Full band calculations of the intrinsic lower limit of contact resistivity

J. Maassen,1,a) C. Jeong,1 A. Baraskar,2 M. Rodwell,3 and M. Lundstrom1
1Network for Computational Nanotechnology, School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907, USA
2GlobalFoundries, Yorktown Heights, New York 10598, USA
3Department of Electrical and Computer Engineering, University of California, Santa Barbara, California 93106, USA

(Received 21 November 2012; accepted 8 March 2013; published online 21 March 2013)

The intrinsic lower limit of contact resistivity ($\rho_{c}^{LL}$) for InAs, In$_{0.53}$Ga$_{0.47}$As, GaSb, and Si is calculated using a full band ballistic quantum transport approach. Surprisingly, our results show that $\rho_{c}^{LL}$ is almost independent of the semiconductor. An analytical model, derived for 1D, 2D, and 3D, correctly reproduces the numerical results and explains why $\rho_{c}^{LL}$ is very similar in all cases. Our analysis sets a minimal carrier density required to meet the International Technology Roadmap for Semiconductors call for $\rho_{c} = 10^{-9}$ $\Omega$-cm$^2$ by 2023. Comparison with experiments shows there is room for improvement, which will come from optimizing interfacial properties. © 2013 American Institute of Physics. [http://dx.doi.org/10.1063/1.4798238]

Achieving ultra-low contact resistivities for metal/semiconductor junctions represents a significant obstacle for the scaling of nano-scale devices and the development of THz bandwidth high-speed transistors.1 Much experimental effort has been devoted to producing high quality material interfaces, utilizing in-situ techniques for sample fabrication, with the goal of reducing the specific contact resistivity ($\rho_{c}$).2 The International Technology Roadmap for Semiconductors (ITRS) has called for a contact resistivity of $10^{-9}$ $\Omega$-cm$^2$ for 2023.3 This leads us to ask the following question: What is the ultimate lower limit of Ohmic contact resistivity?

A recent theoretical study based on the parabolic approximation and realistic band profiles demonstrated contact resistivity values in excellent agreement with experiment.4 This study also computed the case of perfect reflection-less contacts, and found that this resistance was within an order of magnitude of the lowest measured contact resistivity to In$_{0.53}$Ga$_{0.47}$As. To achieve lower $\rho_{c}$ values, high doping concentrations are required, which can push the Fermi level far from the band edge where the electronic dispersion can be highly non-parabolic. In this work, we address this important issue by calculating the theoretical lower limit of contact resistivity ($\rho_{c}^{LL}$) for several semiconductors by combining accurate full band electronic structure calculations with the Landauer quantum transport formalism. By naturally taking into account the role of valley degeneracy, band anisotropy, higher energy bands, and the highly non-parabolic shape of the bands, our results allow us to quantitatively assess how much improvement in $\rho_{c}$ is possible given the present experimental values and verify the validity of parabolic-based $\rho_{c}$ models. An analytical model, applicable to 1D, 2D, and 3D structures, is found to adequately reproduce the full band numerical results and is utilized to provide an answer to what material properties are desired to achieve the lowest possible $\rho_{c}^{LL}$.

In this work, the intrinsic lower limit of metal/semiconductor contact resistivity is calculated assuming that (i) transport across the junction is ballistic, i.e., no scattering and no Schottky barrier for carriers to tunnel through, and that (ii) the metal is ideal meaning it contains more than enough conducting channels to supply the semiconductor. To compute $\rho_{c}^{LL}$, we use the Landauer approach, which is naturally suited to treat ballistic transport. The contact resistivity (or inverse of the conductance per unit area), defined in the limit of zero applied bias, can be expressed as:

$$\frac{1}{\rho_{c}^{LL}} = \frac{4q^2}{h} \int_{-\infty}^{+\infty} M(\varepsilon) \left[ -\frac{\partial f}{\partial \varepsilon} \right] \mathrm{d}\varepsilon, \quad [\Omega^{-1}\text{m}^{-2}], \quad (1)$$

where $q$ is the electron charge, $h$ is Planck’s constant, $M(\varepsilon)$ is the semiconductor distribution of modes (DOMs) or number of conducting channels at energy $\varepsilon$, and $f$ is the Fermi-Dirac distribution. The function $[-\partial f / \partial \varepsilon]$ is centered at the Fermi level $E_F$ and is strongly suppressed a few $k_B T$ away from $E_F$, where $k_B$ is Boltzmann’s constant and $T$ is the temperature. Thus, $(\rho_{c}^{LL})^{-1}$ can be simply interpreted as the number of “active” conducting channels times the quantum of conductance $2q^2/h$; an extra factor of two appears because half of the resistance is associated with each of the two contacts.

In this study, we consider four of the most important semiconductors for nano-scale and high-frequency electronic devices: InAs, In$_{0.53}$Ga$_{0.47}$As, GaSb, and Si. The first three are low-effective mass ($m^*$) III-V semiconductors with a single isotropic conduction band (CB) located at $\Gamma$, while Si has a six-fold degenerate large $m^*$ anisotropic CB. By studying these much different semiconductors, our goal is to understand and identify characteristics that can lead to a reduction in contact resistivity. Fig. 1 presents the calculated band structures of InAs, In$_{0.53}$Ga$_{0.47}$As, GaSb, and Si using tight-binding and first principles frameworks (details found in figure caption). Most calculations of contact resistivity assume a parabolic-type form for the electronic bands, however at high doping concentrations, $E_F$ may reside far from the band edge where features can be highly non-parabolic.
Electronic band structure of InAs (a), In_{0.53}Ga_{0.47}As (b), GaSb (c), and Si (d) along the high-symmetry points. (a)–(c) were calculated using a sp^3d^5s^* tight-binding model as implemented in NEMO, and (d) was obtained from VASP, a density functional theory software package, using the generalized gradient approximation, and applying the scissor technique to obtain the experimental band gap.

FIG. 1. Electronic band structure of InAs (a), In_{0.53}Ga_{0.47}As (b), GaSb (c), and Si (d) along the high-symmetry points. (a)–(c) were calculated using a sp^3d^5s^* tight-binding model as implemented in NEMO, and (d) was obtained from VASP, a density functional theory software package, using the generalized gradient approximation, and applying the scissor technique to obtain the experimental band gap.

(is as shown in Fig. 1) and higher energy bands may contribute to transport.

The full electronic dispersion serves as the input for calculating the DOM, which can be efficiently computed using the so-called “band-counting” method. The number of conducting channels per unit area, at a given energy \( \epsilon \), is expressed as

\[
M(\epsilon) = \frac{1}{(2\pi)^2} \int_{BZ} \sum_{m^+} \Theta(\epsilon - \epsilon_{k_m}) \, dk_m, \quad [m^{-2}],
\]

where \( \Theta(\epsilon - \epsilon_{k_m}) \) is the Heaviside step function, \( k_m \) is the reciprocal vector in the plane perpendicular to transport, and \( n^+ \) is the number of forward moving states with energy \( \epsilon \) and wavenumber \( k_m \). The integrand of Eq. (2) is an integer corresponding to the number of conducting channels, at a fixed \( \epsilon \) and \( k_m \). For a non-parabolic Kane dispersion model, where \( \epsilon_k (1 + \epsilon_k) = \hbar^2 k^2 / 2m^* \), the distribution of modes (for a 3D material) can be written analytically as

\[
M(\epsilon) = m_{DOM}^p \frac{1}{(2\pi)^2} \int_{BZ} \sqrt{m_0} \left( 1 + \epsilon_k / \hbar^2 k^2 / 2m^* \right)^2 / 2, \quad (2)
\]

where \( \epsilon_k \) is the non-parabolicity parameter, \( m_{DOM}^p \) is the DOM effective mass, \( g_v \) is the valley degeneracy, and \( \epsilon \) is defined relative to the band edge. Therefore, the number of conducting channels scales linearly with energy for parabolic bands \( (\epsilon = 0) \) and quadratically with energy for non-parabolic bands \( (\epsilon \neq 0) \).

Fig. 2 presents the distribution of modes, \( M(\epsilon) \), for the valence and conduction states of InAs, In_{0.53}Ga_{0.47}As, GaSb, and Si. The p-type \( M(\epsilon) \) is roughly the same for all semiconductors, with Si showing slightly larger values due to its comparatively smaller split-off energy and larger \( m^* \). Significant differences in \( M(\epsilon) \) appear for the conduction states. Si and GaSb have a much larger number of modes compared to InAs and In_{0.53}Ga_{0.47}As. For energies less than 1 eV, InAs and In_{0.53}Ga_{0.47}As have a single low \( m^* \) band contributing to \( M(\epsilon) \). According to the analytic expression shown above, \( M(\epsilon) \) will scale linearly with \( m_{DOM}^p \), which is simply \( m^* \) for an isotropic band. For InAs and In_{0.53}Ga_{0.47}As, \( m^* \) is 0.024 \( m_0 \) and 0.042 \( m_0 \), respectively, where \( m_0 \) is the electron mass. While for Si, the longitudinal and transverse effective masses extracted from Fig. 1(d) are \( m_l = 0.96 \) \( m_0 \) and \( m_t = 0.2 \) \( m_0 \), leading to a \( m_{DOM}^p \) value of 2.15 \( m_0 \), considering a valley degeneracy of six. The difference in \( m_{DOM}^p \) for InAs and In_{0.53}Ga_{0.47}As versus Si, is the source of the roughly factor of 10^2 difference in DOM.

To better illustrate why small effective mass bands lead to a small DOM, in Figs. 2(c) and 2(d), we plot the \( k_x \)-resolved DOM, i.e., the integrand in Eq. (2), for the conduction bands of InAs and Si. Such plots provide information on the projected iso-energy surface of the bands: a single small sphere for InAs and six large ellipsoids for Si. For a given energy, the InAs CB takes up only a small region in \( k \)-space, compared to the six bands of Si, which once projected along \( z \) and integrated leads to a small DOM. GaSb has a small effective mass CB with a \( m^* \) of 0.041 \( m_0 \), however a second large \( m^* \) band appears at 0.1 eV, which dramatically increases the DOM, as shown in the inset of Fig. 2(b). InAs and In_{0.53}Ga_{0.47}As also have higher energy bands located at 1.17 eV and 0.98 eV above the band edge (according to Fig. 1), respectively, which result in sharp increases in \( M(\epsilon) \).

In Figs. 3(a) and 3(b), contact resistivity \( \rho_{xx}^{LL} \) versus carrier concentration \( n \) is presented for p-type and n-type InAs, In_{0.53}Ga_{0.47}As, GaSb, and Si. As the carrier concentration increases, the contact resistivity decreases. The roughly linear trend on the log-log plot indicates that \( \rho_{xx}^{LL} \) is related to \( n \) via a power-law, \( \rho_{xx}^{LL} \propto n^x \), with \( x \approx -2/3 \) extracted from the slope. The main feature in Figs. 3(a) and 3(b) is that, for a fixed carrier concentration, \( \rho_{xx}^{LL} \) is similar for all semiconductors (i.e., within a factor of 2–3). This is an unexpected result considering the very different DOM of each material, particularly for the conduction states.
as shown in Figs. 3(c) and 3(d), confirms that semiconductors with a larger number of channels yield smaller $\rho_{c}^{LL}$ values (for fixed $E_F$). The markers in Figs. 3(c) and 3(d), corresponding to carrier concentrations of $10^{18}$, $10^{19}$, $10^{20}$, and $10^{21}$ cm$^{-3}$, indicate that it is the relationship between $n$ and $E_F$, which leads to similar $\rho_{c}^{LL}$ values for fixed $n$. In order to better understand the correspondence between $\rho_{c}^{LL}$ and $n$, and why all the studied semiconductors have nearly the same $\rho_{c}^{LL}$, we will utilize an approximate analytical formula.

Lowering the contact resistivity requires high doping levels, which means the semiconductors are degenerate. In this case, one can safely evaluate $\rho_{c}^{LL}$ (Eq. (1)) and the carrier concentration $n = \int_{-\infty}^{\infty} D(x)f(\epsilon - E_F) d\epsilon$ in the $T=0$ K limit. Assuming a non-parabolic Kane model for the electronic dispersion in 3D, 2D, and 1D respectively, we find

$$\rho_{c}^{LL} = \frac{h}{4q^2} \frac{m_{DOS}^3}{m_{DOM}^3} \left(3\sqrt{nR^3}\right)^{2/3}, \quad [\Omega - m^2],$$  \hfill (3)

$$\rho_{c}^{LL} = \frac{h}{4q} \sqrt{\frac{2\pi}{3}} \frac{m_{DOS}^3}{m_{DOM}^3} \sqrt{\frac{1}{nR^3}}, \quad [\Omega - m],$$  \hfill (4)

$$\rho_{c}^{LL} = \frac{h}{4q^2} \sqrt{\frac{nR^3}{g}}, \quad [\Omega],$$  \hfill (5)

where $m_{DOS}^3 = \sum_{\epsilon_i} \sqrt{m_{\epsilon_i} m_x m_y}$, $m_{DOM}^3 = \sum_{\epsilon_i} \sqrt{m_{\epsilon_i} m_x m_z}$, $m_{DOS}^2 = \sum_{\epsilon_i} \sqrt{m_{\epsilon_i} m_y}$, and $m_{DOM}^2 = \sum_{\epsilon_i} \sqrt{m_{\epsilon_i} m_z}$ (assuming transport along $z$ and confinement along $x$ for the 2D case). The 3D and 2D carrier densities have units of m$^{-3}$ and m$^{-2}$, respectively. Unless otherwise mentioned, we consider the case of 3D. Equation (3) predicts that $\rho_{c}^{LL}$ should depend on $n$ to the power $-2/3$ in the degenerate limit, as is observed with the numerical results. Note that the non-parabolicity parameter $\alpha$ does not appear in Eqs. (3)-(5), and that the ratio $m_{DOS}^3/m_{DOM}^3$ is one for isotropic bands. Equation (3) is plotted in Figs. 3(a) and 3(b) as a thin yellow solid line, using $m_{DOS}^3/m_{DOM}^3 = 1$, and is found to provide an excellent fit to the numerical results considering this simple equation depends only on fundamental constants. Thus, this model explains why the curves in Figs. 3(a) and 3(b) nearly overlap: any decrease in $\rho_{c}^{LL}$ due to an increase in $M(\epsilon)$ is offset by a small $E_F$ value relative to the band edge. Hence, a good conductor (large $m_{DOM}^3$ and $m_{DOS}^3$) has a small $E_F$ and a poor conductor (small $m_{DOM}^3$ and $m_{DOS}^3$) has a large $E_F$, such that both yield a very similar $\rho_{c}^{LL}$. For Si, we have $m_{DOS}^3/m_{DOM}^3 = 0.52$, which reduces $\rho_{c}^{LL}$ compared to isotropic bands. As expected, comparing Si to InAs and In$_{0.53}$Ga$_{0.47}$As, we observe a roughly a factor of two difference in $\rho_{c}^{LL}$. Finally, we highlight that the analytical expressions in Eqs. (3)-(5) predict that $\rho_{c}^{LL}$ will have a gradually weaker dependence on carrier density, $m_{DOM}^3$ and $m_{DOS}^3$ as dimensionality is reduced.

We may ask how does valley degeneracy, band anisotropy, and higher energy bands affect the lower limit of $\rho_{c}$ and why do such features, present in the full band calculations, do not induce significant differences in $\rho_{c}^{LL}$ from one material to another. Assuming there are $g_v$ degenerate ellipsoidal bands, Eq. (3) can be rewritten as

$$\rho_{c}^{LL} = \frac{h}{4q^2 \sqrt{g_v}} \left(\frac{nR^3}{g_v}\right)^{1/3} \left(\frac{m_{\epsilon_i}}{\sqrt{m_{\epsilon_i} m_x}}\right)^{1/3} \left(\frac{4}{3\sqrt{nR^3}}\right)^{2/3}, \quad [\Omega - m^2].$$  \hfill (6)

We note that the same result, within a factor of two, was reported elsewhere.\textsuperscript{3} Equation (6) indicates that although the lower limit of $\rho_{c}$ can be decreased with increasing band degeneracy, this dependence is weak. For example, with $g_v = 6$, the contact resistivity is reduced by a factor of 1.82. Moreover, Eq. (6) also demonstrates that band anisotropy can lower $\rho_{c}^{LL}$ by decreasing $m_{\epsilon_i}$ (farther from the band edge) and increasing $m_{\epsilon_i}$ (larger number of conducting channels). Due to the effective mass power of $1/3$, anisotropic effects are also weak. Finally, we verify the role of higher
energy bands. As an example, we consider a single lower energy (low \( m' \)) band with six degenerate higher energy (large \( m' \)) bands. In this case, only a factor of 2 reduction in \( \rho_{c}^{LL} \) is observed (see supplemental material for more details).

Hence, valley degeneracy, band anisotropy, and higher energy bands have only a modest impact on \( \rho_{c}^{LL} \), leading to a reduction of 2–3 \times 10^3 \) under optimal conditions. This explains why such effects do not induce significant differences in the numerically calculated \( \rho_{c}^{LL} \).

The markers in Figs. 3(a) and 3(b) show recent experimental values. InAs is known to produce very low \( \rho_{c} \) values, due to Fermi level pinning in the CB, but our theoretical calculations indicate that it is possible to decrease \( \rho_{c} \) by another order of magnitude before reaching its intrinsic lower limit. Thus, there is room for improvement, which will be achieved by optimizing the physical properties at the metal/semiconductor interface to increase the transmission probability. The ITRS calls for \( \rho_{c} = 10^{-9} \) \( \Omega \)-cm\(^2\) for 2023 and is plotted as the dashed horizontal line in Figs. 3(a) and 3(b). Our simulations demonstrate that the ITRS value is indeed possible, but sets a minimal carrier concentration and is plotted as the dashed horizontal line in Figs. 3(a) and 3(b).

In this work, we have assumed a transmission \( T(\epsilon) \) of one, however experimentally \( T(\epsilon) < 1 \) due to tunneling through a Schottky barrier and other non-idealities. When \( T(\epsilon) \neq 1 \), \( \rho_{c}^{LL} \) is calculated by introducing a factor \( T(\epsilon) \) into the integral of Eq. (1). Using measured contact resistivity values, \( \rho_{c}^{exp} \), the energy-averaged transmission is easily extracted using \( T(\epsilon) = \rho_{c}^{LL}/\rho_{c}^{exp} \), where \( \rho_{c}^{LL} \) is simply calculated using Eq. (3). From the data shown in Figs. 3(a) and 3(b), we find \( T(\epsilon) \) values in the range 1%–4% for In\(_{0.53}\)Ga\(_{0.47}\)As and 10% for InAs. In practice, the properties of the interface determine the specific contact resistivity not the properties of the semiconductors alone.

In summary, the intrinsic lower limit for contact resistivity \( \rho_{c}^{LL} \) was calculated for InAs, In\(_{0.53}\)Ga\(_{0.47}\)As, GaSb, and Si using a full band ballistic quantum transport approach. Surprisingly, all calculated \( \rho_{c}^{LL} \) for fixed carrier density are found to be within a factor of 2–3 of each other, with Si and GaSb providing the smallest values. An analytical model, which is found to provide a very good fit to the numerical results, demonstrates that any reduction in \( \rho_{c}^{LL} \) achieved by increasing \( M(\epsilon) \) is compensated by the fact that \( \epsilon_{F} \) will reside closer to the band edge. We also find that the value of \( \rho_{c} = 10^{-3} \) \( \Omega \)-cm\(^2\) called by the ITRS for 2023 is in principle possible, but sets a minimal carrier concentration ranging from \( 2 \times 10^{19} \)\( \text{cm}^{-3} \) to \( 5 \times 10^{19} \)\( \text{cm}^{-3} \). In reality, concentrations above this ideal value will be required. Experimentally, it is the specific interfacial properties of the metal/semiconductor junction, which determines how close \( \rho_{c} \) gets to its intrinsic lower limit. The contribution of this paper is to show how much improvement is possible through optimization of the metal/semiconductor interface.

J.M., C.J., and M.L. acknowledge support from the MSD focus center of the SRC and the Network for Computational Nanotechnology. Work at UCSB supported by the ONR under contract N00014-09-1-0231. J.M. thanks Professor Hong Guo for additional computational resources.

14See supplementary material at http://dx.doi.org/10.1063/1.4798238 for an assessment of the parabolic and non-parabolic approximations, the zero temperature approximation and the effect of higher energy bands using the analytical contact resistivity.