Analysis of Si, InAs, and Si-InAs Tunnel Diodes and Tunnel FETs Using Different Transport Models

Andreas Schenk*, Reto Rhyner*, Mathieu Luisier†, and Cedric Bessire‡
*Integrated Systems Laboratory, ETH Zurich, CH-8092 Zürich, Switzerland
†Purdue University, 207 South Martin Jischke Drive, Indiana 47907, USA
‡IBM Research-Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland
Email: schenk@iis.ee.ethz.ch

***Abstract*—This paper presents a TCAD study on the performance of Si, InAs, and Si-InAs tunnel diodes and tunnel FETs. Comparative NEGF simulations of short InAs homo-diodes and experimental data on Si homo-diodes serve to calibrate the tunnel models for InAs and Si. Two workarounds for the case of Si-InAs hetero devices are found which give similar results. The crucial difference between in-junction and off-junction band-to-band tunneling is pointed out. Whereas the former cannot yield a sub-thermal slope, the latter can eventually produce a point slope of 25 mV/dec, albeit at extremely small current levels. The TCAD prediction for the maximum on-current of a Si-InAs hetero TFET is 3e-6 A/μm, about 3 orders of magnitude less than world-record CMOS.**

I. Introduction

Tunnel FETs (TFETs) are still considered as post-CMOC candidates with potential sub-thermal slope. Hetero structures like Si-InAs can be used to suppress ambipolarity. There is neither an analytical theory nor a TCAD model of band-to-band tunneling (BTBT) in hetero-junctions between a direct and an indirect semiconductor. In this paper it is shown that BTBT in TFETs is either in-junction or off-junction tunneling. When using calibrated models for the border materials of an Si-InAs TFET it turns out that the tunnel path is in either case contained in the small-gap material. This enables the qualified use of TCAD. To do so, a Kane model [1] available in Sentaurus-Device of Synopsys [2] is compared with ballistic tight-binding NEGF results for bulk-like InAs homo diodes. A remarkable agreement is found in the higher doping range. If this Kane model is combined in two different ways with a calibrated model for Si [3], similar IV characteristics are obtained with both methods in the case of Si-InAs hetero diodes as well as in the case of Si-InAs TFETs. The doping values at the InAs side of Si-InAs wire Esaki diodes produced at IBM Zurich [4] are determined by reverse modeling. Envisaged hetero wire TFETs based on these diodes are studied in detail varying certain design parameters. Their performance is contrasted with pure Si DG TFETs. The latter serve to study the circumstances under which a sub-thermal slope can occur. It is demonstrated that in-junction BTBT in Si TFETs can never give a slope steeper than 60 mV/dec because of the drop of the quasi Fermi levels over the depletion region. This is in contrast to simplified ballistic simulations [5] which predicted super-steep slopes due to an effective cut of the Boltzmann tail of the distribution function. Nevertheless, artificial field singularities or strong vertical built-in fields are shown to result in off-junction BTBT with a sub-thermal "point slope" minimum of 25 mV/dec. This is because of the sudden onset of BTBT when conduction and valence band edges overlap energetically beneath the drain-side gate corner [6]. Unfortunately, the current level over the voltage range where sub-thermal slope occurs is in the atto to pico Ampere/μm range and hence of no practical relevance.

II. Simulators and tunnel models

In the ballistic NEGF simulations of the bulk-like InAs homo diodes a nearest neighbor sp^{3}s* TB method with spin-orbit coupling is used (for details and further references see [7]). The InAs tight-binding parameters are taken from [8]. In the "*dynamic nonlocal path BTBT model*" of Sentaurus-Device (called "*Kane model*" here) a tunnel path across the hetero interface can either belong to a direct (zero-phonon) or to a phonon-assisted tunnel process. Combinations are excluded in accordance with the missing theoretical understanding of this case. Two workarounds are applied in this paper: (1) The Kane model for direct material is also used on the silicon side, fitted to experimental data of [9]. (2) The calibrated model for Si [2] (called "*Schenk model*" there – the same name will be used as shorthand term here) is also applied to the InAs side, but setting the phonon energy to a very small value and adjusting the pre-exponential factor for the best fit to the NEGF characteristics. The main difference then is a higher power of the field strength in the pre-exponential factor which results in a slightly different curvature in the generation branch of the IV curve. Another difference is the position of the electron-hole pair generation: With the Kane model the rates of electron and hole generation are separated by the tunnel length, with the Schenk model they coincide (at the position of maximum wave function overlap). It can be checked that the electrostatic impact of this difference is negligible even for the highest rates.

Both workarounds are justified by the belated observation that only very small parts of the tunnel path lie on the silicon side.

III. Simulation results

Fig. 1 shows the band edge profile at $V_{\rm GS}$ = 0 V and $V_{\rm DS}$ = 2 V of a pure Si DG TFET with a body thickness of 10 nm and

N_A = 3e18 cm^{-3} at the onset of in-junction BTBT. The high-resistivity n-side which is needed to transfer the gate voltage to the junction causes a huge tunnel length and, therefore, a tiny BTBT current. The zoom reveals that the energy of the optimal tunnel path (red arrow) lies a few k_BT above the crossing point of conduction band edge and electron quasi Fermi level.

Hence, the tunnel current is carried by thermally excited

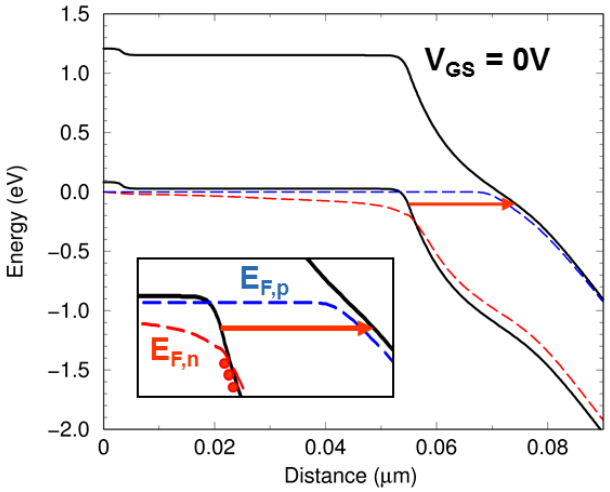

Fig. 1. Band edges and quasi Fermi levels in a Si DG TFET at the onset of in-junction BTBT. The red arrow marks the energy of the optimal tunnel path.

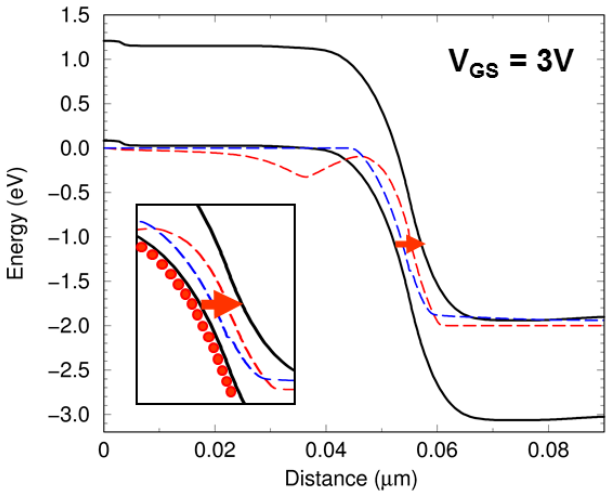

Fig. 2. At high V_{GS} the tunnel length is short and a thin triangular barrier determines the BTBT rate.

carriers, which prevents that the sub-threshold swing (SS) can fall short of the thermodynamic limit of 60 mV/dec. Here, the steepest slope becomes 65 mV/dec due to the excellent electrostatic control (not shown). At high V_{GS} (Fig. 2), the occupation factors are either 1 or 0 in the entire energy interval of significant tunneling rates. The curvature of the I_DV_G curve is determined by the BTBT probability through a thin triangular barrier. Fig. 3 demonstrates the difference between

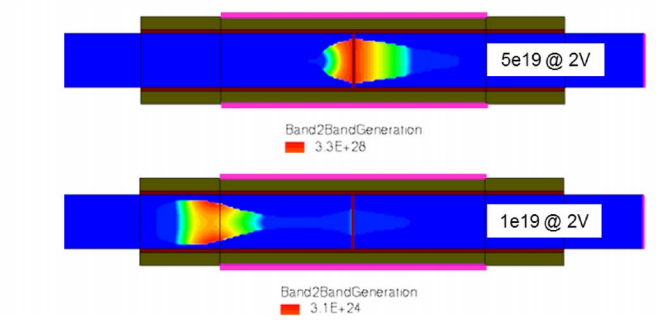

Fig. 3. BTBT rate distribution at different doping levels.

in-junction (upper panel) and off-junction BTBT (lower panel). A lowering of N_A from 5e19 cm^{-3} to 1e19 cm^{-3} moves the spot of the BTBT rate from the pn-junction to the drain-side gate corners where the electric field is now higher than in the junction. However, the maximum rate decreases by 4 orders of magnitude. Note the dark-space effect due to unavailable final states close to the oxide interfaces (non-local version of Schenk model used). As mentioned above, the low SS values obtainable with off-junction tunneling are due to the sudden onset of BTBT when conduction and valence band edges overlap energetically beneath the drain-side gate corner. It is important to note that this sharpness also depends on the assumption of sharp band edges. DOS tails caused by heavy doping would smooth the onset and hence degrade the point slope.

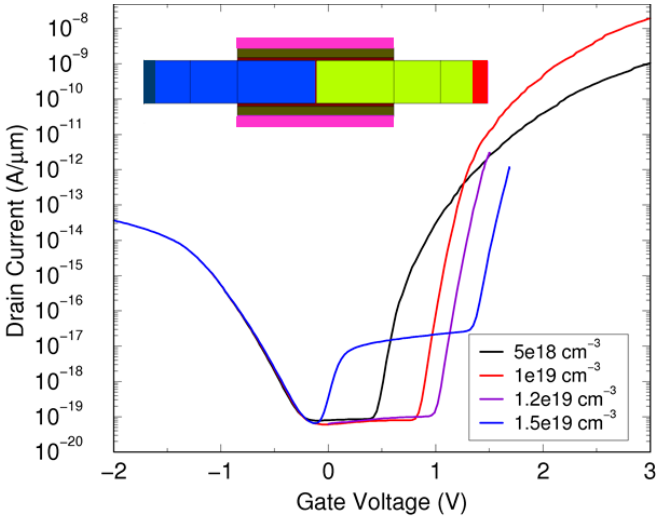

Fig. 4. Artificial increase of off-junction tunneling by etched gate corners for various doping concentration.

Off-junction tunneling can be artificially increased by e.g. etched gate corners (Fig. 4). The arising field singularities result in a best SS of 25 mV/dec (Fig. 5). The currents remain below the pico Ampere/μm level in the interval where the SS is sub-thermal.

Decreasing the tunnel length for in-junction tunneling by doping of the intrinsic region is limited from two reasons: (1)

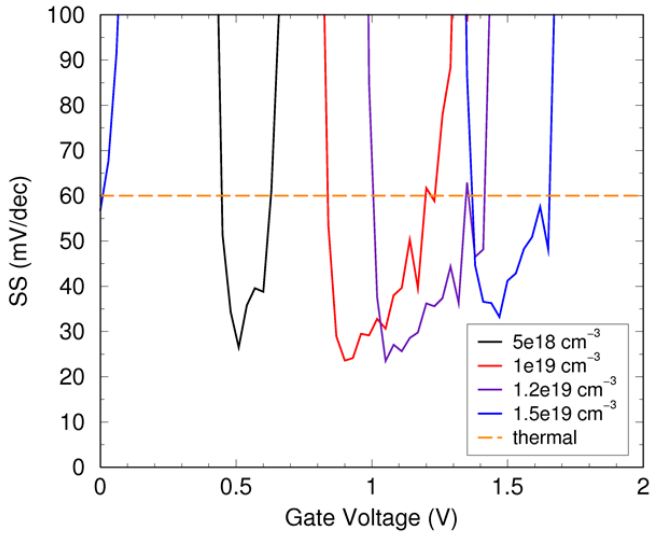

Fig. 5. The sub-threshold swing has a best value of 25 mV/dec independent of the doping concentration.

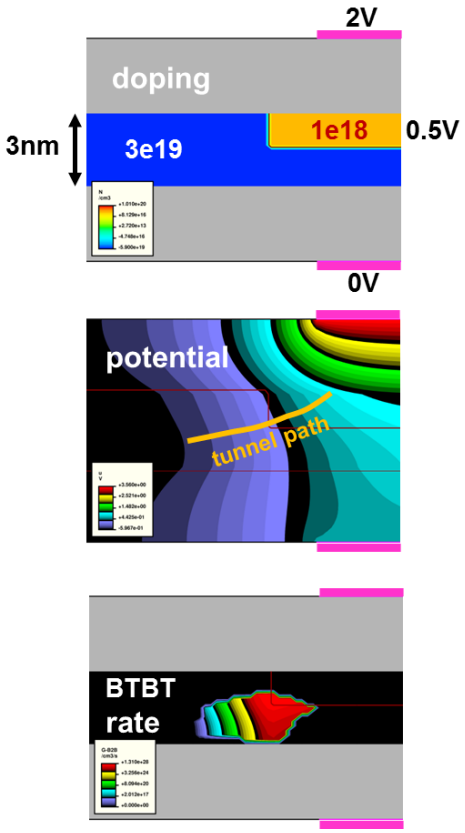

Fig. 6. Tunnel path in a Si DG TFET with ultra-thin body of 3 nm thickness, high-k gate oxides with 1 nm EOT, and asymmetrical gate voltages.

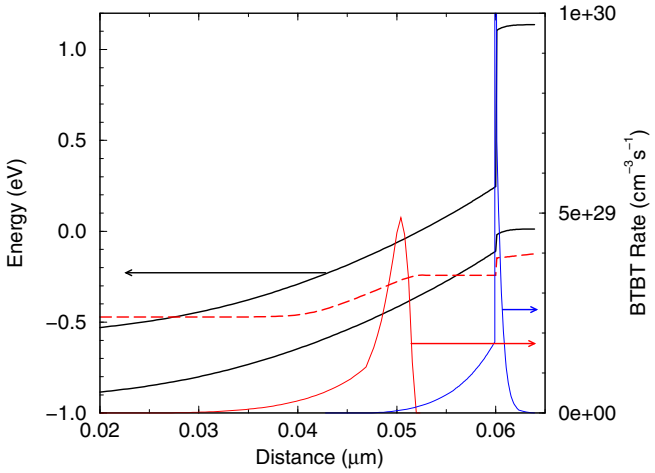

Fig. 7. Electron (red) and hole (blue) generation rates in the Si-InAs Esaki diode described in the text. Red dashed curve: electron quasi Fermi level.

It increases the off-current, and (2) it degrades the electrostatic gate control (transconductance). It has been verified that the intrinsic Si region can be doped up to 1e17 cm^{-3} without any visible effect. Fig. 6 illustrates the idea of limiting the tunnel length geometrically instead. This could be done by an ultra-thin body of 3 nm thickness and the application of asymmetrical gate voltages. (A tunnel length of 3 nm in Si is linked to a very high BTBT rate.) Furthermore, for a larger vertical component of the tunnel path a "pocket was implanted" to induce a strong vertical built-in field. However, the grounded source is always the reason why the tunnel path is forced to bend towards the source side. No advantage was found from such a structure.

Fig. 7 proves that in a Si-InAs diode the tunnel path is basically contained in the InAs region. The abrupt jump of the band gap leads to a delta-like peaked hole generation rate directly at the interface. Only with a doping level higher than 1e17 cm^{-3} in Si, the hole generation peak would start to move into the silicon region. However, the more symmetrical the pn-junction becomes, the higher the off-current which originates from BTBT at the Si-side gate corners. Therefore, a unified BTBT TCAD model for a direct and an indirect semiconductor is only needed for the unwanted case of a high off-current.

The comparison between ballistic NEGF and TCAD (Kane model and modified Schenk model) for ultra-short bulk-like symmetrically doped InAs Esaki diodes is shown in Fig. 8. The maximum length (for the lowest doping considered) is 80 nm which is at the memory limit. Surprisingly good agreement is obtained at least in the higher doping range, although the WKB approximation has long since broken down there. The NEGF simulations were performed on a cluster of 256 CPUs, and the current for one bias point was obtained in about 1 h.

The two workarounds were used to determine the doping values at the InAs side of Si-InAs wire Esaki diodes produced at IBM Zurich [4] by reverse modeling (Fig. 9). Boron doping in Si of 4e19 cm^{-3} and negligible series resistance were assumed. A series resistance would alter the $F_{\max}(V_{\text{appl}})$-dependence and flatten the generation curves. Fig. 10 shows simulated $I_D V_G$ curves of hypothetical Si-InAs hetero wire TFETs (p-type) with 60 nm diameter, 1 nm EOT, 4.05 eV work

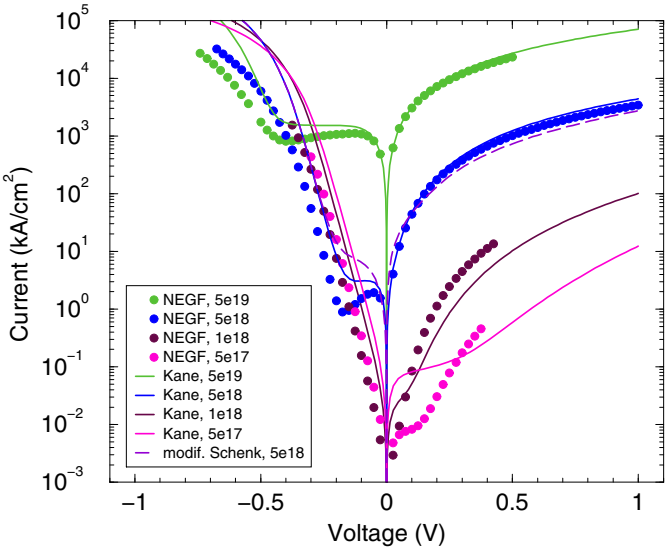

Fig. 8. Comparison between NEGF and TCAD models for ultra-short bulk-like symmetrically doped InAs Esaki diodes.

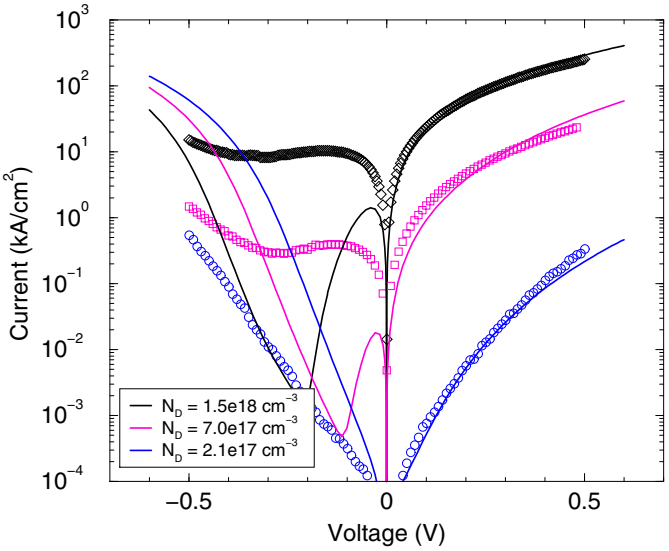

Fig. 9. Fitted doping values at the InAs side of Si-InAs wire Esaki diodes produced at IBM Zurich [4].

function, and 150 nm gate length for various doping levels on the InAs side. Here, workaround (1) is shown. The on-current is limited by (roughly) 3e−6 A/μm. The best SS is found to be 35 mV/dec for a doping of 5e18 cm^{-3} in InAs. However, here the current is of the order of attoAmp/μm only. Variation of the InAs doping level (not shown) results in an optimal doping value of 1e18 cm^{-3} which leads to a point slope of 42 mV/dec and a sub-thermal-slope interval with a maximum current of $\sim$ 5e−8 A/μm. Variation of the gate overlap (not shown) reveals that the best performance is obtained with zero gate overlap, resulting in a point slope of 25 mV/dec at 5e−9 A/μm and a sub-thermal-slope interval with a maximum current of $\sim$ 1e−7 A/μm.

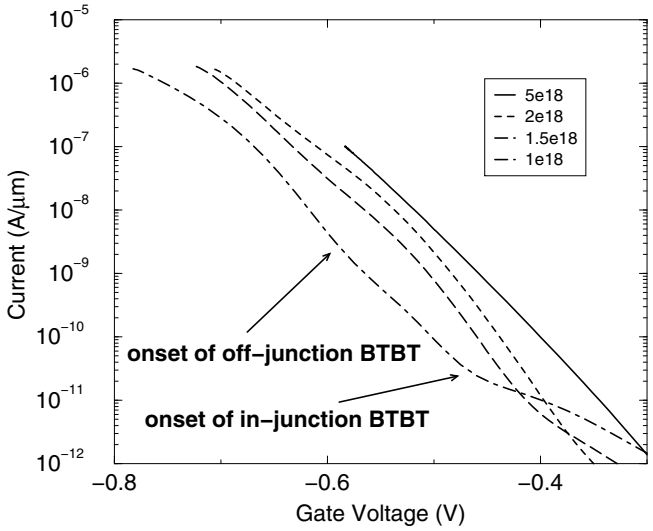

Fig. 10. Simulated $I_D V_G$ curves of the hypothetical Si-InAs hetero wire TFETs described in the text for various doping levels in InAs.

IV. Conclusion

When ramping the gate voltage of a TFET, there is a transition from in-junction (BTBT rate located within the pn-junction) to off-junction tunneling (BTBT rate located under the gate corners). At this transition the slope becomes minimal and can be further downsized to about 25 mV/dec by creating field singularities or by positioning the gate corner at the metallurgical junction (zero gate overlap). However, the corresponding currents are in the atto to pico Ampere/μm range only. In a Si-InAs TFET the tunnel path is basically contained in the InAs region. This is always true for off-junction tunneling, but also for in-junction tunneling as long as the pn-junction is highly asymmetrical (one-sided). It allows to use existing BTBT TCAD models. The on-current limitation of the Si-InAs TFET (3e−6 A/μm) is related to the principle upper boundary of the BTBT rate in a semiconductor.

Acknowledgment

This work was financially supported by the Swiss NanoTera project ENABLER and by Toshiba Corporation.

References

[1] E. O. Kane, J. Phys. Chem. Solids **12**, 181 (1959).
[2] Synopsys Inc., Sentaurus Device User Guide, Version 2009.06, Mountain View, California, (2009).
[3] A. Schenk, Solid-State Electronics **36** (1), 19 - 34 (1993).
[4] M. T. Björk, H. Schmid, C. D. Bessire, K. E. Moselund, H. Ghonheim, S. Karg, E. Lörtscher, and H. Riel, Appl. Phys. Lett. **97**, 163501 (2010).
[5] J. Knoch, S. Mantl, and J. Appenzeller, Solid-State Electronics **51**, 572 (2007).
[6] C. Hu, D. Chou, P. Patel, and A. Bowonder, *Proc. Int. Symp VLSI-TSA*, Apr. 2008, pp. 14 - 15.
[7] M. Luisier and G. Klimeck, J. Appl. Phys. **107**, 084507 (2010).
[8] G. Klimeck, R. C. Bowen, T. B. Boykin, and T. A. Cwikc, Superlattices and Microstructures **27** (5-6), 519 - 524 (May 2000).
[9] P. M. Solomon, J. Jopling, D. J. Frank, C D'Emic, O. Dokumaci, P. Ronsheim, and W. E. Haensch, J. Appl. Phys. **95** (10), 5800 (2004).