Limits of Specific Contact Resistivity to Si, Ge and III-V Semiconductors Using Interfacial Layers

Gautam Shine and Krishna C. Saraswat
Department of Electrical Engineering
Stanford University
Stanford, CA 94305 USA
gshine@stanford.edu

Abstract—**Specific contact resistivities of source/drain contacts employing interfacial layers are calculated with simulations of tunneling transport. Fermi level depinning, dipoles, and other techniques for barrier lowering are explored. Interfacial materials with the potential to meet future contact resistivity requirements are identified for silicon and high-mobility alternatives.**

I. INTRODUCTION

Scaling requirements call for steadily decreasing specific contact resistivity (ρ_c) for source/drain contacts to nanoscale MOSFETs. Insertion of thin dielectric layers between metals and semiconductors has experimentally demonstrated reduced contact resistivity in cases where Fermi level pinning causes formation of large Schottky barriers [1]. Pinning is especially problematic in high-mobility materials due to the low dopant activation currently achievable in them.

The interfacial material diminishes Fermi level pinning by spatially separating the metal and semiconductor, thus attenuating the metal wavefunctions. The fundamental tradeoff is the additional series resistance introduced by the layer. Other barrier-reducing mechanisms include dipole-induced potential shifts, fixed charge in the dielectric, and transfer of the pinning to another interface. The relative strength of each mechanism depends on the specific materials stack.

We have developed a contact simulator to explore the materials space and predict the best choices. The materials Si, Ge, and GaSb are the focus of this work because of their adequate mobility for both types of carriers, permitting use in both nFETs and pFETs. Contacts to p-type Ge and GaSb are trivial due to pinning at the valence band, but n-type contacts are poor for the same reason and exacerbated by low donor activation. Large improvements can be expected in such cases, but even Si and InGaAs can benefit from the method.
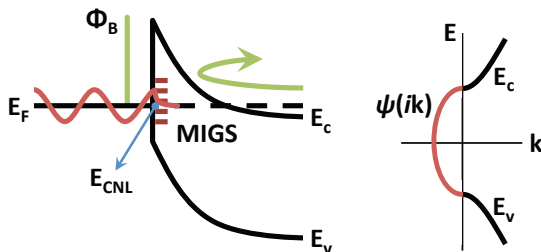

Fig. 1: Left: Schottky barrier formation due to Fermi level pinning caused by metal induced gap states (MIGS). Right: The complex bands within the band gap responsible for charge neutrality levels at semiconductor interfaces.
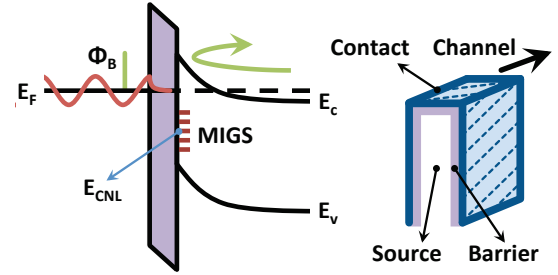

Fig. 2: Left: Schottky barrier height reduction through MIGS blocking. Right: The contact area advantage when barriers are used instead of silicides.

II. METHODOLOGY

We calculate the current density at the interface near equilibrium using the Tsu-Esaki tunneling model [2], which is a special case of the general Landauer formalism with the assumptions of parallel wavevector conservation and effective masses for each parabolic band of interest. Parallel is defined as lying in the interface plane. Current density is given by

$$J = \frac{q}{h} \int dE \ T(E) \ (f_1 - f_2)$$
$$= \frac{q m^* kT}{2\pi^2 \hbar^3} \int d\epsilon_\perp \ T(\epsilon_\perp) \ log \left(\frac{1 + exp(\frac{E_{F_1} - \epsilon_\perp}{kT})}{1 + exp(\frac{E_{F_2} - \epsilon_\perp}{kT})} \right)$$

$T(\epsilon_\perp)$ is the transmission probability at transverse energy $\epsilon_\perp = \hbar^2 k_\perp^2 / 2m^*$, i.e. the carrier's energy in the direction of transport. $T(\epsilon_\perp)$ is evaluated using numerical solutions to the Schrödinger equation obtained from transfer matrices. Thermionic and field emission do not need to be distinguished within this framework. Note that density of states (DOS) is a function of total energy. The logarithmic factor appearing in the current density relates the dependence on total energy to transverse energy, allowing integration in a single variable $\epsilon_\perp$.

Three different semiconductor effective masses enter into the calculation. The tunneling effective mass is used for $T(\epsilon_\perp)$ evaluation. The m^* appearing in the current density prefactor is related to the density of states. This is identical to the usual m^* for isotropic energy surfaces such as III-V Γ-valleys. For anisotropic valleys such as Si and Ge's conduction band minima, the DOS is projected on to the plane perpendicular to transport so that $m^* = \Sigma_{g_v} \sqrt{m_x m_y}$. For example, the six-fold degenerate Si X-valley has $m^* = 2m_t + 4\sqrt{m_t m_l} = 2.05 m_0$ ($m_t = 0.19$, $m_l = 0.92$), rather than the usual value of $1.08 m_0$.

However, for determining the position of the Fermi level in the semiconductor, the conventional m^*_{DOS} is the relevant value. In some cases, highly degenerate doping levels cause the Fermi level in III-Vs to cross another band with a much higher density of states. GaSb is such a case, with the L-valley only about 1 kT above the Γ-valley and possessing a much higher m^*_{DOS} of $0.57m_0$ compared to $0.041m_0$ at Γ. Thus for $N_D > 10^{18}$ cm^{-3}, the L-valley dominates conductance in GaSb and is treated similarly to Ge.

The potential profile is determined by solving the 1D Poisson equation in equilibrium, including fixed bulk and interface charge. Fermi level pinning (i.e. interface states) and interfacial dipoles are accounted for through shifts in the vacuum level and semiconductor surface potential. The overall effect after achieving self-consistency is that the semiconductor sees an effective metal work function different from the actual value. This modulates the Schottky barrier height and thus has dramatic effects on ρ_c through the transmission probability, which has an exponential dependence on barrier height.

Fermi level pinning is modeled using the pinning factor S [3][4] from the theory of metal-induced gap states (MIGS). It is estimated from experimental fits or by using the empirical trend established by Mönch [5] that relates it to the strength of dielectric screening. S and effective ϕ_M are given by

$$S = \frac{\partial \phi_B}{\partial \phi_M} = \frac{1}{1 + 0.1(\epsilon_\infty - 1)^2}$$

$$\phi_{M,eff} = S\phi_M + (1 - S)\phi_{CNL}$$

In the limit of zero interfacial layer thickness, S is determined by the semiconductor alone. In the limit of high thickness, it is determined by the interfacial material alone. We can interpolate based on the exponential decay of MIGS density, with the decay approximated using band gaps and lattice constants [6]

$$|\psi|^2 \propto exp(-\beta t), \quad \beta = \frac{m_0 a E_g}{\pi \hbar^2}$$

The energy ϕ_{CNL} is the charge neutrality level of the insulator or semiconductor referenced to vacuum. It is a feature of the material's complex band structure and lies within the band gap. Complex wavevectors are valid solutions to the Schrödinger equation when crystal periodicity is terminated. At ϕ_{CNL} these gap states switch between donor-like (similar to conduction band orbitals) and acceptor-like (similar to valence band orbitals). Fermi level deviation from this energy rapidly causes charge accumulation so the Fermi level is "pinned" to maintain neutrality. This diminishes the influence of ϕ_M considerably and gives rise to unavoidable Schottky barriers.
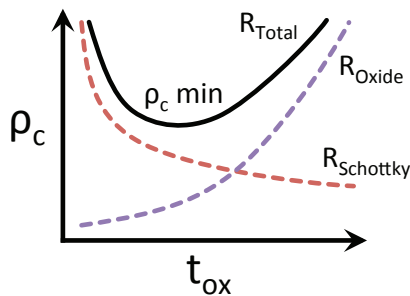
Fig. 3: General curve for specific contact resistivity v. barrier thickness.

III. RESULTS

The basic goal of interfacial layers is to alter $\phi_{M,eff}$, either by depinning the metal Fermi level so that the semiconductor is responsive to the metal's actual ϕ_M or by "re-pinning" the metal Fermi level to a more favorable energy level. In either case, it is critical that this interfacial layer itself be highly conductive. Altering pinning away from ϕ_{CNL} requires spatial separation of the metal and semiconductor in order to block metal wavefunctions from penetrating into the semiconductor. The interfacial layer achieves this, at which point conventional MOS electrostatics determines the band bending rather than interface effects such as MIGS or dipoles.

Previous work, both experimental and theoretical [2][7][8], has generally found that the lowest contact resistivities are obtained with materials that have a low band offset with the semiconductor. Although higher band gaps attenuate the metal wavefunctions faster, the tunneling resistance is much higher for the same reason and presents a poor tradeoff. Almost all work has focused on n-type contacts and of particular interest have been TiO_2 and ZnO, which have succeeded on Si [9], Ge [10][11], GaSb [12], and GaAs [13]. These are n-type wide band gap semiconductors rather than conventional insulators. With an electron affinity of $\sim$4 eV, their conduction bands line up with low offset to most common semiconductors. Moreover, their tendency to dope n-type decreases ohmic resistance when injected electrons drift-diffuse in the conduction band rather than tunnel through the band gap.
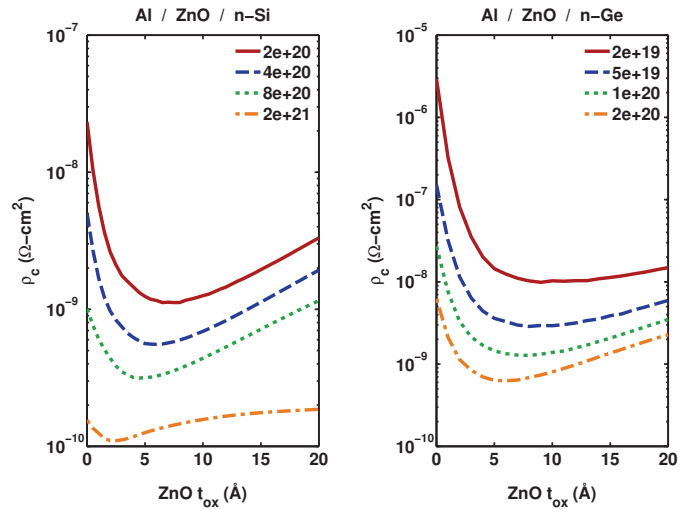
Fig. 4: Al-ZnO contacts to n-Si and n-Ge. Considerable improvement is seen in Ge because of valence band pinning and lower doping levels. Improvement in Si is less dramatic, especially at doping levels near the solid solubility limit.

Al was chosen for contacting to ZnO because the interface is reactive and Al atoms serve as a donor in ZnO, producing n^{++}-ZnO. Though TiO_2 has been common in experiments, the model suggests that ZnO is a superior alternative. This is consistent with what one experiment has found for contacts to n-Ge [11]. The primary reasons are its greater depinning (S $\approx$ 0.6 v. 0.2) and a ϕ_{CNL} close to its conduction band, resulting in small dipoles at both interfaces. Moreover, the typical metal to contact to TiO_2 is Ti because it is known to deplete the oxide of O atoms and thus dope it. Al has a more favorable work function for $\chi \approx 4$ eV semiconductors

($\chi \approx 4.1$ eV v. 4.3 eV). ZnO also dopes more heavily than TiO_2, a favorable effect not simulated here. The benefit of the interfacial layer diminishes at higher doping levels since the full Schottky barrier is already quite transparent due to thinness, negating the need for barrier height lowering.

While low-offset materials for n-type contacts have become well-established, a good p-type analogue remains to be found. Band line-ups suggest that NiO and $CuAlO_2$ are possible candidates [14] for contacts to p-type Si and $In_{0.53}Ga_{0.47}As$ (for THz HBTs). Favorable p-type alignments are much rarer than n-type because both χ and E_g affect the valence band offset, whereas matching χ alone is sufficient for the conduction band. Common experimental values for χ and E_g are 1.5 and 3.7 eV for NiO and 2.0 and 3.0 eV for $CuAlO_2$, respectively. Both possess a ϕ_{CNL} close to 1 eV above the valence band.

$CuAlO_2$ (Fig. 5) was found to be significantly more effective than NiO due to a negative valence band offset. As with the n-type case, the benefit diminishes at higher doping levels. Moreover, the $10^{-9} \ \Omega - cm^2$ target cannot be reached for Si at these doping levels, either with or without these oxides. Nickel-based silicides appear to provide superior performance [15]. Continuing improvement is seen with higher thicknesses due to the oxide potential drop lessening band bending. However, real structures would be limited by ohmic series resistance above 1-2 nm [9], particularly since these oxides have poor hole mobility due to d-orbitals comprising the valence bands.
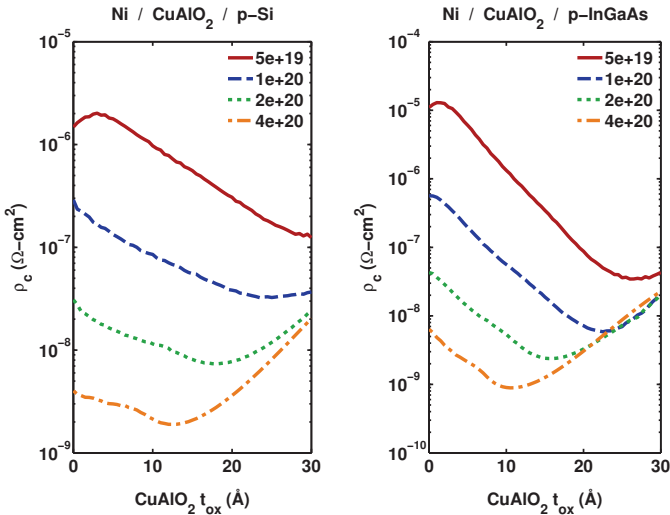

Fig. 5: Ni-CuAlO$_2$ contacts to p-Si and p-In$_{0.53}$Ga$_{0.47}$As. Target ρ_c of $10^{-9} \ \Omega - cm^2$ is not met. The continuing improvement seen with increasing layer thickness cannot be expected in real contacts due to ohmic resistance.

An alternative approach is to deliberately utilize pinning in an interfacial layer rather than attempt to depin. A useful candidate is InAs, which has a *negative* conduction band offset with metals. Given its low band gap and high dielectric constant, InAs is quite strongly pinned and inherently forms excellent n-type contacts that exhibit $\rho_c \approx 10^{-9} \ \Omega - cm^2$ without any additional layers or processing. A possible strategy is to use InAs itself as an interfacial layer for contacts to n-type III-Vs. InAs has mean free paths on the order of 100 nm at room temperature and thus conducts ballistically at the length scales needed here. This method was applied to n-GaAs early on with some success [16] and n-GaSb more recently [17].
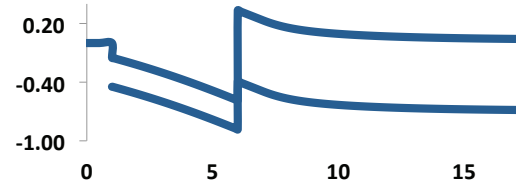

Fig. 6: Metal / InAs / n-GaSb band diagram.

The interface with III-Vs will form a heterojunction that resembles a Schottky barrier and presents a new resistance component. In the case of InAs-GaSb, a triangular potential well is formed at the interface due to the type III 'broken gap' line-up (Fig. 6). This creates resonant tunneling states which cause oscillations in the $T(\epsilon_\perp)$ spectrum (Fig. 8). Minima in $\rho_c(t_s)$ occur whenever a transmission resonance approaches the Fermi energy. The maximum barrier height is markedly lower with the formation of the heterojunction and decreases as InAs thickness increases. Nevertheless, Al-ZnO contacts show better results without the drawbacks of InAs due to the highly favorable band line-up with GaSb's $\chi \approx 4.07$ eV.
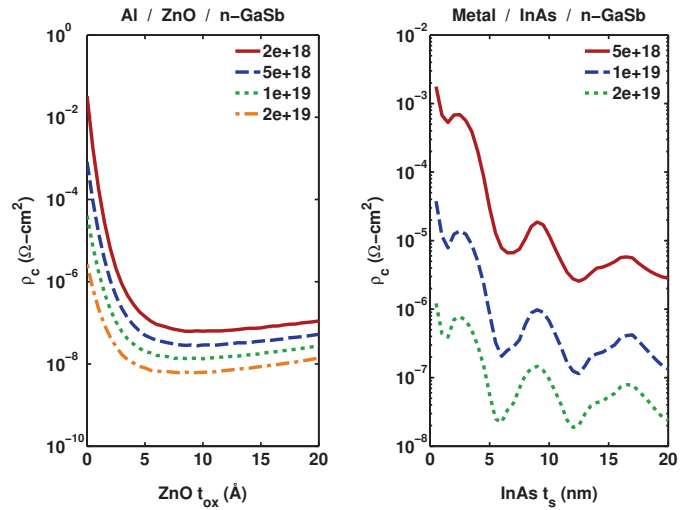

Fig. 7: Al-ZnO and InAs contacts to n-GaSb. Al/ZnO reduces ρ_c by several orders, but attainable doping levels in GaSb still cannot meet the requirement of $\rho_c \approx 10^{-9} \ \Omega - cm^2$. InAs shows poorer results and exhibits oscillations.
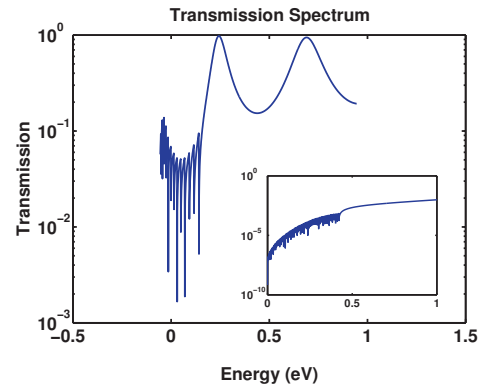

Fig. 8: Transmission spectrum for InAs contacts to GaSb shows a series of resonances. Inset shows the T(E) spectrum for a typical tunnel barrier contact.

In principle this abnormal behavior of ρ_c is not a detriment since contact interfaces ideally drop no potential during device operation. With precise thickness control, a resonance could reliably be placed near the Fermi energy to always maximize conductance. In practice such thickness sensitivity is highly problematic due to process variation. Another issue is that for momentum-conserving transport, the electrons injected from the GaSb L-valley would face a high barrier in InAs due to its high Γ-L separation. Therefore the calculated minima likely represent a lower bound on achievable ρ_c with this method. However, experimental achievement of $\sim$3 x 10^{-6} $\Omega - cm^2$ [17] at much lower doping (N_D = 5 x 10^{17} cm^{-3}) suggests that the method is viable even with incoherent transport.

One notable feature of InAs contacts is that the heterojunction barrier decreases in both height *and* thickness as doping increases, while Schottky barriers maintain their height. Furthermore, the heterojunction barrier could be removed entirely by grading the InAs composition with a suitable III-V source/drain. This novel approach also adds process complexity and there is little motivation to pursue it when a single thin layer of conductive oxides such as ZnO and TiO$_2$ offer comparable or superior results.

Dipoles at dielectric-dielectric interfaces in bilayers also modulate Schottky barrier height and offer another degree of freedom for ρ_c improvement. They have recently shown excellent results on n- and p-Si [18]. An Al$_2$O$_3$/TiO$_2$ structure on n-Ge is shown in Fig. 9, along with Al$_2$O$_3$/ZnO for comparison since ZnO was determined to be the best single layer choice. Using oxygen areal density arguments [19], the Al$_2$O$_3$/TiO$_2$ interface likely has a $\sim$0.4V dipole due to the migration of oxygen ions, leaving positive charge (oxygen vacancies) on the Al$_2$O$_3$ side. However, experimental data is not always consistent [13], likely due to the process-dependency of oxide densities. Thus we show the general case of varying magnitude and polarity to demonstrate the effect on ρ_c. The dipole's modulation of effective ϕ_M depends on its distance from the metal and must be calculated self-consistently with the induced semiconductor surface potential.

As seen in Fig. 9, a dipole can lower overall ρ_c initially but the tunneling resistance introduced by the additional dielectric quickly overtakes the tunneling resistance of the Schottky barrier. TiO$_2$ shows initial improvement with the introduction of Al$_2$O$_3$ while ZnO does not because its superior depinning ability diminishes the benefit of the dipole. At the doping levels necessary for 10^{-9} $\Omega - cm^2$, bilayers are a hindrance unless both constituents have low band offset. Such a materials stack has not yet been experimentally realized.

IV. CONCLUSION

We've simulated source/drain contacts to several materials with interfacial layers for Fermi level depinning. While initial work on this technique focused on conventional insulators, the transparent conducting oxides such as ZnO now seem to be the best performing materials due to their low band offsets. Our results show that the target ρ_c of 10^{-9} $\Omega - cm^2$ is within reach for n-Ge while n-GaSb requires advances in doping to meet future requirements. Direct Schottky contacts to these two materials currently have ρ_c on the order of 10^{-3} to 10^{-6} $\Omega - cm^2$ at best so improvements with depinning layers are large. The method shows smaller improvements for n-Si due to the high doping levels already achievable. Any benefit at all for p-Si remains to be seen.

Some possible complementary solutions for further ρ_c reduction include increasing leakage with intentional defects and electrical breakdown or further depinning with surface passivation treatments. Our work has focused on lowering contact resistivity by bringing up the transmission probability, but it is worth noting that to reach the ultimate lower limits, there is no substitute for higher doping [20].

REFERENCES

[1] D. Connelly et al., *Appl. Phys. Lett.*, 88, 012105 (2006).
[2] A. M. Roy et al., *Electron Dev. Lett.*, 31, 10, 1077 (2010).
[3] J. Robertson, *J. Vac. Sci. Technol. B*, 18, 1785 (2000).
[4] J. F. Wager et al., *J. Appl. Phys.*, 109, 094501 (2011).
[5] W. Mönch, *Phys. Rev. Lett.*, 58, 12601263 (1987).
[6] W. Mönch, *Rep. Prog. Phys.*, 53, 221 (1990).
[7] A. Agrawal et al., *Appl. Phys. Lett.*, 101, 042108 (2012).
[8] S. Gupta et al., *J. Appl. Phys.*, 113, 234505 (2013).
[9] A. Agrawal et al., *VLSI Symp.* (2013).
[10] J.-Y. Lin et al., *Electron Dev. Lett.*, 33, 11, 1541 (2012).
[11] P. Paramahans et al., *VLSI Symp.* (2012).
[12] Z. Yuan et al., *Appl. Phys. Lett.*, 98, 172106 (2011).
[13] J. Hu et al., *Appl. Phys. Lett.*, 99, 092107 (2011).
[14] J. Robertson et al., *Phys. Rev. B*, 83, 075205 (2011).
[15] N. Stavitski et al., *Electron Dev. Lett.*, 29, 4, 378 (2008).
[16] S. L. Wright et al., *Appl. Phys. Lett.*, 49, 1545 (1986).
[17] C. Lauer et al., *Semicond. Sci. Technol.* 21, 1274 (2006).
[18] K.-W. Ang et al., *IEDM*, 18.6.2 (2012).
[19] K. Kita et al., *Appl. Phys. Lett.*, 94, 132902 (2009).
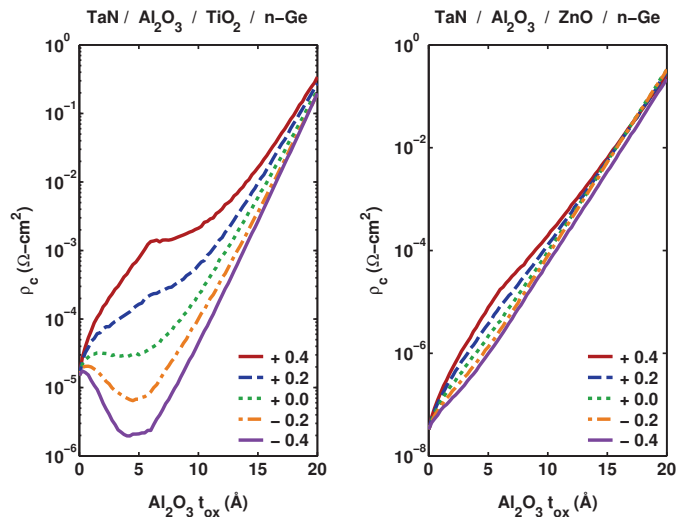[20] J. Maassen et al., *Appl. Phys. Lett.*, 102, 111605 (2013).

Fig. 9: Bilayer contacts to n-Ge with $N_D = 10^{20}$ cm^{-3}. A favorable dipole polarity counteracts the additional tunneling resistance in Al$_2$O$_3$/TiO$_2$. However, ZnO is effective enough by itself that Al$_2$O$_3$/ZnO is always worse.