



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Surface roughness scattering in multisubband accumulation layers

Han Fu, K. V. Reich, and B. I. Shklovskii

Phys. Rev. B **93**, 235312 — Published 23 June 2016

DOI: [10.1103/PhysRevB.93.235312](https://doi.org/10.1103/PhysRevB.93.235312)

Surface roughness scattering in multisubband accumulation layers

Han Fu,^{1,*} K. V. Reich,^{1,2} and B. I. Shklovskii¹

¹*Fine Theoretical Physics Institute, University of Minnesota, Minneapolis, MN 55455, USA*

²*Ioffe Institute, St Petersburg, 194021, Russia*

(Dated: May 18, 2016)

Accumulation layers with very large concentrations of electrons where many subbands are filled became recently available due to ionic liquid and other new methods of gating. The low temperature mobility in such layers is limited by the surface roughness scattering. However theories of roughness scattering so far dealt only with the small-density single subband two-dimensional electron gas (2DEG). Here we develop a theory of roughness-scattering limited mobility for the multisubband large concentration case. We show that with growing 2D electron concentration n the surface dimensionless conductivity $\sigma/(2e^2/h)$ first decreases as $\propto n^{-6/5}$ and then saturates as $\sim (da_B/\Delta^2) \gg 1$, where d and Δ are the characteristic length and height of the surface roughness, a_B is the effective Bohr radius. This means that in spite of the shrinkage of the 2DEG thickness and the related increase of the scattering rate, the 2DEG remains a good metal.

I. INTRODUCTION

The electron mobility is a very important parameter of electronic devices. In heavily doped bulk semiconductors, the low temperature mobility is determined by the electron scattering on ionized donors and is relatively small. Larger low temperature mobilities can be achieved near the surface of a lightly doped say n -type semiconductor, where the electron accumulation layer is induced by the applied surface electric field. In such devices the mobility becomes sensitive to the semiconductor surface roughness, which can be imagined as a collection of atomic-size steps of total height Δ and characteristic size $d \gg \Delta$ along the surface. The roughness scattering dominates at high electric fields E when electrons are squeezed closer to the surface. In this case¹⁻⁶ the mobility μ limited by the surface roughness scattering behaves as $\mu \propto 1/E^2$. For a large enough field E the two-dimensional (2D) concentration of electrons $n \propto E$ so that $\mu \propto 1/n^2$ and the surface conductivity $\sigma = ne\mu \propto 1/n$. This result holds for an inversion layer in a lightly doped p -type semiconductor when the electric field E is larger than the electric field of the depletion layer. The low temperature mobility is also extensively studied in quantum wells, where it is limited by the surface roughness of both interfaces. This mobility strongly depends on the width of the quantum well⁷⁻⁹.

Because of the dominating interest in higher mobilities, the surface roughness scattering was studied theoretically only for relatively small concentrations n , when only the first energy quantization subband is filled at low temperatures²⁻⁹. Also, it was difficult to induce large electron concentrations n (higher than 10^{13} cm^{-2} in Si). So the the roughness scattering in the case of large concentrations n when many subbands are filled at low temperatures has remained unexplored.

The last decade, however, witnessed growing interest in accumulation layers with large n which allow to achieve qualitatively new properties of the electron gas, such as superconductivity or magnetism. New methods to create

large electron concentrations were developed. One of them is based on ion gating with help of an electrolyte or a room temperature ionic liquid, which does not need an insulator layer and, therefore, makes a double layer with a very large capacitance. In Si concentrations $n \sim 5 \cdot 10^{13} \text{ cm}^{-2}$ were achieved using gating by an electrolyte¹⁰ and by an ionic liquid¹¹. Even larger concentrations $\sim 10^{14} \text{ cm}^{-2}$ were induced in single and bilayer graphene^{12,13}, ZnO¹⁴, MoS₂¹⁵ and SrTiO₃^{16,17} with this method.

Another important method is based on heterojunctions of polar and nonpolar perovskites such as GdTIO₃ and SrTiO₃, which accumulate $3 \cdot 10^{14} \text{ cm}^{-2}$ electrons¹⁸. Concentrations n up to 10^{15} cm^{-2} were obtained combining this effect with the electron spill-out for a special band alignment¹⁹. Similar physics takes place in GaN-based heterojunctions where concentrations up to $4.4 \times 10^{13} \text{ cm}^{-2}$ were achieved²⁰⁻²².

At $na_B^2 > 1$ the dimensionless parameter $r_s < 1$ and many electron subbands are filled. Here $r_s = (\pi na_B^2)^{-1/2}$, the effective Bohr radius $a_B = \kappa \hbar^2 / m^* e^2$ in Gaussian units, κ is the dielectric constant, m^* is the effective electron mass. In the cases of ZnO and MoS₂ mentioned above na_B^2 reaches 5. In semiconductors with relatively large a_B such as GaAs, InAs, InSb, and PbTe, it should be easy to reach $na_B^2 \gg 1$.

As we said above the roughness scattering limited mobility for the multisubband case has not been theoretically studied. In this paper, we fill this gap and study the low temperature mobility limited by surface roughness in an accumulation layer with large n . Our multisubband result for the dimensionless conductivity $\sigma(n)/(2e^2/h)$ at $d < a_B$ is shown on Fig. 1 as a function of the dimensionless concentration na_B^2 for the exponential model of roughness²³ by the thick solid line (black). The conductivity first decreases with n as $\propto n^{-6/5}$ and then saturates at the level

$$\frac{\sigma}{2e^2/h} \simeq \frac{1.2da_B}{\Delta^2}, \quad (1)$$

which is much larger than unity assuming that both $d \gg$

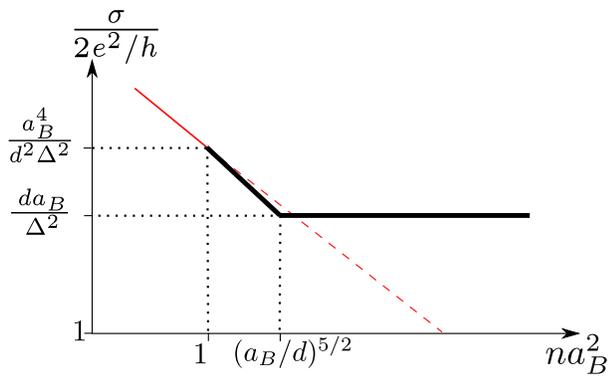


FIG. 1. (Color online) The schematic log-log plot of the dimensionless conductivity of an accumulation layer $\sigma/(2e^2/h)$ limited by the surface roughness scattering as a function of the dimensionless 2D electron concentration na_B^2 at $d < a_B$. The thick solid line (black) shows the conductivity for a multisubband accumulation layer. It first decreases as $(na_B^2)^{-6/5}(a_B^4/d^2\Delta^2)$ and then saturates at $na_B^2 \sim (a_B/d)^{5/2}$, where the wavelength $k_F^{-1} \sim d$. The thin solid line (red) represents the $1/n$ dependence derived for a single subband³. Conjectured extrapolation²⁴ of this dependence to larger concentrations is shown by the thin dashed thin (red).

Δ and $a_B \gg \Delta$. The thin solid line (red) schematically shows the $1/n$ dependence of the conductivity derived for a single subband by previous work³.

We see that with growing n the multisubband conductivity at first approximately continues the single subband dependence $1/n$, but then saturates. The saturation happens when the electron wavelength k_F^{-1} is equal to the size d of the roughness. Before this point, the roughness felt by electrons is averaged over all irregularities within the region of size k_F^{-1} . As the concentration increases, k_F increases and fewer irregularities are averaged over making the surface “rougher” for electrons. This results in decreasing $\sigma(n)$. When k_F^{-1} gets below d , the electron “hits” only a single irregularity and the level of “roughness” is fixed leading to the saturation of the conductivity.

Our results show that at large n where many subbands are filled an accumulation layer remains metallic. This agrees with decent mobilities observed in Refs. 10, 14, 15, and 25.

The organization of our paper is as follows. In Sec. II, we explain the structure of an accumulation layer where electrons occupy many energy quantization subbands. In Sec. III, we introduce two models of the surface roughness including the exponential one assumed in Fig. 1 and the Gaussian widely used in earlier studies. In Sec. IV, we present an intuitive quasi-classical interpretation of our mobility results for the exponential roughness. In Sec. V, we introduce the more formal quantum-mechanical approach starting from the case of a single subband connecting to previous studies. In Sec. VI, we discuss the multisubband case, take into account the scattering of electrons between

different subbands, and give the final scattering rate and mobility. We conclude in Sec. VII. In the main text of the paper we use the scaling approach and drop numerical coefficients. In Appendix A we estimate the coefficients of the conductivity for the most interesting case of the exponential roughness.

II. ELECTRONIC STRUCTURE OF AN ACCUMULATION LAYER

The accumulation layer is created near the surface of an n -type semiconductor when the orthogonal-to-surface electric field E induces a large 2D concentration of excessive surface electrons $n = E/4\pi e$ in the layer of width $\sim L$. We assume that n/L is much larger than the bulk concentration of electrons N in the semiconductor. This means that the bulk Fermi level at low temperatures is either below the conduction band bottom or slightly above it. In Fig. 2 illustrating the accumulation layer we actually assumed that the bulk Fermi level coincides with the conduction band bottom, which in this paper serves as the reference point of the electron energy. Our description of the electron accumulation layer is applicable also to an inversion layer in the very lightly doped p -type semiconductor, where the 2D concentration n_{depl} of ionized acceptors forming the depletion layer is much smaller than the 2D concentration n of the electrons so that we can still use $n = E/4\pi e$.

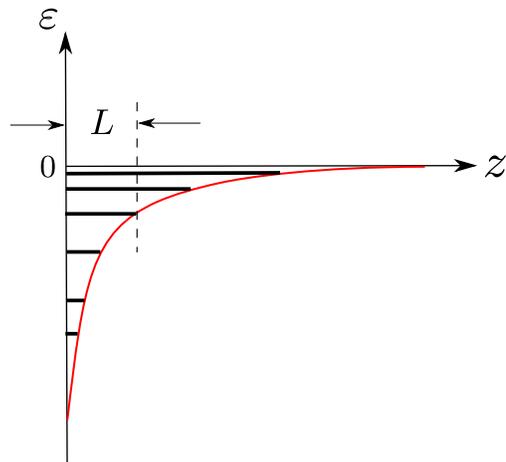


FIG. 2. (Color online) Schematic plot of the subbands electrons occupy at low temperatures in the accumulation layer. The ground states of subbands are shown by the black thick horizontal lines. The surface potential well $-e\varphi(z)$ is shown by the grey thin line (red), where $\varphi(z)$ is given by Eq. (2). The reference energy level $\varepsilon = 0$ is the conduction band bottom. In a lightly n -doped bulk semiconductor the electron Fermi level is close to zero. Most electrons are located within $z \lesssim L$ where L is the decay length of the electron concentration given by Eq. (4).

In this paper we are focused on large E and n

cases when the one-dimensional potential well created by the electric field E along the z -axis normal to the surface is so deep that it has several quantized levels (see Fig. 2). Each of such levels forms a subband with a two-dimensional (2D) Fermi gas moving freely parallel to the surface. At large n all these 2D gases with the same Fermi level form a three-dimensional (3D) degenerate gas, which can be described in the Thomas-Fermi approximation neglecting the discreteness of subbands. The nonlinear screening of the electric field E by such a gas was studied by Frenkel²⁶ via solving the Poisson-Thomas-Fermi equation for the self-consistent potential $\varphi(z)$ and the 3D electron density $N(z)$. The result is^{26,27}

$$\varphi(z) = C_1 \frac{e}{\kappa a_B} \left(\frac{a_B}{z+L} \right)^4, \quad (2)$$

$$N(z) = C_2 \frac{1}{a_B^3} \left(\frac{a_B}{z+L} \right)^6, \quad (3)$$

where z is the distance from the interface, and the width of the electron gas is

$$L = C_3 \frac{a_B}{(na_B^2)^{1/5}}. \quad (4)$$

Here, n is the total 2D concentration of electrons inside the accumulation layer, $C_1 = 225\pi^2/8 \simeq 278$, $C_2 = 1125\pi/8 \simeq 442$, $C_3 = (225\pi/8)^{1/5} \simeq 2.4$. This solution is valid for $na_B^2 \gg 1$. In this case $N(0)a_B^3 \sim (na_B^2)^{6/5} \gg 1$ and $k_F L \sim (na_B^2)^{1/5} \gg 1$ where $N(0) = n/L$ and $k_F = N(0)^{1/3}$. These inequalities confirm that we deal with a 3D degenerate gas and that the Thomas-Fermi approximation is valid and many subbands are filled. At smaller concentrations when $na_B^2 < 1$ only one subband is filled and we go back to the 2D case studied for the inversion layer or narrow quantum wells.

III. MODELS OF SURFACE ROUGHNESS

The surface roughness is a random shift of the interface $\Delta(\vec{r})$ from $z = 0$ so that $\langle \Delta(\vec{r}) \rangle = 0$, where $\vec{r} = (x, y)$ is the coordinate in $z = 0$ interface plane (see Fig. 3). The roughness is described by the height correlator and its Fourier transform

$$\begin{aligned} \langle \Delta(\vec{r})\Delta(\vec{r}') \rangle &= W(\vec{r} - \vec{r}'), \\ \langle |\Delta(q)|^2 \rangle &= W(q). \end{aligned} \quad (5)$$

Two main models of roughness are used in literature. One is Gaussian

$$\begin{aligned} W(\vec{r} - \vec{r}') &= \Delta^2 e^{-(\vec{r} - \vec{r}')^2/d^2}, \\ W(q) &= \pi \Delta^2 d^2 e^{-q^2 d^2/4}, \end{aligned} \quad (6)$$

where d, Δ are the characteristic size and height of the roughness, $\Delta \ll d, L$. This model was widely used in earlier studies²⁻⁸.

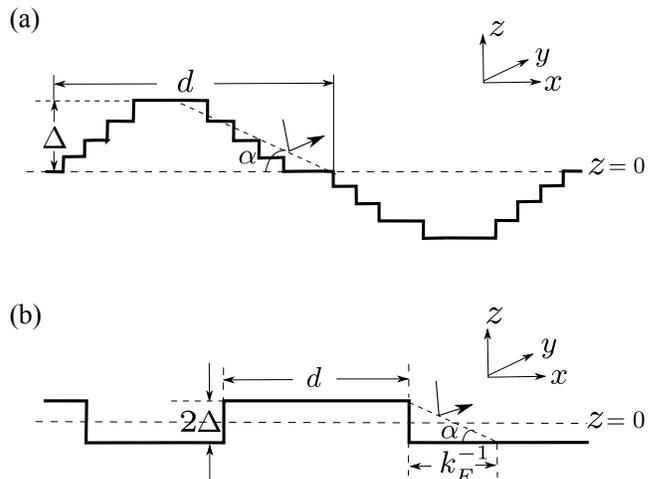


FIG. 3. Two types of surface roughness. (a) The Gaussian type of roughness. Here the lattice discreteness can be ignored where $\Delta \gg a$ and a is the lattice constant. (b) The exponential kind of roughness. The size of the roughness is d and the height fluctuates as $\pm\Delta$ with respect to the average interface ($z = 0$), where $2\Delta = a$. When the electron wavelength $k_F^{-1} \gg d$, the roughness is effectively averaged as $\Delta/\sqrt{1/k_F^2 d^2} = \Delta k_F d$ on a length scale of k_F^{-1} for both types of surface roughness. At $k_F^{-1} \ll d$, the incident electron feels only a single hill/valley or island. For Gaussian roughness, the electron is scattered by the slope with the angle $\alpha \sim \Delta/d$ shown in (a). For the exponential roughness, however, the electron is scattered by the island edge which has a height Δ and an effective length k_F^{-1} , and thus the effective angle is $\alpha \simeq k_F \Delta$ as shown in (b).

However, later experimental observations found that the spacial correlations are more likely to follow an exponential behavior^{23,28}

$$\begin{aligned} W(\vec{r} - \vec{r}') &= \Delta^2 e^{-\sqrt{2}|\vec{r} - \vec{r}'|/d}, \\ W(q) &= \pi \Delta^2 d^2 (1 + q^2 d^2/2)^{-3/2}. \end{aligned} \quad (7)$$

The important difference from the Gaussian case is that here $W(q)$ decays as q^{-3} at large q . This leads to a stronger scattering at large n . One way to envision this kind of roughness is to think about randomly distributed flat islands of an additional lattice layer with typical diameter d on the top of the last complete layer of the crystal. Our calculations of the roughness-limited mobility for accumulation layers are focused on this type of surface roughness. However, to make a connection with earlier studies²⁻⁸, we will also calculate the mobility for the Gaussian model and compare the results of these two models.

IV. QUASI-CLASSICAL PICTURE

Inspired by Ref. 6, in this section we start from an intuitive quasi-classical picture of the electron scattering

by the surface roughness and get the scaling result shown in Fig. 1. Electrons are scattered when they hit the rough “hard wall” surface. The time between two consecutive collisions of electrons with the surface is $\sim L/v_F \sim m^*L/\hbar k_F$. For each bounce, the reflection is specular with respect to the tangential plane of the hitting point and therefore adds a random angle α to the direction of the reflected momentum. Due to this angular diffusion, the total relaxation of the momentum direction requires α^{-2} times collisions. Thus the relaxation time is

$$\tau = \frac{m^*L}{\hbar k_F \alpha^2}. \quad (8)$$

Below we are going to investigate the deviation angle α at different values of k_F and thus find τ .

For the exponential surface roughness, one can imagine the irregularities as islands going up or down. Each island is flat on a scale d and drops or rises abruptly by a height Δ on the edges. An electron can be regarded as a particle only on length scales larger than the wavelength k_F^{-1} . At $d \ll k_F^{-1}$, electrons can only feel an averaged roughness of all islands within the region of size k_F^{-1} whose number is $(k_F^{-1})^2/d^2 = 1/k_F^2 d^2$. Due to the randomness of the distribution of these islands, the resulting height or depth has a magnitude $\sim \Delta/\sqrt{1/k_F^2 d^2} = k_F d \Delta$. Such a height/depth on a length scale k_F^{-1} effectively results in $\alpha = \Delta k_F d / k_F^{-1} = \Delta d k_F^2$. Using Eq. (8), we get for $k_F d \ll 1$

$$\tau \sim \frac{1/(\Delta d k_F^2)^2}{\hbar k_F / m^* L} = \frac{m^* L}{\hbar \Delta^2 d^2 k_F^5}. \quad (9)$$

In the opposite case at $k_F^{-1} \ll d$, the electron hits a single island each time it bounces off the surface. However, when electrons hit the flat middle plane of the island, there is no momentum relaxation. Only when electrons happen to hit the sharp edges can they get a “scattering” reflection. Since only on a scale k_F^{-1} can electrons be seen as quasi-classical particles, the scattering edge is then estimated to be of height Δ and size k_F^{-1} which gives rise to $\alpha = \Delta k_F$ (see Fig. 3b). The probability to hit one such edge is proportional to its area fraction $k_F^{-1} d / d^2 = 1/k_F d$. This gives

$$\tau \sim \frac{1/(\Delta k_F)^2 m^* L}{1/k_F d \hbar k_F} = \frac{m^* L d}{\hbar \Delta^2 k_F^2}. \quad (10)$$

Since the mobility μ is $e\tau/m^*$ and the conductivity is $ne\mu$, we then get the expressions of μ and σ in terms of k_F , Δ , d and L . Since $L \sim a_B/(na_B^2)^{1/5}$, $k_F \sim (na_B^2)^{2/5}/a_B$ as we said in Sec. II, we then get the mobility as

$$\mu \sim \frac{e}{\hbar} \frac{1}{\Delta^2} \times \begin{cases} \frac{a_B^{8/5}}{d^2 n^{11/5}}, & na_B^2 \ll (a_B/d)^{5/2}, \\ \frac{a_B d}{n}, & na_B^2 \gg (a_B/d)^{5/2}, \end{cases} \quad (11)$$

and the 2D conductivity $\sigma = ne\mu$ as shown in Fig. 1, where $k_F d \sim 1$ at $na_B^2 \sim (a_B/d)^{5/2}$. When $d > a_B$, in the 3D regime where $na_B^2 > 1$ there is no range of $k_F d \ll 1$ and the mobility always decreases as $\propto 1/n$. One should note that the 2D conductivity saturation $\sigma/(2e^2/h) \sim da_B/\Delta^2$ at large concentration n is usually much larger than unity as $d, a_B \gg \Delta$ and implies that the accumulation layer remains metallic at large concentrations.

V. SINGLE SUBBAND CASE

Let us turn to the more formal quantum-mechanical approach now. The scattering rate $1/\tau$ of an electron at the Fermi level with the wave vector \vec{k}' can be found according to Fermi's golden rule:

$$\frac{1}{\tau} = \frac{2\pi}{\hbar} \int \frac{d^3 \vec{k}}{(2\pi)^3} \frac{|U(q)|^2}{\epsilon(q)^2} \delta(\epsilon - \epsilon_F) (1 - \cos \theta') \quad (12)$$

where $\epsilon = \hbar^2 k^2 / 2m^*$, $\epsilon_F = \hbar^2 k'^2 / 2m^* = \hbar^2 k_F^2 / 2m^*$ are the final and initial (Fermi level) energies of an electron, \vec{k} is the final electron momentum, k_F is the Fermi wavenumber, θ' is the angle between initial and final electron momenta and q is the magnitude of the projection of the transferred momentum $\vec{q} = \vec{k} - \vec{k}'$ in the $x - y$ plane. Due to the electronic screening inside a single subband, the Fourier transform of the scattering potential $U(q)$ is reduced by the dielectric function $\epsilon(q)$.

Let us start from the single subband case where the scattering occurs within the same subband and $q = 2k_F \sin(\theta'/2)$. The dielectric function is⁵

$$\epsilon(q) \simeq 1 + 2/a_B q. \quad (13)$$

Let us first derive the scattering potential resulting from the surface roughness. We know that electrons are confined near the interface and thus have a quantization kinetic energy $E_z = \hbar^2 k_z^2 / 2m^*$ where k_z is a multiple of π/L (for the first subband, $k_z = \pi/L$). Due to the surface roughness, the confinement width L fluctuates by $\Delta(\vec{r})$ at position r . The kinetic energy then varies by $(dE_z/dL)\Delta(\vec{r}) \sim E_z \Delta(\vec{r})/L$. These fluctuations of the quantization kinetic energy act as a fluctuating potential $U(\vec{r})$ for the 2-dimensional motion of confined electrons. Its scattering matrix element for 2D Bloch states $U(q)$ within a given subband then satisfies

$$|U(q)|^2 = \left(\frac{E_z}{L}\right)^2 W(q) \delta(k_z - \pi/L), \quad (14)$$

As a result we get for the scattering rate the following expression:

$$\frac{1}{\tau} \sim \frac{\hbar}{m^*} \frac{k_z^4}{L^2} \int d\theta' \frac{W(q)}{\epsilon(q)^2} (1 - \cos \theta'). \quad (15)$$

At $k_F d \ll 1$, according to Eqs. (6) and (7), two models of roughness give the same $W(q) \sim \Delta^2 d^2$. For the one

subband case, $k_F a_B \simeq 1$ and we get the scattering rate as

$$\frac{1}{\tau} \sim \frac{1}{\hbar} \frac{m^*}{\hbar^2} \left(\frac{E_z}{L} \right)^2 \Delta^2 d^2 \sim \frac{\hbar}{m^*} \frac{k_z^4 \Delta^2 d^2}{L^2}. \quad (16)$$

Since $k_z \sim 1/L$, the mobility is then

$$\mu \sim \frac{e}{\hbar} \frac{L^6}{\Delta^2 d^2}. \quad (17)$$

Since in the single subband case $r_s \gtrsim 1$ or $na_B^2 \lesssim 1$, the condition of validity of Eq. (17) $k_F d \ll 1$ is fulfilled for the roughness with $d < a_B$. The case $d < a_B$ was studied for silicon inversion layers in Ref. 3 and discussed in Introduction above. For the 2D inversion layers, the width L is determined by the applied electric field as $L \propto E^{-1/3}$ and one gets^{2,3,6} $\mu \propto 1/E^2$. As the electron concentration increases, the interfacial electric field $E \propto n$ and mobility $\mu \propto 1/n^2$. Such a dependence was obtained in Ref. 3 and used in Ref. 24 for extrapolation to the multisubband case as shown in Fig. 1. Note that for the 2D quantum wells, the width L of the electron gas is the same as the well width so that Eq. (17) agrees with the well known result of Ref. 7.

However, if $d > a_B$, for the single subband case, there is also a range of concentrations that satisfies $k_F d \gg 1$. The results of two roughness models are different. For the Gaussian case, the typical q is $\sim 1/d$ and the typical θ' is $\sim q/k_F \sim 1/k_F d$, $W(q) \sim \Delta^2 d^2$, $\epsilon(q) \sim d/a_B$. The scattering rate is then

$$\frac{1}{\tau} \sim \frac{\hbar}{m^*} \frac{k_z^4 \Delta^2 a_B^2}{L^2 k_F^3 d^3}. \quad (18)$$

Putting $k_z \sim 1/L$ for the single subband case, we get the mobility as

$$\mu \sim \frac{e}{\hbar} \frac{L^6 d^3 k_F^3}{\Delta^2 a_B^2}, \quad (19)$$

which can also be obtained from results in Refs. 6 and 8.

For the exponential case, $W(q)$ decays in a much milder way as $\propto 1/q^3$ at large q . This leads to the large angle scattering. Indeed, let us consider the contribution to the integral in Eq. (15) from the small angles $\theta' \sim (k_F d)^{-1}$ and large angles $\theta' \sim 1$. In the first case, $W(q \sim 1/d) \sim \Delta^2 d^2$, $\epsilon(q) \sim d/a_B$ and

$$\int_0^{(k_F d)^{-1}} d\theta' (1 - \cos \theta') \simeq \theta'^3 \simeq \frac{1}{(k_F d)^3}.$$

As a result the contribution from the small angles to the scattering rate is the same as Eq. (18). For the large-angle scattering $W(q \sim k_F) \simeq \Delta^2 d^2 [k_F d \sin(\theta'/2)]^{-3}$ is $(k_F d)^3$ times smaller than that of the small-angle scattering, $\epsilon(q) \simeq 1$ and the integral over the angle is

$$\int_{(k_F d)^{-1}}^{\pi} d\theta' (1 - \cos \theta') \frac{1}{[\sin(\theta'/2)]^3} \sim \ln(k_F d).$$

If we ignore the logarithmic factor, the angle integral is $(k_F d)^3$ times larger than that from the small angles, and the scattering rate is then

$$\frac{1}{\tau} \sim \frac{\hbar}{m^*} \frac{k_z^4 \Delta^2}{L^2 k_F^3 d}, \quad (20)$$

which due to the absence of screening is $(d/a_B)^2$ times larger than that from the small-angle scattering, so that

$$\mu \sim \frac{e}{\hbar} \frac{L^6 d k_F^3}{\Delta^2}. \quad (21)$$

The dominance of the large-angle scattering is a unique feature of the exponential roughness.

VI. INTERSUBBAND SCATTERING IN MULTISUBBAND ACCUMULATION LAYERS

In Sec. V we have calculated the mobility limited by the surface roughness scattering of a single subband. We not only have recovered the results for the Gaussian type of roughness obtained by previous studies but also have got the results for the relatively unexplored exponential roughness. For multisubband accumulation layers, the situation is different from the single subband case. First, k_z is not $\sim 1/L$ but typically is $\sim k_F$, where $k_F = (n/L)^{1/3}$ is the 3D Fermi wavenumber of the electron gas. Second, in addition to the intrasubband scattering, there is also intersubband scattering.

Typically, the intersubband scattering rate scales in the same way as the intrasubband scattering. Therefore, as shown in Appendix A, the final scattering rate is approximately the intrasubband scattering rate times the number of subbands which the initial electron can be scattered into. The typical transferred momentum in the z -direction is of the order of the typical transferred momentum q in the $x-y$ plane. At $1/d \gg k_F$ where $q \sim k_F$, all subbands “communicate” with each other and the total number is $k_F L$. Multiplying by $k_F L$ the intrasubband result Eq. (16) with $k_z \sim k_F$, we arrive at the final scattering rate given by Eq. (9)²⁹. This is a universal result for both Gaussian and exponential models. At $1/d \ll k_F$, for the exponential roughness, the typical transferred momentum is $q \sim 1/d$ for the small-angle scattering and $q \sim k_F$ for the large-angle one. So for the former, the number of subbands involved in the scattering process is $\sim L/d$, much smaller than the number $k_F L$ for the latter. Therefore, the large-angle scattering mechanism dominates. Using the intrasubband scattering rate as given by Eq. (20) and $k_z \sim k_F$, we arrive at the total scattering rate at $k_F d \gg 1$ for the exponential roughness given by Eq. (10). Thereby, we get the same expressions for the mobility as in Eq. (11). The corresponding 2D conductivity at large n saturates as $\sigma/(2e^2/h) \sim da_B/\Delta^2$ (see Fig. 1), which means that there is no re-entrant metal-insulator transition. More careful check of this statement requires

TABLE I. Mobility μ in units of $(e/\hbar\Delta^2)$ as a function of the 2D electron concentration n at different values of d for two types of surface roughness, the Gaussian model (G) and the exponential one (E). Since the Fermi wavenumber $k_F \simeq (n/L)^{1/3}$, and the width of the 2D electron gas $L \sim a_B/(na_B^2)^{1/5}$, we get $k_F^{-1} = d$ at $na_B^2 \sim (a_B/d)^{5/2}$ and $L = d$ at $na_B^2 \sim (a_B/d)^5$.

	$d < k_F^{-1}$	$k_F^{-1} < d < L$	$L < d < a_B$	$a_B < d$
G	$a_B^{8/5}/n^{11/5}d^2$	$a_B^{4/5}d^2/n^{3/5}$	$a_B^{7/5}d/n^{4/5}$	$d^3/a_B^{3/5}n^{4/5}$
E	$a_B^{8/5}/n^{11/5}d^2$	$a_B d/n$	$a_B d/n$	$a_B d/n$

accurate evaluations of the numerical coefficients which are done in Appendix A and confirm our argument that the saturation value of the conductivity is much larger than the critical value $2e^2/h$.

Compared to the simple and neat result in the exponential case, the mobility for the Gaussian type of roughness is more complicated. At $k_F d \gg 1$, since for accumulation layers $na_B^2 > 1$, $d < a_B$ is possible in this range. At this value of d , the intrasubband scattering is unscreened for the Gaussian case and the rate becomes the same as the unscreened result of the exponential roughness given by Eq. (20). The typical momentum transfer is $q \sim 1/d$, so only a few subbands whose number is $\sim q/L^{-1} = L/d$ out of all participate in the intersubband scattering. The total scattering rate is then the scattering rate in Eq. (20) times L/d . At even larger $d \gg L \gg k_F^{-1}$, the typical momentum transferred $1/d$ is smaller than the z -direction momentum quantization $1/L$. No intersubband scattering is possible and the total scattering rate is given by Eq. (20). This situation resembles the single subband case and is natural since when $L \ll d$ we actually are dealing with a 2D system. When d exceeds the 2D screening radius $a_B \gg L$, not only the scattering is restricted within the same subband but also the potential is screened. One should then use Eq. (18) for the total (intrasubband) scattering. The result is summarized as follows

$$\frac{1}{\tau} \sim \begin{cases} \frac{\hbar k_F \Delta^2}{m^* L d^2}, & k_F^{-1} \ll d \ll L, & (22a) \\ \frac{\hbar k_F \Delta^2}{m^* L^2 d}, & L \ll d \ll a_B, & (22b) \\ \frac{\hbar k_F \Delta^2 a_B^2}{m^* L^2 d^3}, & a_B \ll d. & (22c) \end{cases}$$

$k_F^{-1} = d$ is reached at $na_B^2 \sim (a_B/d)^{5/2}$ and $d = L$ is achieved at $na_B^2 \sim (a_B/d)^5$. By expressing k_F and L in terms of n , one can then get the mobility as a function of the 2D electron concentration. The corresponding results together with that for the exponential case are listed in Table. I.

At $d \ll a_B$, the obtained $\mu(n)$ dependence is presented in Fig. 4. For the exponential roughness, the corresponding scaling behavior of the 2D conductivity σ is shown in Fig. 1. At $d \gg a_B$, both $(a_B/d)^{5/2}$ and $(a_B/d)^5$ is much smaller than 1. Throughout the 3D metallic range where $na_B^2 > 1$, there is only one result for the Gaussian and exponential models respectively as given by the last column in Table. I. The smallest conductivity for the exponential case is larger than $2e^2/h$ as mentioned above. For the Gaussian model, at $d \ll a_B$, the smallest $\sigma/(2e^2/h)$ is also $\sim da_B/\Delta^2$. Since in reality, $d > \Delta$, we get $\sigma/(2e^2/h) \gg 1$. At $d \gg a_B$, $\sigma/(2e^2/h) \geq d^3/a_B\Delta^2 \gg 1$. So, the smallest conductivity for the Gaussian roughness is also always above the critical value and no re-entrant metal-insulator transition will happen in realistic situations.

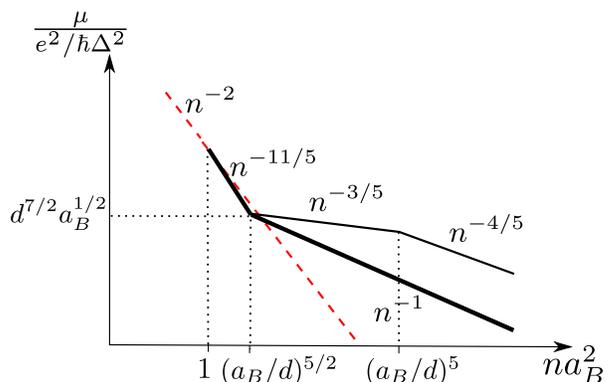


FIG. 4. The scaling behavior of the mobility μ in units of $e^2/\hbar\Delta^2$ as a function of the electron concentration n in units of a_B^{-2} at $d < a_B$ in a double logarithmic scale. The thick solid line (black) denotes the mobility of the accumulation layer for the exponential roughness. The thin solid line (black) represents the mobility for the Gaussian roughness which also decreases. Here only powers of the n dependence are shown while the complete scaling formulae are presented in Table. I. The thin dashed line (red) represents the $1/n^2$ dependence derived for a single subband³.

In Sec. IV, we gave a quasi-classical explanation of the mobility limited by the exponential surface roughness. Inspired by Ref. 6, we can interpret the Gaussian roughness results quasi-classically as well. Below we again start from Eq. (8) and find α for the Gaussian roughness in different situations. At $k_F d \ll 1$, the roughness relief is averaged over the electron wavelength and the resulting relaxation time is the same as in the exponential case given by Eq. (9). At $k_F^{-1} \ll d \ll L$, the wavelength k_F^{-1} is smaller than the size of the roughness hill. The electron then collides with a single hill (valley) each time it hits the surface and the deviation angle α is the slope of each single hill (valley) $\sim \Delta/d$ (see Fig. 3a). The relaxation time is given by Eq. (22a). At $L \ll d \ll a_B$, the size of each single hill is so large that the electron can hit the same hill consecutively for several times during which it has traveled back and

forth for $\sim d/L$ times within the accumulation layer. Since these consecutive hits are on the same slope, the scattered angle is the same and α accumulates, different from uncorrelated random collisions on different hills (valleys). After the electron finishes colliding with the same slope, the accumulated angle is $(\Delta/d)(d/L) = (\Delta/L) \ll 1$ and the time of such a series of collisions is $\sim (L/k_F)(d/L) = d/k_F$. One then gets the relaxation time as in Eq. (22b). At $d \gg a_B$, the electronic screening changes the scattering potential by a factor a_B/d , which can effectively be regarded as reducing the roughness height from Δ to $\Delta a_B/d$. The resulting relaxation time grows and is given by Eq. (22c). Thus we obtained quasi-classically the same scaling behavior of the mobility limited by the Gaussian roughness as by the quantum-mechanical approach.

VII. CONCLUSION

In this paper, we have studied the surface-roughness limited mobility in inversion and multisubband accumulation layers as a function of the 2D electron concentration n for two models of the surface roughness both quantum-mechanically and quasi-classically. For the more realistic exponential roughness, the mobility decreases as $\propto 1/n$ at large n and results in a 2D conductivity saturation as $\sigma/(2e^2/h) \simeq 1.2da_B/\Delta^2 \gg 1$ since the characteristic roughness size d and the effective Bohr radius a_B are larger than the characteristic roughness height $\Delta \simeq a/2$ where a is the lattice constant. For the Gaussian roughness which was widely used in earlier studies, the minimum conductivity is found to be larger than the critical value as well. So there is no reason to expect the re-entrant metal-insulator transition²⁴ at large concentrations. Indeed, decent conductivities were observed in large concentration accumulation layers in Refs. 10, 14, 15, and 25. By measuring the conductivity saturation value $\sigma/(2e^2/h) \simeq 1.2da_B/\Delta^2$ of the exponential roughness, one can probe the geometry of the roughness obtaining values of d , similarly to what was done for the Gaussian roughness in previous works^{3,7}.

Acknowledgments.

We are grateful to S. A. Campbell, A. V. Chaplik, M. V. Entin, B. Jalan, M. J. Manfra, M. Sammon, M. Shur, and R. A. Suris for helpful discussions. This work was supported primarily by the National Science Foundation through the University of Minnesota MRSEC under Award No. DMR-1420013.

Appendix A: Numerical coefficients in $\sigma(n)$ dependence in the exponential model

In Eq. (3), it shows that about 90% of electrons are located within a distance $L/2$ from the interface. So it is a good approximation to assume that electrons inside the accumulation layer are more accurately confined within a width of $D = L/2$. Then, the wave function of each subband can be approximated as

$$\xi(r, z) \simeq \sqrt{\frac{2}{D}} \exp(i\vec{k}_r \cdot \vec{r}) \sin(k_z z) \quad (\text{A1})$$

where \vec{k}_r , $k_z = m\pi/D$ with m being a positive integer are respectively the $x-y$ plane and z -direction momenta of electrons in this subband and different subbands correspond to different values of k_z . Therefore, similar to that in Ref. 3, the matrix element $U(q)$ satisfies

$$\langle |U|^2 \rangle = \frac{8}{D} \varepsilon_F^2 \left(\frac{k'_z k_z}{k_F^2} \right)^2 W(q) \quad (\text{A2})$$

where an isotropic mass spectrum is assumed, k'_z, k_z are the initial and final z -components of the electron wavevector.

At $k_F d \ll 1$, for both models of roughness, $W(q) = \pi \Delta^2 d^2$. So the scattering rate in Eq. (12) can be rewritten as

$$\begin{aligned} \frac{1}{\tau} &= \frac{2}{\pi \hbar} \frac{\varepsilon_F^2 \Delta^2 d^2}{D} \int k_r dk_r d\phi dk_z \left(\frac{k'_z k_z}{k_F^2} \right)^2 \\ &\quad \left(1 - \frac{k'_z k_z + k'_r k_r \cos \phi}{k_F^2} \right) \delta \left[\frac{\hbar^2 k_F^2}{2m^*} - \frac{\hbar^2 (k_z^2 + k_r^2)}{2m^*} \right] \\ &= \frac{2\varepsilon_F \Delta^2 d^2 k_F^3}{\hbar D} \int d(\cos \theta) (\cos \theta \cos \theta_0)^2 (1 - \cos \theta \cos \theta_0) \end{aligned} \quad (\text{A3})$$

where $\cos \theta = k_z/k_F$, $\cos \theta_0 = k'_z/k_F$, k'_r is the $x-y$ component magnitude of the initial momentum \vec{k}' . One should note that this scattering rate is for one specific direction of the initial momentum k' . To get the averaged result, one should average over all θ_0 and get

$$\frac{1}{\tau} \simeq \frac{0.1\varepsilon_F \Delta^2 d^2 k_F^3}{\hbar D}. \quad (\text{A4})$$

At $k_F d \gg 1$, for the exponential case, using Eq. (7), we have

$$\begin{aligned} \frac{1}{\tau} &= \frac{2}{\pi \hbar} \frac{\varepsilon_F^2 \Delta^2 d^2}{D} \int k_r dk_r d\phi dk_z d(k'_z/k_F) \left(\frac{k'_z k_z}{k_F^2} \right)^2 \\ &\quad \left\{ 1 + \left[k_r^2 + (k'_r)^2 - 2k_r k'_r \cos \phi \right] d^2/2 \right\}^{-3/2} \\ &\quad \left(1 - \frac{k'_z k_z + k'_r k_r \cos \phi}{k_F^2} \right) \delta \left[\frac{\hbar^2 k_F^2}{2m^*} - \frac{\hbar^2 (k_z^2 + k_r^2)}{2m^*} \right] \\ &\simeq \frac{0.73\varepsilon_F \Delta^2}{\hbar D d}, \end{aligned} \quad (\text{A5})$$

which matches Eq. (A4) at $k_F d \simeq 2$. Substituting $k_F = (3\pi^2 n/D)^{1/3}$ into Eqs. (A4) and (A5), where

$D = L/2$ and L is given by Eq. (4), we obtain numerical coefficients for $\sigma(n)$ mentioned in Introduction and the corresponding saturation point $na_B^2 \simeq 0.4(a_B/d)^{5/2}$.

-
- * fuxxx254@umn.edu
- ¹ R. E. Prange and T.-W. Nee, Phys. Rev. **168**, 779 (1968).
 - ² A. V. Chaplik and M. V. Entin, Soviet Physics - JETP **28**, 514 (1969); M. Entin, Soviet physics - Solid state **1**, 781 (1969).
 - ³ T. Ando, Journal of the Physical Society of Japan **43**, 1616 (1977).
 - ⁴ S. Mori and T. Ando, Phys. Rev. B **19**, 6433 (1979).
 - ⁵ T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
 - ⁶ M. Krylov and R. Suris, Soviet Physics - JETP **56**, 1316 (1982).
 - ⁷ H. Sakaki, T. Noda, K. Hirakawa, M. Tanaka, and T. Matsusue, Applied Physics Letters **51**, 1934 (1987).
 - ⁸ A. Gold, Phys. Rev. B **35**, 723 (1987).
 - ⁹ U. Penner, H. Rcker, and I. N. Yassievich, Semiconductor Science and Technology **13**, 709 (1998).
 - ¹⁰ A. Tardella and J.-N. Chazalviel, Phys. Rev. B **32**, 2439 (1985).
 - ¹¹ J. Nelson and A. M. Goldman, Phys. Rev. B **91**, 241304 (2015).
 - ¹² D. K. Efetov and P. Kim, Phys. Rev. Lett. **105**, 256805 (2010).
 - ¹³ D. K. Efetov, P. Maher, S. Glinskis, and P. Kim, Phys. Rev. B **84**, 161412 (2011).
 - ¹⁴ H. Yuan, H. Shimotani, A. Tsukazaki, A. Ohtomo, M. Kawasaki, and Y. Iwasa, Advanced Functional Materials **19**, 1046 (2009).
 - ¹⁵ Y. Zhang, J. Ye, Y. Matsushashi, and Y. Iwasa, Nano Letters **12**, 1136 (2012).
 - ¹⁶ K. Ueno, S. Nakamura, H. Shimotani, A. Ohtomo, N. Kimura, T. Nojima, H. Aoki, Y. Iwasa, and M. Kawasaki, Nat. Mater. **7**, 855 (2008).
 - ¹⁷ P. Gallagher, M. Lee, J. R. Williams, and D. Goldhaber-Gordon, Nat. Phys. **10**, 748 (2014).
 - ¹⁸ J. Y. Zhang, C. A. Jackson, R. Chen, S. Raghavan, P. Moetakef, L. Balents, and S. Stemmer, Phys. Rev. B **89**, 075140 (2014).
 - ¹⁹ P. Xu, T. C. Droubay, J. S. Jeong, K. A. Mkhoyan, P. V. Sushko, S. A. Chambers, and B. Jalan, Advanced Materials Interfaces **3**, 1500432 (2016).
 - ²⁰ A. Bykhovski, B. Gelmont, and M. Shur, Journal of Applied Physics **74**, 6734 (1993).
 - ²¹ J. S. Speck and S. F. Chichibu, MRS Bulletin **34**, 304 (2009).
 - ²² J. Yang, X. Hu, J. Deng, R. Gaska, M. Shur, and G. Simin, in *Semiconductor Device Research Symposium, 2009. ISDRS '09. International* (2009) pp. 1–2.
 - ²³ S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, Phys. Rev. B **32**, 8171 (1985).
 - ²⁴ S. Das Sarma and E. H. Hwang, Phys. Rev. B **89**, 121413 (2014).
 - ²⁵ Y. Takahide, H. Okazaki, K. Deguchi, S. Uji, H. Takeya, Y. Takano, H. Tsuboi, and H. Kawarada, Phys. Rev. B **89**, 235304 (2014).
 - ²⁶ J. Frenkel, Zeitschrift für Physik **51**, 232 (1928).
 - ²⁷ K. V. Reich, M. Schechter, and B. I. Shklovskii, Phys. Rev. B **91**, 115303 (2015).
 - ²⁸ R. M. Feenstra, D. A. Collins, D. Z. Y. Ting, M. W. Wang, and T. C. McGill, Phys. Rev. Lett. **72**, 2749 (1994).
 - ²⁹ One could worry about the use of single subband results based on the 2D dielectric function for the 3D multisubband accumulation layer with the width L . Indeed, it is known that in a metallic film with width L at the distance $q^{-1} \gg L$ the screening is two-dimensional, while at $q^{-1} \ll L$ it becomes three-dimensional. In our case, the 3D screening radius inside the layer is $r_D = a_B/(Na_B^3)^{1/6} \sim a_B/(na_B^2)^{1/5} \sim L \ll a_B$. Thus, for $q^{-1} < L$ the 3D dielectric function is ~ 1 . At these q values it coincides with the large q limit of Eq. (13) because $L < a_B$. Thus, the 3D screening may be ignored.