Nanowire Transistor Solutions for 5nm and Beyond

A. Asenov^{1,2}, Y. Wang³, B. Cheng², X. Wang¹, P. Asenov², T. Al-Ameri¹, V. P. Georgiev¹

¹Device Modelling Group, School of Engineering, University of Glasgow, Glasgow G12 8LT, UK

²Gold Standard Simulations Ltd., 11 Somerset Place, Glasgow G3 7JT, UK

³Institute of Microelectronics, Peking University, Beijing, 100871, China

¹E-mail: asen.asenov@glasgow.ac.uk

Abstract

In this paper we present a comprehensive computational study of silicon nanowire transistor (SNT) and a SNM SRAM cell based on advanced design technology co-optimization (DTCO) TCAD tools. Utilizing this methodology, we provide guidelines and solutions for 5 nm and beyond in CMOS technology. At first, drift-diffusion (DD) results are fully calibrated against a Poisson-Schrodinger (PS) solution to calibrate density-gradient quantum corrections, and ensemble Monte Carlo (EMC) simulations to calibrate transport models. The calibrated DD gives us the capability to simulate statistical variability in nanowire transistors of the 5nm node and beyond accurately and efficiently. Various SNT structures are evaluated in terms of device figures of merit, and optimization of SNTs in terms of electrostatics driven performance is carried out. A variability-aware hierarchical compact model approach for SNT is adopted and used for statistical SRAM simulation near the “scaling limit”. The scaling of SNTs beyond the 5 nm is also discussed.

Keywords

Compact model, Monte Carlo, nanowire transistor, Poisson-Schrodinger, SRAM, statistical variability

1. Introduction

The three-dimensional tri-gate bulk FinFET was introduced last decade as a technology designed to tackle the challenges of the current leakage, short-channel effects, and the degraded performance in extremely scaled bulk planar MOSFETs. It enables significant current leakage reduction and effectively manages variability through multi-gate action and the tolerance to low channel doping. The FinFET adoption is in the process of ramping up at the 16/14 nm technology node, with mass production already starting at major world players of semiconductor manufacturing [1]. Currently the 10 nm technology generation is under intensive development. Technology advance is also pushed by market needs. For example, the booming handset market further requires the application designs with ultra-low power and high performance in minimized chips.

Simultaneously, variability is becoming increasingly severe with device scaling, significantly affecting performance, power, area and yield (PPAY). Random discrete dopants (RDD), gate line edge roughness (GER), polysilicon gate granularity (PSG) and metal gate granularity (MGG) are the major statistical variability sources. One of the main drivers in the introduction of novel device structures is the high statistical variability in latter bulk nodes (28nm/20nm). Advanced FinFET (and also fully-depleted SOI MOSFET)

Figure 1: The simulation domain of a basic silicon nanowire transistor of 5 nm technology.

technologies do not completely resolve the issue of statistical variation [2], promoting the need for the more advanced technology.

One solution to these challenges is to add more gates controlling the channel in the MOS architecture. The gate-all-around nanowire transistor is proposed as a FinFET successor at 5 nm and beyond to boost the performance while reducing the leakage and maintaining the variability at more manageable levels [3, 4]. However, in order to ensure that a correct architecture is chosen, it is important to take into account key circuit performance metrics as well as pure device figures of merit. [5]. In this paper we will use our advanced TCAD tools to investigate the possible optimal nanowire structure in the perspective of device and circuit performance – specifically an SRAM cell.

2. The nanowire transistors and simulation method

As required by electrostatic integrity, the 14 nm FinFET fin-width is critically below 8 nm [1]. To follow the pace of device scaling, nanowire transistor dimensions for the 5 nm technology node are necessarily less than 8nm, and silicon remains the most convenient channel material. Contrary to FinFETs and the planar architectures, in a nanowire transistor the quantum mechanical (QM) effects are very well pronounced due to the small cross section dimensions and surrounding potential barrier coming from the oxide. Hence, in such ultra-scaled devices it is mandatory to carefully take into account the quantum mechanical nature of the charge distribution. As a result, the transport phenomena, especially the ballistic and quasi-ballistic effects, require precise QM simulations.

In our simulation scope, the dimension, shape, orientation, and source/drain are varying in order to establish the link between device characteristics and optimal performance. Figure 1 shows a basic SNT simulation structure. It features an 8×7 elliptical cross-section and a gate-length of 18 nm. The channel is aligned along the <110> crystalline orientation on a (001) wafer. The channel is undoped and

source/drain regions are with a doping concentration of $2 \times 10^{20} \text{ cm}^{-3}$.

The simulation methodology is illustrated in the flow chart in Figure 2(a). Firstly, the Poisson-Schrodinger (PS) self-consistent solution of charge density in a 2D cross section in the channel is obtained. Thereafter, it is used for ensemble Monte Carlo (EMC) simulations in the device at selected biases. The obtained simulation quantities such as channel charge density profiles, carrier velocities, current-voltage (I_D - V_G) characteristics from the coupled PS and EMC simulations are utilized for gradient density (DG) quantum correction, and mobility model corrections in the drift-diffusion (DD) module, respectively. Thus, the calibrated DD module simultaneously features accuracy and simulation efficiency. This is critical for the following study of device performance of a range of different SNTs, and statistical variability study of microscopically different devices. The statistical simulations are carried out on a statistical ensemble of 1000, including the RDD, GER, and MGG, which are described in detail elsewhere [6-9]. A new type of variation related to nanowire edge roughness is described and illustrated later in this paper.

The TCAD study of nanowire transistor solutions for 5nm and beyond, as shown in Figure 2(b), includes early device design, design of experiments, and statistical compact modeling and circuit simulations, which in turn provide the technical feedback to improve technology.

Figure 2: (a) The simulation tool calibration flow chart. (b) The optimization flow used in this TCAD study.

3. Performance driven SNT selection

3.1. PS simulations of SNT scenarios

In order to select the optimal SNT solutions subject to realistic manufacturing processes, we consider numerous design and material parameters such as the cross-sectional shape, 3D geometry device configurations and channel orientation.. Figure 3 shows the first factor that we consider which is the 2D cross-sectional shape of the wire. We report simulation of SNTs with square, rectangular, circular, and elliptical shapes of the cross-section and $\langle 110 \rangle$ -channel

Figure 3: The carrier density in the various cross sections by PS simulations, DD simulations coupled DG and without DG.

orientation. For direct comparison, all devices have identical area of the cross-section. Also in the same figure we compare three different types of simulations – PS, “DD +DG” and “classical simulations” where we don’t consider any QM effects. The following important conclusions can be obtained from Fig. 3. Firstly, the carrier density is significantly different when QM effects are taken into account. This is very clearly visible in PS simulations where the charge is not uniformly distributed in the 2D cross-section. Also the QM charge profile has peaks away from channel/oxide interface, known as “volume inversion”, which is very well correlated to the particular SNT’s shape. Such non-uniform charge distribution is critical for gate coupling and also carrier transport. Therefore it is very important to consider QM effects in such ultra-scaled devices.

Secondly, it is important to point out that PS and DD+DG calculations reveal very similar charge profiles. This is a significant result, enabling us to calibrate semi-classical, computationally light approaches such as DD+DG to more complex and accurate QM simulations such as PS. Indeed it proves to be the fact that such calibration is possible in Fig.4. Fig. 4 shows 1D orthogonal cut lines of cross section for PS and calibrated DG. The calibration is achieved by adjusting the DG effective masses. Fig. 3 compares the 2D charge profile for each cross-section and it is clear that the calibrated DD +DG results are capable to reproduce the complex QM nature of the charge obtained from PS.

The solved electrostatic performance metrics show that all nanowire transistors have the potential to achieve the targets of the 5 nm technology. The SS drops into the range of 60-70mV/dec at room temperature, and DIBL is less than 45mV/V for all shapes at gate length of 10-20nm [10]. When aligning the same charge density at the off state, the elliptical shapes (8x7, 10x5.6) especially the case of 10x5.6 have more total mobile charge than others. When the gate-length is varied, it shows that the charge density at gate length of ~18nm is larger than others due to the source/drain impact. Therefore in the following sections, with the consideration of the realistic manufacturing the 8x7 elliptical nanowire transistor of 18nm gate-length is studied as example if without explicit descriptions.

Figure 4: The DG calibration of carrier density against PS simulation.

3.2. EMC simulation of SNT

Accurately simulating the carrier transport in the nanowire transistors of ultra-short gate-lengths is extremely challenging for DD simulation without calibration. However, EMC can provide predictive and accurate results, therefore EMC simulations are carried out at selected biases and used to for mobility calibrations of DD module. Fig. 5 shows the comparison between EMC and DD simulations of the cross section mobile charge in circular nanowire. By adjusting the mobility models in DD module especially including the saturation model to account for the injection ballistics, the DD module also provides the accurate device characteristics shown in Figure 6.

Figure 5: The mobile charge profiles in EMC simulation and DD simulation.

Figure 6: The DD mobility calibration in terms of I_D - V_G characteristics of EMC and DD simulations.

From Fig. 6 it is clear that without such calibration and incorporation of the accurate mobility module the default DD results could be significantly different.

4. Variability study

Unavoidably, variability is becoming important in nanoscale transistors due to process deviation and intrinsic material properties, also due to increasing sensitivities of device performance to small dimensions. RDD, GER and MGG remain as main sources of statistical variability (SV) due to the inherited processes in doping and gate patterning. Also there is a new type of variation worth considering the direction and the cross section of nanowire, called nanowire edge roughness (WER). All variability sources mentioned above are shown in Fig. 7.

Figure 7: The statistical variability sources in SNT.

Various sources of SV have a distinguished impact on the device transfer characteristics in the magnitude and in the physical parameters. Shown in Fig. 8, at first the threshold voltage fluctuation distributions are compared. Among others the MGG renders the largest V_T -variation source, followed by WER. However the relative variation magnitude in on-current is changed. Although MGG still is the responsible for the largest I_{ON} variation, WER causes the largest relative on-current variation ($\sigma I_{ON}/\sigma V_T$). This is

because the WER brings strong interface scattering variation in the transport besides affecting the subthreshold regions.

Figure 8: The figure of merit distributions due to various SV sources.

4. Design technology co-optimization

4.1. Statistical compact modelling

In order to investigate the impact of variability on critical circuit metrics, a statistical compact model (CM) should be available for circuit simulations. Unfortunately, there is no ready-for-use statistical model, and various methods can be used to generate such models based on TCAD simulations or measurements of an ensemble of devices. Our methodology features a comprehensive “base” CM extraction, following by the statistical CM extraction based on the statistical TCAD data [11]. Those original statistical compact models are later used for statistical CM generation in the circuit simulations.

BSIM-CMG is used for the base CM extraction of a SNT, and, as is shown in Figure 9, the compact model manages to accurately reproduce the TCAD data in both I_D - V_G and C-V characteristics. For SV modeling, a subset of CM parameters are selected and re-extracted for each device in the statistical ensemble. To ensure the statistical CM quality, the device figures of merit (FoM) are monitored in terms of their distributions and correlations. Figure 10 shows the results of one of the statistical SNT ensembles. The results show that critical device FoMs are captured, as well as their correlations.

Figure 9: The compact model extraction of (a) I_D - V_G characteristics, and (b) C-V characteristics of SNT.

Figure 10: The scatter plots of figures of merit of SNT obtained from statistical TCAD simulations and extracted statistical CM.

4.2. Statistical circuit simulations

Once statistical compact models are available, statistical circuit simulation can be carried out to investigate the impact of variability on circuit performance and circuit yield. Here we take a SNT based SRAM example to investigate the SNT variability impact on the corresponding stability. The 6T-SRAM schematic is shown in Fig. 11. The mismatch between two complementary inverters in the SRAM can cause the unsymmetrical states and cause instability and ultimately yield loss.

Figure 11: The schematic of 6T-SRAM.

Figure 12: The scatter plots of SNM's and read currents at fast/slow/typical corners.

Similar to the FinFET technology the SNT is discrete in “width” as it is limited to multiples of individual nanowires. Stack-nanowire transistors have been observed to have less variability than single-nanowire transistors, and this can also affect the SRAM stability. In this case we consider a minimal SRAM cell arrangement of 1:1:1 for pull-up, pass-gate and pull-down transistors, as an “ultra-high density” type design we expect it to perform the worst in terms of variability but its increased density is extremely appealing in on-chip memory design. Over the design of experiments of gate-length, cross-section size, and equivalent oxide thickness, the fast/slow/typical corners of SNT are identified. The statistical circuit simulations of 1:1:1 SRAM with SNT at fast/slow/typical corners are carried out. The corresponding static noise margins (SNM’s) and read currents are extracted, and plotted in Figure 12. As expected there is anti-correlation is observed between SNM variation and read current variation, this is due to the fact that a stronger pulldown/passgate current will lead to a higher voltage on the internal SRAM cell node – causing a reduction in overall stability. Although the read current reduces by a factor of two across corners it is the extremely low SNM at the fast corner which is of most concern, indicating that the 1:1:1 SRAM cell may have some stability issues.

5. Nanowire transistor scaling beyond 5nm

In order to evaluate the performance of ultra-scaled “beyond 5 nm” nanowires, in this section we report PS solutions of three SNT’s with the circular diameter of 8nm, 6nm, and 4nm. Shown below, in Figure 13, is the carrier density profile in the cross section for the SNT in the <110> channel direction. Importantly, the mobile charge peaks move from the interface towards the center of the channel when the cross-section decreases. Also with the decreasing of the diameter the distribution of the charge in the channel becomes more uniform. This is more evident from the cut-line plot in Fig. 14. For example at 4 nm wire the cut lines along the two principal axes of the devices are almost identical. This is not valid for the wires with 6 nm and 8 nm in the case of <110> channel direction.

Figure 13: The electron density of circular SNT’s with diameters of 8nm (left), 6nm (middle) and 4nm (right) at the same high gate bias.

In addition, in Fig. 14 we compare nanowires with two channel directions of <110> and <100>. In the <100> case, the vertical and lateral profile are identical and the curves are overlapped. This is the reason why on the plot only one is visible. However, in the <110> channel direction, there is a clear difference of the charge profile along the principal axes. The main reason for this is the different effective

masses along the vertical and lateral direction of the picture. Also, for example in 8 nm case, the vertical charge distribution shows higher peaks which are closer to the interface in comparison to the lateral profile. Also with decreasing of the size of the wire both curves become more similar. For example, at 4 nm both vertical and lateral profiles become almost identical. In future work, we will investigate the interplay between these new physical features, the variability and reliability issues in such ultra scaled nanowire transistors.

Figure 14: 1D charge distributions for the circular cross-section NWT of 4nm, 6nm, and 8nm for both the <100> and <110> devices. The dashed line represents the plot over the vertical line of the wire in the “height” (z direction) and the solid line is along the “width” (y direction). For <100> wires the solid and dashed lines overlapped.

6. Conclusions

In this paper we carried out a comprehensive TCAD simulation study on nanowire transistor solutions for 5nm and beyond. The comprehensive calibration procedure is implemented to DD module using PS and EMC simulations over nanowire transistors. The electrostatic performance driven SNT selection is carried out. It is followed by simulations of nanowire transistors with various sources of statistical variability. The results reveal that wire edge roughness is a new and important variability source which needs to be considered. The statistical compact models are built and used in the statistical SRAM simulations. The nanowire transistor proves to be a solid technology solution for 5nm SRAM design. The scaling of nanowire transistors for beyond 5nm technologies illustrates new and challenging physical features, which requires more intensive TCAD early study.

7. References

- [1] S. Natarajan, et al., “A 14nm logic technology featuring 2nd-generation FinFET transistors, air-gapped interconnects, self-aligned double patterning and a 0.0588um² SRAM cell size,” in Proc. IEDM, 2014, pp.3.7.1-3.7.4
- [2] X. Wang, et al., “Interplay Between Process-Induced and Statistical Variability in 14-nm CMOS Technology Double-Gate SOI FinFETs,” IEEE

Transactions on Electron Devices, Vol. 60, No. 8, pp. 2485–2492, Aug. 2013.

- [3] S. R. Mehrotra, K. SungGeun, T. Kubis, M. Povolotskyi, M. S. Lundstrom, and G. Klimeck, "Engineering Nanowire n-MOSFETs at $L_g < 8$ nm," Electron Devices, IEEE Transactions on, vol. 60, pp. 2171–2177, 2013.
- [4] S. D. Suk, S.-Y. Lee, E.-J. Yoon, S.-M. Kim, M. Li, et al., "High performance 5nm radius Twin Silicon Nanowire MOSFET (TSNWFET) : fabrication on bulk si wafer, characteristics, and reliability," in Electron Devices Meeting, 2005. IEDM Technical Digest, 2005, pp. 717–720.
- [5] A. Asenov, et al., "Variability Aware Simulation Based Design-Technology Cooptimization (DTCO) Flow in 14 nm FinFET/SRAM Cooptimization," IEEE Transactions on Electron Devices, Vol. 62, No. 6, pp. 1682–1690, June 2015.
- [6] A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub-0.1 μm MOSFET's: A 3-D 'atomistic' simulation study," IEEE Trans. Electron Devices, vol. 45, no. 12, pp. 2505–2513, Dec. 1998.
- [7] A. Asenov, A. R. Brown, J. H. Davies, S. Kaya, and G. Slavcheva, "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs," IEEE Trans. Electron Devices, vol. 50, no. 9, pp. 1837–1852, Sep. 2003.
- [8] G. Roy, A. R. Brown, F. Adamu-Lema, S. Roy and A. Asenov, "Simulation study of individual and combined sources of intrinsic parameter fluctuations in conventional nano-MOSFETs," IEEE Transactions on Electron Devices, Vol. 53, No. 12, pp. 3063–3070, 2006.
- [9] A. R. Brown, N. M. Idris, J. R. Watling, and A. Asenov, "Impact of metal gate granularity on threshold voltage variability: A full-scale three dimensional statistical simulation study," IEEE Electron Device Lett., vol. 31, no. 11, pp. 1199–1201, Nov. 2010.
- [10] Y. Wang, et al., "Simulation Study of the Impact of Quantum Confinement on the Electrostatically Driven Performance of n-type Nanowire Transistors," IEEE Transactions on Electron Devices, Vol. 62, No. 10, pp. 3229–3236, Oct. 2015.
- [11] X. Wang, B. Cheng, D. Reid, A. Pender, P. Asenov, C. Millar and A. Asenov, "FinFET Centric Variability-Aware Compact Model Extraction and Generation Technology Supporting DTCO," IEEE Transactions on Electron Devices, Vol. 62, No. 10, pp. 3139–3146, Oct. 2015.