Adamu-Lema, F., Duan, M. , Berrada, S., Lee, J., Al-Ameri, T. , Georgiev, V. and Asenov, A. (2017) Modelling and simulation of advanced semiconductor devices. *ECS Transactions*, 80(4), pp. 33-42. (doi:10.1149/08004.0033ecst)

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

http://eprints.gla.ac.uk/159057/

Deposited on: 15 June 2018

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk

Modelling and Simulation of Advanced Semiconductor Devices

F. Adamu-Lema, M. Duan, S. Berrada, J. Lee, T. Al-Ameri,
V. P. Georgiev, A. Asenov

School of Engineering, University of Glasgow, G12 8LT, UK
fikru.adamu-lema@glasgow.ac.uk

Abstract: This paper presents a modelling and simulation study of advanced semiconductor devices. Different Technology Computer Aided Design approaches and models, used in nowadays research are described here. Our discussions are based on numerous theoretical approaches starting from first principle methods and continuing with discussions based on more well stablished methods such as Drift-Diffusion, Monte Carlo and Non-Equilibrium Green's Function formalism.

I. Introduction

Modelling and simulation of semiconductor devices has been driving the development of the semiconductor technology and contributed significantly to comprehensive understanding of their operation. Moreover, it is estimated that the Technology Computer Aided Design (TCAD) can reduce the cost of the technology, circuit and system development and hence reducing the time-to-market. Analysis using state-of-the-art device modelling helped the process of miniaturization of MOSFETs through the past decades in characterizing their parameters and predicting their behaviour in system level. Moreover, modelling and simulation provide a deeper understanding of the critical issues surrounding the device scaling and its impact on circuit system level, including functionality and reliability of its operation. This helps to deliver better device/design solutions [1]. To illustrate the complexity and the importance of the modern TCAD technology, we present in this paper numerous simulations techniques based on various theoretical approaches and models. As test cases, we have used nowadays devices architectures such as nanowires, FinFETs and Bulk transistors.

II. Model description

The state-of-the-art in semiconductor modelling and simulation covers a wide range of specialist areas related to the entire process of integrated semiconductor device manufacturing. The main areas of modelling and simulation [2] that are typically covered by state of the art TCAD simulation tools are the following. The first is process modelling: this is so called front-end process modelling that is used to characterizes the essential steps during the device manufacturing such as deposition, etching and other manufacturing processes. Secondly, there is a device modelling section of the TCAD which the main aim is to understand how the devices operate by studying their electrical characteristics and intrinsic physical properties. And lastly there is the compact modelling which is used to perform circuit level device simulation that characterize the behaviour of the integrated circuit component as a module. However, a detailed discussion of all these physical models and simulation methodologies is not the aim of this paper. Therefore, the emphasis has been limited to cover only device modelling and simulation of advanced semiconductor devices.

The modelling of semiconductor devices involves numerical solutions of mathematical equation that describe their fundamental physical properties and electrical characteristics. In general, there are three methods that have been used in the TCAD world to perform device modelling. By far the most popular and traditionally well-established way of modelling the carriers' transport and electrical behaviour in transistors, is the Drift Diffusion (DD) method. This method can be combined with density-gradient correction which is designed to consider the quantum mechanical effects which are playing important role in the modern semiconductor devices. The second approach of modelling semiconductor devices is based on Monte Carlo (MC) method. MC is more computationally demanding in comparison to DD because it solves Boltzmann transport equation for each electrical carrier such as holes and electrons. Finally, the most recent and the most computationally intensive method is based on the Non-Equilibrium Green Function formalism in combination with first principle methods such as Density Functional Theory (DFT). In the section, we will provide a brief of all three methods – DD, MC and NEGF applied to current transistor's architecture.

A. Drift Diffusion Model

Historically, the working horse of semiconductor modelling and simulation, is the Drift Diffusion (DD) method. It is based on the drift and diffusion properties of carrier transport phenomena [3,4]. In the presence of electric field in the semiconductor, the drift velocity is proportional to the created electric field expressed as $v_d = \mu E$ where μ [cm^2/V*s] is the proportionality termed as mobility. This relation is true for systems at low electric field. Moreover, carrier mobility is affected by various carrier scattering mechanisms like phonon, ionized impurity, and temperature. At high electric field the dependence of drift velocity of carriers is no more linear and it behaves independent of the electric field and converges to a saturation point of ~10^7 V/cm for silicon as the field increases. The carrier drift current is then described using $J_n = q\mu_n nE$ for electrons.

On the other hand, the introduction of extra dopant into the silicon, will create uneven carrier distribution [4]. Unevenness in the carrier concentration leads to the migration of electrons/holes from high concentration region to the lower hence the flow of carriers creates the diffusion current. Diffusion current is proportional to gradient of the carrier concentration as $J_n = qD_n\,(dn/dx)$, where D_n is the diffusion coefficient. Combining the drift and diffusion mechanisms of carrier transport one can model the flow of current in semiconductors using the drift-diffusion method, which is the combination of the two currents described in the previous paragraph. DD equations can describe the transport properties of carriers by solving the current continuity equation which is coupled to the Poisson equation ($\nabla^2\phi = -\rho/\varepsilon$) self consistently.

The DD engine, which is used to simulate the devices in this paper, is a three-dimensional, quantum corrected, statistical device simulator which self-consistently solves the carrier concentration and potential distribution coupled with current continuity throughout the simulation domain to characterize the electrical properties of the device. A simplified version of the simulation flow of our numerical solver with transport modules based on DD, MC and NEGF methods is presented in Fig. 1 below.

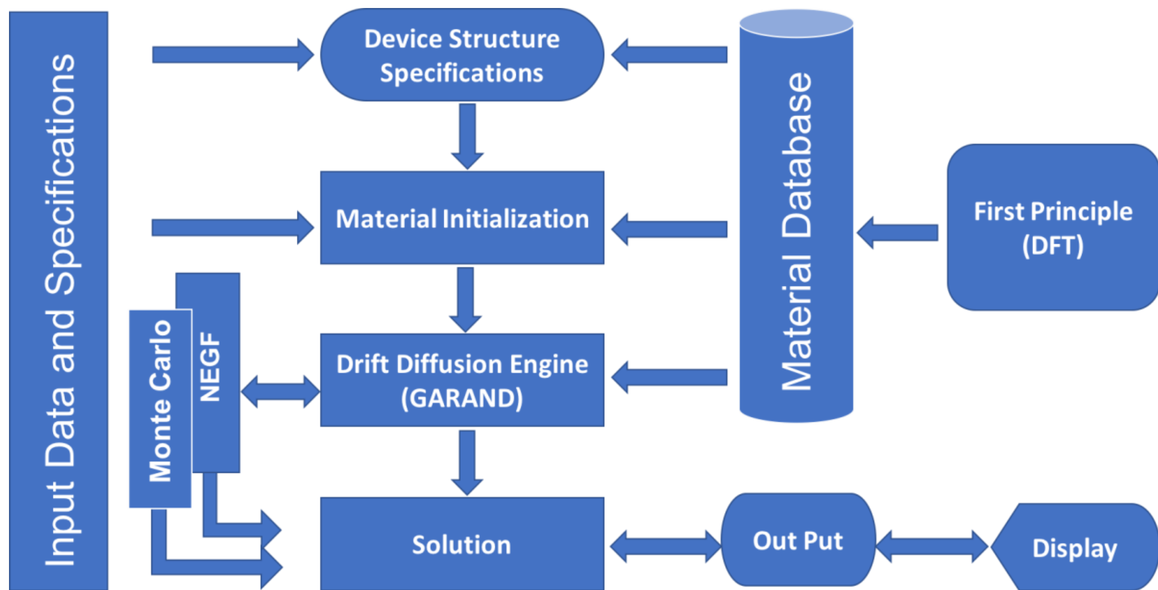

Fig. 1 An example of flow chart of showing an integrated device simulator with Drift Diffusion engine coupled with Non-Equilibrium Green Function and Monte Carlo simulator.

B. Monte Carlo Method

Monte Carlo (MC) carrier transport modelling is a particle approach that combines the period of deterministic free-flight times with stochastic scattering process implemented to estimate carrier dynamics (transport) in advanced semiconductor devices. It uses the random process to calculate the stochastic motions of the particles and their scattering mechanisms. The bigger the statistical sampling of the particle motion the better MC simulation approximates the distribution function. There are different types of MC simulation methods (single particle, ensemble, and self-consistent MC). For device simulations, the self-consistent MC approach is suitable, since it solves the Poisson equation self consistently to determine electrostatic potential distribution in the semiconductor device. MC simulation does not make any assumption about the distribution function to approximate the carrier transport [5-8] and the carriers are considered as a particle rather than as a fluid, which makes the MC approach more general and accurate in approximating the carrier transport than the DD methods.

The MC simulator is integrated in to the hierarchical device simulation tool as shown in Fig. 1. In our simulator, a MC simulation takes as input a solution from the DD engine or, optionally, the solution from a prior MC simulation. The input to the MC engine completely specifies the simulation domain, it's boundary conditions and the initial state. Default material models are defined equivalently for DD and MC engines and are inputs to both. The solution from Monte Carlo simulation is provided through the collection of statistical data taken directly, or indirectly, from particle distributions representing the electron/hole carrier distributions within the simulation domain. The particle distribution evolves over the course of the simulation via the repeated application of periods of deterministic free flight and instantaneous stochastic scattering events. Carriers move through the real-space simulation domain during the free flight and, given a local field, accelerate, while scattering, given a list of mechanisms, instantaneously affects the carrier's momentum [8].

C. Non-Equilibrum Green Function Method

The Non-Equilibrium Green Functions (NEGF) formalism is the most suitable approach, which can capture accurately important quantum effects in charge and transport for present-day nano-scale devices. For example, due to ultra-small cross-section in current transistors the charge distribution shows strong quantum confinement behavior. Additionally, the gate length for future technological nodes will be comparable to the mean free path of electrons in silicon at room temperature. Thus, the electrons will exhibit a strong wave-like behavior in these ultra-short devices.

As a result, the source-to-drain tunneling resulting from this wave-like behavior has important consequences on the device operation in terms of I_{on}/I_{off} ratio, threshold voltage, sub-threshold swing. Therefore, such quantum mechanical effects are mandatory to be considered in the simulation process to accurately capture the observed physical processes and estimate their electrical properties. Consequently, it is of prime importance to develop simulators based on a quantum mechanical formalism such as NEGF [8,9] to capture the tunneling effects related to the wave-like behavior of the carriers. Like the MC approach described above, in our simulator, a NEGF simulation takes as input a solution from the DD engine or, optionally, the solution from a prior NEGF simulation. This is coupled in self-consistent loop with Poisson equation (Fig.1).

D. First Principle method

First principle methods are used to describe the electrical and optical material properties of solids and molecules. Currently, the most common and widely used approach is based on Density Functional Theory (DFT) [10,11]. Coupling DFT with NEGF formalism could provide more accurate description of the quantum confinement in current electronic devices. Also, DFT could provide understanding of the carrier's transport in devices, energy levels in isolated molecules and electronic structure of various materials such as semiconductors and oxides. In our case, we primarily use DFT to compute effective mass of electrons and holes, which are extracted from the band structure. Those effective mases are used as input parameters in all other models illustrated in Fig. 1 and discussed above.

III. Results and discussions

Nowadays advanced devices such as FinFET, Silicon Nano Wire transistors, SOI MOSFETs and others non-conventional semiconductor devices were born out of the necessity of continuing the scaling beyond the conventional MOSFET architecture. In this section, we present simulation and analysis on some of the most recent and important devices structures.

A. FinFET

FinFET technology is massively adopted as the advanced logic node technology since 22nm generation. Modelling and simulation of FinFETs and accurate predictions can significantly speed up the next technology node development. 7nm technology node is a critical point due to its extremely complicated physics and technical difficulty in fabrication. Employing advanced quantum corrected drift diffusion model we present the simulations results for 7nm FinFET featured by the extremely scaled fin width of 5nm and height of 25 nm as shown in the Figure 2(a). Accurate capturing of the charge distribution in the nanoscale channel is mandatory since the quantum effect plays a determined role with regards to the performance. Garand [12] is used to simulate the full 3D device structure. Also we activated the 2D Poisson-Schrodinger solution which accurately captures the

quantum confinement in the cross-section of the device. The wave-functions in the corresponding sub-bands of each valley is calculated and the charge distribution is obtained for different gate biases. This charge distribution serves as the calibration target in the TCAD calibration tool Enigma for calibrating density-gradient masses which are used in the DD simulation. The calibration results are shown in Fig. 2(b). Calibrating the DG effective masses to the 2D PS solution provides an accurate charge distribution based on quantum mechanics for the DD method. The calibrated simulation deck then will be used for the further DD simulations and MC simulations and analysis of the electrical characteristics.

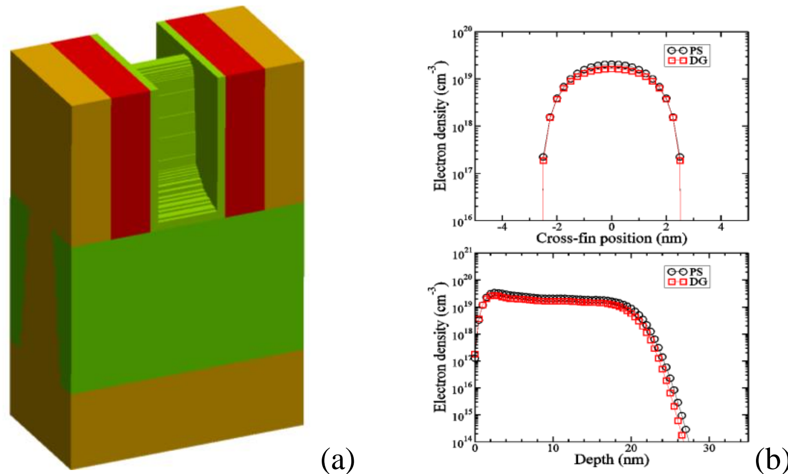

Fig.2 The bulk FinFET structure (a), and Poisson-Schrodinger simulation of fin cross-section and corresponding density gradient calibration (b).

B. Z2FET

Z^2FET has very complex structures, and is one of the most promising devices to be used for application in memory cell, without connecting to extremal charge storage component (i.e. 1T-DRAM). Significant progress has been made in characterizing its memory behaviour and analyzing its physical properties. These qualities can be exploited in the implementation of Z^2FET in embedded 1T-DRAM application. To understand the complexity of this promising device, TCAD model is invaluable asset. Although experimental method can help the characterization process, the use of numerical modelling is equally important in understanding the dynamics of carriers, e.g., how the DC operation of this device can be characterized in the different sections of the device. In this paper, we investigate the DC operation of Z2FET in conjunction with its memory behaviour and how it can be used in 1-T DRAM memory.

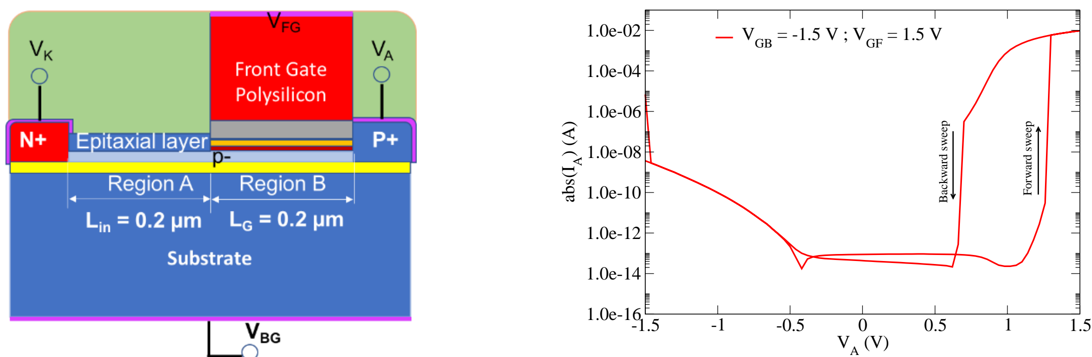

Fig. 3 Z^2FET device stricture (left) showing a typical I_d-V_d characteristics when operating as a memory device (right) [13].

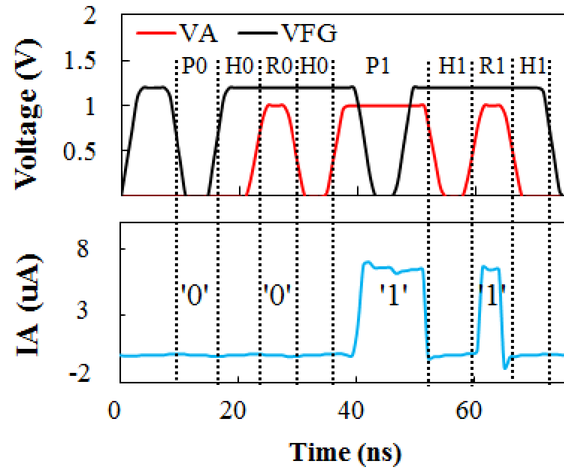

Fig. 4 Z^2FET memory operation. When data '0' is programmed, followed read out is '0'. Similarly, data '1' is read out when '1' is programmed.

C. Silicon Nano Wire Transistors (SNWT)

Figure 5 below shows a plot the LDOS in the two first delta valleys obtained in Off-state (Fig.5a-b) and On-state (Fig.5c-d) for a square NWT oriented along <100> direction. Top figures show the results in Off-state for (a) the first delta valley and (b) the second one. Bottom figures are obtained in On-State for (c) the first delta valley and (d) the second one. The results for the third valley are identical to the ones of the second valley (for symmetry reasons).

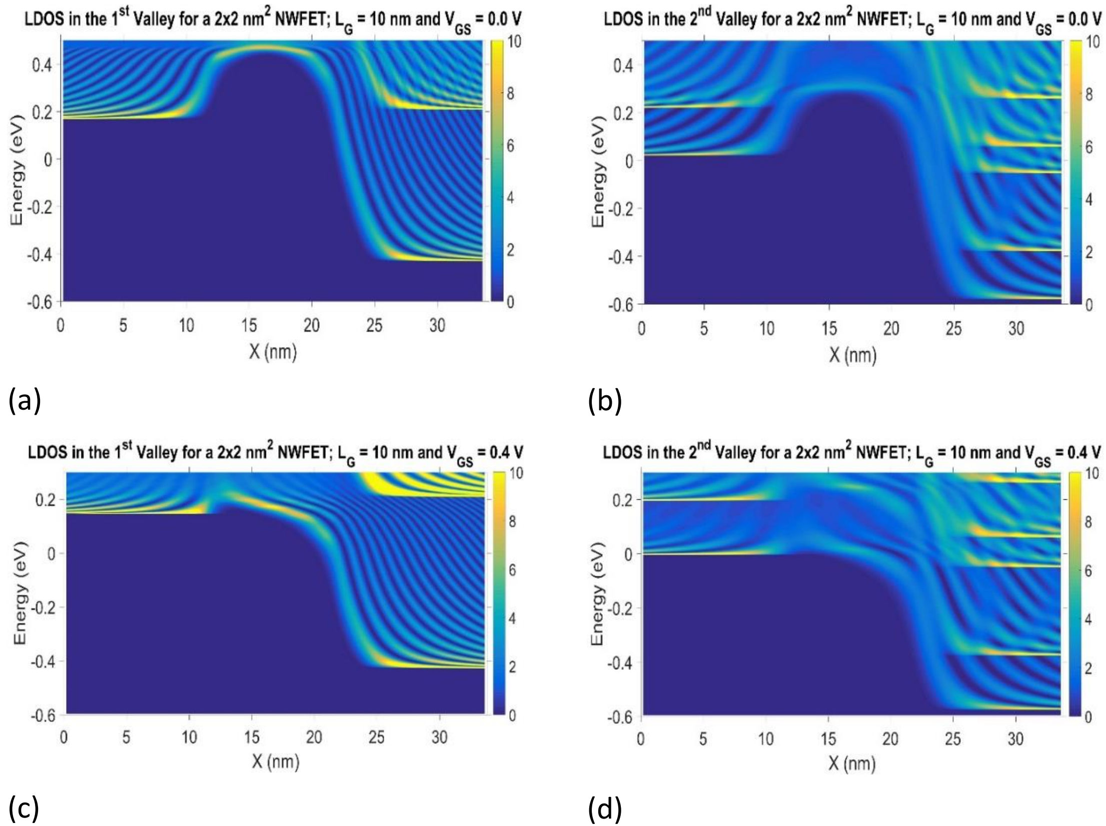

Fig. 5 Local density of states obtained by NGEF after a self-consistency with Poisson equation is achieved.

The local density of states and the tunneling from the source to the drain have been computed thanks to our quantum simulator based on NEGF formalism. Also, the SNWT described in Fig.5 are with ultra-scaled dimension such as 2nm x 2nm cross section and 10 nm gate length. In such nanowires, without capturing the quantum mechanical effects the device simulations will be simply wrong.

D. Vertically stacked multiple SNWT

As an example, we present a simulation result of vertically stacked NWTs. Fig. 6 (right-hand side) compares the I_D-V_G curves for NWT with a single, double, and triple channel configuration. The three-channel device has higher I_{ON} current when compared the other structures with one and two channels transistors. However, the current in the three-channel transistor does not provide a triple current value in respect of the single channel as we may expect is the case.

The reason for this is clearly visible on the left hand-side of Fig. 6, which shows that the 3D view of the current density for NWTs with single, double, and triple channels devices. The transistor with two channels has almost identical current density in top and bottom channels. However, the device with three-channels show significant difference between the current density of the top nanowire and the bottom two channels. The main reason for this is that the contacts are on the top of the source and drain region which makes it easier for the current to flow through the closest nanowire channel in relation to the contacts. The performance results obtained from the preliminary analysis of multi-vertically stack laterals NWTs in Fig. 6 could exceeds the scaling targets set by the industry.

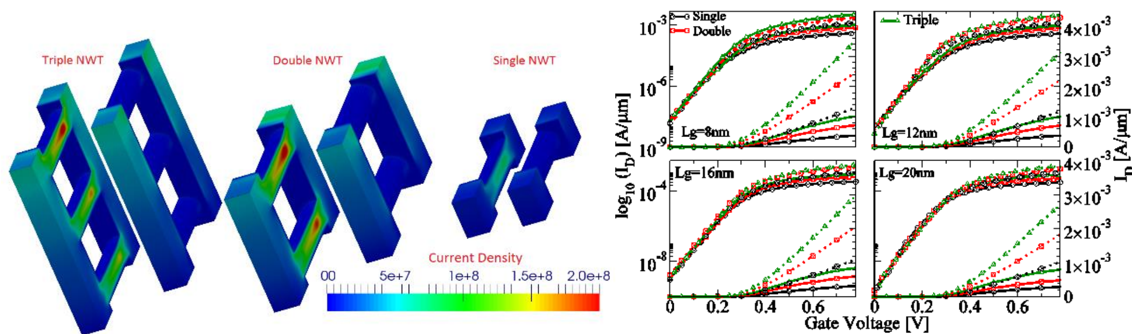

Fig. 6 Shows a 3D view of the current density for vertically stacked NWTs with single, duple, and triple channels with L_G=12nm (left). The I_D-V_G curves of single, double, triple NWTs where with four different gate lengths L_G=8nm, 12nm, 16nm and 20nm. The dashed lines are at V_D=700 mV and the solid lines are at V_D=50mV.

E. First Principle Simulation

As the device size has continued to shrink to a nanoscale regime, the quantum confinement causes the important physical material parameters such as band gap and effective masses to change. For this reason, we need to estimate them precisely through the first principle calculation for example DFT. Fig. 7 (a) shows the cross-section of the atomic structure of SNW with 2.2×2.2 nm^2. Dangling bonds at the surface were passivated with hydrogen atoms. All atoms are relaxed until the maximum force of atoms becomes less than 0.01 eV/Ang [15]. Fig. 7 (b) and (c) show the band structures of SNWs with 2.2×2.2 nm^2 and 3.9×3.9 nm^2, respectively. To calibrate the band gap with experiment (1.12 eV), we have used the Meta-Generalized Gradient Approximation (MGGA) with a

parameter of 1.072. We have found that the SNWs with 2.2×2.2 nm^2 and 3.9×3.9 nm^2 have larger band gap (1.71 and 1.30 eV) and heavier transport effective masses (0.617 and 0.285 m_e, m_e is the electron rest mass) than bulk Si. These new effective masses are used to replace the default bulk Si effective masses in our DD and NEGF simulator.

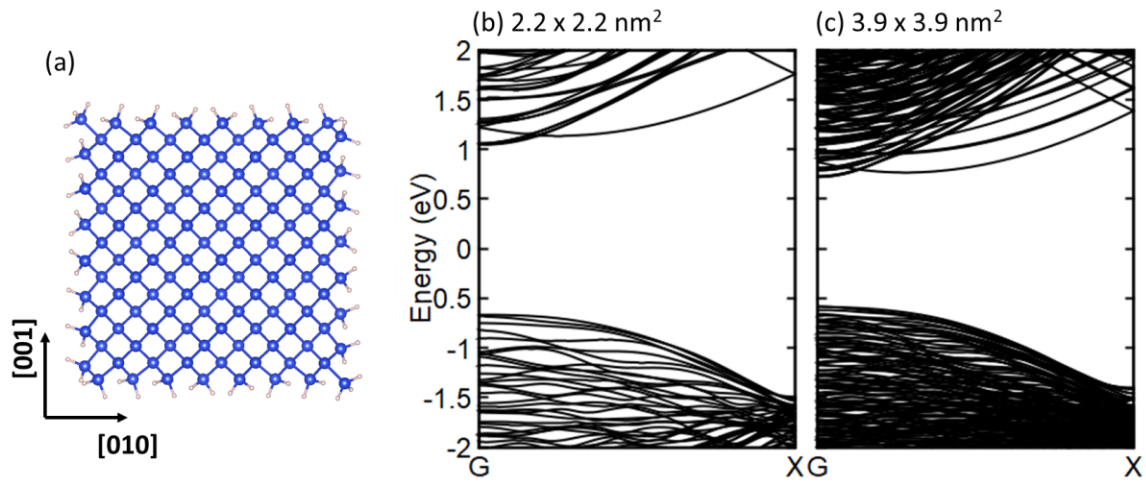

Fig.7 (a) Atomic structure of H-passivated Si nanowire (Width = Height = 2.2 nm). The band structure of (b) $2.2 \times 2.2\ \mathrm{nm}^2$ and (c) $3.9 \times 3.9\ \mathrm{nm}^2$. 0.0 eV is set to Fermi-level.

F. Statistical Variability

As device dimension is getting in the region below 20-30nm statistical variability will affect main figures of merit of devices significantly. Random discrete dopants (RDD) [15,16], Line Edge Roughness (LER) [17] and Metal Grain Granularity (MGG) [18,19] are the main sources of statistical variability. Presenting detailed examples of simulation results is beyond the scope of this paper. For completeness, we briefly discuss RDD, MGG and LER.

Even if the fluctuations associated with lithographic dimensions and layer thicknesses are well controlled, random fluctuation of the relatively small number of dopants and their discrete microscopic arrangement in the channel of sub-0.02 μm MOSFET's will lead to significant variations in the threshold voltage and the drive current. Similarly, the line edge roughness (LER) caused by tolerances inherent to materials and tools used in the lithography processes is yet another source of intrinsic parameter fluctuations. Although the introduction of metal as a gate material to replace polysilicon, the metal grains (MGG) with different crystallographic orientations have different work functions leading to a threshold voltage variation. Fig. 8 shows the 3D potential distribution in the channel given all sources of statistical variability.

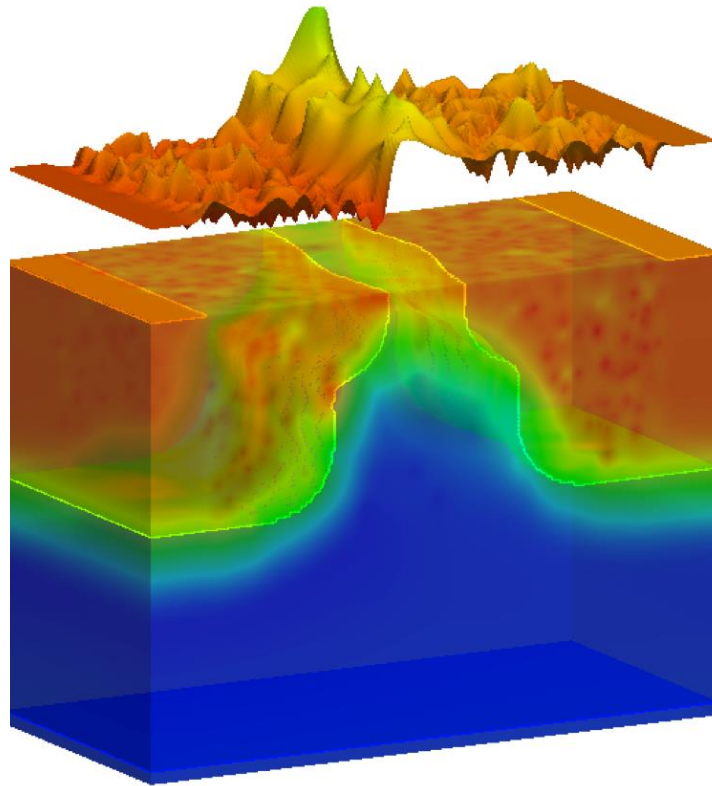

Fig. 8 3-D 25 nm MOSFET structure showing the potential distribution in the channel (top) and the electron density in the active area (bottom) at threshold when including all three sources of statistical variability.

IV. Conclusions

In this paper, we have described the current state of the art of modeling and simulation of advanced semiconductor devices. We have discussed the three most widely used models and methods in the TCAD area such as DD, MC and NEGF. The DD methods with quantum corrections is still the main simulation tool in the field. However, more advanced transport modules such as MC methods and the NEGF approach becomes compellingly attractive for nanoscale transistors research and development. Although that they are computationally more expensive than the DD, MC and NEGF are mandatory to accurately capture the complex physical picture in modern ultra-scaled devices such as FinFET and nanowires. Moreover, first principle methods such as DFT are required for more precise description of the material properties in the novel electronic devices. In conclusion, due to nanoscale dimension of the nowadays transistors, the current TCAD state of the art approach is based on combination of numerous models and theoretical approaches.

References

1. A. Asenov, B. Cheng, X. Wang, A. R. Brown, C. Millar, C. Alexander, S. M. Amoroso, J. B. Kuang and S. Nassif, IEEE Transactions on Electron Devices, Vol. 62, No. 6, pp. 1682–1690, June 2015.
2. Semiconductor Industry Association, International Technology Roadmap for Semiconductor Industry: ITRS'03 edition, Austin, Texas: 2003

3. J. P. Colinge and C. A. Colinge, “Physics of semiconductor devices”. Springer 2005
4. S. M. Sze, Kwok K. Ng, “Physics of Semiconductor Devices”, Third edition, John Wiley & Sons, Inc., 2007
5. U. Ravaioli, Semicond. Sci. Technol., Vol. 13, pp. 1-10, 1998
6. M. V. Fischetti and S. E. Laux, Physical Review B, 38:9721–9745, November 1988.
7. C. Jacoboni and P. Lugli. Springer-Verlag Wien New York, 1989.
8. K. Tomizawa, Artech. House, Boston, London, 1993. Z. Ren, R. Venugopal, S., Datta, and M. Lundstrom,, IEDM Tech. Dig., pp, 107-110, December 2001
9. S. Datta,, IEDM Tech. Dig., pp, 703-706, December 2002.
10. Cuevas, Heurich, Pauly, Wenzel and Schon, Nanotechnology, vol.14, 29 (2003).
11. R.O. Jones and O. Gunnarsson, Rev. Mod. Phys. 61, 689 (1989)
12. GARAND atomistic device simulator manual, Version 2015.1.0
13. F. Adamu-Lema, *et al*., SISPAD 2017
14. Atomistix Tool Kit (2016.4), QuantumWise A/
15. A. Asenov, IEEE Trans. Electron Devices, vol. 45, pp. 2505–2513, Dec. 1998.
16. A. Asenov, A. R. Brown, J. H. Davies, S. Kaya, and G. Slavcheva, IEEE Trans. Electron Devices, vol. 50, pp. 1837–1852, Sep. 2003.
17. A. Asenov, S. Kaya, and A. R. Brown, IEEE Trans. Electron Devices, vol. 50, pp. 1254–1260, May 2003.
18. A. R. Brown, N. M. Idris, J. R. Watling and A. Asenov, IEEE Electron Device Letters, Vol. 31, No. 11, pp. 1199–1201, Nov. 2010.
19. X. Wang, A. R. Brown, N. M. Idris, S. Markov, G. Roy and A. Asenov, IEEE Transactions on Electron Devices, Vol. 58, No. 8, pp. 2293–2301, Aug. 2011.