Analysis of InAs-Si Heterojunction Nanowire Tunnel FETs: Extreme Confinement vs. Bulk

Hamilton Carrillo-Nuñez, Mathieu Luisier, Andreas Schenk
Integrated Systems Laboratory ETH Zürich, Gloriastrasse 35, 8092 Zürich, Switzerland

Abstract
Extremely narrow and bulk-like p-type InAs-Si nanowire TFETs are studied using (i) a full-band and atomistic quantum transport simulator based on the $sp^3d^5s^1$ tight-binding model and (ii) a drift-diffusion TCAD tool. As (iii) option, a two-band model and the WKB approximation have been adapted to work in heterostructures through a careful choice of the imaginary dispersion. It is found that for ultra-scaled InAs-Si nanowire TFETs, the WKB approximation and the quantum transport results agree very well, suggesting that the former could be applied to larger hetero-TFET structures and considerably reduce the simulation time while keeping a high accuracy.

1. Introduction

Among the new generation of devices, band-to-band tunneling (BTBT) FETs (TFETs) are considered as very promising candidates on the road towards energy-efficient transistors [1]. However, the performance of TFETs is still limited, especially the ON-current remains very low. Heterostructures made of III-V/Si materials have been proposed to address this issue [2, 3]. Recently, InAs nanowires with diameters of few tens of nanometers have been grown on Si substrates [3]. Although these devices exhibit a good electrostatics control due to their gate-all-around configuration, they can still be considered as bulk-like. In case of InAs, it has been demonstrated that quantum confinement effects manifest themselves when the nanowire diameter does not exceed 20 nm [4]. In order to shed light on the impact of confinement, a realistic bulk-like and an extremely...
narrow InAs-Si nanowire TFET are studied in this paper. A TCAD tool and an atomistic tight-binding quantum transport simulator called OMEN [5] are employed for that purpose. OMEN is a multidimensional atomistic simulator based on a $sp^3d^5s^*$ tight-binding representation of the band structure of various semiconductors [5]. Quantum transport equations are solved within the Non-Equilibrium Green’s Function (NEGF) or Wave Function (WF) formalism together with Poisson’s equation self-consistently. OMEN has been successfully applied to the simulation of homo-TFETs with direct and phonon-assisted band-to-band tunneling [6]. Despite a high computational efficiency, the size of the structures OMEN can handle is very limited. This does not apply to drift-diffusion TCAD tools, e.g. the Sentaurus-Device (S-Device) simulator from Synopsys [7], which offer local and non-local analytical BTBT models.

2. Test device: bulk-like InAs-Si Esaki diode

For heterostructures as the bulk-like InAs-Si Esaki diode depicted in Fig. 1(a), the applicability of analytical BTBT models such as the “dynamical non-local path BTBT” (Kane’s model) [8] might be called into question because no analytical theory exists for BTBT between a direct (InAs) and an indirect (Si) semiconductor. A very simple work-around is either to consider a zero-phonon (direct BTBT) or a phonon-assisted (indirect BTBT) tunneling path through the interface of the two materials as required by S-Device. As will be discussed below, in an InAs-Si heterostructure the first case is more appropriate and can be achieved by using Kane’s model as given on the InAs side by the well-established values for gap and effective masses, but with calibrated parameters (called “A” and “B” in S-Device) for the silicon side using the experimental data of Solomon et al. [9]. Likewise, the analytical model developed for phonon-assisted tunneling in silicon [10] could be used on both sides provided that the phonon energy is set to a very small value on the InAs side and the critical field and prefactor are carefully adjusted to match the InAs BTBT rate [11]. Both variants give similar results, the first one being used here. It is labeled “TCAD, bulk parameters” in Fig. 2 because the parameters of the bulk materials are applied piece-wise. Note that the described complication only occurs for inter-material tunneling (“point tunneling”) but not for under-the-gate tunneling (“line tunneling”), as shown in Fig. 1. The latter has tunneling paths fully contained in InAs and dominates the ON-current in a bulk-like InAs-Si TFET as will
Assuming that inter-material tunneling only takes place between the Si valence band (VB) and the InAs conduction band (CB), BTBT in the InAs-Si heterostructure is expected to be mainly direct. This is illustrated in Fig. 2 where the ON-current density of a bulk-like InAs-Si Esaki diode (circular geometry, n-type InAs nanowire with radius of $R \approx 60 \text{ nm}$ grown on a p-type ⟨111⟩ Si substrate with $N_A = 10^{20} \text{ cm}^{-3}$) is plotted as a function of the donor concentration. The device dimensions and the acceptor concentration were adjusted according to experimental results reported in Ref. [3]. It has been verified that the tunneling current computed with OMEN does not depend on the presence of electron-phonon scattering, indicating that the inter-material tunneling through a InAs-Si junction is direct. It can also be observed in Fig. 2 that the ballistic S-Device current simulated by the aforementioned work-around (labeled “TCAD, bulk parameters”) agrees with
well with OMEN up to a concentration of $N_D \approx 4 \times 10^{19}$ cm$^{-3}$. For higher donor concentrations the pn-junction becomes one-sided. Therefore, the tunneling path more and more penetrates into Si leading to a decreasing tunnel current since the BTBT rate in bulk silicon is much smaller than in InAs. This behavior is not found with OMEN! The discrepancy is a clear indication of direct tunneling that can be included in S-Device using a “model semiconductor”, i.e. a homo-junction with a common gap called “tunnel gap” and a reduced effective mass built from the Si light-hole and the InAs $\Gamma$ mass. The Si light-hole mass along the $\langle 111 \rangle$-direction was extracted from OMEN. While the CB of the model semiconductor is clearly given by the InAs CB edge, the choice of the VB is somewhat ambiguous. The hole generation rate has the form of a delta-like peak directly at the interface, as a consequence of the large Si band gap. In the model of an abrupt heterointerface this means that a tunnel path from Si can start at any occupied energy level within the VB offset interval (compare Fig. 1). When varying the parameter “tunnel gap” between $E_{g,\text{InAs}}$
and $E_{g,\text{InAs}} - \Delta E_v$, the current at the highest doping concentration changes by less than a factor of 2. The reason for this relatively small change is that the transmission of the tunnel barrier is already close to 1 at the highest doping level. For the comparison in Fig. 2 the tunnel gap was chosen as the InAs gap reduced by half of the VB offset ($\Delta E_v = 80$ meV used here based on the Anderson rule). This results in the open symbols in Fig. 2. With the described approach the current does not drop anymore. Unfortunately, this method cannot be easily applied to hetero TFETs as it requires to identify three types of tunnel paths - bulk-InAs, bulk-Si, and inter-material - each with its own model. Finally, TCAD allows to include the measured mobility of the InAs wires as well as the extracted contact resistances. These parasitic components have a strong influence on the ON-current, as shown in Fig. 2 by the single simulation point compared with the experimental data.

*Imaginary dispersion models*

BTBT can be interpreted as the transfer of electrons between the valence and conduction bands of one (homo) or two (hetero) semiconductor(s). In such a process, the VB and CB edges are bridged within the forbidden energy gap by an imaginary dispersion. Fig. 3 shows the bulk imaginary dispersion of InAs and Si along the $\langle111\rangle$ crystal direction. The full-band imaginary dispersion is obtained by a $sp^3d^5s^*$ tight-binding calculation with OMEN. Different analytical models are compared to this reference curve. As expected for a direct small-gap semiconductor, all 2-band models, e.g. Kane[8] and Flietner[12] model, are equally suitable for InAs. The situation is different for Si where the one-band effective-mass-approximation (EMA) model fits much better its full-band dispersion. The Flietner imaginary dispersion has the following analytical form

$$\kappa = \sqrt{\frac{2m_cE_g}{h^2}} \frac{\kappa_c\kappa_v}{\kappa_c^2 + (1 - \alpha)\kappa_v^2},$$

where $m_v(c)$ and $\kappa_v(c) = \sqrt{2m_{v(c)}|E - E_{v(c)}|/h^2}$ are the hole (electron) effective mass and the one-band imaginary dispersion for the valence (conduction) band, respectively. The variable $E_g = E_c - E_v$ is the energy gap and the parameter $\alpha$ is defined as $1 - \sqrt{m_c/m_v}$. The Flietner model can be extended to the case of inter-material tunneling processes by treating the effective masses and
energy gaps as position-dependent quantities, i.e. \( m \rightarrow m(x, y, z) \) and \( E_g \rightarrow E_g(x,y,z) \), respectively.

The Flietner two-band model for heterostructures is first tested on a bulk-like InAs-Si Esaki diode. The current can be computed by the Landauer formula once the transmission probability is known. The transmission, or equivalently, the BTBT rate, is derived from the solution of the envelope equation

\[
\left[ E_{k_z} - \frac{\hbar^2}{2m_0} \frac{\partial^2}{\partial x^2} \right] \phi_{k_z}(x) = \frac{E_g (E - U_c) (E - U_v)}{\sqrt{m_0/m_c} (E - U_v) - \sqrt{m_0/m_v} (E - U_c)^2} \phi_{k_z}(x),
\]

where the energy expression on the right-hand-side (r.h.s) is \( \hbar^2 k^2 / 2m_0 \), \( \kappa \) is defined in Eq. (1), and \( U_{v(c)} = E_{v(c)} + U_e \) represents the position-dependent conduction (valence) band edge due to the electrostatic potential \( U_e(x) \). The energy related to the transverse modes is given by \( E_{k_z} = \hbar^2 k_z^2 / 2m_0 \) with the free electron mass \( m_0 \). The coordinate \( x \) on the material parameters has been omitted for brevity.

Although Eq. (2) has not been derived rigorously from an envelope function method, it is straightforward to show that it reduces to the one-electron envelope equation with the correct effective mass \( m_{e(v)} \) in the limit where the electron energy approaches the conduction (valence) band edge, i.e. \( E \rightarrow U_{c(v)} \). Furthermore, Eq. (2) contains non-parabolic corrections for the conduction
Figure 4: (a) OMEN and Flietner model $I_D - V_D$ characteristics of a bulk-like InAs-Si Esaki diode. (b) The spectral currents computed with OMEN and the Flietner model are compared for $V_D = 0.375 \text{ V}$. The dashed blue curves are the hole (left) and electron (right) generation rates extracted from Eq. (3). The solid lines are the conduction (top) and valence (bottom) band edge profiles obtained with OMEN.

and valence bands. In the case of InAs, the full-band structure is well reproduced up to 400 meV above the CB minimum and up to 150 meV below the valence band maximum. If the effective masses $m_c$ and $m_v$ are equal the two-band imaginary dispersion used for carbon nanotube [14] and graphene nanoribbons FETs [15] is obtained. Hence, Eq. (2) ensures that electrons are well described by their appropriate envelope equation on the real branch of the band dispersion, but also on the imaginary branch where the hole and electron envelope equations are coupled.

Following [16], the BTBT is computed by numerically solving Eq. (2) via a recursive Green’s function approach with open boundary conditions for each given transverse energy $E_{k_\perp}$ [17]. Then, by replacing the $E$-variable with the position based on the semi-classical expression of the BTBT current, $I_{x-c} = q \int GdV$, and by using the Landauer formula, a relation between the generation rate $G$ and the transmission probability $T(U_{v(c)})$ can be established:

$$G_{v(c)} = \frac{1}{\pi \hbar A} \sum_{k_\parallel} T(k_\perp, U_{v(c)}) \{ f_L(U_{v(c)}) - f_R(U_{v(c)}) \} \nabla_x U_e,$$

with $A$ being the cross section area perpendicular to the transport direction, i.e. the $x$-direction.

Fig. 4 compares the $I_D - V_D$ characteristics and the spectral current of a bulk-like InAs-Si Esaki diode computed with OMEN and using the Flietner model. Both results agree well over the entire $V_D$ range shown here. However, by implementing and using the Flietner model the
computational cost associated with an atomistic tool such as OMEN significantly decreases as long as the appropriate band edge profiles $U_c$ and $U_v$ are available to solve Eq. (2). Fig. 4(b) shows that the hole and electron generation, calculated as in Eq. (3) with the Flietner model, mainly occur at the heterojunction interface.

In an extremely narrow nanowire TFET, quantum confinement occurs and a ladder of sub-band energies appears due to energy quantization. The full-band structure of this device reveals that the lowest conduction and highest valence sub-band edges are connected by an imaginary band. In the presence of an external electric field, the sub-band edges can be expressed as three-dimensional (3D) position-dependent quantities $U_{v(c)}(r)$ due to the non-uniform 3D electrostatic potential $U_e(r)$. Solving a 3D version of Eq. (2) and its corresponding quantum transport problem is very time consuming. Alternatively, the position-dependent imaginary dispersion in Eq. (1) can be used to evaluate a modified WKB transmission probability according to

$$T_{WKB}(E) = \frac{1}{A_{NW}} \int_{A_{NW}} d^2 r_\perp \exp \left\{ -2 \int \kappa(x, r_\perp) dx \right\},$$

where $A_{NW}$ is the cross section area of the nanowire and $r_\perp = (y, z)$. Note that in Eq. (4) the transmission is averaged over the nanowire cross section. Finally, Eq. (4) has been used to compute the BTBT transmission probability in narrow InAs-Si nanowire TFETs.

3. InAs-Si TFET simulations

In Fig. 5(a) a schematic of the IBM p-TFET studied in this work is presented: a $n^+-$InAs nanowire (source) is grown on a $\langle 111 \rangle$ $p^+-$Si layer (drain). This realistic device can still be considered as bulk-like since the nanowire diameter exceeds 100 nm on both the InAs and Si sides, with a total length of 250 nm. The doping concentrations have been estimated to be $N_A = 5 \times 10^{19}$ cm$^{-3}$ in Si and $N_D = 5 \times 10^{17}$ cm$^{-3}$ in InAs. The ultra-scaled TFET depicted in Fig. 5(b) has a diameter equal to 3.5 nm and a total length of 60 nm. In this case higher doping concentrations, such as $N_A = 2 \times 10^{20}$ cm$^{-3}$ and $N_D = 10^{19}$ cm$^{-3}$, are necessary for flat potential at the contacts to avoid artifact effects in the Poisson-Schrödinger self-consistency solution and to achieve a lower ON-set voltage for the point tunneling. Band gap narrowing (BGN) in partially depleted regions with geometrical and channel confinement constitutes a physical problem beyond the scope of
this study. Instead of using a simplistic model as function of doping, BGN was disregarded in all regions.

In both devices, the gate is all-around the channel composed of an intrinsic Si region with an overlap covering a part of the InAs side. The role of an intrinsic Si channel is to widen the tunneling barrier in order to achieve low OFF-state currents and hence to minimize the ambipolar behavior. The effective oxide thickness is set to 3.5 nm for the bulk device and 0.46 nm for the extremely narrow device with the same oxide permittivity $\epsilon_{\text{high}} = 9$.

Device simulations are performed with two different TCAD tools: S-Device to solve the transport problem in the bulk device and OMEN to take into account the atomistic and full-band structure of the InAs-Si nanowire. Once the quantum transport problem has been solved with OMEN, in a post-processing step, the converged electrostatic potential is used as an input for Eq. (1) and Eq. (4) to compute the BTBT current based on the WKB approximation and the Flietner 2-band model for the imaginary dispersion. Since the band-structure is no longer bulk-like, the electron effective masses must be adjusted according to the tight-binding calculations. For a nanowire with a diameter equal to 3.5 nm, the effective masses are: $m_v = 0.095 m_0$ for Si and $m_v = m_c = 0.058 m_0$ for InAs. Similarly, the energy gap increases due to the strong confinement and it is found to be $E_g = 1.41$ eV and $E_g = 0.98$ eV for Si and InAs, respectively. These material parameters agree very well with those reported in Ref [4].

Fig. 6 shows the $I_D - V_G$ characteristics and the BTBT transmission probability computed using OMEN and the WKB approximation for an extremely narrow InAs-Si nanowire p-TFET.
Figure 6: (a) OMEN and WKB $I_D$ vs $V_G$ characteristics of an InAs-Si nanowire p-TFET with drain-to-source voltage $V_{DS} = -1$ V. The diameter is 3.5 nm and no gate overlap is considered. Transmission probabilities computed with OMEN and the WKB approximation are shown for (b) $V_G = -0.50$ V and (c) $V_G = -0.80$ V. The imaginary dispersion is described by the Flietner model for heterostructures.

with no gate overlap. It can be seen that by combining the WKB approximation and Flietner model the OMEN results are reproduced with a great accuracy. As mentioned earlier, this requires the knowledge of the appropriate electrostatic potential and material parameters. In Fig. 7, the $I_D$ vs $V_G$ characteristics of the two devices considered in this work are plotted, with and without gate overlap on the InAs side. Comparing Figs. 7(a) and 7(b), it appears that the ultra-scaled TFET exhibits a behavior opposite to the bulk-like device; the ON-current of the device with overlap is lower in Fig. 7(a) whereas it is higher in Fig. 7(b). Moreover, Fig. 7(a) shows that in the case of the extremely narrow nanowire TFET with gate overlap the current can be less than $10^{-14}$ A, suggesting that there is less than one electron per second flowing through the device. This can be attributed to the fact that at very low gate voltages, electrons must tunnel through a higher and wider potential barrier to transfer from InAs to Si, as compared to the case without gate overlap. Therefore, only few electrons reach the InAs CB and the OFF-current is significantly reduced in extremely narrow InAs-Si nanowire TFETs. In the bulk-like TFET with a gate overlap region, there is an additional tunneling path perpendicular to the transport direction ("line tunneling"), as shown in Fig. 8, where the generation rate in a real-size InAs-Si p-TFET without and with gate
overlap is reported. Fig. 8(c) shows that a high BTBT generation occurs under the gate overlap region at $V_{GS} = -0.5\, V$ and $V_{DS} = -1\, V$. Because these tunneling components are dominant in the ON-state, the ON-current is higher, as illustrated in Fig. 7(b). However, in strongly confined systems, the gate voltage required to align the highest valence and lowest conduction sub-bands under the gate overlap is much higher than in bulk systems[19]. This inversion condition is never reached here and only BTBT through the InAs-Si heterojunction (“point tunneling”) contributes to the current in Fig. 7(a). This can be further seen in the BTBT generation rate computed from Eqs. (3)-(4) and shown in Fig. 9(a): at low gate voltages BTBT is due to electron transitions between the Si and InAs sides. At higher gate voltages in Fig. 9(b), BTBT is mainly located on the InAs side. No BTBT perpendicular to the gate under the overlap region can be identified.
Figure 8: Generation rate distribution in a real-size InAs-Si p-TFET shown in Fig. 5 without (a) and with (b)-(c) gate overlap. Similar results are reported in Ref. [20].

4. Conclusion

We have simulated ultra-scaled and realistic InAs-Si hetero-TFETs. Our findings suggest that BTBT is mainly direct in this material combination. We have shown that by providing the proper material parameters and electrostatic potential, the Flietner dispersion model, together with the WKB approximation, offers an accurate way to study extremely narrow InAs-Si nanowire TFETs. The material parameters have been extracted from the nanowire band-structure obtained from a $sp^3d^5s^1$ tight-binding approach.

It has also been found that the gate overlap plays an important role in large-scale InAs-Si TFETs by allowing line tunneling under the gate contact, while it has no influence in extremely narrow devices. Based on the results presented here, we anticipate that the electrostatic potential calculated by a TCAD tool such as S-Device could be combined with the material parameters extracted from full-band atomistic simulations to investigate large InAs-Si hetero-TFETs with a
Figure 9: Generation rate distribution in an extremely narrow InAs-Si p-TFET with 10 nm gate overlap at (a) $V_G = -0.35 \text{ V}$ and (b) $V_G = -0.55 \text{ V}$ with an applied $V_{DS} = -1.0 \text{ V}$. The electrostatic potential obtained from OMEN is used as input in Eq. (3). Horizontal and vertical white lines delimit the nanowire/oxide and hetero-junction interfaces, respectively. The top and bottom red lines show the gate location.

5. Acknowledgment

Funding from the European Community’s Seventh Framework Programme under grant agreement No. 619509 (Project E$^2$SWITCH) is acknowledged.

References


