^2 W(q) \quad (20)$$

for the gas confined to the lowest radial subband where $W(q) = \Delta^2 d^2$ at $k_F d \ll 1$. Setting $Z = R$ in region G and $Z = D_0$ in region D, we arrive at the mobilities given by Eqs. (17) and (18).

We see in Fig. 1 that region G is located at small R and small η and extends until the line $R = d$. Beyond this point, the characteristic size of the islands d becomes larger than the radius of the wire R and the model of isotropically distributed islands on the wire surface breaks down.

So far we have dealt with the mobility of quantum wires that are cylindrically symmetric. A stripe-like wire along the x axis can be made out of a narrow single subband GaAs/AlGaAs quantum well by the etching or split-gate techniques [40].

The mobility of such a modulation-doped stripe of 2DEG was calculated [22] for $k_F d \ll 1$ under the assumption that all scattering happens on the one-dimensional rough $y = 0$, R edges of the stripe and that the stripe has many y -direction subbands filled. Although our undoped wires studied in regions H and I are different from wires of Ref. [22], they share an important feature with them, i.e., the conduction is determined by the lowest subband. This can be easily understood quasiclassically, because the lowest subband electrons have most of its kinetic energy in the x direction and run approximately parallel to the surfaces or edges, and thus rarely get scattered.

V. BALLISTIC-DIFFUSIVE BOUNDARY AND THE CONDUCTANCE OF A WIRE WITH LENGTH \mathcal{L}

In the Introduction we explained that, due to the 1D nature of the wire, the transport properties differ greatly across the different regions of Fig. 1. Specifically, in the multisubband regions the wire of characteristic size R undergoes a transition between a ballistic conductor and a diffusive metal as a function of concentration. We will now explain why such a transition occurs and calculate the conductance G within these regions.

Let us first review what we know about the Drude conductance and show where it fails. In Table I we give the Drude mobility for the various regions of Fig. 1. By using these formulas one can calculate the dimensionless Drude conductance $G_D = (h/e)\eta\mu/\mathcal{L}$ per spin for a wire with length \mathcal{L} and linear concentration η . One can then define the metal-insulator transition by the condition $G_D = 1$. For example, in region H we find that $G_D = R^{14/3}/[\eta^{1/3}R_c(\mathcal{L})^5]$, where $R_c(\mathcal{L}) = (\Delta^2 d^2 \mathcal{L})^{1/5}$ is defined in the Introduction. Using the requirement $G_D = 1$, we find the MIT border within region H to be $R_{MI}(\eta) = \eta^{1/14} R_c(\mathcal{L})^{15/14}$. Similar calculations for regions G, E, and B lead to the $R_{MI}(\eta)$ in Table II.

The dimensionless Drude conductance is valid in all regions where $G_D > 1$, but the mean-free path $l < \mathcal{L}$. In region G where there is a single subband occupied, $G_D = 1$ and $l = \mathcal{L}$ are the same as long as $\eta a_B > 0.5$ where we can safely ignore electron-electron interactions. However in the multisubband regions B, H, and E the conditions are different. This can be understood by realizing that the condition $G_D = 1$ is equivalent to $\zeta = \mathcal{L}$, where ζ is the localization length. When multiple subbands are occupied, ζ grows larger than l , so in the multisubband region we can satisfy the conditions $l \ll \mathcal{L} \ll \zeta$ required for diffusive transport.

Let us begin with the simplest region B where all subbands have the same l . We define the mean-free path as $l = v_F \tau$, where τ is the relaxation time and $v_F = \hbar k_F/m^*$ is the Fermi velocity. The relaxation time τ can be calculated from the mobility in Table I and we find that in region B,

$$l = \frac{a_B^{7/5} R^{9/5}}{\Delta^2 d^2 \eta^{9/5}}. \quad (21)$$

The border equation is defined by the condition $l = \mathcal{L}$ and is found to be

$$R_{BD}(\eta) = \frac{\eta R_c(\mathcal{L})^{25/9}}{a_B^{7/9}}, \quad (22)$$

as shown in Table II.

In regions E and H there are radial subbands which are geometrically confined. As we showed in Sec. III, subbands that are geometrically confined will have different relaxation times, with higher subbands having smaller relaxation times. As a result G_D in these regions is determined by the lowest radial subbands where τ and l are largest. In region E the bottom M radial subbands are confined electrostatically, while the higher subbands are confined geometrically. Similar to region B the subbands that are electrostatically confined have the same mean-free path

$$l_{1-M} = \frac{a_B R^{7/3}}{\Delta^2 d^2 \eta^{5/3}}. \quad (23)$$

These are the lowest subbands that determine G_D and thus setting $l_{1-M} = \mathcal{L}$ leads to $R_{BD}(\eta)$ in Table II.

Finally, in region H all radial subbands are geometrically confined and therefore have different mean-free paths. The mean-free path of the N th subband l_N is given by

$$l_N = \frac{R^{13/3}}{\Delta^2 d^2 \eta^{2/3} N^2}. \quad (24)$$

We see that $l_N \propto N^{-2}$ and the conductance is determined by the lowest radial subband where $N = 1$. We can define the diffusive border by the condition that $l_1 = \mathcal{L}$, leading to the border equation in Table II.

Let us now use these results to determine the number K of ballistic subbands at the border. Recall that in the ballistic regions, the dimensionless conductance of the wire is $G_B = K$. At the border $G_B = G_D$, and so using our results of the Drude conductance we can self-consistently find the number of ballistic subbands. In region H, we find $K = k_F R$, in region E we find that $K = M k_F R$, and in region B we find that $K = k_F^2 R D$, where D is given by Eq. (6) with $n = \eta/R$. These results can be easily understood. For each radial subband there are $k_F R$ azimuthal subbands that contribute equally to the conductance. Then we can generically set $K = (k_F R) K_r$ where K_r will be the number of ballistic radial subbands at the border. In region H only one radial subband is ballistic, in region E there are M ballistic radial subbands, and finally in region B there are $k_F D$ radial subbands which are ballistic. Beyond the border, K_r increases as $(l_1/L)^{1/2}$ until K reaches the total number of subbands given in Table II, where l_1 is given by Eq. (24) for $N = 1$. The condition $K_r = k_F R$ defines a final border

$$R(\eta) = \eta^{4/11} R_c(\mathcal{L})^{15/11} \quad (25)$$

in regions H and E, beyond which all subbands are ballistic.

VI. VARIABLE-RADIUS MODEL OF A NANOWIRE

Previously, we considered a model of the surface roughness as flat islands of size $d \ll R$ and height Δ randomly distributed over the surface of the crystal. For the case of the nanowire, however, one can imagine another model of roughness in which the radius of the wire varies along its length but is independent of the azimuthal direction. We may consider these variations as ring-like steps of typical length d and thickness Δ . The step-like nature of the roughness means that we can describe this new model from our old one by restricting the spatial correlator given in Eq. (3) to variations in the x direction. The

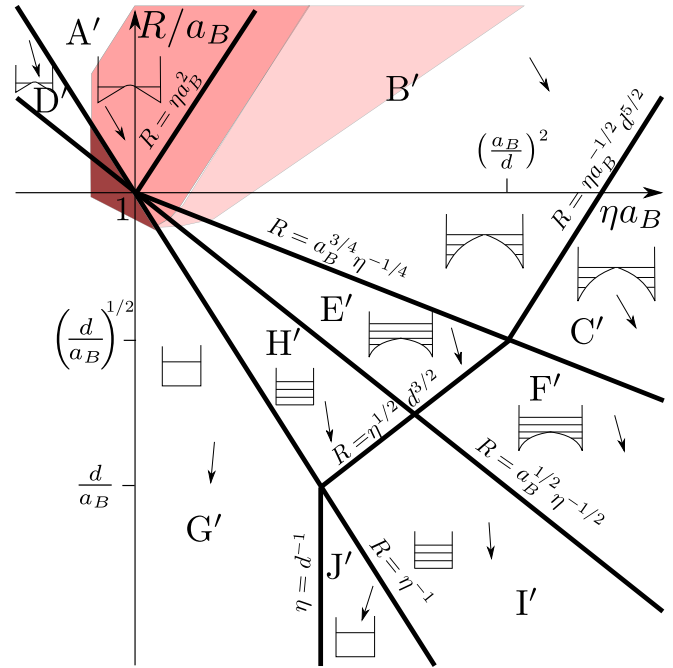


FIG. 5. The scaling phase diagram of roughness limited electron mobility of a quantum wire for the variable-radius model (VRM) plotted as a function of radius R and linear electron concentration η for $d < a_B$ on a log-log scale. Different phases or regions are denoted by capital letters. Mobility expressions corresponding to these regions are given in Table IV. Region boundaries are given by the equations next to them. The schematic self-consistent electron potential-energy profile along the the wire diameter and subbands occupied by electrons are shown for each region. Small arrows show the direction of mobility decrease in each region. All regions and the borders have the same definitions as Fig. 1, with the exception of a new region J' that was previously forbidden. The dark-red, light-red, and pink regions correspond respectively to the single-subband ballistic conductor, many-subband ballistic conductor, and diffusive metal defined by the same conditions as the isotropic model for $\mathcal{L} = a_B^{7/2} \Delta^{-2} d^{-1/2}$. We see that, for the same \mathcal{L} , the metallic window in the VRM is much smaller than the isotropic model. Electrons are localized in all colorless regions at $T = 0$.

corresponding Fourier transform of the correlator is then given by

$$W(q_x, q_y) = 2\sqrt{2}\pi \Delta^2 d (1 + q_x^2 d^2/2)^{-1} \delta(q_y), \quad (26)$$

where q_x is the momentum along the wire's length and q_y is the momentum in the azimuthal direction. We call this model the variable-radius model (VRM).

The new phase diagram for the VRM is shown in Fig. 5. It should not be surprising that most of the regions and borders are identical to those in Fig. 1, because these are set either by the number of subbands occupied, the type of confinement, or comparison between the island size d and the wavelength k_F^{-1} . Because none of these properties depend on the details of the correlator, the regions and borders remain the same as Fig. 5. However, there is a new region J' that emerges in Fig. 5 that did not appear in Fig. 1. This region is the geometrically confined 1DEG under the condition $k_F d \gg 1$. We see that this region occurs in the limit $R \ll d$, which was forbidden for the

TABLE IV. Mobility μ in units of $(e/\hbar)(d^4/\Delta^2)$ as a function of the linear electron concentration η at $d < a_B$ for different regions of Fig. 5.

| A' | B' | C' |
|----------------------------------|--------------------------------------|--------------------------------------|
| $a_B^2 R^{3/2} / \eta^{3/2} d^5$ | $a_B^{7/5} R^{9/5} / d^5 \eta^{9/5}$ | $a_B R / \eta d^3$ |
| D' | E' | F' |
| $a_B^2 R^2 / \eta d^5$ | $a_B^{1/2} R^3 / \eta^{3/2} d^5$ | $a_B^{1/2} R^{5/3} / \eta^{5/6} d^3$ |
| G' | H' | I' |
| $R^6 \eta / d^5$ | $R^4 / \eta d^5$ | $R^{8/3} / \eta^{1/3} d^3$ |
| | | J' |
| | | $\eta^3 R^6 / d^3$ |

previous model of roughness. No such restriction is necessary for the VRM, and so the new region emerges.

The mobility of these regions are given in Table IV. We notice immediately that the mobility expressions in regions C', F', and I' are identical to the same lettered regions in Fig. 1. The reason is that, in these regions, $k_F d \gg 1$, and the scattering is dominated by large-angle scattering at the edge of a single island, rather than an effect averaged over many islands. The lack of averaging eliminates the differences between the two models in this region, and so the mobility expressions are the same. When $k_F d \ll 1$, the electrons feel instead an averaged effect, and so we see differences emerge between the two models. The effect of averaging results in a reduction of the scattering rate by the number of scattering centers which are typically seen. In the model considered previously, the variations are two dimensional and so the electrons average along both the x direction and the azimuthal direction. This leads to an average number of islands that contribute to scattering given by the factor $1/(k_F d)^2$ in the 2DEG and 3DEG regions, and $R/(k_F d^2)$ in the 1DEG limit. In the VRM the variations only occur in the x direction and so we do not average in the azimuthal direction. This reduces the number of islands averaged over to be $1/(k_F d)$ in all regions. Knowing this, we may easily obtain the new mobilities of most regions by multiplying the expressions in Table I by the ratio of the new number of islands to the old number of islands. This ratio is $k_F d$ in the 2DEG and 3DEG and d/R in the 1DEG. The results are shown in Table IV.

While we can understand the changes in mobility in the VRM as due to a difference in averaging, we may also derive these changes from the correlator in Eq. (26). All the differences between the two models occur in the regions where $k_F d \ll 1$, where the correlator is simply $\sqrt{2} \Delta^2 d \delta(q_y)$. We see that the major difference from Eq. (3) is that $d^2 \rightarrow d \delta(q_y)$, and so it must be true that this difference is responsible for the change in the mobility between the two models. Indeed, when calculating the scattering rate, we integrate the correlator over the possible final states k' , so that it appears in the scattering rate as a factor $\int d^2 k' W(q)$. In our previous model this provided to the scattering rate an overall factor of $k_F^2 \Delta^2 d^2$ for the 2DEG and 3DEG regions, and $k_F \Delta^2 d^2 / R$ in the 1DEG. In the VRM the presence of a δ function for the azimuthal momentum means that these factors change to $k_F \Delta^2 d$ in all regions. From here it is clear that the change in the correlator leads to a difference in the mobility between the two models by a factor

of $k_F d$ in 2DEG and 3DEG regions and d/R in the 1DEG regions as we described above.

We have shown that all regions in Fig. 5 can be obtained from Fig. 1 except for the region J'. In this region $k_F d \gg 1$, where the scattering rate is determined by large-angle scattering. As was discussed in Sec. III, the large-angle scattering reduces the correlator, and thus the scattering rate, by a factor of $(k_F d)^3$ in the denominator. This allowed us to obtain the mobility for $k_F d \gg 1$ from the corresponding region with $k_F d \ll 1$ by multiplying the expression by the factor $(k_F d)^3$. The same logic may be applied in the VRM, but with a small change. The correlator for the VRM has a different power in the denominator than the previous model. The large-angle scattering then reduces the correlator by a factor of $(k_F d)^2$ in the denominator, rather than $(k_F d)^3$. This means that we may obtain the mobility of J' from that of G' by multiplying by the factor $(k_F d)^2 = (\eta d)^2$, and this value is shown in Table IV.

The results presented in Sec. V about the conductance and ballistic-diffusive border can easily be generalized to the VRM model. As the results are quite similar, we do not repeat the discussion here.

VII. DISCUSSION

Here we would like to estimate the critical value $R_c(\mathcal{L})$ in which the metallic window opens for InAs and InSb nanowires. To obtain an accurate estimate of $R_c(\mathcal{L})$, we first need the proper numerical coefficient beyond the scaling approach. Fortunately, the simple single-subband structure of regions G and G' allows this number to be determined analytically if we ignore electron-electron interactions. We have calculated these coefficients for a cylindrical wire in Appendix A and found that the mobility in region G of the isotropic model is

$$\mu = 0.047 \frac{e}{\hbar} \frac{\eta R^7}{\Delta^2 d^2}, \quad (27)$$

while for region G' of the VRM we find the mobility to be

$$\mu = 0.017 \frac{e}{\hbar} \frac{\eta R^6}{\Delta^2 d}. \quad (28)$$

$R_c(\mathcal{L})$ is defined to be the radius in which the dimensionless conductance $G = 1$. Using Eqs. (27) and (28), and assuming we are on the border $\eta R = 1$ between regions G and H (or G' and H'), we find the value of $R_c(\mathcal{L})$ in the isotropic roughness model to be

$$R_c(\mathcal{L}) = 1.8(\Delta^2 d^2 \mathcal{L})^{1/5}, \quad (29)$$

while for the VRM we find

$$R_c(\mathcal{L}) = 2.8(\Delta^2 d \mathcal{L})^{1/4}. \quad (30)$$

Now let us see what our theory predicts for a wire with $\mathcal{L} = 1 \mu\text{m}$. If we assume that $\Delta = 1 \text{ nm}$ and $d = 10 \text{ nm}$, then using Eq. (29) we find that $R_c(\mathcal{L}) = 18 \text{ nm}$ for the isotropic model, while using Eq. (30) for the VRM we find $R_c(\mathcal{L}) = 28 \text{ nm}$. We see that $R_c(\mathcal{L}) < a_B$ in both InAs ($a_B \approx 34 \text{ nm}$) [18] and in InSb ($a_B = 64 \text{ nm}$) [41], so that the ballistic single subband region exists. Recent experiments [7] have demonstrated ballistic transport in InSb nanowires with $\mathcal{L} \leq 1 \mu\text{m}$ and R in the range of 40–50 nm. These R satisfy the

condition $R_c(\mathcal{L}) < R < a_B$ from our estimates, and thus our theory is consistent with their observation of ballistic transport.

In the above estimate we used the condition $G = 1$ so that the conductance per spin was e^2/h . One could use a different condition in which $R_c(\mathcal{L})$ is defined to be the R such that $l = \mathcal{L}$. This different definition alters $R_c(\mathcal{L})$ by a factor 1.1 in the isotropic model and 1.2 in the VRM, and so our prediction for $R_c(\mathcal{L})$ is only slightly different between the two definitions.

VIII. CONCLUSION

In this paper, we have studied the surface-roughness-limited mobility in quantum wells and wires for single-subband and multisubband cases. In these systems, electrons are either electrostatically confined by the surface electric field E or geometrically confined by the surface barriers. The mobility is found to be a function of the electron concentration and well width L or wire radius R . Both quantum wells and wires are studied for the exponential model of roughness. For the wires, another model of variable radius (VRM) where there is exponential roughness only in the direction of the wire axis is also discussed. We have presented phase diagrams summarizing the rich collection of mobility scaling regions and found that in quantum wires there exists a critical size $R_c(\mathcal{L})$ so that wires with $R > R_c(\mathcal{L})$ have a window of concentrations where the wire is metallic, while for $R < R_c(\mathcal{L})$ electrons are localized at $T = 0$.

So far we have ignored the spin-orbit coupling of electrons. In InAs and InSb nanowires studied for the purpose of quantum computations [5,6,11,12], the spin-orbit interaction is quite strong. However, the experimentally relevant Rashba spin-orbit interaction [42] just shifts two electron bands of opposite spin polarizations away from each other in the Brillouin zone. Therefore, electrons in each spin-polarized band move independently of the other band and the mobility is the same as the case without the spin-orbit coupling.

ACKNOWLEDGMENTS

We are grateful to A. Kamenev, V. Pribiag, X. Ying, and K.V. Reich for helpful discussions. Han Fu was supported by the University of Minnesota Doctoral Dissertation Fellowship. M. Sammon was supported primarily by the National Science Foundation through the University of Minnesota MRSEC under Award No. DMR-1420013.

APPENDIX A: COEFFICIENTS OF MOBILITY FOR GEOMETRICALLY CONFINED ONE-DIMENSIONAL ELECTRON GAS IN CYLINDRICAL NANOWIRES IN REGIONS G AND G'

In the Discussion, we used the coefficient of the mobility and thus the mean-free path of electrons in narrow nanowires of cylindrical cross sections at low electron concentrations (regions G of Fig. 1 and G' of Fig. 5). In this appendix, we derive this coefficient.

For a narrow nanowire at low electron concentrations, electrons occupy only the first subband in the wire cross section forming a 1DEG which is geometrically confined. If we ignore

correlation effects, the wave function of the lowest subband in a cylindrical nanowire of radius R is

$$\xi(r, \phi, x) = \frac{J_0(v_0 r/R) e^{ikx}}{\sqrt{\pi} R J_1(v_0)}, \quad (\text{A1})$$

where x is directed along the wire axis, r is the distance from the wire center, ϕ is the azimuthal angle in the cross section of the wire, J_0 and J_1 are the zeroth- and first-order Bessel functions of the first kind, and $v_0 \approx 2.4$ is the first zero of J_0 .

It can be easily derived that, for a 1DEG, the scattering rate is

$$\frac{1}{\tau} = \frac{2\pi}{\hbar} |V|^2 \rho (1 - \cos \theta), \quad (\text{A2})$$

where $|V|$ is the scattering matrix element due to roughness, $\theta = \pi$ is the angle between initial and final electron momenta, $\rho = m^*/2\pi \hbar^2 k_F$ is the density of states into which the backscattering can happen, and k_F is the Fermi wave number of the 1DEG. For 1D scattering, only backscattering can cause momentum relaxation, and so the angle between the initial and final momenta is π .

Similar to Eq. (8), according to Ref. [25], one can obtain the scattering potential in the cylindrical geometry to be

$$V(\phi, z) = \frac{\hbar^2}{2m^*} \Delta(\phi, z) \left. \frac{\partial \xi}{\partial r} \frac{\partial \xi'}{\partial r} \right|_{r=R}, \quad (\text{A3})$$

and the scattering matrix element for $R \gg d$ is

$$\langle |V(q)|^2 \rangle = \frac{v_0}{2\pi} \frac{\hbar^4}{m^{*2} R^7} W(q), \quad (\text{A4})$$

where $q = 2k_F$ is the transferred momentum along the wire axis for backscattering of electrons at the Fermi level.

If we combine Eqs. (A2) and (A4), set $k_F = (\pi/2)\eta$ for a 1D gas, and use $k_F d \ll 1$ for the correlator given in Eq. (3), we find the mobility $\mu = e\tau/m^*$ to be

$$\mu = \frac{\pi}{2v_0^4} \frac{e}{\hbar} \frac{\eta R^7}{\Delta^2 d^2}. \quad (\text{A5})$$

If instead we consider the VRM model described in Sec. VI, then we use the correlator given in Eq. (26) instead. As a result, the mobility in region G' in the VRM is

$$\mu = \frac{\pi}{4\sqrt{2}v_0^4} \frac{e}{\hbar} \frac{\eta R^6}{\Delta^2 d}. \quad (\text{A6})$$

APPENDIX B: LARGE-ANGLE-SCATTERING DOMINANCE IN SCATTERING RATE FOR QUANTUM WELL

Here by using Eq. (7) we prove that the scattering rate in $k_F d \gg 1$ regions V and VII is dominated by the large-angle scattering, i.e., scattering events with large $q \simeq k_F$. One might expect that, because the correlator $W(q) \sim \Delta^2 d^3 / (k_F d)^3$ for large-angle scattering with $q \sim k_F$ is much smaller than that for scattering into small angles with $q \sim 1/d$ by a factor of $(k_F d)^3$ in the denominator, that the scattering is dominated by the small-angle regime. However as we show below, the limited number of final subbands that electrons can scatter into for $q \leq 1/d$, the small value of the angular integral $\int d\phi (1 - \vec{v}_k' \cdot \vec{E} \tau_N / v_k \cdot \vec{E} \tau_N)$, and in certain cases the smaller z -direction

momentum of final states $k'_z \ll k_F$, act to suppress the small-angle scattering rate so that the scattering is determined by the large-angle scattering. We show this below for three cases: $L < d$ (in some part of regions V and VII); $L > d$ and $M/L < 1/d$ (for the rest of region V and some part of region VII); $L > d$ and $M/L > 1/d$ (for the rest of region VII).

First let us consider the case when $L < d$. From energy conservation, the total magnitude of the momentum is fixed, and so any difference in magnitude of the in-plane momenta follows from the difference $|k_z - k'_z| \sim 1/L$ of their z momentum. When $L < d$, $q \leq 1/d \ll 1/L$ and the scattering happens only within the same subband. This means that the DFR term $\vec{v}_{k'} \cdot \vec{E} \tau'_N / \vec{v}_k \cdot \vec{E} \tau_N$ reduces to the usual $\cos \phi$ for 2D scattering, where ϕ is the angle between $\vec{v}_{k'}$, \vec{v}_k . The final angular integral for the small-angle scattering is $\int (1 - \cos \phi) d\phi \sim \phi^3 \sim (k_F d)^{-3}$, while it is of order unity for the large-angle scattering. This cancels the advantage of larger $W(q)$ in the small-angle scattering. Moreover, the small-angle scattering has only one final subband to scatter into while the large-angle scattering covers all $k_F L$ subbands. This combined with the small-angular integral means that the small-angle scattering rate is $k_F L \gg 1$ times smaller than that of the large-angle scattering when $L < d$.

Now let us look at the second case where $L > d$ and $M/L < 1/d$. For simplicity, we focus on the lowest subbands with $k_z < M/L \ll k_F$ because these dominate the conductivity in regions V and VII. In the limit $L > d$ and $M/L < 1/d$, there will always exist $L/d > M$ subbands with $k'_z < 1/d$ so that the scattering now involves intersubband scattering. As a result the DFR term is not reduced to $\cos \phi$ and the term $(1 - \vec{v}_{k'} \cdot \vec{E} \tau'_N / \vec{v}_k \cdot \vec{E} \tau_N)$ is of order unity instead of being infinitesimal for small q and thus small ϕ . The angular integral of this term would just give $1/(k_F d)$ from the small angle $\int d\phi \simeq 1/dk_F$ and does not compensate the $(k_F d)^3$ reduction of the correlator. However, we must consider the importance of k'_z in the scattering matrix element according to Eq. (8). For the small-angle scattering $k'_z < 1/d$, while for the large-angle regime $k'_z \sim k_F$. This gives an extra factor $1/(k_F d)^2$ to $|V(q)|^2$ in small-angle regime relative to the large-angle scattering. This additional factor combined with the small angular integral compensates the $1/(k_F d)^3$ reduction of the correlator. Considering also the accessible number of final subbands L/d in the small-angle limit is $k_F d$ times smaller than the $k_F L$ available subbands for large-angle scattering, we find

that the small-angle scattering rate is $k_F d$ times smaller than that of the large-angle scattering.

Finally, we must consider the intermediate case when $L > d$ and $M/L > 1/d$. For small-angle scattering the number of subbands L/d that may be scattered into is small, and so we expect that the DFR term is near the 2D limit $\cos \phi$. Expanding around this value, we find that the DFR term is approximately

$$\frac{\vec{v}_{k'} \cdot \vec{E} \tau'_N}{\vec{v}_k \cdot \vec{E} \tau_N} = \cos \phi \left(1 - \frac{\delta v_k}{v_k} - \frac{\delta \tau_N}{\tau_N} \right), \quad (\text{B1})$$

where $\delta v_k = |v_k - v'_k|$ and $\delta \tau_N = |\tau_N - \tau'_N|$. Let us examine these correction terms, beginning with $\delta v_k/v_k$. The allowed difference in k and k' is $1/d$ for the small-angle scattering. Since $k \simeq k_F$ for the lowest subbands, the velocity difference ratio $|v_{k'} - v_k|/v_k = |k' - k|/k$ is then $1/k_F d$. In considering the other correction term $\delta \tau_N/\tau_N$, let us assume that the scattering rate of each subband is always determined by their large-angle scattering into typical subbands and show that this assumption self-consistent. With this assumption the difference in relaxation times $\delta \tau_N$ is solely caused by the different z -direction momenta and subband widths as seen from Eq. (8). Again we focus on the the lowest M subbands because these determine the conductivity. For the bottommost subbands, all subbands within $q \sim 1/d$ are electrostatically confined and $\delta \tau_N = 0$ because $k_z^2/Z' = k_z^2/Z = n/a_B$ [see Eq. (8)]. For the higher subbands with $k_z \sim M/L$, there are bands within $q \sim 1/d$ which are instead geometrically confined and the correction is nonvanishing. Indeed, we find that $\delta \tau_N \sim \tau_N \delta k_z/k_z$ and so the correction is given by $(1/d)/(M/L)$. We find then that the leading contribution to the DFR term in Eq. (B1) in the small-angle regime is approximately $1 - (1/d)/(M/L)$, where we have used the fact that $1/(k_F d) \ll L/Md$ in the limits being considered.

Using the DFR term above, the angular integral now gives a factor $(1/k_F d)(1/d)/(M/L)$ to the scattering rate, while the integral is of order unity for the large-angle limit. Combined with the fact that the final state in the small-angle regime has $k_z^2/Z' \simeq (M/L)^2/L$, we find that these terms give an extra factor $(Md/L)/(k_F d)^3$ compared with the same terms for the large-angle limit. We see then that there is a factor of $1/(k_F d)^3$ term that compensates the suppression of the correlator in the large-angle limit. Adding the fact that the small-angle scattering can only scatter into $L/d \ll k_F L$, we find that the ratio of scattering rates in the small- and large-angle regimes is $M/(k_F L) \ll 1$ and indeed the large-angle limit dominates.

-
- [1] C. M. Lieber, *MRS Bull.* **28**, 486 (2003).
[2] P. Yang, *MRS Bull.* **30**, 85 (2005).
[3] R. M. Lutchyn, E. Bakkers, L. Kouwenhoven, P. Krogstrup, C. Marcus, and Y. Oreg, [arXiv:1707.04899](https://arxiv.org/abs/1707.04899).
[4] L. Hofstetter, S. Csonka, J. Nygard, and C. Schonenberger, *Nature (London)* **461**, 960 (2009).
[5] Y.-J. Doh, J. A. van Dam, A. L. Roest, E. P. A. M. Bakkers, L. P. Kouwenhoven, and S. De Franceschi, *Science* **309**, 272 (2005).
[6] S. Nadj-Perge, S. M. Frolov, E. P. A. M. Bakkers, and L. P. Kouwenhoven, *Nature (London)* **468**, 1084 (2010).
[7] E. M. Fadaly, H. Zhang, S. Conesa-Boj, D. Car, Ö. Gül, S. R. Plissard, R. L. Veld, S. Kölling, L. P. Kouwenhoven, and E. P. Bakkers, *Nano Lett.* **17**, 6511 (2017).
[8] R. M. Lutchyn, J. D. Sau, and S. Das Sarma, *Phys. Rev. Lett.* **105**, 077001 (2010).
[9] Y. Oreg, G. Refael, and F. von Oppen, *Phys. Rev. Lett.* **105**, 177002 (2010).
[10] E. Majorana, *Nuovo Cimento* **14**, 171 (2008).
[11] V. Mourik, K. Zuo, S. M. Frolov, S. R. Plissard, E. P. A. M. Bakkers, and L. P. Kouwenhoven, *Science* **336**, 1003 (2012).
[12] M. T. Deng, C. L. Yu, G. Y. Huang, M. Larsson, P. Caroff, and H. Q. Xu, *Nano Lett.* **12**, 6414 (2012).
[13] H. O. H. Churchill, V. Fatemi, K. Grove-Rasmussen, M. T. Deng, P. Caroff, H. Q. Xu, and C. M. Marcus, *Phys. Rev. B* **87**, 241401 (2013).

- [14] J. D. Sau, S. Tewari, and S. Das Sarma, *Phys. Rev. B* **85**, 064512 (2012).
- [15] Önder Gül *et al.*, *Nanotechnol.* **26**, 215202 (2015).
- [16] M. Scheffler, S. Nadj-Perge, L. P. Kouwenhoven, M. T. Borgström, and E. P. A. M. Bakkers, *J. Appl. Phys.* **106**, 124303 (2009).
- [17] J. J. Hou, F. Wang, N. Han, H. Zhu, K. Fok, W. Lam, S. Yip, T. Hung, J. E.-Y. Lee, and J. C. Ho, *Appl. Phys. Lett.* **102**, 093112 (2013).
- [18] A. C. Ford, J. C. Ho, Y.-L. Chueh, Y.-C. Tseng, Z. Fan, J. Guo, J. Bokor, and A. Javey, *Nano Lett.* **9**, 360 (2009).
- [19] S. Poli, M. G. Pala, T. Poiroux, S. Deleonibus, and G. Baccarani, *IEEE Trans. Electron Devices* **55**, 2968 (2008).
- [20] J. Chen, T. Saraya, K. Miyaji, K. Shimizu, and T. Hiramoto, in *2008 Symposium on VLSI Technology* (IEEE, 2008), pp. 32–33.
- [21] F. Wang, S. Yip, N. Han, K. Fok, H. Lin, J. J. Hou, G. Dong, T. Hung, K. S. Chan, and J. C. Ho, *Nanotechnol.* **24**, 375202 (2013).
- [22] H. Akera and T. Ando, *Phys. Rev. B* **43**, 11676 (1991).
- [23] S. Jin, M. V. Fischetti, and T. wei Tang, *J. Appl. Phys.* **102**, 083715 (2007).
- [24] J. Motohisa and H. Sakaki, *Appl. Phys. Lett.* **60**, 1315 (1992).
- [25] T. Ando, A. B. Fowler, and F. Stern, *Rev. Mod. Phys.* **54**, 437 (1982).
- [26] H. Sakaki, T. Noda, K. Hirakawa, M. Tanaka, and T. Matsusue, *Appl. Phys. Lett.* **51**, 1934 (1987).
- [27] A. Gold, *Phys. Rev. B* **35**, 723 (1987).
- [28] M. Entin, *Sov. Phys.-Solid State* **1**, 781 (1969).
- [29] M. Krylov and R. Suris, *Sov. Phys.-JETP* **56**, 1316 (1982).
- [30] S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, *Phys. Rev. B* **32**, 8171 (1985).
- [31] R. M. Feenstra, D. A. Collins, D. Z.-Y. Ting, M. W. Wang, and T. C. McGill, *Phys. Rev. Lett.* **72**, 2749 (1994).
- [32] H. Fu, K. V. Reich, and B. I. Shklovskii, *Phys. Rev. B* **93**, 235312 (2016).
- [33] L. I. Glazman, I. M. Ruzin, and B. I. Shklovskii, *Phys. Rev. B* **45**, 8454 (1992).
- [34] C. L. Kane and M. P. A. Fisher, *Phys. Rev. Lett.* **68**, 1220 (1992).
- [35] M. P. Fisher and L. I. Glazman, in *Mesoscopic Electron Transport* (Springer, Berlin, 1997), pp. 331–373.
- [36] S. Das Sarma and E. H. Hwang, *Phys. Rev. B* **89**, 121413 (2014).
- [37] J. Frenkel, *Eur. Phys. J. A* **51**, 232 (1928).
- [38] K. V. Reich, M. Schechter, and B. I. Shklovskii, *Phys. Rev. B* **91**, 115303 (2015).
- [39] E. D. Siggia and P. C. Kwok, *Phys. Rev. B* **2**, 1024 (1970).
- [40] H. J. Suominen, M. Kjaergaard, A. R. Hamilton, J. Shabani, C. J. Palmstrøm, C. M. Marcus, and F. Nichele, *Phys. Rev. Lett.* **119**, 176805 (2017).
- [41] S. G. Pandya and M. E. Kordesch, *Nanoscale Res. Lett.* **10**, 258 (2015).
- [42] I. Van Weperen, Ph.D. thesis, Delft University of Technology, 2014 (unpublished).