A comparison between a fully-3D real-space versus coupled mode-space NEGF in the study of variability in gate-all-around Si nanowire MOSFET

A. Martinez, A. R. Brown and A. Asenov
Dept. Electronics & Electrical Engineering
University of Glasgow
Glasgow, UK
antonio@elec.gla.ac.uk

Natalia Seoane
Dept of Electronics & Computer Science
University of Santiago de Compostela
Santiago de Compostela, Spain

Abstract—**In this work a comparison between the fully-3D (F3D) real-space approach and the Couple Mode Space (CMS) approach in solving the Non-Equilibrium Green Function (NEGF) quantum transport equations is carried out. The CMS approach, as with every mode decomposition technique, is inherently an approximate method, because of the finite number of modes used. This method, due to its quasi-1D nature, is less computationally expensive compared to the fully 3D one. In the simulation of devices, the correct magnitude of the electron current and its electrostatic self-consistency are important issues. We use the current as an indicator for the accuracy of different CMS implementations. The F3D and the CMS approach are compared in the simulation of thin gate-all-around Si Nanowire transistor. The comparison is carried out for devices with different types of non-uniformities including: (i) smooth SiO_2 interfaces and continuous doping as a reference, (ii) discrete charge in the channel (iii) random discrete dopants in the S/D and (iv) surface roughness. The focus is on the performance and accuracy of the CMS simulations as a function of the number of coupled modes in comparison with the F3D simulation results for the current. Because the CMS approach separates the confinement in the transversal directions from the propagation longitudinal direction, this simplifies the dissection of the underlying transport physics. The transmission dependence on energy can be explained as the interaction between different modes. Also, the sub-band energies allow us to visualise the resonances, when superimposed on the LDOS.**

Keywords- Fully-3D real-space Non-Equillibrium Green Functions ; Couple Mode space; Gate-All-Around Si Nanowire.

I. INTRODUCTION

The electrostatic integrity and performance related issues associated with the scaling of Si bulk MOSFETs to a sub 10 nm channel length favours research into new device architectures such as SOI, double gate and nanowire MOSFETs [1]. At sub 10 nm channel length all of these devices show a high degree of ballisticity and onset of source-to drain tunnelling [2]. As a result, Drift-Diffusion and Monte Carlo (quasi-classical) tools progressively lose predictive capability, favouring quantum transport simulations. Fabricated nanowire transistors [2] start to support these claims. Fully three dimensional (3D) quantum transport (QT) simulations are essential when dealing with variability in nanowire transistors introduced by discrete charges and interface roughness [3]. However fully-3D (F3D) QT simulations are very computationally expensive, particularly when using the Non-Equilibrium Green's Function (NEGF) algorithm that is favoured when considering the inclusion of scattering in the simulations. Additionally, the study of variability requires simulation of large statistical samples of microscopically different devices, which transforms the 3D problem into a four-dimensional one. An alternative to the fully-3D real-space NEGF simulations is the coupled mode-space (CMS) approach [4]. The CMS approach substantially reduces the computational load and provides a tool to understand the physics in terms of the modal decomposition and their interactions. Tunnelling currents can be easily assessed due to the transverse-longitudinal decomposition of the Hamiltonian. However it is difficult to predict in advance how many modes will be required to produce an accurate current voltage characteristic in the presence of variability sources. The next section of this paper highlights the major differences between the F3D real-space method and the CMS approach. In Section III, the two methods are applied to the simulations of different inhomogeneous devices in order to compare them in terms of speed and accuracy. The last section reviews the main finding of this work.

II. FULLY-3D REAL-SPACE VS COUPLED MODE SPACE

A. *Fully-3D Real-Space*

The F3D real-space method as its name suggest calculates

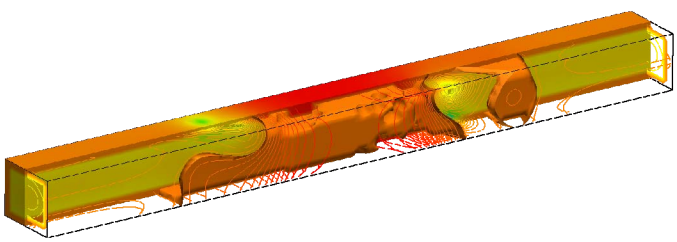

Fig. 1 Electrostatic potential of the Si nanowire transistor with rough surface and random dopants in the source and drain.

This work is supported by the UK EPSRC through Platform grant EP/E038344/1 "Atomic Scale Simulation of Nanoelectronic Devices".

the 3D electron density directly from the 3D real-space Green function matrix. This calculation requires the inversion of the 3D retarded Green function G^R matrix in order to compute $G^<$, the imaginary part of which is proportional to the carrier density. The off-diagonal terms of $G^<$ are used to compute the electron current. Even with the use of recursive algorithms [3] that avoid the inversion of the full matrices, the calculations are computationally very expensive. The 3D spatial $G^<$ matrix, with dimensions $(n_{t1} \times n_{t2} \times n_l)^2$ (where n_l is the number of nodes in the transport, and n_{t1} and n_{t2} are the number of nodes in the transverse directions) needs to be calculated for every energy point, imposing further computational overheads.

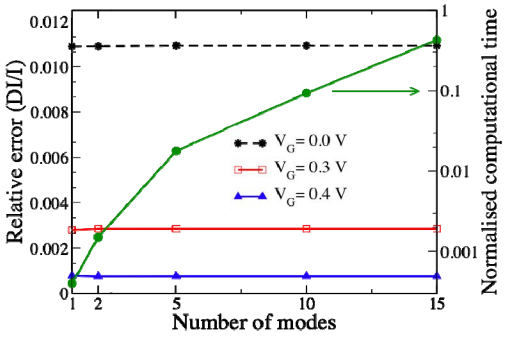

Fig. 2 Relative error in the current, and the computational time for CMS. A smooth device is considered.

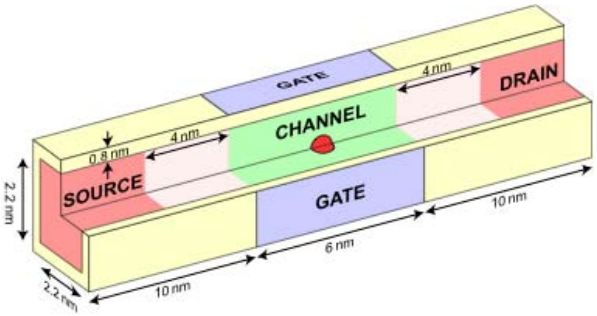

Fig. 3 Schematic view of the nanowire with a discrete charge in the channel.

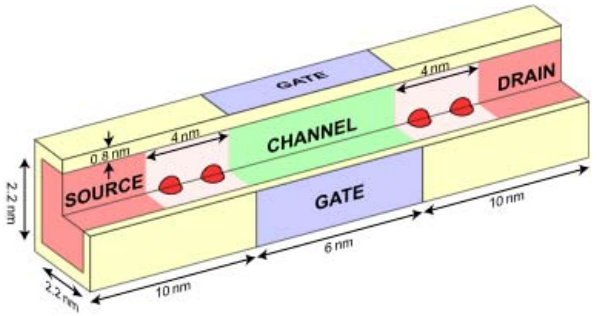

Fig. 4 Schematic view of the nanowire with dopants aligned in the S/D.

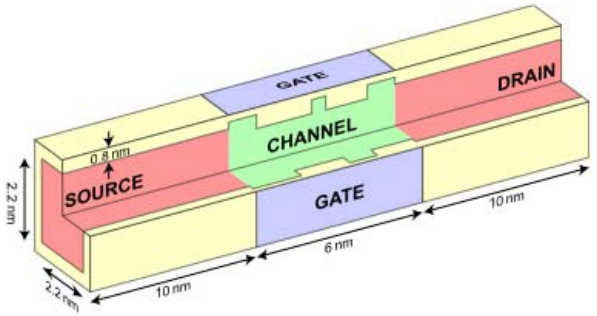

Fig. 5 Schematic view of the nanowire with rough interface in the channel.

B. *Coupled Mode Space*

The CMS method splits the problem into the transversal and longitudinal spaces. The transversal space provides the cross-sectional wave functions and sub-band energies. The transport is solved in the product space of the longitudinal space with the mode space. The dimension of this space is $(n_l \times N)$ where N is the number of modes, which is substantially smaller than the discretised 3D real-space of dimensions $(n_{t1} \times n_{t2} \times n_l)$. In this work 15 modes have been used to achieve a good agreement between the CMS and the F3D simulations. For larger cross-sections more modes will probably be required.

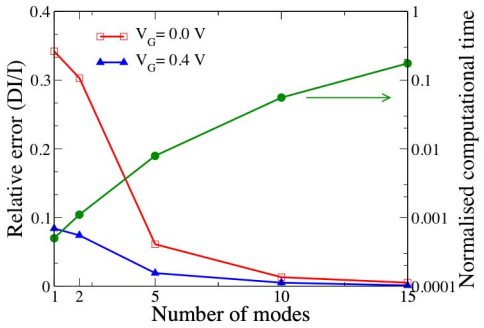

Fig. 6 Relative error in the current, and the computational time for CMS. A device with a channel dopant is considered.

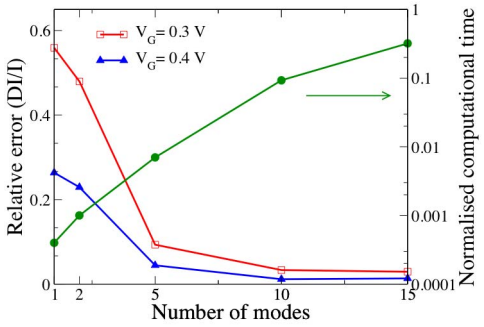

Fig. 7 Relative error in the current, and the computational time for CMS. A device with source/drain dopants is considered.

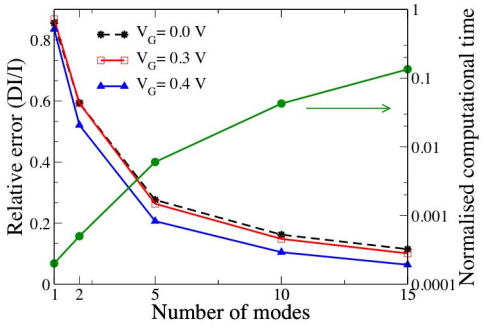

Fig. 8 Relative error in the current, and the computational time for CMS. A rough device is considered.

III. Results and Discussion

In this section we have compared results from F3D real-space NEGF and CMS simulations for different types of nanowire devices with: (i) smooth SiO_2 interfaces and continuous doping, (ii) discrete charge in the channel (iii) random dopants in the S/D and (iv) surface roughness. The focus is on the relative error and the performance of the CMS simulations as a function of the number of coupled modes.

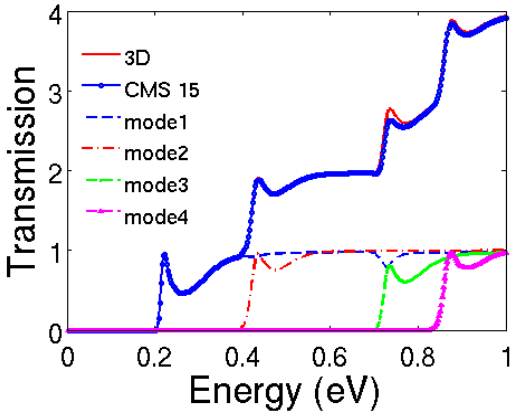

Fig. 9. CMS, the first four modes, and 3D transmission coefficients for the device with a channel dopant (V_G=0.5V,V_D=1mV).

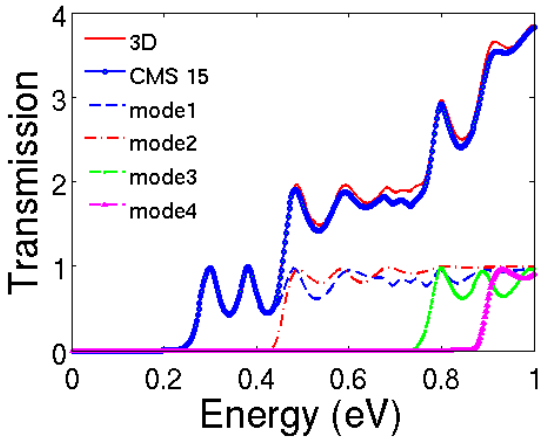

Fig. 10 CMS, the first four modes, and 3D transmission coefficients for the device with dopants in S/D (V_G=0.4V,V_D=50mV).

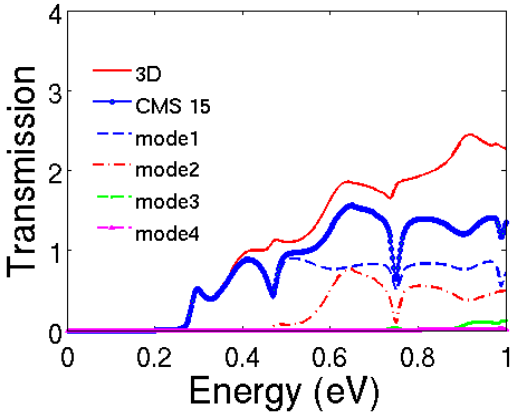

Fig. 11 CMS, the first four modes, and 3D transmission coefficients for the device with rough surface (V_G=0.4V,V_D=50mV).

In order to visualise the spatially inhomogeneous nature of the self-consistent potential we present in Fig. 1 the equipotential distribution in a thin Si nanowire transistor with random dopants in the S/D and rough SiO_2 interface. In this case, the inhomogeneous potential will mix many modes in the same cross-section when matching the wave function from cross-section to cross-section.

In the case of the smooth device there is practically no mode-mixing because of the smooth behaviour of the potential and the expectation is that the uncoupled and the coupled mode approaches will produce similar results. Also, size quantisation induces a shift between the first and second sub-bands that is of the order of several kT. This means that the dominant current flows only in the first sub-band. Fig. 2 shows how, for the smooth device, the relative error in the current (REC) and the normalised computational time (NCT) depend on the number of coupled modes. The NCT has been normalised to the value obtained in the F3D simulations. The REC slightly decreases with an increase in the gate bias due to the exponential sensitivity of the current on small discrepancies in the potential distribution. Importantly REC is independent of the number of modes, which for a smooth device makes the F3D NEGF simulations unnecessary.

Figs 3, 4 and 5 show the schematics of three simulated devices representing three important source of variability that introduce different types of inhomogeneities in the potential. The REC and the NCT for the three cases are shown in Figs 6, 7 and 8 respectively. With the introduction of variability sources the number of coupled modes needed to match the F3D NEGF results increases. In general there is a sub log-linear increase in the NCT with the increase in the number of the modes, but even for 15 modes the simulation time is less than half of the F3D NEGF simulation time. For the discrete channel charge and random dopants only, approximately 5 modes are needed to achieve a REC of less than 0.2 at 0.01 of the computational time required for the F3D simulation.

The roughness case (Fig. 6) requires more modes in order to achieve the same accuracy. This is due to the significant change in confinement induced by the surface roughness from cross-section to cross-section.

The corresponding transmission coefficients at V_G=0.4V are presented in Figs 9, 10 and 11 respectively. These show

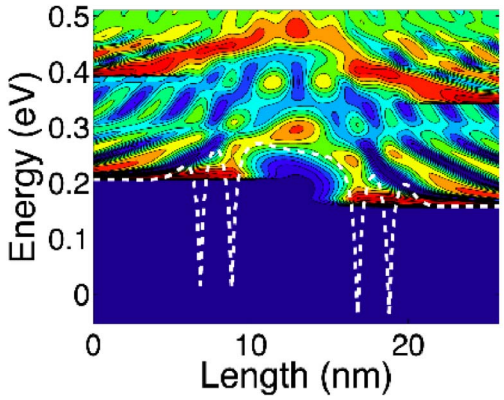

Fig. 12 Density of states along the wire axis for a device with random dopants in the source and drain

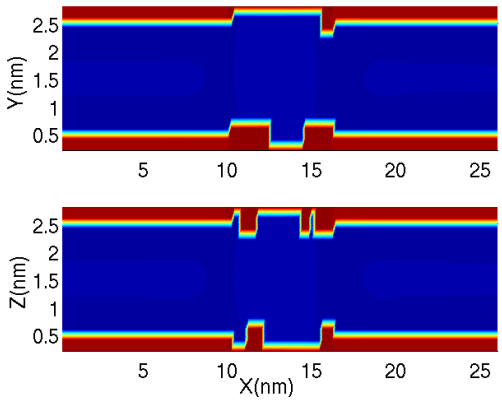

Fig. 13 Interface roughness for two perpendicular planes along the wire for one of the devices simulated

the relative contributions of the different modes to the current, and the interaction between the different modes. We also show the 3D transmission for comparison as it provides an understanding of the behaviour of the total 3D transmission coefficients.

The presence of resonances in the transmission is clear (Fig. 9) for the case of a donor in the middle of the channel (Fig. 2). The second mode shows the same type of resonance as the first mode but this is diminished due to the poor overlap of the second mode with the impurity potential. There is an interaction between the first and the third modes due to the high overlap between them and the impurity potential in the middle of the channel.

The transmission for the case of the dopants in the Source/Drain is more complicated, and therefore it is helpful to study the Local Density of States (LDOS). The LDOS of the device from Fig. 3 with 2 discrete dopants in the source and 2 in the drain regions is shown in Fig. 12. Note the localisation of charge in the coulomb wells due to resonant states (quasi bound states). The first two resonances in the transmission (Fig. 10) appear approximately at 0.30 and 0.38 eV respectively, which indicates that the first resonant levels of the dopants are not involved in the current flow directly. The resonant states relevant for the current are of more global character. They carry the signature of the dopants and of the whole longitudinal electrostatic potential through the channel. This could be observed in the LDOS of Fig. 12.

The transmission coefficients for the surface roughness case, shown in Fig. 11, shows a good agreement between the CMS and the F3D simulations at low energies but differs at high energies. This discrepancy comes from the fact that, due to the strong surface roughness variations, the longitudinal spatial resolution is not sufficient to successfully couple the first and second subbands. This effect occurs at energies substantially larger than kT and therefore does not affect the value of the current significantly.

The interface roughness could create resonant cavities inside the device. Figure 13 shows the electrostatic potential in two perpendicular planes for a device with interface roughness that creates a resonant cavity in the channel. This cavity produces resonances that can be observed in the transmission coefficient (Fig. 14). There is a strong coupling between the first and second modes that induces dips and peaks in the transmission. The LDOS shown in Fig. 15 confirms the

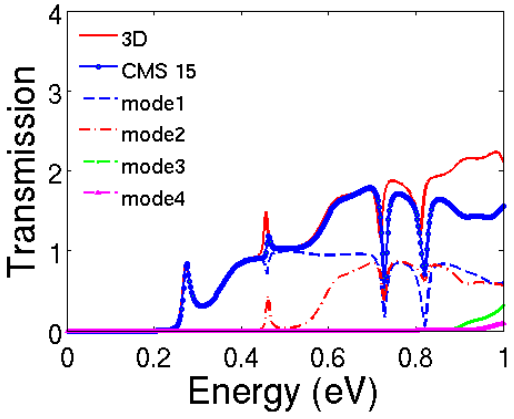

Fig. 14 CMS, the first four modes, and 3D transmission coefficients for the rough device shown in Fig.13 (V_G=0.4V,V_D=50mV).

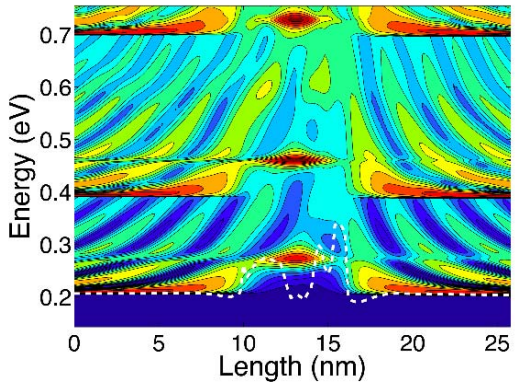

Fig. 15 Density of states along the wire axis for the device shown in Fig. 13 (V_G=0.4V,V_D=50mV)

existence of resonant states in the middle of the channel; the white dashed line shows the first sub-band, making explicit the longitudinal-formed cavity.

IV. Conclusions

The NEGF equations describing the quantum electron transport have been solved using the CMS approach and the F3D real-space approach in the simulation of Si nanowire transistors with different types of inhomogeneities including discrete dopants and rough interfaces. Just 5 modes are sufficient in order to achieve a relative error of less than 10% for the discrete dopant cases. For the rough interface devices the convergence of the CMS approach as a function of the mode numbers is slower and more than 10 modes are needed to achieve similar accuracy. A detailed study of resonances in terms of the transmission coefficients and LDOS has been carried out. This work will facilitate future study of variability including dissipation in a quantum transport framework due to the lower computational cost of the CMS compared to F3D.

References

[1] N. Singh, *et al*., IEDM Tech. Dig., pp 547-550 (2006).
[2] K. H. Cho, *et al*., Appl. Phys. Lett. **92**, 052102 (2008)
[3] A. Martinez, *et al*. IEEE Trans. Electron Devices, **54**, 2213 (2007)
[4] R. Venugopal, et al J. App. Phys. **92**. 292, (2004)