# First-Principles Simulations of Nanoscale Transistors

Anders Blom, Umberto Martinez Pozzoni, Troels Markussen, and Kurt Stokbro QuantumWise A/S Copenhagen, Denmark anders.blom@quantumwise.com

Abstract—We describe the transport characteristics of a 50 nm (gate length) 2D InAs tunnel field-effect n-i-n transistor in a double-gate fin-like geometry (fin width 2.3 nm) by means of atomic-scale simulations. In particular, we compare results from density functional theory (DFT) using the Meta-GGA exchange correlation potential with those from a tight-binding Hamiltonian. For the first time we show that the two methods give comparable results, proving the predictive power of atomic-scale simulations for this type of devices, and that it is possible to accurately study realistic ultrascaled devices with first-principles methods.

# Keywords—density-functional theory; transistor; simulation; atomistic; first-principles; ultrascaled device; FinFET

### I. INTRODUCTION

The principle method to achieve higher performance in electronic devices has for a very long time been to scale down the feature sizes. Although there is a natural atomic limit to this pathway, which soon will become a significant issue, each node in the development of faster, more energy-efficient and tighter packed transistors is still marked by an ever-decreasing gate length and pitch size.

Now, as transistors evolve, so do the tools used to model them. In previous work [1] we have shown how an ultrascaled FinFET-like InAs double-gate structure could be modeled using first-principles methods (DFT). Here we will revisit the same structure, more or less, but with several updates: the gate length is considerably longer (50 nm vs. 16 nm), the fin is two times wider, and we are using a much more robust DFT model that gives a correct band gap without resorting to crude approximations. In addition, we will consider a more accurate way of treating the dangling bonds at the surface, and the doping can now be extended farther into the channel (and if desired, also under the gate). All these improvements are naturally designed to make the simulations correspond closer to actual experimental devices, and are enabled by recent developments in the latest version of our software package Atomistix ToolKit (ATK) [2].

After a presentation of the methodology, results for the DFT models will be shown, including comparisons with tightbinding models, with comments on the similarities and differences.

# II. METHODOLOGY

#### A. Accurate InAs Band Structure from DFT

For the main part of this study we have chosen to work with DFT to describe the electronic structure of the materials. While significant progress has been made in simulating nanoscale devices using tight-binding (TB) models, these methods have clear disadvantages which DFT can improve on. Most notably, DFT can quite generally be applied to structures containing multiple chemical elements, including explicit dopants like boron and phosphorous. TB models also have some advantage in calculation speed over DFT, but as we will see below, the actual bottleneck lies in the electrostatic model.

A traditional claim is that DFT cannot accurately describe the band gap of semiconductors, but with the introduction of modern exchange-correlation functionals this is no longer the case. Using hybrid functionals, which often give good band gaps, would however increase the computational burden to a point where it's completely impractical for relevant device sizes. Instead, we employ a Meta-GGA (MGGA) exchangecorrelation potential TB09 [3], which incurs a minor penalty on the simulation time, and yet is able to accurately describe the band structures of many semiconductor materials.

The TB09 potential contains a fitting parameter c, which in principle can be self-consistently determined in the calculation, via Eq. (3) in Ref. 3. However, that expression diverges if the system contains vacuum, where the electron density goes to zero. Therefore, we tune c to obtain the correct band gap for the bulk crystal (InAs in our case) and then use a fixed value for confined structures such as fins and nanowires.

By this procedure, we can obtain a band structure of bulk InAs that exactly matches the "recommended" direct gap 0.42 eV [4]. Notably, at the same value of *c*, also the electron effective mass at the  $\Gamma$  point comes out very close to  $0.026m_e$ . The indirect gaps  $\Gamma$ -L and  $\Gamma$ -X predicted by our calculation are 1.2 and 1.9 eV, respectively. This is in fair agreement with Ref. 4, however that doesn't really mean much; in fact, there are no experimental references for these values. A well-cited reference (often referred to as "experimental") is Adachi [5] who gives the two gaps as 1.07 and 1.37 eV, but if one traces the origin of these numbers back to its source, it turns out they are results obtained from a tight-binding simulation performed by in 1982 [6].



Fig. 1. Band structure of InAs, computed with TB09 MGGA, using c=0.936 and a lattice constant of 6.0583 Å. The basis set was here DoubleZetaPolarized, thus a larger set than used for the main calculations in this work, but by adjusting c at least the direct gap can be adjusted to fit the experimental value for each basis set and pseudopotential.

More interesting is to look at what happens when we confine the InAs crystal to effectively 2D (a slab, or a "fin") geometry. In the effective mass model, the confinement effect on the effective mass is typically ignored, and the gap is taken to scale as

$$E_g = \frac{\hbar^2}{2m^*} \left( k_x^2 + k_y^2 + \frac{\pi^2}{W^2} \right), \tag{1}$$

where W is the effective width of the slab. However, this assumes that the band structure is parabolic,  $E = \hbar^2 k^2 / 2m^*$ , which is far from the case in InAs, and it also means that the effective mass of the slab just attains the bulk value  $m^*$ . If instead we use a non-parabolic dispersion  $E(1+\alpha E) = \hbar^2 k^2 / 2m^*$ , we arrive at another relation for the band gap of the 2D slab:

$$E_{g,2D} = \frac{-1 + \sqrt{1 + \frac{2\alpha\hbar^2}{m^*} \left(k_x^2 + k_y^2 + \frac{\pi^2}{W^2}\right)}}{2\alpha}, \qquad (2)$$

and correspondingly for the effective mass

$$m_{2D} = m^* \sqrt{1 + \frac{2\alpha\hbar^2 \pi^2}{m^* W^2}} \,. \tag{3}$$

In these expressions,  $m^*$  is still the bulk effective mass, and  $\alpha$  is also fitted using the bulk band structure. However, the 2D effective mass now scales parametrically with W. We can compare these expressions to the band gap and effective mass for explicit atomistic models of confined systems, computed with the same DFT model as used for bulk InAs above. The results are shown in Fig. 2.

The surface of an unpassivated nanocrystal consists of dangling bonds which will introduce gap states. A common trick for eliminating these is to passivate the surface with hydrogen. For polar materials like III-V binaries and ternary alloys, it is however necessary to use "pseudo-hydrogen" atoms with a fractional pseudopotential core charge [7], in order to shift the surface states away from the Fermi energy. If



Fig. 2. The band gap and effective mass of the non-parabolic model from (1) and (2), compared to a parabolic effective mass model for InAs, using the values  $m^*=0.023$  and 3.13 eV<sup>-1</sup>.

a surface atom has *m* valence electrons, this atom will provide m/4 electrons to each of its four bonds in a tetrahedral  $(sp^3)$  crystal. To pair these m/4 electrons in each dangling bond, a passivating agent should provide (8-m)/4 additional electrons. Since As has m=5, it should be passivated with pseudo-hydrogen containing 0.75 electrons. Likewise, In has m=3 valence electrons, and should therefore be passivated with pseudo-hydrogen with 1.25 electrons. Alternatively, we can assign the H atoms binding to In (As) a permanent surplus (deficit) of 0.25 electrons in the self-consistent model; this approach also works in tight-binding.

## B. InAs FinFET Device Model

The 2D fin-like model used for the device simulations is shown in Fig. 3. The system consists of *n*-doped electrodes with an intrinsic region in between. It has a width of 2.3 nm and a total length of 50 nm. In the central part of the system we add a 10 nm long top and bottom electrostatic gate.

The electrodes and electrode extensions are doped with an effective density of  $5 \cdot 10^{19}$  cm<sup>-3</sup>. The doping is modeled by adding a small positive (compensation) charge of 0.00139e at each In and As atom in the doped regions. At the same time, a negative charge corresponding to the sum of all atomic compensation charges (with the opposite sign) is added to the

electron density. This makes it possible to model any doping concentration without including explicit dopant atoms (of course this means that dopant scattering is not accounted for), which would require a very large simulation volume.



Fig. 3. The InAs device structure used in the simulations.

There are in total 2,300 atoms in the scattering region (compared to just over 800 atoms in Ref. 1), and we use a numerical  $sp^3d^5$  LCAO basis set (Ref. 1 used only  $sp^3$ ). For comparison, we will also run the same system using a Slater-Koster TB  $sp^3d^5s^*$  parameter set [8]. All calculations are spin unpolarized and spin-orbit coupling is not included. We use 7 k-points in the transverse direction for the self-consistent electron density calculations, and 21 k-points for the transmission calculations ion. For all simulations, we use the NEGF+DFT (and TB) model as implemented in Atomistix ToolKit version 2015 [2].

The gate is separated from the InAs by a 1 nm thick dielectric region with dielectric constant  $3.8\epsilon_0$ . In order to obtain a self-consistent solution, the electrostatic problem is set up with Neumann boundary conditions (BC) in the direction perpendicular to the slab (B axis), periodic BC in-plane (A axis), perpendicular to the transport direction, and with Dirichlet BC in the transport direction (C axis) at the interface between central region and electrodes.

The Poisson equation is solved with a direct real-space solver which is parallel in memory, and as it turns out, this is really the bottleneck for the calculations. It would be quite possible to go to a larger number of atoms (a wider slab or longer gate length), but the memory requirement for the direct solver is quite large. For simple tight-binding models this is much less of an issue, since the electrostatic grids used there are about 10-20 times coarser (in each of the 3 dimensions of space) than for DFT, so with the first-principles model we are dealing with 1,000–10,000 times more grid points.

In spite of the hydrogen atoms preventing surface states, as described above, we still observe a significant charge transfer from the In/As atoms to the surface H atoms (Fig. 4, left). This charge transfer leads to the formation of a surface dipole, which results in a pronounced potential jump of about 1.3 eV at the InAs surface (Fig 4, right). We therefore adjust the gate electrode potential such that  $V_g = -1.3$  V corresponds to "zero". This is equivalent to adjusting the metal work function, as is done in other simulations.

#### III. RESULTS

We now apply a source-drain bias voltage of  $V_b = 0.2$  V and calculate the current *I* vs. the gate voltage  $V_g$ . Fig. 5 shows the *I*- $V_g$  curves obtained with ATK-DFT and ATK-TB, together with TB results from OMEN [9].



Fig. 5. Current vs. gate voltage at a source-drain voltage of 0.2 V. The subthredshold swing (SS) is shown in the plot for each method.

The three calculations give very similar results in the offstate. The two TB calculations agree on a quantitative level across all gate voltages. In particular, they predict almost the same subthreshold swing (SS): 98 mV/dec. with ATK and 103 mV/dec. with OMEN. This agreement might be expected, since the InAs TB-parameters are the same [8]. However, the electrostatic problem is treated very differently: ATK uses charges at the atoms, which are smeared out around each atom, while OMEN uses point charges. Another significant differrence is the treatment of the InAs surfaces, where OMEN increases the energy of dangling-bond states, instead of attaching hydrogen atoms. Given the good agreement between the two results, we may conclude that both passivation schemes work equally well.

Comparing the DFT results with the TB calculations we note that the results in the off-state agree rather well, although DFT-MGGA predicts a lower SS of 82 mV/dec. In the onstate, DFT-MGGA gives currents which are essentially the same as the TB results. While ATK-TB agrees very well with OMEN in the off-state, we note that slightly lower on-state currents are obtained with ATK. This might be related to differences in the treatment of the electrostatic potential and the use of Dirichlet BC in the transport direction.

The difference between the DFT and TB results is most likely primarily rooted in small differences in the electrode band structures, as indicated by Fig. 6. It is evident that DFT predicts a larger conduction band effective mass  $(0.107m_e)$  than the TB model  $(0.080m_e)$  does. This larger DFT effective mass results in a smaller tunneling probability through the barrier, which is the dominating process in the off-state. We believe the DFT result to be more reliable, since the TB model has much fewer degrees of freedom to describe the changes to the electron structure that take place in a confined structure, compared to bulk for which the TB model was fitted.

Another reason for the difference between DFT and TBand indeed between the two TB approaches-might be found in the electrostatic difference potentials. Fig. 7 shows the potentials in the off-state (left) and on-state (right) as calculated with DFT-MGGA (black) and TB (red). In the off-



Fig 6. Electrode band structures calculated with DFT-MGGA (black) and TB (dashed red). The band structures have been shifted such that the conduction band minimum is at zero energy.

state, both potentials are well converged with respect to the electrode length, since the potentials are flat (apart from the rapid fluctuations, which are due to the charge transfer taking place between In and As atoms). However, the TB potential barrier appears to be significantly higher than the DFT potential. In the on-state (right panel), the potentials are actually not well converged, since there are clear slopes close to the electrodes (at C=0 and C=1). This is a consequence of the Dirichlet BC condition, discussed above. In this case, Neumann BC in the transport direction would be more appropriate. We note that the DFT potential has larger slopes close to especially the left edge, than what is observed for TB. This might be an explanation of the smaller on-current in DFT.

Both the DFT and TB calculations predict a relatively large SS. The reason for this is that the off-state current is dominated by tunneling through the potential barrier instead of thermionic ("over the barrier") emission. This is illustrated in Fig. 8 obtained for TB calculations. The left panel again shows the electrostatic potential along the transport direction, while the right panel shows the spectral current. The energy scale of the spectral current plot has been aligned such that the left Fermi level is aligned with the electrostatic potential of the left electrode. It is clearly seen that the current is dominated by electrons with energies below the barrier top–i.e. by tunneling electrons. The large tunneling current is a consequence of the small electron mass of InAs. Since the effective mass decreases with increasing slab width (due to the non-parabolic bulk InAs



Fig. 7. Electrostatic (difference) potential in the center of the unit cell along the transport direction. The left panel show the off-state potential, while the right panel show the on-state potential. In both panels, the DFT potential is shown as black, while the TB potential is shown as red.



Fig. 8. Off-state tunneling. The left panel shows the electrostatic (Hartree) potential along the transport direction for a bias voltage of 0.2 V. The right panel shows the spectral current aligned such that the left Fermi energy corresponds to electrostatic potential of the left electrode.

conduction band), the tunneling current might increase if the slab width is increased. On the other hand, an InAs nanowire with quantum confinement in two directions would have a larger mass, and decreased tunneling can be expected.

As also discussed above, it it likely that the smaller SS obtained with DFT can be explained by the larger effective mass, resulting in a lower tunneling current.

### IV. CONCLUSIONS

We have shown results for InAs device calculations performed with DFT-MGGA and with TB. The DFT-MGGA is fitted to reproduce the bulk conduction band effective mass. The two approaches are generally in very good agree, with only quantitative differences in the  $I-V_g$  curves, which are most likely due to different effective conduction band masses of the confined InAs slabs. In the on-state, the electrostatic potential is not properly described with the Dirichlet BCs and can be another cause of differences between DFT and TB currents. Relatively large subthredshold swing (SS) is found with both DFT (SS=82 mV/dec.) and with TB (98 mV/dec.).

#### ACKNOWLEDGMENT

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement III-V-MOS Project n619326). We thank Mathieu Luisier for fruitful discussions and the sharing of simulation results from OMEN.

#### REFERENCES

- [1] K. Stokbro, A. Blom, S. Smidstrup, SISPAD 2013.
- [2] Atomistix ToolKit (ATK); <u>http://www.quantumwise.com</u>.
- [3] F. Tran, and P. Blaha, Phys. Rev. Lett., vol. 102, p. 226401, 2009.
- [4] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, J. Appl. Phys., vol. 89, p. 5815, 2001.
- [5] S. Adachi, J. Appl. Phys. vol. 61, p. 4869, 1987.
- [6] W. Porod, D. K. Ferry, and K. A. Jones, J. Vac. Sci. Tech., vol 21, p. 965, 1982.
- [7] J. Li, and L.-W. Wang, Phys. Rev. B, 72, 125325, 2005.
- [8] Boykin et al., Phys Rev. B. 66, 125207, 2002.
- [9] M. Luisier, private communications.