

Two-dimensional metal-insulator transition as a potential fluctuation driven semiclassical transport phenomenon

S. Das Sarma,¹ E. H. Hwang,^{1,2} and Qiuzi Li¹¹*Condensed Matter Theory Center, Department of Physics, University of Maryland, College Park, Maryland 20742, USA*²*SKKU Advanced Institute of Nanotechnology and Department of Physics, Sungkyunkwan University, Suwon 440-746, Korea*

(Received 11 June 2013; published 10 October 2013)

We theoretically consider the carrier density tuned (apparent) two-dimensional (2D) metal-insulator transition (MIT) in semiconductor heterostructure-based 2D carrier systems as arising from a classical percolation phenomenon in the inhomogeneous density landscape created by the long-range potential fluctuations induced by random quenched charged impurities in the environment. The long-range Coulomb disorder inherent in semiconductors produces strong potential fluctuations in the 2D system where a fraction of the carriers gets trapped or classically localized, leading to a mixed two-component semiclassical transport behavior at intermediate densities where a fraction of the carriers is mobile and another fraction immobile. At high carrier density, all the carriers are essentially mobile, whereas at low carrier density, all the carriers are essentially trapped since there is no possible percolating transport path through the lake-and-mountain inhomogeneous potential landscape. The low-density situation with no percolation would mimic an insulator, whereas the high-density situation with allowed percolating paths through the lake-and-mountain energy landscape would mimic a metal with the system manifesting an apparent MIT in-between. We calculate the transport properties as a function of carrier density, impurity density, impurity location, and temperature using a two-component (trapped and mobile carriers) effective medium theory. Our theoretically calculated transport properties are in good qualitative agreement with the experimentally observed 2D MIT phenomenology in 2D electron and hole systems. We find a high-(low-) density metallic (insulating) temperature dependence of the 2D resistivity, and an intermediate-density crossover behavior which could be identified with the experimentally observed 2D MIT. The calculated density- and temperature-dependent resistivity in our theory mimics the phenomenology of 2D MIT experiments with reasonable parameter values for the background disorder.

DOI: [10.1103/PhysRevB.88.155310](https://doi.org/10.1103/PhysRevB.88.155310)

PACS number(s): 73.40.-c, 71.30.+h

I. INTRODUCTION

The carrier density tuned two-dimensional (2D) metal-insulator transition (MIT) phenomenon is ubiquitous in semiconductor heterostructures at low temperatures.¹⁻⁶ At “higher” carrier density, the measured resistivity manifests “metallic” temperature dependence, whereas at “lower” carrier density, the 2D resistivity manifests “insulating” temperature dependence with a complex density-dependent crossover behavior in the “intermediate” density regime where the system makes a “transition” from being a high-density effective metal to a low-density effective insulator. In Fig. 1, we show, purely to provide a context for the current theoretical work, a set of representative 2D MIT experimental transport data⁷⁻⁹ for 2D electrons in Si MOSFETs, 2D electrons in GaAs heterostructures, and 2D holes in GaAs quantum wells, taken from the experimental publications of several different groups over the years. (These three systems are by far the most extensively studied experimental systems in the 2D MIT studies over the last 15 years.⁷⁻²⁸)

The current theoretical work presented in this paper describes a simple and intuitively appealing semiclassical model as a plausible explanation for the 2D MIT phenomenon. After developing the model in depth, we carry out numerical calculations for the density- and temperature-dependent 2D resistivity which can be compared with the experimental observations (Fig. 1). The qualitative agreement between our theoretical results and the experimental 2D MIT data is *prima facie* evidence that our proposed physical mechanism is likely to be playing a role (perhaps even a major role) in the underlying

physics of 2D MIT phenomena.^{7-12,15,18,19,21-23,26-29} We can not, however, rule out (certainly, not conclusively) the possibility of alternate physical mechanisms also playing some role in the 2D MIT phenomena since our model is based only on the single physical mechanism described in this work.

The physical mechanism we consider is disorder-induced density inhomogeneity necessarily present in semiconductors at low carrier densities where screening effects are sufficiently weak so that the long-range Coulomb disorder introduced by random quenched charged impurities invariably present in the semiconductor environment leads to strong potential fluctuations in the system.³⁰⁻³² These potential fluctuations are screened out at high carrier densities, but at low carrier densities they create highly inhomogeneous “lakes-and-hills” type potential landscape in which the carriers move around within the physical sample.³³ In addition to creating the lakes-and-hills (or equivalently “valleys-and-mountains”) inhomogeneous potential landscape, the charged impurities, of course, also cause momentum scattering of the carriers, leading to the measured resistivity. Since 2D MIT typically is a low-density phenomenon where screening is weak, the “lakes-and-hills” potential landscape is expected to play a key role in the phenomenon.

The key ingredient of our model is that we treat the carrier system as an effective two-component (or equivalently two-phase) system: bound and unbound carriers or, equivalently, trapped (i.e., immobile) and mobile (i.e., free) carriers (we use the latter terminology throughout), where all classically trapped carriers in the potential well (or valley) regions are

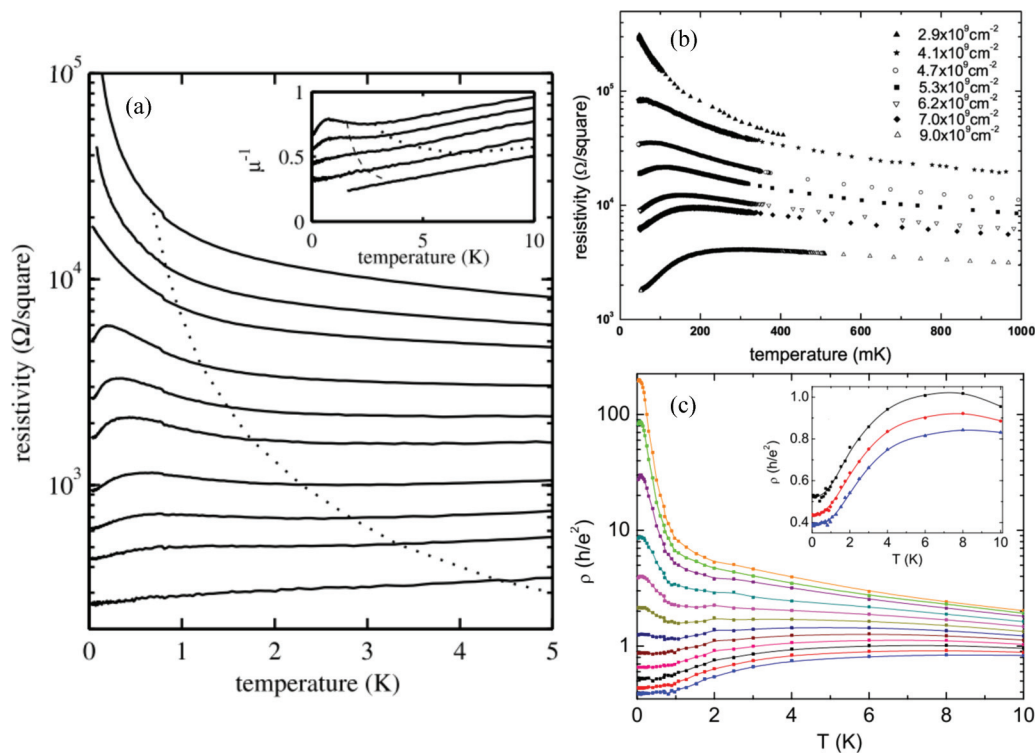


FIG. 1. (Color online) (a) Temperature dependence of the observed resistivity for 2D electrons in GaAs heterostructure. Adapted from Ref. 7. (b) Temperature dependence of the observed resistivity for p -type GaAs quantum well system. Adapted from Ref. 8. (c) Experimental resistivity as a function of temperature for Si MOSFETs. Adapted from Ref. 9.

assumed to be localized without contributing to the conductivity (\equiv the inverse of resistivity) at zero temperature ($T = 0$). Thus, the “trapped” carriers in the classically forbidden region (with $E \leq 0$) do not contribute to the $T = 0$ conductivity (or, equivalently, have infinite resistivity at $T = 0$), and the $T = 0$ conductivity is determined entirely by the mobile carriers (with $E \geq 0$) which are classically free. Our basic two-component transport model is thus a classical (or semiclassical) model with the carriers being divided into bound or trapped or immobile carriers and unbound or free or mobile carriers because of the “lakes-and-hills” potential fluctuation landscape induced by the random charged impurities. It is obvious that in the absence of Anderson localization (which we neglect) and in the absence of background potential fluctuations (or for very small fluctuations), all the carriers contribute to the $T = 0$ “metallic” conductivity since there are no effective potential wells or barriers to trap the carriers; the same is essentially true when the net carrier density is very high so that the chemical potential or the Fermi energy is very large compared with the typical magnitude of the potential fluctuations. On the other hand, when the Fermi energy is very small (i.e., low carrier density), most of the carriers are immobile, leading to exponentially small (and “insulating” or activated) conductivity.

We assume that the mobile carriers undergo standard disorder scattering limited diffusive transport and as such can be treated by the semiclassical Boltzmann transport theory (taking into account the screening of the Coulomb disorder by the carriers themselves so that the resistive scattering mechanism is the finite-temperature screened Coulomb disorder arising from the quenched random charged impurities^{34–37}).

At finite temperatures (but not at $T = 0$), the trapped carriers also contribute to the conductivity through the activation process as the bound carriers can be thermally excited over the disorder-induced potential fluctuations to become effectively mobile. (We neglect all quantum tunneling and phonon-assisted hopping effects.) Thus, there are two independent transport channels in the problem: Diffusive transport by the mobile electrons/holes and activated transport by the trapped electrons/holes. We use a two-component effective medium theory (EMT),^{38–42} the two components being the fractions of trapped (“activated transport”) and mobile (“diffusive transport”) carriers, to describe the transport behavior in the system. Obviously, high-density transport, when the fraction of mobile carriers is very high, would appear diffusive and metallic, and low-density transport, when the fraction of trapped carriers is very high, would appear activated insulating with a crossover at some disorder-dependent intermediate carrier density.

We emphasize that our basic two-component (trapped and mobile carriers) model is classical as is our two-component EMT transport calculation³⁸ in the sense that quantum interference (i.e., localization) effects are neglected. We calculate the carrier resistivity itself using the semiclassical Boltzmann transport theory which treats carrier momentum scattering by disorder in a quantum mechanical manner. In this sense, our theory is semiclassical. All quantum tunneling and quantum interference (and electron-electron interaction beyond screening) effects are ignored in our theory since the problem becomes intractable otherwise. Also, the two-component classical transport model is inconsistent with

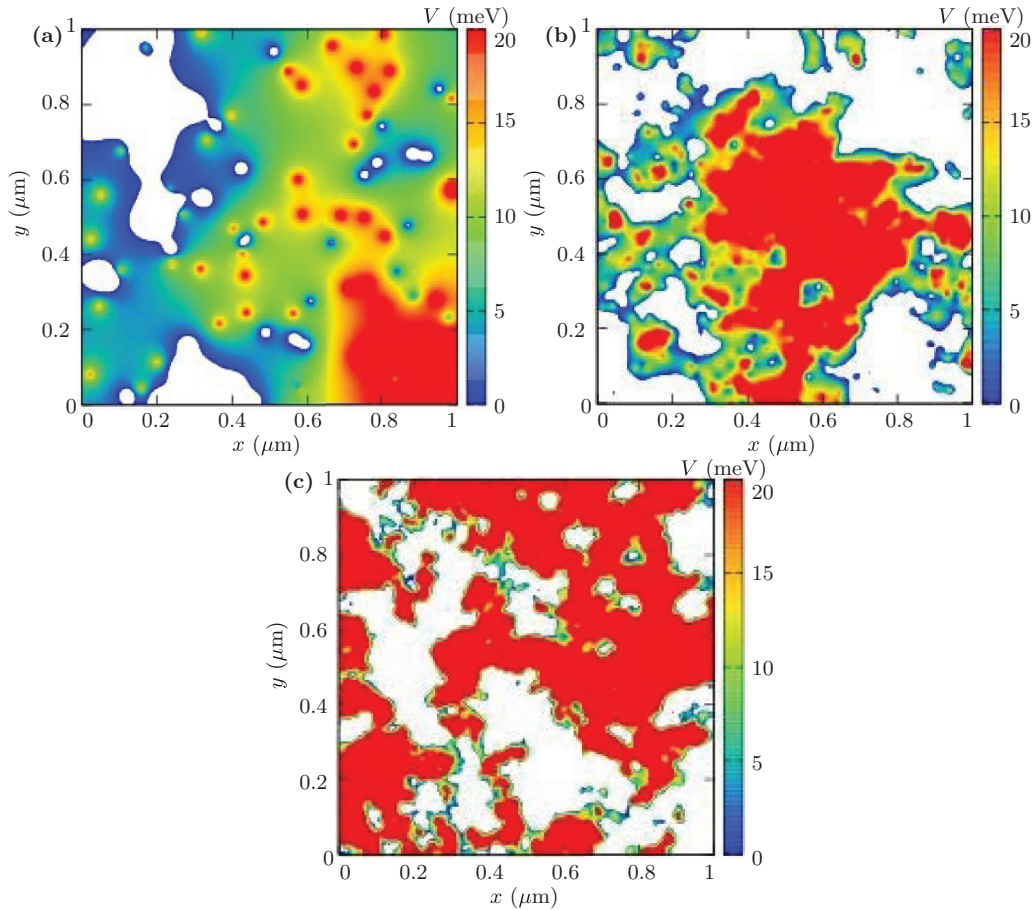


FIG. 2. (Color online) 2D density plot of our simulated $V(\mathbf{r})$ measured from E_F for randomly distributed impurities with $z_0 = 10$ nm and $\kappa = 10$. (a) $n_i = 10^{10}$ cm $^{-2}$. (b) $n_i = 10^{11}$ cm $^{-2}$. (c) $n_i = 10^{12}$ cm $^{-2}$. Since this figure is just for schematic purposes, we do not include screening effects in our simulations.

quantum tunneling and interference (since in the presence of tunneling, the two-component concept does not apply), and our goal here is to study the physically motivated semiclassical two-component transport model as completely as possible without the very difficult additional complications of quantum tunneling and interference. The basic idea here is that the strong potential fluctuations in the semiconductor system arising from random charged impurity disorder and the associated lakes-and-hills energy landscape necessarily lead to the zeroth-order carrier transport properties being dominated by the two-component behavior through the diffusion of mobile carriers and the activation of trapped carriers with all other effects (tunneling, interference, interaction) being weak perturbations in the strongly fluctuating highly inhomogeneous 2D system. The two-component transport is relevant only when the Fermi energy of the carrier system and the typical scale of the potential fluctuations are comparable so that a highly inhomogeneous density and potential landscape dominates the transport properties (Figs. 2 and 3). This is the typical situation for 2D semiconductor systems in the presence of random charged impurity disorder where 2D MIT phenomena occur.

We mention that there is a very long history^{38,43–49} in the literature discussing transport in disordered semiconductor systems using the inhomogeneous two-component model of

trapped and mobile carriers in the lakes-and-hills landscape of potential fluctuations. The problem has a strong formal similarity to the problem of classical percolation in a highly inhomogeneous medium,^{38,50–57} and occasionally the metal-insulator transition in electronic materials, particularly in disordered semiconductors, has been studied using the classical percolation approach.⁵⁸ The key physical point here is that the strong potential fluctuations induce strong density inhomogeneity with the electrons (or holes) forming spatial puddles separated by potential barriers.⁵⁹ We mention here that a large number of 2D experimental^{7–10,18–24,26–29} and theoretical^{39–42,54–64} works in the literature have already suggested the two-component percolation transport as the underlying mechanism for the 2D MIT phenomena.

In Sec. II, we describe our theory in details, providing the numerical results for the calculated 2D transport properties for various situations in Sec. III. We provide a thorough discussion, emphasizing limitations of our theory and corrections to existing theoretical and experimental results in Sec. IV, concluding in Sec. V with a discussion on the comparison with experiment and a summary.

II. THEORY

We use a minimal model to describe the 2D semiconductor system as a 2D electron (or hole) gas (2DEG) of 2D density n

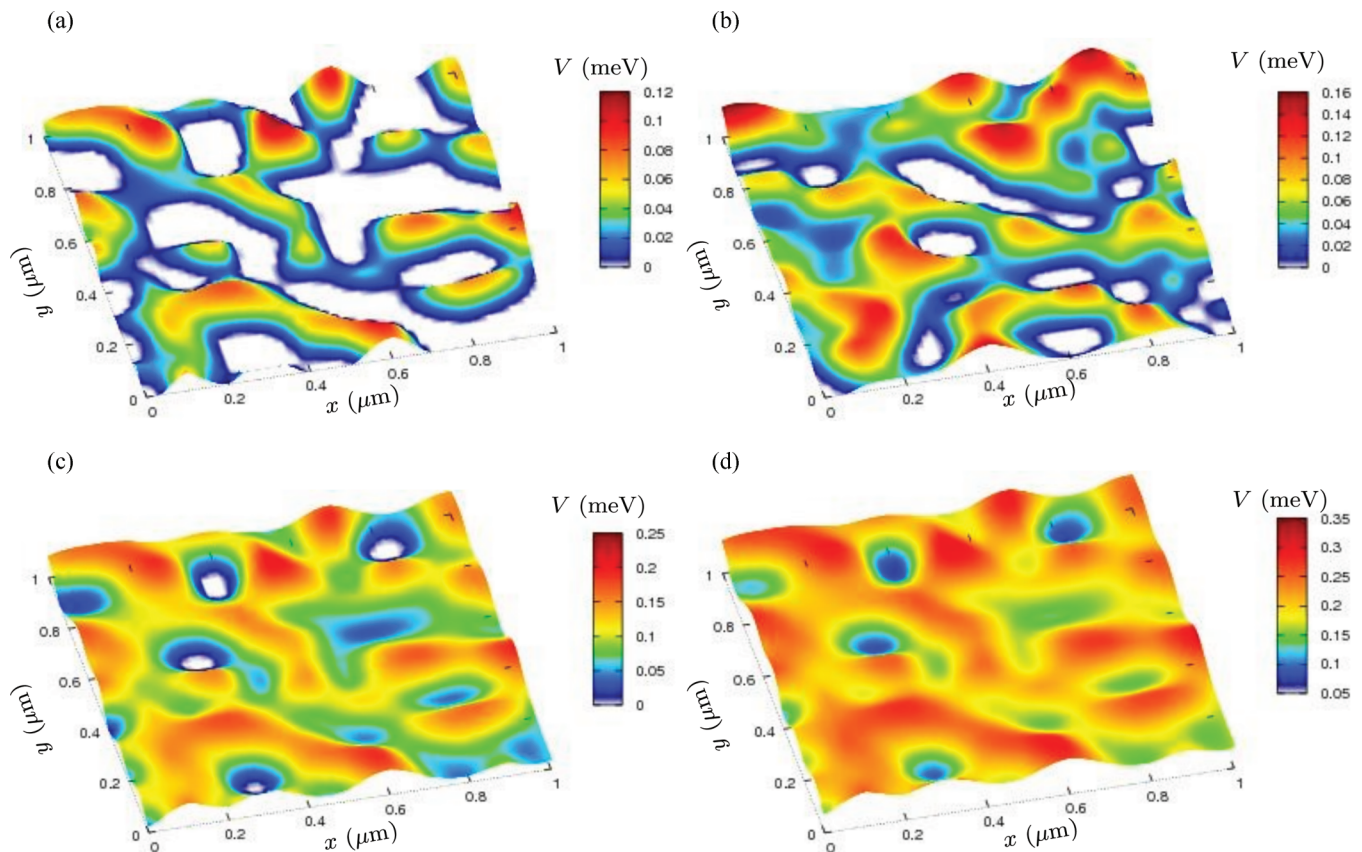


FIG. 3. (Color online) Schematic diagram of $V(\mathbf{r})$ for $s = 0.05$ meV to show the lakes-and-hills inhomogeneous potential landscape. (a) $E_F = 0.01$ meV. (b) $E_F = 0.05$ meV. (c) $E_F = 0.1$ meV. (d) $E_F = 0.2$ meV.

characterized by an effective mass of m , a valley degeneracy factor of g_v , and a background effective dielectric constant κ (which is taken as the average of the lattice dielectric constant for the semiconductor and the insulator defining the 2D heterostructure, i.e., Si-SiO₂, GaAs-AlAs, etc.). We assume a spin-degenerate system with a spin degeneracy $g_s = 2$ throughout (except when discussing the effect of an in-phase magnetic field which could lift the spin degeneracy changing g_s from two at zero magnetic field to one at high magnetic field). We neglect the thickness (in the z direction perpendicular to the 2D layer) of the 2DEG in describing our theory and equations since this has only quantitative, but no qualitative, effect for the 2D MIT physics of interest in this work. Adding a finite 2D layer thickness effect is straightforward, and would only quantitatively modify the relative strength of the disorder in the theory. All our numerical results shown in the figures include the full effects of quasi-2D finite thickness in a realistic manner.

Disorder is included in our model through random charged impurities of concentration (i.e., 2D density) n_i which are located a distance z_0 from the 2DEG in the z direction perpendicular to the layer. We assume that n_i and z_0 together define completely the resistive scattering of the carriers with the disorder potential being modeled by the screened Coulomb interaction.

Finally, we need to specify the potential fluctuations dominating the low-density behavior of the carrier system leading to the basic two-component/two-phase model

underlying our theoretical approach. In principle, the charged impurity distribution characterized by n_i and z_0 and the 2D carrier system (characterized by n , m , g_v , g_s , κ) should suffice to define the disorder potential fluctuation distribution $P(V)$ through the self-consistent (and simultaneous) solutions of the Poisson equation for the charged impurity induced electrostatic potential $V(\mathbf{r})$ and the many-electron Schrödinger equation for the response of the interacting 2D electron gas to the external disorder potential. Conceptually, one needs to use a general density functional approach where the disorder-induced potential $V(\mathbf{r})$ serves as the inhomogeneous potential and the ground-state inhomogeneous carrier density $n(\mathbf{r})$ must minimize the appropriate (free) energy functional $E[n(\mathbf{r})]$. Given that the charged impurities are modeled by averaged quantities (e.g., n_i , z_0) since their precise locations are unknown, one must average over many different configurations of the impurity distribution after carrying out the full density functional self-consistent calculations for each configuration. Obviously, such a numerical density functional solution of the self-consistent problem [which would not only be nonlinear,^{60,61} but also nonlocal since there is no *a priori* reason to assume a local density approximation (LDA) to apply] is completely out of question because it will be computationally far too prohibitive (particularly since our goal is to obtain theoretical results for the 2D resistivity as a function of carrier density and temperature for different values of n_i , z_0 , m , g_v , κ , etc.). In fact, even the ground-state problem of obtaining just the inhomogeneous $n(\mathbf{r})$ for the disordered 2D

system at $T = 0$ in the presence of random quenched charged impurities in the background has only been attempted only a few times in the literature^{62–68} within restrictive approximation schemes such as Hartree, LDA, and Thomas-Fermi theories. In spite of the simplified approximation schemes used in these theoretical works, it would be impossible to utilize any of these completely numerical nonlinear self-consistent approaches to obtain the potential fluctuation distribution function for our purpose because combining such self-consistent calculations of disorder potential distributions with the requisite two-component/two-phase transport calculations is simply beyond the computational power of currently available machines.

Instead, we use the theoretically appealing approach of using a parametrized distribution function for the potential fluctuation function $P(V)$, which has recently been successful in developing a theoretical description for monolayer and bilayer graphene transport^{39–41,69,70} at low carrier densities around the Dirac point in the presence of electron-hole puddles induced by charged impurity disorder. In this theoretical approach, the disorder-induced spatial potential fluctuation is approximated by a Gaussian function $P(V)$ which is assumed to be the same throughout the sample. Thus, $V(\mathbf{r})$ at different \mathbf{r} is assumed to be uncorrelated and described by a Gaussian function. The great advantage of this theoretical approach is its simplicity and conceptual clarity: the potential distribution $P(V)$ can be completely characterized by a single energy “ s ” which is the variance or the root-mean-square fluctuation in the impurity induced disorder potential.⁷¹ We assume, with no loss of generality, equal numbers of positive and negative random quenched point charges ($\pm e$) to be contributing to the impurity disorder so that there is no average (or net) potential contributed by disorder. Assuming uncorrelated random Poisson distribution in the impurity locations (all impurities located at random 2D positions in a layer z_0 away from the 2DEG), it is straightforward (but somewhat messy) to show that “ s ” is related to the impurity and the 2DEG parameters [see Eq. (2)] if linear Thomas-Fermi screening holds in the problem (which certainly does not hold at low carrier density). We do not, however, assume any relationship between n_i , z_0 on the one hand and s on the other hand, taking n_i , z_0 , and s together to define a minimal three-parameter description for the random disorder, where s (relative to the Fermi energy E_F) defines the strength of the potential fluctuations and n_i , z_0 together give the strength of diffusive resistive scattering of the mobile carriers. For $s = 0$, all the carriers are mobile, and for infinite s all the carriers are trapped.

The probability $P(V)dV$ of finding the local electronic potential energy within a range dV around V is a Gaussian form

$$P(V) = \frac{1}{\sqrt{2\pi}s^2} \exp(-V^2/2s^2), \quad (1)$$

where s denotes the standard deviation of potential fluctuations. Generally, larger s means more impurity disorder (and more inhomogeneity) in the system. The potential fluctuation is directly related to the charged impurity density,^{49,72,73} and

within the linear Thomas-Fermi screening theory

$$s = \frac{e^2 \sqrt{\pi n_i}}{\sqrt{2\kappa z_0 q_{\text{TF}}}}, \quad (2)$$

where n_i is the density of charged impurity density, κ is the arithmetic average of the semiconductor and insulator dielectric constant, $q_{\text{TF}} = \frac{g_s g_v m e^2}{\kappa \hbar^2}$ (g_s and g_v are the spin and valley degeneracies, respectively) is the Thomas-Fermi 2D linear screening wave vector, m is the effective mass of the carriers, and z_0 is the average distance of the charged impurity to the interface. We choose s as a free parameter in the calculations (along with n_i and z_0) below as explained above, the reason being that linear screening in general breaks down at low carrier density.

The density of states $D(E)$ in the presence of potential fluctuations $P(V)$ is given in terms of the ideal 2D density of states D_0 in the absence of disorder:

$$D(E) = \int_{-\infty}^E \frac{g_s g_v m}{2\pi \hbar^2} P(V) dV = \frac{D_0}{2} \operatorname{erfc}\left(-\frac{E}{\sqrt{2}s}\right), \quad (3)$$

where D_0 is the ideal disorder-free 2D density of states given by

$$D_0 = \frac{g_s g_v m}{2\pi \hbar^2}. \quad (4)$$

At $T = 0$, the carrier density in the band tail, i.e., the trapped or immobile carrier density, is given by the following formula:

$$n_t(T = 0) = D_0 \int_{-\infty}^0 \left[\frac{1}{2} \operatorname{erfc}\left(-\frac{E}{\sqrt{2}s}\right) \right] dE = D_0 \frac{s}{\sqrt{2\pi}}. \quad (5)$$

Note that Eq. (5) defines the classically trapped fraction of the carrier density being in the classically forbidden negative energy states.

The total carrier density is given by [with μ being the finite- T chemical potential with $E_F = \mu(T = 0)$]

$$n = \int_{-\infty}^{\infty} D(E) \frac{dE}{\exp[\beta(E - \mu)] + 1}, \quad (6)$$

where μ is determined by the conservation of the carrier density:

$$\begin{aligned} n &= \int_{-\infty}^{\infty} D(E) \frac{dE}{\exp[\beta(E - \mu)] + 1} \\ &= \int_{-\infty}^{E_F} D(E) dE \quad \text{at } T = 0 \\ &= D_0 \left[\frac{E_F}{2} \operatorname{erfc}\left(-\frac{E_F}{\sqrt{2}s}\right) + \frac{s}{\sqrt{2\pi}} \exp\left(-\frac{E_F^2}{2s^2}\right) \right]. \quad (7) \end{aligned}$$

The trapped electron density at finite temperatures $n_t(T)$ is determined by the potential fluctuation strength, but also depends on the temperature and is given by

$$n_t = D_0 \int_{-\infty}^0 \left[\frac{1}{2} \operatorname{erfc}\left(-\frac{E}{\sqrt{2}s}\right) \right] \frac{dE}{\exp[\beta(E - \mu)] + 1}. \quad (8)$$

The effective mobile or free carrier density $n_m = n - n_t$ depends on the temperature, the Fermi energy E_F , and the standard deviation of potential fluctuation s . At zero

temperature $T = 0$, the effective mobile carrier density is given by

$$n_m(T = 0) = D_0 \left[\frac{E_F}{2} \operatorname{erfc} \left(-\frac{E_F}{\sqrt{2}s} \right) + \frac{s}{\sqrt{2\pi}} \right] \times \exp \left(-\frac{E_F^2}{2s^2} - \frac{s}{\sqrt{2\pi}} \right). \quad (9)$$

Thus, $n = n_m + n_t$ is the total carrier density with n_m , n_t being the temperature-dependent fractions of mobile and trapped carrier densities. These are the two phases (n_m and n_t) coexisting in the disordered 2DEG, and we must now develop a two-component transport theory to obtain the 2D conductivity. The activated conductivity of the trapped immobile carriers is given by

$$\sigma_a(V) = \sigma \exp[\beta(E_F - V)], \quad (10)$$

where $\sigma = \frac{ne^2\langle\tau\rangle}{m}$ with τ as the transport scattering time. We note that the variable range hopping should also be included in the conductivity of the trapped carriers, but we neglect it for two reasons. First, the variable range hopping is a phonon-assisted quantum tunneling phenomenon for the localized carrier, and our neglecting this is consistent with our neglect of all quantum tunneling processes in the theory. Second, including a hopping conduction would simply add more unknown parameters to the theory without adding any physical clarity.

For the mobile carriers, we simply have the diffusive conductivity from the Boltzmann theory

$$\sigma_d = \frac{(n - n_t)e^2\langle\tau\rangle}{m} = \frac{n_m e^2\langle\tau\rangle}{m}, \quad (11)$$

where the finite-temperature averaged $\langle\tau\rangle$ is given by^{35,37,74}

$$\langle\tau\rangle = \frac{\int d\epsilon D(\epsilon)\epsilon\tau(\epsilon)(-\partial f/\partial\epsilon)}{\int d\epsilon D(\epsilon)f(\epsilon)}. \quad (12)$$

Here, we use $D(\epsilon) = D_0$ for the mobile carriers. $f(\epsilon)$ is the Fermi-Dirac distribution function

$$f(\epsilon) = \frac{1}{1 + e^{\beta(\epsilon - \mu_0)}} \quad (13)$$

with μ_0 as the finite-temperature chemical potential for the homogeneous system, which is determined by

$$\mu_0 = E_F + k_B T \ln \left(1 - e^{-\frac{E_F}{k_B T}} \right). \quad (14)$$

The energy-dependent scattering time is given by

$$\frac{\hbar}{\tau(\epsilon_{\mathbf{k}})} = 2\pi n_i \int \frac{d^2k'}{(2\pi)^2} \left| \frac{V(q)}{\epsilon(q)} \right|^2 [1 - \cos\theta] \delta(\epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}}), \quad (15)$$

where $V(q) = \frac{2\pi e^2}{\kappa q} e^{-qz_0}$ is the scattering potential strength and $\epsilon(q) = 1 - \frac{2\pi e^2}{\kappa q} \Pi(q, T)$ is the finite-temperature 2D screening function.^{1,75-77}

We define the fractional occupancy as given by

$$p = \int_{-\infty}^{E_F} P(V) dV. \quad (16)$$

In the two-component effective medium theory, we have the conductivity σ_1 and σ_2 for the two phases “1” and “2”. In our

case, phase 1 consists of mobile carrier

$$\sigma_1 = \sigma_d. \quad (17)$$

The conductivity ($\sigma_2 = \sigma_a$) of phase 2 corresponds to the activated conduction of the trapped carriers

$$\begin{aligned} \sigma_2 &= \frac{1}{q} \int_{E_F}^{\infty} \sigma_a(V) P(V) dV \\ &= \frac{\sigma}{2q} e^{\frac{\beta^2 s^2}{2} + \beta E_F} \operatorname{erfc} \left(\frac{E_F}{\sqrt{2}s} + \frac{\beta s}{\sqrt{2}} \right) = \sigma_a, \end{aligned} \quad (18)$$

where $q = 1 - p$.

The total conductivity is given by (according to the effective medium theory^{38,78})

$$\sigma = \frac{2p - 1}{2} \left[(\sigma_1 - \sigma_2) + \sqrt{(\sigma_1 - \sigma_2)^2 + \frac{4}{(2p - 1)^2} \sigma_1 \sigma_2} \right] \quad (19)$$

with $\sigma_1 = \sigma_d$ and $\sigma_2 = \sigma_a$.

In the next section, we give numerical results for $\rho = 1/\sigma$ in terms of system parameters.

Before presenting our numerical results for the carrier density and temperature-dependent 2D resistivity in terms of s , n_i , z_0 , we mention that in spite of the conceptual simplicity of the two-component effective medium theory description of the two-phase transport of the 2DEG in the presence of potential fluctuations, the results arise from a complex interplay of many distinct physical mechanisms which are not easy to grasp because of the large number of parameters defining the 2D MIT phenomena. First, there is a strong density and temperature dependence of σ_d , i.e., the diffusive conductivity of mobile carriers, itself arising from the variation in $q_{TF}/2k_F$ (“screening”) and $k_B T/E_F$ (“degeneracy”) as discussed already in details in the literature.^{35-37,74} This screening and degeneracy dependence of 2D “metallic” transport is included in our theory and is actually enhanced by the fact that $p < 1$ by virtue of $s \neq 0$. Thus, the potential fluctuations enhance the temperature dependence of σ_d (or σ_1) indirectly. More importantly, however, the activated conduction inherent in $\sigma_2 (= \sigma_a)$ strongly enhances the temperature dependence of the conductivity, particularly at lower carrier densities (i.e., larger values of s/E_F) where the trapped carrier fraction is large. Since the two contributions to the conductivity (i.e., σ_d and σ_a or equivalently σ_1 and σ_2) have opposite temperature dependencies at low temperatures, it is possible for the system to exhibit a very weak temperature dependence at some intermediate carrier density mimicking a quantum critical density between a metal and an insulator, whereas in reality, within our theory, it is simply a two-phase crossover behavior where the 2D metallic phase at high density (and large E_F) consisting mainly of mobile/diffusive carriers is crossing over to a low-density (and small E_F) 2D system consisting primarily of trapped/insulating 2D carriers. Our theory by construction is a crossover two-component theory.

III. TRANSPORT IN THE TWO-COMPONENT MODEL

In providing the numerical results for the 2D transport properties in the two-component effective medium theory

approximation developed in Sec. II, we mention that the quantitative results would depend strongly on the 2D material (since the relevant system parameters such as q_{TF} and E_F depend on materials parameters such as m , g_v , and κ) as well as on the nature of the 2D impurity disorder. In particular, whether the dominant disorder arises from near impurities (relatively small values of z_0) or far impurities (large z_0) has a strong effect on the results. Another possibility we will consider is a background 3D impurity distribution in the 2D quantum well itself for which an integration would have to be carried out over a range of z_0 values, and the theory of Sec. II has to be generalized for a finite 2D layer thickness (“ a ”) which is straightforward to do assuming a square-well quantum 2D confinement in the z direction with a layer thickness of a .

We provide our numerical results only for 2D GaAs-AlGaAs hole system (corresponding to Ref. 8) since typically this system exhibits the strongest 2D MIT behaviors although we have produced transport results for many different 2D systems used in various experimental studies obtaining qualitatively similar results. Following, we discuss and present our numerical transport results corresponding to Ref. 8 by considering individual models for impurity disorder arising from 2D interface impurities (Sec. III A), remote 2D impurities (Sec. III B), background 3D impurities (Sec. III C), and remote 3D impurities (Sec. III D). We mention that, in reality, all four disorder mechanisms are likely to be present with varying quantitative magnitudes.

A. Interface 2D impurity model

We show in Fig. 4 our calculated $\rho(T)$ for various values of total carrier density n for $s = 0.05$ meV for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å assuming 2D interface impurity density of 1.5×10^8 cm $^{-2}$. This is a rather clean system corresponding to $n_i(T = 0) = 0.25 \times 10^{10}$ cm $^{-2}$. The system corresponds to the 2D hole system

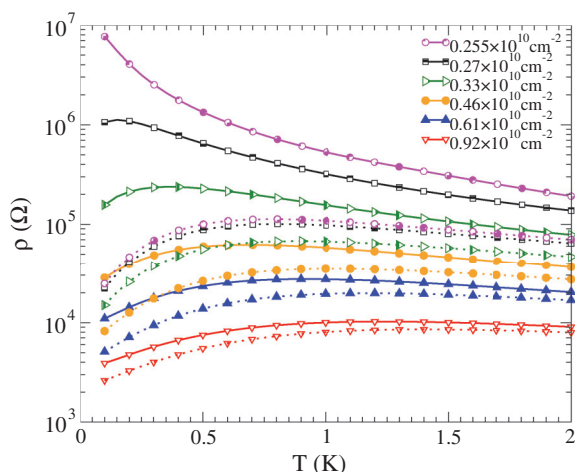


FIG. 4. (Color online) Calculated $\rho(T)$ for various values of total carrier density n for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å assuming 2D interface impurity density of 1.5×10^8 cm $^{-2}$. The dotted and solid lines are for $s = 0$ and 0.05 meV [corresponding to $n_i(T = 0) = 0.25 \times 10^{10}$ cm $^{-2}$], respectively. The theoretical parameters in Figs. 4–20 (i.e., all our presented results) are chosen to correspond to the experimental data in Ref. 8.

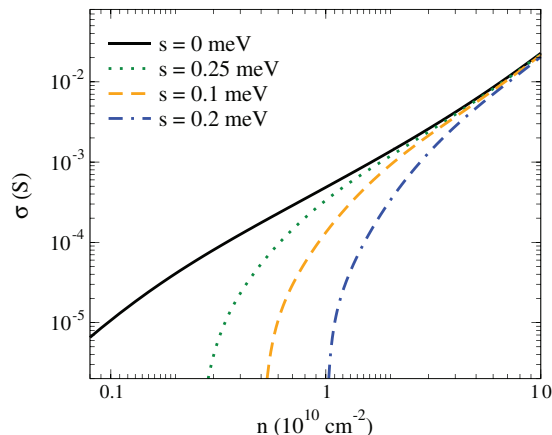


FIG. 5. (Color online) Calculated $\sigma(n)$ at $T = 50$ mK for different values of potential fluctuations s for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å assuming 2D interface impurity density of 1.5×10^8 cm $^{-2}$.

studied in Ref. 8. In Fig. 5, we show the corresponding $\sigma(T)$ plots (with $\sigma \equiv \rho^{-1}$ by definition) at $T = 50$ mK for different values of $s = 0, 0.05, 0.1,$ and 0.2 meV. Results shown in Figs. 4 and 5 indicate that the two-component effective medium approximation, assuming conduction by diffusive mobile carriers and activated trapped carriers, qualitatively reproduce the observed features of the 2D MIT phenomena. In particular, $\rho(T)$ mimics a metal-insulator transition around n_i , and the critical density for the 2D MIT crossover depends strongly on the magnitude of the potential fluctuations with the crossover “critical” density increasing with increasing “ s .”

In Fig. 6, we plot the numerical value of the effective conductivity exponent “ δ ” defined as either $\delta = \frac{\partial \ln \sigma}{\partial \ln n}$ or $\delta' = \frac{\partial \ln \sigma}{\partial \ln(n - n_i)}$ as a function of carrier density for $s = 0.05$ meV, showing that δ and δ' are also consistent with experimental findings.⁸ In particular, δ itself becomes very large as density decreases but $\delta' \sim 1.5$ consistent with a percolation picture.^{8,9,21,23}

We emphasize that there is no point in demanding strict quantitative agreement between theory and experiment since n_i , z_0 , and s are all unknown experimentally (and a minimal model involving only three parameters is unlikely to capture all the complexity of the actual disorder in the realistic samples).

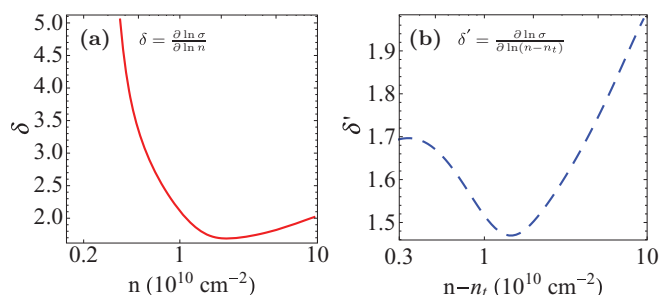


FIG. 6. (Color online) Numerically calculated effective conductivity exponent “ δ ” at $T = 50$ mK for $s = 0.05$ meV for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å assuming 2D interface impurity density of 1.5×10^8 cm $^{-2}$.

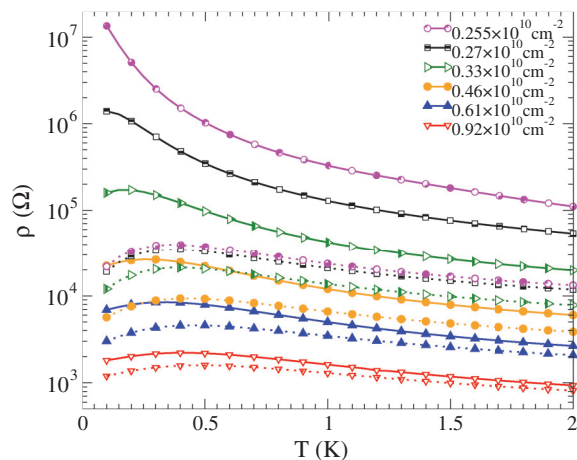


FIG. 7. (Color online) Calculated $\rho(T)$ for various values of total carrier density n for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å assuming remote 2D impurity density of 3.5×10^8 cm $^{-2}$ and $z_0 = 300$ Å. The dotted and solid lines are for $s = 0$ and 0.05 meV [corresponding to $n_t(T = 0) = 0.25 \times 10^{10}$ cm $^{-2}$], respectively.

B. Remote 2D impurity model

In Figs. 7–9, we show our calculated results for remote 2D impurity scattering, again for 2D holes in GaAs/AlGaAs quantum wells. Results are very similar for the near impurity case (Figs. 4–6) except that the temperature dependence in the metallic phase for $n \gg n_t$ is weaker than the interface impurity case of the last section because $2k_F$ scattering is suppressed in the remote impurity case.

C. Background 3D impurity model

Here, the impurities are distributed randomly throughout the quantum well with some 3D average impurity density, often in very high mobility ultrapure 2D systems, unintentional 3D background charged impurities at some low concentration dominate transport properties. In Figs. 10–12, we show our

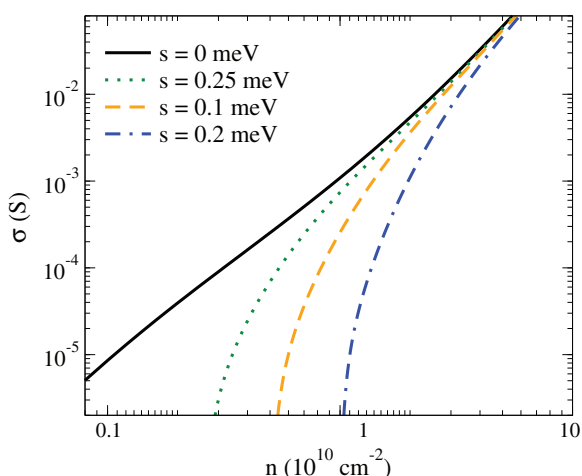


FIG. 8. (Color online) Calculated $\sigma(n)$ at $T = 50$ mK for different values of potential fluctuations s for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å assuming remote 2D impurity density of 3.5×10^8 cm $^{-2}$ and $z_0 = 300$ Å.

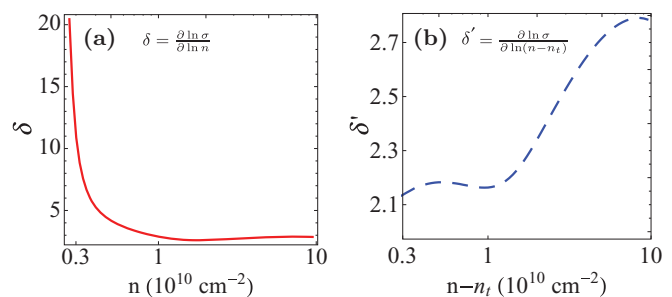


FIG. 9. (Color online) Numerically calculated effective conductivity exponent “ δ ” at $T = 50$ mK for $s = 0.05$ meV for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å assuming remote 2D impurity density of 3.5×10^8 cm $^{-2}$ and $z_0 = 300$ Å.

numerical results for transport limited by 3D background impurities. Again, the results are qualitatively similar to those in Figs. 4–9 except that the metallic temperature dependence of the high-density resistivity is somewhat stronger in this case since the impurities reside in the 2D layer itself.

D. Remote 3D impurity model

Here, the charged impurities are distributed randomly in three dimensions over a remote spacer layer located away from the 2D quantum well. These results are shown in Figs. 13–15, and are similar to the results in Figs. 4–12 except for quantitative differences.

We conclude this section on our numerical results by emphasizing that we can easily produce results for transport limited by different combinations of impurity scattering (i.e., combinations of 2D near and far scatters and/or of 3D background and remote scatters), but the qualitative behavior would remain the same. Therefore, we do not see any compelling reason to present additional numerical results including all four disorder models together which will, in principle, involve 12 unknown disorder parameters. For the same reason, we

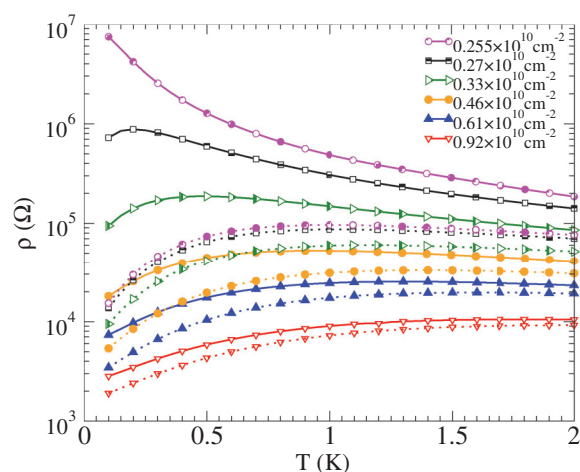


FIG. 10. (Color online) Calculated $\rho(T)$ for various values of total carrier density n for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å. The impurities are distributed randomly throughout the quantum well with 3D average density 7.5×10^{13} cm $^{-3}$. The dotted and solid lines are for $s = 0$ and 0.05 meV [corresponding to $n_t(T = 0) = 0.25 \times 10^{10}$ cm $^{-2}$], respectively.

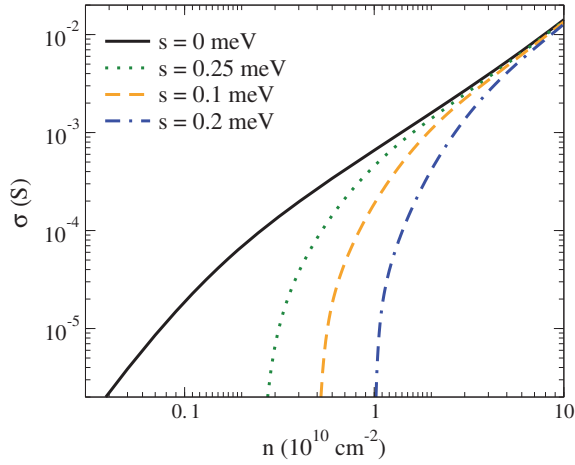


FIG. 11. (Color online) Calculated $\sigma(n)$ at $T = 50$ mK for different values of potential fluctuations s for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å. The impurities are distributed randomly throughout the quantum well with 3D average density $7.5 \times 10^{13} \text{ cm}^{-3}$.

refrain from providing numerical transport results for other 2D systems since the qualitative physics would remain the same and the results would look very similar with obvious quantitative differences.

Instead, in Figs. 16–18, we show the effect of having different values of potential fluctuations (by varying “ s ”) on our calculated 2D MIT phenomenology. These figures show that increasing s , i.e., having stronger disorder, suppresses the metallic temperature dependence, explaining why the strong metallicity (i.e., strong metallic temperature dependence of the 2D resistivity) necessitates having cleaner samples which presumably have smaller values of s .

IV. DISCUSSION

The main point established through our numerical results (Figs. 4–18) presented in this paper is that a theory based on a two-component transport mechanism involving a two-phase coexistence of both mobile (diffusive) and trapped (activated) carriers in the inhomogeneous “lakes-and-hills” landscape of disorder-induced potential fluctuations provides a reasonable qualitative description of the phenomenology associated with

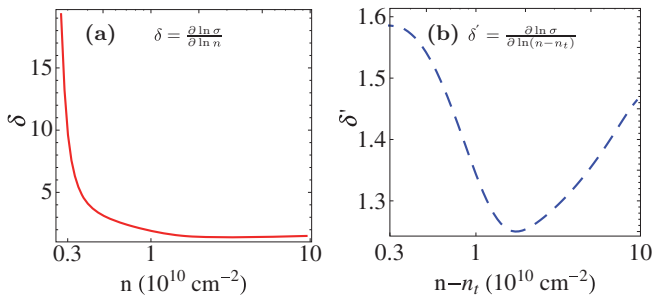


FIG. 12. (Color online) Numerically calculated effective conductivity exponent “ δ ” at $T = 50$ mK for $s = 0.05$ meV for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å. The impurities are distributed randomly throughout the quantum well with 3D average density $7.5 \times 10^{13} \text{ cm}^{-3}$.

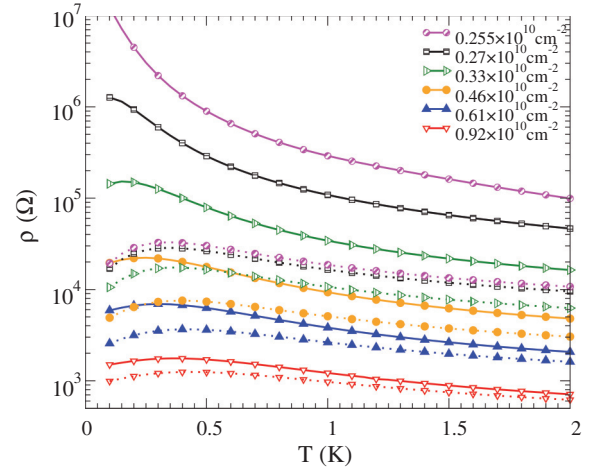


FIG. 13. (Color online) Calculated $\rho(T)$ for various values of total carrier density n for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å. The 3D remote impurities are distributed randomly from $z_1 = 300$ Å to $z_2 = 350$ Å with 3D average density $7 \times 10^{14} \text{ cm}^{-3}$. The dotted and solid lines are for $s = 0$ and 0.05 meV [corresponding to $n_t(T = 0) = 0.25 \times 10^{10} \text{ cm}^{-2}$], respectively.

the experimental 2D MIT observations. In particular, the metal-to-insulator low-density crossover and the associated temperature dependence of the resistivity in the experimental observations are similar to our theoretical results.

Following, we provide some critical remarks in the context of a series of questions which might arise with respect to any theory purporting to provide a complete qualitative answer to the substantial (and long-standing) puzzle which goes by the name of 2D MIT. The answers to these questions are all qualitative, but our theory, as shown in Sec. III, is capable of providing detailed quantitative transport results if the impurity disorder is known in quantitative depth (which is, of course, an impossibility in real laboratory samples).

(i) Is there a true phase transition in the theory? The answer is “no”: our theory is by construction a pure crossover

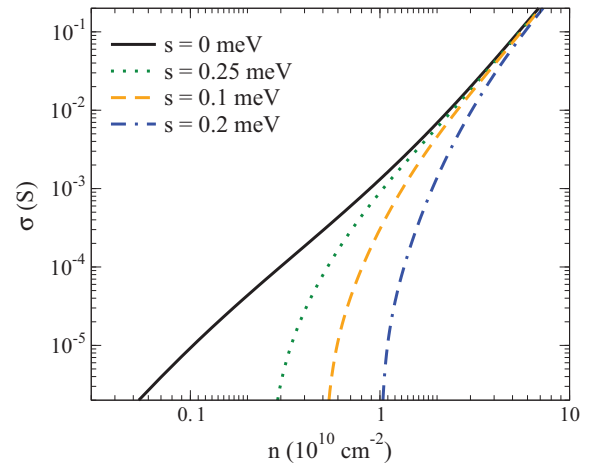


FIG. 14. (Color online) Calculated $\sigma(n)$ at $T = 50$ mK for different values of potential fluctuations s for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å. The 3D remote impurities are distributed randomly from $z_1 = 300$ Å to $z_2 = 350$ Å with 3D average density $7 \times 10^{14} \text{ cm}^{-3}$.

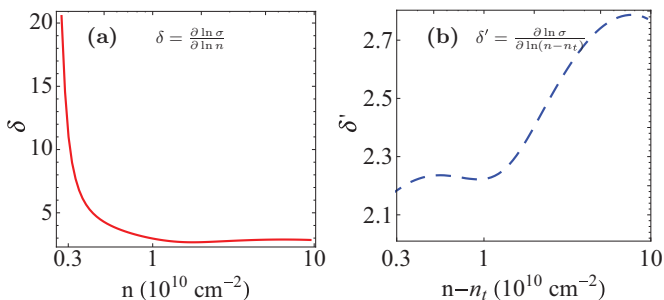


FIG. 15. (Color online) Numerically calculated effective conductivity exponent “ δ ” at $T = 50$ mK for $s = 0.05$ meV for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å. The 3D remote impurities are distributed randomly from $z_1 = 300$ Å to $z_2 = 350$ Å with 3D average density 7×10^{14} cm $^{-3}$.

theory where in the “higher-density” metallic phase diffusive conduction by the mobile carriers dominates, whereas in the “low-density” insulating phase-activated conduction by the trapped carriers dominates. But, the crossover can be very sharp, and depending on the situation, could easily mimic a sharp metal-to-insulator phase transition. In Fig. 19, we show our numerical results for $\sigma(n)$ for different values of T for a fixed s , clearly showing that the $\sigma(n)$ curves for different temperatures do not go through a single value of critical transition density, but are bunched in a small crossover density regime.

(ii) How does the crossover (or transition) density depend on the disorder strength? The transition density always increases with increasing the potential fluctuation strength s . Thus, more disordered systems will behave insulating at higher carrier density. This is obvious from the results shown in Sec. III. A corollary of this finding, which is also apparent in the results of Sec. III, is that for larger disorder with higher- s values, the temperature dependence of the resistivity in the metallic phase will be strongly suppressed since the

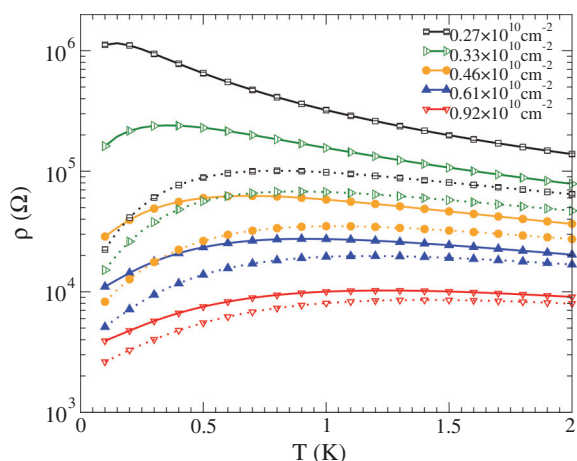


FIG. 16. (Color online) Calculated $\rho(T)$ for various values of total carrier density n for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å. The 2D interface impurity density is 1.5×10^8 cm $^{-2}$. The 3D remote impurities are distributed randomly from $z_1 = 1600$ Å to $z_2 = 1650$ Å with 3D average density 7×10^{14} cm $^{-3}$. The dotted and solid lines are for $s = 0$ and 0.05 meV [corresponding to $n_i(T = 0) = 0.25 \times 10^{10}$ cm $^{-2}$], respectively.

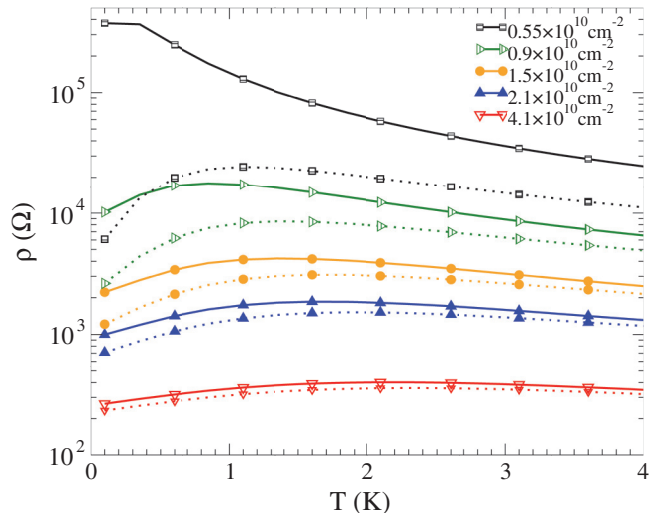


FIG. 17. (Color online) Calculated $\rho(T)$ for various values of total carrier density n for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å. The 2D interface impurity density is 1.5×10^8 cm $^{-2}$. The 3D remote impurities are distributed randomly from $z_1 = 1600$ Å to $z_2 = 1650$ Å with 3D average density 7×10^{14} cm $^{-3}$. The dotted and solid lines are for $s = 0$ and 0.1 meV [corresponding to $n_i(T = 0) = 0.5 \times 10^{10}$ cm $^{-2}$], respectively.

temperature dependence of the diffusive $\rho(T)$ for the mobile carriers is determined by the parameters T/T_F and q_{TF}/k_F , both of which decrease with increasing density. Thus, very low-disorder 2D systems with small values of s are necessary for the experimental observation of the 2D MIT phenomena.

(iii) Is there a temperature dependence of the crossover density? The answers to questions (i) and (ii) and a careful reading of our presented results in Sec. III show that indeed there is a weak temperature dependence of the crossover density, thus, the crossover density depends both on the disorder strength and the temperature scale of the experiment.

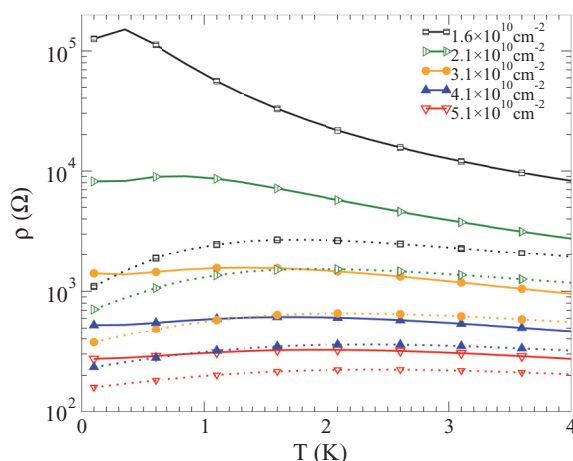


FIG. 18. (Color online) Calculated $\rho(T)$ for various values of total carrier density n for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å. The 2D interface impurity density is 1.5×10^8 cm $^{-2}$. The 3D remote impurities are distributed randomly from $z_1 = 1600$ Å to $z_2 = 1650$ Å with 3D average density 7×10^{14} cm $^{-3}$. The dotted and solid lines are for $s = 0$ and 0.3 meV [corresponding to $n_i(T = 0) = 1.5 \times 10^{10}$ cm $^{-2}$], respectively.

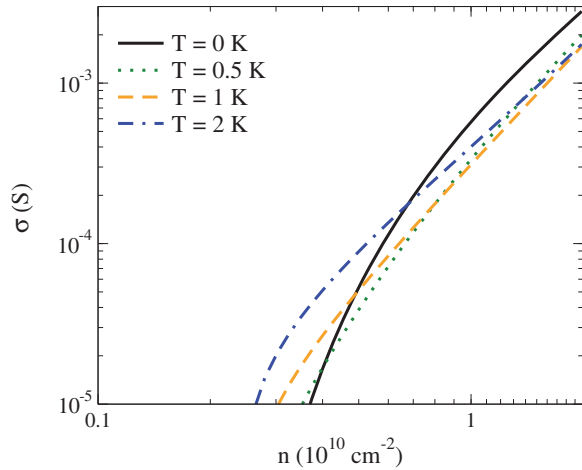


FIG. 19. (Color online) Calculated $\sigma(n)$ for different values of T with a fixed $s = 0.05$ meV for 2D holes in GaAs/AlGaAs quantum wells of width $a = 200$ Å. The 2D interface impurity density is 1.5×10^8 cm $^{-2}$. The 3D remote impurities are distributed randomly from $z_1 = 1600$ Å to $z_2 = 1650$ Å with 3D average density 7×10^{14} cm $^{-3}$.

In particular, the transition density increases with increasing temperature (and it also increases with increasing disorder). The reason for this is a bit trivial since it arises from the fact that at higher temperatures (and in the absence of phonon scattering), $\rho(T)$ shows monotonically slowly decreasing behavior with increasing temperature up to rather high density, thus making it appear that the metallic phase is pushed up to higher density as temperature increases since activated conduction becomes increasingly important at higher T . This trend of an apparent increase of the crossover density with increasing temperature (or, equivalently, an apparent decrease of the “critical” density with decreasing temperature) has been experimentally observed.⁷⁹ A recent experimental work²⁸ also shows the dominance of activated conduction at higher temperatures leading to a $\rho(T)$ decreasing at higher temperatures consistent with our two-phase coexistence model.

(iv) What about the in-plane magnetic field dependence of the 2D resistivity? We now comment on an important class of experiments where an in-plane magnetic field is applied parallel to the 2D system^{80–82} with the consequent finding that the effective metallic (insulating) phase is strongly suppressed (enhanced), i.e., the crossover density increases with increasing in-plane magnetic field. This phenomenon can be explained in our theory simply as a magnetic-field-induced enhancement of the trapped immobile carrier density in the system through an increase of the potential fluctuation parameter s due to the increasing spin polarization of the system (induced by the applied in-plane field). In particular, for a large enough magnetic field, the 2DEG becomes completely spin polarized, thus reducing the screening wave vector q_{TF} by a factor of 2, which then enhances s by a factor of 2, in turn leading to a factor of 2 increase in the trapped carrier density [see Eq. (5)] even for exactly the same disorder configuration. This factor of 2 increase in the trapped carrier density then leads to a large increase in the crossover density for the metal to the insulator transition. Such an enhancement (roughly by a factor of 2) in the “critical” density for 2D MIT in the presence

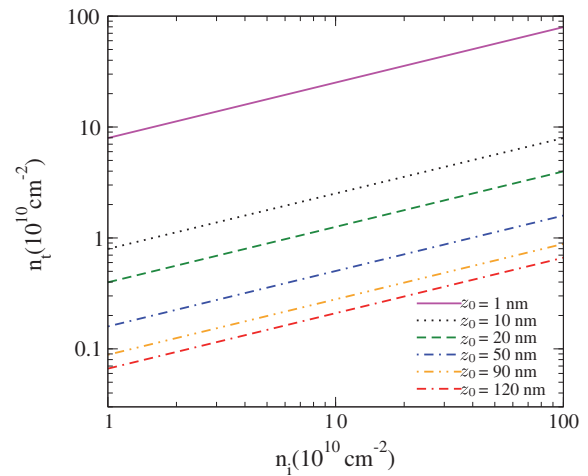


FIG. 20. (Color online) Calculated $n_t = \frac{\sqrt{n_i}}{4\pi z_0}$ as a function of n_i .

of a strong in-plane applied field has been reported in many experiments.^{83–87}

(v) What is the relationship between the random charged impurities (defined by n_i and z_0) and the potential fluctuations [defined by the variance s of the Gaussian probability distribution function $P(V)$ for the disorder-induced potential fluctuations]? In principle, s is related to n_i , z_0 since the random charged impurities give rise to the potential fluctuations. In our theory, we have used a simple minimal model where n_i , z_0 , and s are three independent parameters defining the impurity disorder, but in principle they are related. This relationship is, however, in general unknown since the precise impurity distribution (modeled in our theory by only two parameters n_i and z_0) is unknown (and would, in principle, require an infinite number of parameters to define). In addition, even for a given n_i and z_0 , the real calculation of s is numerically intractable and impractical because one must carry out a nonlinear and self-consistent density functional theory where the carrier density distribution $n(\mathbf{r})$ and the potential fluctuations $V(\mathbf{r})$ determine each other. Within a linear screening theory (which surely does not apply at low carrier density) we find $s = e^2 \sqrt{\pi n_i} / (\sqrt{2} \kappa z_0 q_{TF})$ with $q_{TF} = g_s g_v m e^2 / (\kappa \hbar^2)$, and $n_t \equiv \sqrt{n_i} / (4\pi z_0) = (\frac{g_s g_v m}{2\pi \hbar^2}) (\frac{s}{\sqrt{2\pi}})$. In Fig. 20, we show our calculated n_t as a function of n_i for 2D GaAs holes for different values of z_0 . These relationships also tell us how the 2D material (i.e., Si or GaAs, electrons or holes, etc.) enters the theory through the system parameters m , κ , g_s , g_v . For example, 2D electrons in GaAs quantum wells have $m = 0.07m_e$ in contrast to 2D holes with $m = 0.3m_e$ with all the other parameters being the same. This implies that for the same disorder parameters the effective s in 2D holes is four times smaller than in 2D electrons, leading to much more prominent 2D MIT behavior in 2D GaAs holes than in 2D GaAs electrons in the same carrier density and temperature regime as observed experimentally.^{7,8,23} By contrast, 2D Si electrons have $m = 0.19m_e$ but $g_v = 2$ with a “ $\kappa(\approx 8)$ ” which is somewhat smaller than in GaAs ($\kappa \approx 12$). Thus, for the same disorder parameters, 2D GaAs holes will manifest somewhat stronger 2D MIT phenomena than 2D Si electrons, as observed experimentally.⁹

V. CONCLUSION

Our physically motivated two-component transport theory based on the coexistence of two phases of carriers in the 2D disordered system, trapped/activated carriers, and mobile/diffusive carriers provides a qualitative description of the experimentally observed features of the 2D MIT phenomenology. The basic model of the two-phase coexistence (“trapped” and “mobile”) is classical, somewhat akin to coexisting solid (“trapped”) and liquid (“mobile”) phases in the lakes-and-hills fluctuating potential landscape, although we calculate the 2D conductivity using a semiclassical Boltzmann transport theory where carrier scattering is treated quantum mechanically. In our theory, the basic metal-to-insulator transition is a crossover phenomenon where with lowering carrier density, the relative activated contribution of the trapped carriers to the net conductivity in the two-component model increases with respect to the diffusive contribution of the mobile carriers, reflecting eventually in an apparent change in sign of $d\rho/dT$ at low temperatures and densities. The overall temperature and density (as well as in-plane magnetic field) dependence of the transport properties arises from a complex interplay among potential fluctuations, screening, and activated transport.

The main effects left out of the theory are quantum tunneling, quantum interference, and electron-electron interaction [beyond random phase approximation (RPA) screening]. We do not see an easy way of including these effects in our model. Tunneling can be included^{88,89} within a Landauer-Buttiker-type approach, but then many realistic effects such as screening and diffusive impurity scattering must be left out. In principle, at $T = 0$ all states in 2D systems become localized, but experimentally this limit seems difficult to reach. The advantage of our theory is that one can do concrete microscopic calculations for the density-, temperature-, and disorder-dependent 2D conductivity for different 2D materials. The disadvantage is that it is not easy to see how to include quantum effects such as localization/tunneling and interaction, which must eventually play a role at low enough temperatures.

It may be useful to emphasize two aspects of our theory. First, the results presented in this work are not only in excellent qualitative agreement of the existing experimental results, they are, in fact, in good quantitative agreement with the experimental 2D MIT data in the literature. In particular, the theoretically calculated temperature dependence of the 2D metallic resistivity at higher carrier density agrees well with the existing experimental data by virtue of our inclusion of the full temperature- and density-dependent screening of the underlying Coulomb disorder as already discussed extensively in the earlier works^{35–37,74,80–82} where it was established that the strong linear metallic temperature dependence of the 2D metallic resistivity arises from temperature dependence of the 2D screening function around $2k_F$. Similarly, the strong exponential insulating temperature dependence for lower carrier density is guaranteed in our effective medium approximation by the activated conduction of the trapped carriers. Experimentally, of course, the quantitative details vary enormously from system to system since the relevant disorder parameters (e.g., n_i , z_0 , s , etc.) vary strongly among different 2D systems. For example, in Si MOSFETs, z_0 is

typically around 1 nm whereas n_i is typically 10^{10} cm⁻², which leads to $n_t \sim 10^{11}$ cm⁻², which is in good quantitative agreement^{2,11} with the so-called 2D MIT critical density in Si-based 2D systems. In modulation-doped 2D GaAs systems, where the background charged impurity density is negligibly small, the main scattering mechanism is provided by the remote dopants, leading to $z_0 \sim 100$ nm or larger with $n_i \sim 10^{10}$ cm⁻². This leads to $n_t \sim 10^9$ cm⁻², which is in agreement with the observed critical density^{7,8} in high-mobility 2D GaAs systems. In Fig. 20, we explicitly show the dependence of n_t on n_i and z_0 , which can be directly compared with experiments. One serious problem in this context is, of course, that the precise values of n_i and z_0 are unknown in experimental samples, and therefore any direct quantitative comparison is not very meaningful (which is the reason for emphasizing qualitative rather than quantitative agreement with the data throughout this paper) since the theory depends sensitively on the precise values of these disorder parameters. We emphasize, however, that to the extent the disorder parameters in the experimental samples are known, our results provide a reasonable semiquantitative agreement with the experimental data. Second, our theory should at best be thought of as a zeroth-order semiclassical two-component description of the reality because of our neglect of many important physical effects invariably present in nature (e.g., Anderson localization, electron-electron interaction, quantum tunneling). A complete theory including both the effects of long-range Coulomb disorder (the main ingredient included in our theory) and localization/interaction/tunneling is at this stage not only beyond the scope of our work, it is in fact essentially impossible.

Our work is motivated entirely by the belief that the zeroth-order physics for 2D MIT is captured well by considering the long-range Coulomb disorder (leading to the model of trapped and mobile carriers) effects and leaving out other effects. All we can say at this stage, other than emphasizing the qualitative agreement between our results and experiments, is that earlier works in many different 2D systems^{8–10,18,19,21,23,24,27,29,54,59,62–64,88,89} have already claimed the applicability of a two-component classical percolation model to the 2D MIT phenomenon, and our work simply carries out a quantitatively complete transport calculation using the two-component model so that the actual theoretical transport results can be explicitly obtained using the effective medium theory. More work would obviously be needed to complete the story of 2D MIT phenomenon in figuring out the precise roles played by quantum tunneling Anderson localization, and electron-electron interaction in the transport physics of 2D disordered systems.

A relevant question in this context is what one can learn about the experimental 2D MIT phenomena from our two-component theoretical analysis and our presented transport results in this work. After all, our theory is explicitly constructed as a crossover theory with both mobile and trapped carrier contributing to the conduction (with their relative contributions changing in a complicated manner depending on disorder parameters, carrier density, and temperature), and as such, our work obviously can not conclusively theoretically establish 2D MIT to be a crossover phenomenon since we do not compare

and contrast our crossover theory to any quantum critical theory treating 2D MIT to be a true quantum localization transition. (We comment as an aside that no detailed theoretical transport results are available in the literature based on a quantum critical model of 2D MIT.) The key message of our analyses is that the two-component semiclassical percolation theory requires very reasonable system parameters to theoretically reproduce the experimentally observed 2D resistivity $\rho(n, T)$ over a very wide range of carrier density and temperature without any fine tuning of the generally unknown disorder parameters (n_i , z_0 , and s in our theory for the specific impurity disorder mechanism). In particular, the same set of unique values for (n_i , z_0 , s) reproduces the approximate “critical” carrier density regime for the metal-to-insulator crossover transition in $\rho(T)$ as well as the approximate (and often strong) temperature dependence for $\rho(T)$ in the nominally metallic phase. In addition, the same set of n_i , z_0 , and s (which reproduces the correct crossover density scale) also reproduces the absolute magnitude of $\rho(T, n)$ both in the metallic and in the insulating regimes. This is obviously a rather nontrivial theoretical accomplishment, indicating that the temperature dependence of $\rho(T, n)$ on the metallic side is closely connected with both the actual density for the metal-to-insulator transition as well as the behavior of $\rho(T)$ on the insulating side. This can be easily inferred by comparing our Fig. 4 with Fig. 1 in the corresponding experimental work.⁸ Both our Fig. 4 and Fig. 1 in Ref. 8 indicate a crossover carrier density around $3\text{--}5 \times 10^9 \text{ cm}^{-2}$, and the measured $\rho(n, T)$ in the metallic regime in Ref. 8 is precisely in the same range as our calculated values in Fig. 4. The change in $\rho(T)$ on the metallic side agrees well between our theory and the data in Ref. 8. In addition, our calculated effective conductivity exponent $\delta' \approx 1.6$ in Fig. 6 agrees well with the experimentally measured exponent given in Fig. 5 of Ref. 8. What is important here is not the precise quantitative agreement, but the fact that very reasonable agreement between several different independent experimental quantities and our theoretical calculations are achieved using a single set of reasonable parameter values for n_i , s , and z_0 . This seems to indicate that the qualitative agreement between the two-component crossover theory and experimental data is unlikely to be a coincidence (since we are not fine tuning unknown parameters), and the experimental 2D MIT phenomenon is likely to be strongly affected by the percolation physics presented in this work. This is the key message of our work.

A direct corollary of the above discussion about the agreement between our percolation theory and 2D MIT experiments (without the fine tuning of disorder parameters) is that any experiment finding a sharp separatrix (i.e., a sharp value of a temperature-independent critical density n_c corresponding to a sharp temperature-independent critical resistance ρ_c) distinguishing the low-density ($n < n_c$) insulating phase from the high-density ($n > n_c$) metallic phase with a characteristic critical resistance $\rho_c \equiv \rho(n = n_c)$ is in fundamental qualitative disagreement with our theoretical approach. If a sharp temperature-independent separatrix exists distinguishing the metallic phase from the insulating phase, then our theory is inapplicable and the physics is likely to be a quantum critical phenomenon (and not a crossover

as we have assumed in our model). In the early days of 2D MIT physics, such a sharp separatrix and an associated critical density/resistance for 2D MIT phenomenon were often claimed to be true experimentally.¹¹ But, it is now well established that most 2D MIT experiments do not have any sharp separatrix (or an associated critical resistance), and this was already emphasized by Hanein *et al.* as far back as 1998.¹⁵ In fact, the experimental results presented in Fig. 1 of this paper clearly show that the typical characteristic resistance associated with 2D MIT crossover varies from $\sim 8 \text{ k}\Omega$ in Fig. 1(a) through $\sim 40 \text{ k}\Omega$ in Fig. 1(c) to $\sim 50 \text{ k}\Omega$ in Fig. 1(b), and this “critical resistance” is strongly temperature dependent in each sample. In the literature, the experimentally observed critical resistance varies by more than an order of magnitude^{2–6} in different systems and samples, casting doubts on the early claim of a universal ρ_c characterizing 2D MIT. We must emphasize, however, that our theory is simply inapplicable to any experiment, such as those reported in Ref. 11, where the 2D MIT is characterized by a sharp separatrix and an associated precise critical resistance. Fortunately (for our theory), many experimental observations are consistent with 2D MIT being a crossover phenomenon with no precise critical density (and critical resistance) including the data shown in Fig. 1 of this paper. It is, of course, possible (although unlikely in our opinion) that some specific subsets of 2D MIT experiments are indeed observing a quantum critical transition and others (e.g., the three distinct experimental results depicted in Fig. 1 of our paper) are observing crossover behaviors. This could, in principle, arise from the underlying effective disorder mechanism controlling 2D MIT transport behavior being different in different systems and samples studied in different experiments. For example, GaAs-based 2D systems could be dominated by long-range Coulomb disorder, whereas Si MOSFETs (studied in Ref. 11) could be dominated by short-range disorder due to the fact that the GaAs 2D samples are modulation doped with the ionized dopants being far from the 2D carriers, whereas in Si MOSFETs the main scattering is by nearby impurities in the oxide layer and by interface roughness.

In the unlikely scenario that such a dichotomy exists with some 2D MIT phenomena being quantum critical and others being semiclassical crossover, our theory would apply only to the situation where the transport is dominated by long-range Coulomb disorder (leading to the inhomogeneous potential landscape) and not to the situation dominated by short-range white-noise disorder.

Finally, we conclude with some brief remarks on the precise quantitative comparison between our theoretical numerical results and the experimental data of Ref. 8: (i) The most important discrepancy is that we overestimate the resistivity deep in the insulating phase where our theoretical resistivity is $\sim 10^7 \text{ ohm}$, whereas the corresponding experimental resistivity in Ref. 8 is around $\sim 10^6 \text{ ohm}$. This order of magnitude discrepancy most likely arises from our neglecting the variable range hopping conduction in the insulating phase which would strongly enhance the insulating conductivity. (ii) Our transition density tends to be somewhat smaller ($\sim 10\%$) than the experimental transition, which again could be due to our neglect of the variable range hopping conduction in the insulating phase. (iii) Typically, the calculated resistivity in

our Figs. 7 and 13 is somewhat (by about 30% or so) smaller than the experimental data at higher carrier density, which may be due to our neglect of weak localization corrections. (iv) On the other hand, our calculated resistivity in Figs. 4, 10, and 16 is somewhat larger (by about 50%) than the experimental data. (v) The last two points together indicate that a complete model of disorder, which includes all different disorder mechanisms together (i.e., interface disorder, remote disorder, bulk disorder) that we have treated separately in our numerical results may very well be able to reproduce the experimental results precisely quantitatively [except for

the strongly insulating resistivity discussed in item (i) above]. But, such a theory would have at least 10 independent disorder parameters, and getting quantitative agreement by adjusting 10 independent disorder parameters would essentially be a meaningless data-fitting exercise. Without more quantitative information about the underlying disorder, our current theory and results are probably the best one can do.

ACKNOWLEDGMENTS

This work is supported by LPS-CMTC and Microsoft Q.

- ¹T. Ando, A. B. Fowler, and F. Stern, *Rev. Mod. Phys.* **54**, 437 (1982).
- ²E. Abrahams, S. V. Kravchenko, and M. P. Sarachik, *Rev. Mod. Phys.* **73**, 251 (2001).
- ³S. V. Kravchenko and M. P. Sarachik, *Rep. Prog. Phys.* **67**, 1 (2004).
- ⁴S. Das Sarma and E. H. Hwang, *Solid State Commun.* **135**, 579 (2005).
- ⁵B. Spivak, S. V. Kravchenko, S. A. Kivelson, and X. P. A. Gao, *Rev. Mod. Phys.* **82**, 1743 (2010).
- ⁶S. Das Sarma, S. Adam, E. H. Hwang, and E. Rossi, *Rev. Mod. Phys.* **83**, 407 (2011).
- ⁷M. P. Lilly, J. L. Reno, J. A. Simmons, I. B. Spielman, J. P. Eisenstein, L. N. Pfeiffer, K. W. West, E. H. Hwang, and S. Das Sarma, *Phys. Rev. Lett.* **90**, 056806 (2003).
- ⁸M. J. Manfra, E. H. Hwang, S. Das Sarma, L. N. Pfeiffer, K. W. West, and A. M. Sergent, *Phys. Rev. Lett.* **99**, 236402 (2007).
- ⁹L. A. Tracy, E. H. Hwang, K. Eng, G. A. Ten Eyck, E. P. Nordberg, K. Childs, M. S. Carroll, M. P. Lilly, and S. Das Sarma, *Phys. Rev. B* **79**, 235307 (2009).
- ¹⁰C. Jiang, D. C. Tsui, and G. Weimann, *Appl. Phys. Lett.* **53**, 1533 (1988).
- ¹¹S. V. Kravchenko, G. V. Kravchenko, J. E. Furneaux, V. M. Pudalov, and M. D'Iorio, *Phys. Rev. B* **50**, 8039 (1994); S. V. Kravchenko, W. E. Mason, G. E. Bowker, J. E. Furneaux, V. M. Pudalov, and M. D'Iorio, *ibid.* **51**, 7038 (1995).
- ¹²D. Popovic, A. B. Fowler, and S. Washburn, *Phys. Rev. Lett.* **79**, 1543 (1997).
- ¹³M. Y. Simmons, A. R. Hamilton, M. Pepper, E. H. Linfield, P. D. Rose, D. A. Ritchie, A. K. Savchenko, and T. G. Griffiths, *Phys. Rev. Lett.* **80**, 1292 (1998).
- ¹⁴Y. Hanein, D. Shahar, J. Yoon, C. C. Li, D. C. Tsui, and H. Shtrikman, *Phys. Rev. B* **58**, R13338 (1998).
- ¹⁵Y. Hanein, U. Meirav, D. Shahar, C. C. Li, D. C. Tsui, and H. Shtrikman, *Phys. Rev. Lett.* **80**, 1288 (1998).
- ¹⁶J. Yoon, C. C. Li, D. Shahar, D. C. Tsui, and M. Shayegan, *Phys. Rev. Lett.* **82**, 1744 (1999).
- ¹⁷A. P. Mills, A. P. Ramirez, L. N. Pfeiffer, and K. W. West, *Phys. Rev. Lett.* **83**, 2805 (1999).
- ¹⁸S. Ilani, A. Yacoby, D. Mahalu, and H. Shtrikman, *Phys. Rev. Lett.* **84**, 3133 (2000).
- ¹⁹S. Ilani, A. Yacoby, D. Mahalu, and H. Shtrikman, *Science* **292**, 1354 (2001).
- ²⁰A. Lewalle, M. Pepper, C. J. B. Ford, E. H. Hwang, S. Das Sarma, D. J. Paul, and G. Redmond, *Phys. Rev. B* **66**, 075324 (2002).
- ²¹R. Leturcq, D. L'Hôte, R. Tourbot, C. J. Mellor, and M. Henini, *Phys. Rev. Lett.* **90**, 076402 (2003).
- ²²R. Leturcq, G. Deville, D. L'Hôte, R. Tourbot, C. J. Mellor, and M. Henini, *Proc. SPIE* **5469**, 101 (2004).
- ²³S. Das Sarma, M. P. Lilly, E. H. Hwang, L. N. Pfeiffer, K. W. West, and J. L. Reno, *Phys. Rev. Lett.* **94**, 136401 (2005).
- ²⁴G. Allison, E. A. Galaktionov, A. K. Savchenko, S. S. Safonov, M. M. Fogler, M. Y. Simmons, and D. A. Ritchie, *Phys. Rev. Lett.* **96**, 216407 (2006).
- ²⁵S. Anissimova, S. V. Kravchenko, A. Punnoose, A. M. Finkelstein, and T. M. Klapwijk, *Nature (London)* **3**, 707 (2007).
- ²⁶K. Lai, W. Pan, D. C. Tsui, S. Lyon, M. Mühlberger, and F. Schäffler, *Phys. Rev. B* **75**, 033314 (2007).
- ²⁷S. Adam, S. Cho, M. S. Fuhrer, and S. Das Sarma, *Phys. Rev. Lett.* **101**, 046404 (2008).
- ²⁸R. L. J. Qiu, X. P. A. Gao, L. N. Pfeiffer, and K. W. West, *Phys. Rev. B* **83**, 193301 (2011); N. Gobler, J. Watson, M. Manfra, and X. P. A. Gao, arXiv:1308.3698.
- ²⁹L. A. Tracy, J. P. Eisenstein, M. P. Lilly, L. N. Pfeiffer, and K. W. West, *Solid State Commun.* **137**, 150 (2006).
- ³⁰A. Gold, *Phys. Rev. B* **44**, 8818 (1991).
- ³¹P. T. Coleridge, R. L. Williams, Y. Feng, and P. Zawadzki, *Phys. Rev. B* **56**, R12764 (1997).
- ³²Z. Wilamowski, N. Sandersfeld, W. Jantsch, D. Többen, and F. Schäffler, *Phys. Rev. Lett.* **87**, 026401 (2001).
- ³³F. Stern, *Phys. Rev. B* **9**, 4597 (1974); J. T. C. Chen and R. S. Muller, *J. Appl. Phys.* **45**, 828 (1974); J. R. Brews, *ibid.* **46**, 2181 (1975); **46**, 2193 (1975).
- ³⁴F. Stern, *Phys. Rev. Lett.* **44**, 1469 (1980); F. Stern and S. Das Sarma, *Solid State Electron.* **28**, 211 (1985).
- ³⁵S. Das Sarma and E. H. Hwang, *Phys. Rev. Lett.* **83**, 164 (1999).
- ³⁶S. Das Sarma and E. H. Hwang, *Phys. Rev. B* **61**, R7838 (2000).
- ³⁷S. Das Sarma and E. H. Hwang, *Phys. Rev. B* **69**, 195305 (2004).
- ³⁸S. Kirkpatrick, *Rev. Mod. Phys.* **45**, 574 (1973).
- ³⁹Q. Li, E. H. Hwang, and S. Das Sarma, *Phys. Rev. B* **84**, 115442 (2011); See, for example, E. Hajaj *et al.*, *ibid.* **88**, 045128 (2013) for an explicit experimental verification of the two-component transport theory in monolayer graphene.
- ⁴⁰E. H. Hwang and S. Das Sarma, *Phys. Rev. B* **82**, 081409 (2010).
- ⁴¹S. Das Sarma, E. H. Hwang, and Q. Li, *Phys. Rev. B* **85**, 195451 (2012).
- ⁴²Q. Li, E. Rossi, and S. Das Sarma, *Phys. Rev. B* **86**, 235443 (2012).
- ⁴³C. Herring, *J. Appl. Phys.* **31**, 1939 (1960).
- ⁴⁴E. O. Kane, *Phys. Rev.* **131**, 79 (1963).
- ⁴⁵T. N. Morgan, *Phys. Rev.* **139**, A343 (1965).

- ⁴⁶R. Zallen and H. Scher, *Phys. Rev. B* **4**, 4471 (1971).
- ⁴⁷B. I. Shklovskii and A. L. Efros, *Zh. Eksp. Teor. Fiz.* **61**, 816 (1971) [*Sov. Phys.-JETP* **34**, 435 (1972)]; *Electronic Properties of Doped Semiconductors* (Springer-Verlag, Berlin, 1984).
- ⁴⁸E. Arnold, *Appl. Phys. Lett.* **25**, 705 (1974).
- ⁴⁹E. Arnold, *Surf. Sci.* **58**, 60 (1976).
- ⁵⁰T. P. Eggarter and M. H. Cohen, *Phys. Rev. Lett.* **25**, 807 (1970).
- ⁵¹A. L. Efros and B. I. Shklovskii, *Phys. Status Solidi B* **76**, 475 (1976).
- ⁵²A. Weinrib, *Phys. Rev. B* **26**, 1352 (1982).
- ⁵³A. Bunde, A. Coniglio, D. C. Hong, and H. E. Stanley, *J. Phys. A: Math. Gen.* **18**, L137 (1985).
- ⁵⁴A. L. Efros, *Solid State Commun.* **65**, 1281 (1988); **70**, 253 (1989).
- ⁵⁵D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London, 1992).
- ⁵⁶M. B. Isichenko, *Rev. Mod. Phys.* **64**, 961 (1992).
- ⁵⁷S. He and X. C. Xie, *Phys. Rev. Lett.* **80**, 3324 (1998).
- ⁵⁸É. G. Batyev, *J. Exp. Theor. Phys.* **102**, 776 (2006).
- ⁵⁹F. G. Pikus and A. L. Efros, *Zh. Éksp. Teor. Fiz.* **96**, 985 (1989) [*Sov. Phys.-JETP* **69**, 558 (1989)].
- ⁶⁰A. L. Efros, *Solid State Commun.* **67**, 1019 (1988).
- ⁶¹M. M. Fogler, *Phys. Rev. B* **69**, 121409 (2004).
- ⁶²J. A. Nixon and J. H. Davies, *Phys. Rev. B* **41**, 7929 (1990).
- ⁶³A. L. Efros, F. G. Pikus, and V. G. Burnett, *Phys. Rev. B* **47**, 2233 (1993).
- ⁶⁴J. Shi and X. C. Xie, *Phys. Rev. Lett.* **88**, 086401 (2002); J. Shi, S. He, and X. C. Xie, *Phys. Rev. B* **60**, R13950 (1999).
- ⁶⁵E. Rossi and S. Das Sarma, *Phys. Rev. Lett.* **101**, 166803 (2008).
- ⁶⁶E. Rossi, S. Adam, and S. Das Sarma, *Phys. Rev. B* **79**, 245423 (2009).
- ⁶⁷M. Polini, A. Tomadin, R. Asgari, and A. H. MacDonald, *Phys. Rev. B* **78**, 115426 (2008).
- ⁶⁸E. Rossi and S. Das Sarma, *Phys. Rev. Lett.* **107**, 155502 (2011).
- ⁶⁹Q. Li, E. H. Hwang, E. Rossi, and S. Das Sarma, *Phys. Rev. Lett.* **107**, 156601 (2011).
- ⁷⁰Q. Li, E. Hwang, and E. Rossi, *Solid State Commun.* **152**, 1390 (2012).
- ⁷¹P. Van Mieghem, *Rev. Mod. Phys.* **64**, 755 (1992).
- ⁷²F. Stern and W. E. Howard, *Phys. Rev.* **163**, 816 (1967).
- ⁷³F. Stern, *Surf. Sci.* **58**, 162 (1976).
- ⁷⁴S. Das Sarma and E. H. Hwang, *Phys. Rev. B* **68**, 195315 (2003).
- ⁷⁵F. Stern, *Phys. Rev. Lett.* **18**, 546 (1967).
- ⁷⁶T. Ando, *J. Phys. Soc. Jpn.* **51**, 3215 (1982).
- ⁷⁷S. Das Sarma, *Phys. Rev. B* **33**, 5401 (1986).
- ⁷⁸Y. Meir, *Phys. Rev. B* **63**, 073108 (2001).
- ⁷⁹Y. Hanein, D. Shahar, J. Yoon, C. C. Li, D. C. Tsui, and H. Shtrikman, *Phys. Rev. B* **58**, R7520 (1998).
- ⁸⁰S. Das Sarma and E. H. Hwang, *Phys. Rev. Lett.* **84**, 5596 (2000).
- ⁸¹S. Das Sarma and E. H. Hwang, *Phys. Rev. B* **72**, 205303 (2005).
- ⁸²S. Das Sarma and E. H. Hwang, *Phys. Rev. B* **72**, 035311 (2005).
- ⁸³J. Yoon, C. C. Li, D. Shahar, D. C. Tsui, and M. Shayegan, *Phys. Rev. Lett.* **84**, 4421 (2000).
- ⁸⁴V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, [arXiv:cond-mat/0103087](https://arxiv.org/abs/cond-mat/0103087).
- ⁸⁵R. Pillarisetty, H. Noh, E. Tutuc, E. P. De Poortere, D. C. Tsui, and M. Shayegan, *Phys. Rev. Lett.* **90**, 226801 (2003).
- ⁸⁶Y. Tsui, S. A. Vitkalov, M. P. Sarachik, and T. M. Klapwijk, *Phys. Rev. B* **71**, 113308 (2005).
- ⁸⁷K. Lai, W. Pan, D. C. Tsui, S. A. Lyon, M. Mühlberger, and F. Schäffler, *Phys. Rev. B* **72**, 081313 (2005).
- ⁸⁸Y. Meir, *Phys. Rev. Lett.* **83**, 3506 (1999).
- ⁸⁹Y. Meir, *Phys. Rev. B* **61**, 16470 (2000).