Drain Current Computation in Nanoscale nMOSFETs: Comparison of Transport Models

E. Sangiorgi, C. Alexander, A. Asenov, V. Aubry-Fortuna, G. Baccarani, A. Bournel, M. Braccioli, B. Cheng, P. Dollfus, A. Esposito, D. Esseni, C. Fenouillet-Beranger, C. Fiegna, G. Fiori, A. Ghetti, G. Iannaccone, A. Martinez, B. Majkusiak, S. Monfray, P. Palestri, V. Peikert, S. Reggiani, C. Riddet, J. Saint-Martin, A. Schenk, L. Selmi, L. Silvestri, P. Toniutti, J. Walczak

Abstract - In this paper the modelling approaches for determination of the drain current in nanoscale MOSFETs pursued by various partners in the frame of the European Projects Pullnano and Nanosil are mutually compared in terms of drain current and internal quantities (average velocity and inversion charge). The comparison has been carried out by simulating template devices representative of 22 nm Double-Gate and 32 nm Single-Gate FD-SOI. A large variety of simulation models has been considered, ranging from drift–diffusion to direct solutions of the Boltzmann-Transport-Equation. The predictions of the different approaches for the 32 nm device are quite similar. Simulations of the 22 nm device instead, are much less consistent. Comparison with experimental data for a 32 nm device shows that the modeling approach used to explain the mobility reduction induced by the high-k dielectric is critical.

I INTRODUCTION

In the frame of the European Projects Pullnano and Nanosil, different research groups have developed and mutually compared [1] innovative device simulation models for the deter-mination of the drain current Ids in MOSFETs. One of the reasons driving these modeling efforts is the industry need to understand performance improvements due to quasi-ballistic transport and other technology boosters such as strain, high-j dielectrics and ultra-thin-body Silicon-On-Insulator (SOI) architectures [2]. The possible modeling approaches can be grouped in a few families which range from modifications of the conventional drift–diffusion (DD) model used in commercial TCAD tools to advanced Monte-Carlo [3]

E. Sangiorgi, G. Baccarani, C. Fiegna, S. Reggiani, M. Braccioli, L. Silvestri are with University of Bologna – IUNET
C. Alexander, A. Asenov, B. Cheng, A. Martinez, C. Riddet are with the University of Glasgow
V. Aubry-Fortuna, A. Bournel, P. Dollfus, J. Saint-Martin are with University Paris–Sud
D. Esseni, P. Palestri, L. Selmi, P. Toniutti, , are with Univ. of Udine – IUNET
A. Esposito, V. Peikert , A. Schenk are with ETHZ
C. Fenouillet-Beranger, S. Monfray are with ST Microelectr.
G. Fiori and G. Iannaccone are with Univ. of Pisa – IUNET
A. Ghetti is with Numonyx
B. Majkusiak, J. Walczak are with Warsaw Univ. of Technol.

(MC) and (MC) and Non-Equilibrium-Green's-Function (NEGF) simulators [4] able to handle the strongly off-equilibrium transport taking place in deca-nanometric devices.

A thorough assessment of these models is not trivial because direct comparison with experimental data is often unable to rule out possible model inaccuracies, since many parameters of the experimental devices, such as doping profiles and series resistances, which play a critical role in determining I_{DS}, are not precisely known and are often used as adjusting parameters.

Comparison between simulations of the same devices performed with different models [5-7] represents a simple and sound methodology to identify and quantify the impact of the assumptions taken by the different models. In this work we have first defined template devices: a 32 nm Fully-Depleted-SOI (FDSOI) and a 22 nm Double-Gate (DG) device, both optimized for low-stand-by-power applications. Then we have simulated them with the available modeling approaches, all previously calibrated on the universal mobility curves [8]. Results in terms of low-field mobility, drain current and internal quantities (concentration and velocity) have been compared. This provides us an estimate of the degree of convergence between the different transport models in aggressively scaled devices.

II SIMULATED DEVICES

The 32 nm FDSOI template is sketched in Fig. 1. The channel is lowly doped. The substrate is p-type (N=10^{18} cm^{-3}). The metal work-function is 4.6 eV. The gate stack consists of 2.3 nm of HfO2 on top of 0.8 nm of SiO2 (EOT = 1.2 nm). The spacer is made of Si3N4 and comes in direct contact to the silicon. Doping profiles for the S/D regions have been obtained from process simulations of a realistic 32 nm process.

The 22 nm DG device is an idealized Double-Gate MOSFET with a gate length of 22 nm, a gate stack consisting of 2.4 nm of HfO2 on top of 0.7 nm of SiO2 (EOT = 1.1 nm). The silicon film thickness is 10 nm and the metal work-function is 4.8 eV. The doping profiles are

similar to the ones of the 32 nm template, with all the diffusion lengths scaled by 22/32.

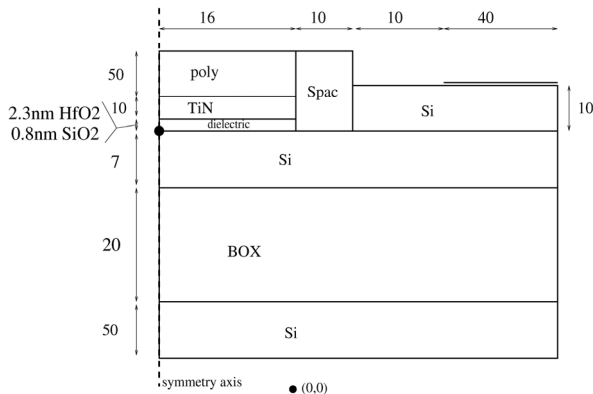Both templates are n-type and feature unstrained Si channels.

Fig. 1 Structure of the 32 nm FDSOI template transistor used in this work. Only one half of the symmetric structure is reported. All dimensions are in nm.

III SIMULATION APPROACHES

In the following, the key features of each model (identified with the acronym of the main developer) are presented. We group the models in two families: the MC family, which collects models based on the direct solution of the Boltzmann-Transport-Equation (BTE) using the Monte-Carlo method, and the DD family, which gathers drift–diffusion-like models where only the first momenta of the BTE are calculated.

A. MC simulators

UD-MSMC. Multi-subband ensemble Monte-Carlo described in [9]. It provides the coupled solution of the effective-mass Schrödinger equation in each section of the device, of the system of coupled BTEs foreach subband in the in version layer and of the 2D Poisson equation. In this way quantization effects such as charge repulsion from the channel/dielectric interface, subband repopulation, dependence of the scattering rate on the size- and bias-induced quantizationare naturally taken into account. An analytical non-parabolic model is used for the energy dispersion of the subbands. Scattering mechanisms included in the solution of the BTE are bulk phonons and surface roughness (SR). Ionized impurity (II) scattering in the S/D extensions is not active in these simulations, but series resistances extracted from DD simulations have been introduced as lumped elements. Vertical S/D contacts are placed just at the end of the spacers .

BO-MC. Full-band ensemble Monte-Carlo (free carrier gas) [10] with quantum corrections (effective potential). Scattering mechanisms include phonons, SR, II as well as carrier-plasmon in the S/D. II scattering in the S/D is calibrated to reproduce bulk mobility data for

doping up to 10^{21} cm^{-3}.ETH-MC. Full-band ensemble Monte-Carlo (free carrier gas) with phonon, II and SR scattering [11].

Numonyx-MC. Full-band ensemble Monte-Carlo [12] (free carrier gas) featuring quantization effects through a quantum mechanical correction of the potential that is computed by solving self-consistently the Schrödinger equation in each section of the device. The silicon anisotropic full-band structure is computed with the Empirical Pseudopotential Method. Scattering mechanisms are assumed to be isotropic and include: elastic acoustic phonon scattering, inelastic optical phonon scattering, II scattering, impact ionization. SR scattering is treated by including both surface roughness and surface phonon scattering mechanisms as a function of the average electric field weighted by the carrier concentration.

IEF-MC. Ensemble Monte-Carlo described in [13]. Quantum corrections are not taken into account here and carriers are treated as a three-dimensional (free) gas in the simulator. We consider an analytical conduction band structure of silicon consisting of six ellipsoidal non-parabolic D valleys located along the [100] directions at 85% of the Brillouin zone edge. All relevant scattering mechanisms are included, i.e. electron–phonon, II and SR scattering,

UGLA-MC. 3D Monte-Carlo simulator [14]. An efficient methodology is used for the fully self-consistent inclusion of 3D density gradient (DG) quantum corrections. Efficient analytic ellipsoidal, nonparabolic band models are employed and all major phonon mechanisms required to calibrate to bulk mobility in silicon are included. Within device simulation, carriers are treated as a free carrier gas and II scattering is included as in bulk via the Brooks–Herring formalism.

B. DD simulators

BO-QDD 1D drift–diffusion solver for SOI-MOSFETs combined with the solution of the coupled Schrödinger–Poisson equations on the device cross-sections normal to the transport direction [15]. The physical model thus accounts for the quantization due to both the structural confinement and the application of the transverse effective field by realistically computing the device electrostatics.

The harmonization of the drift–diffusion model with the Schrödinger equation is pursued by means of Bohm's theory of quantum potential. The model requires the solution of as many drift–diffusion equations as the number of populated subbands. A physically- based unified mobility model has been incorporated in the QDD solver, which is an analytical function of the effective field and doping concentration. The model provides the effective mobility of the 2DEG in a SOI MOSFET channel by averaging the single-valley mobilities weighted with their respective valley populations. Elliptic parabolic bands for the six conduction valleys are assumed. The lowest

subband energies of each valley derived from the Schrödinger–Poisson solver are used to calculate the relative valley population by assuming Boltzmann statistics.

UGLA-aDD. 3D atomistic drift–diffusion simulator. It employs density gradient quantum corrections.

PI-MSDD. Multi-subband DD, i.e. self-consistent solution of the 2D Poisson and Schrödinger equations (in the direction perpendicular to the Si/SiO2 interface), coupled with the solution of the continuity equation along subbands in the DD approximation.

C. Other transport approaches

UGLA-NEGF. Modified version [16] of the fully 2D NEGF simulator initially developed by NASA. In this work scattering is not included, so that carriers move ballistically from source to drain.

WUT. Electron mobility model based on the relaxation time approximation, employing the Matthiessen's rule for different scattering mechanisms. It uses a 1D Poisson-Schrödinger solver which can handle both open and closed boundary conditions for the wave functions, and also different co-existing potential wells (or channels) may be considered independently.

IV MODEL CALIBRATION

The models described above differ in terms of band-structure, scattering models, treatment of non-local transport, etc. For the sake of a fair comparison, all simulators have been first calibrated to reproduce the universal curves in bulk Si devices .

At the time of the comparison, not all models contained all the ingredients to simulate advanced devices as the template transistors previously defined, In particular, some handle strained channels but not high-j stacks. Furthermore, scattering models for options such as high-j dielectrics are not well assessed yet, since there is still a debate about the relative contribution of remote phonons and remote-Coulomb scattering. For these reasons, although the template devices include high-j stacks, and although 32 nm and 22 nm devices are likely to include strained channels, when simulating the template devices we consider unstrained Si and neglect the scattering mechanisms induced by the presence of the high-j dielectric.

V RESULTS

In this section we report the results obtained by simulating the template devices of Section II with the models described in Section III. In all the following figures we have used a consistent set of symbols, so that each model is always identified by the same symbol and type of line.
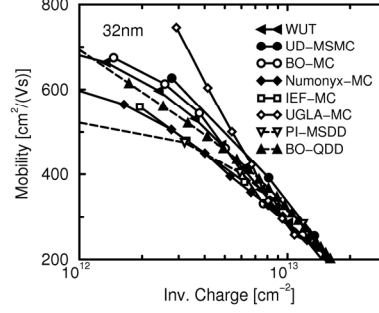
Fig. 2 Simulated low-field mobility of the 32 nm FDSOI template. Being the device undoped, an inversion charge of 10^{13} cm^{-2} corresponds to an effective field of 760 kV/cm.

A. Low field mobility

Figs. 2 and 3 report the low-field mobility as computed in long channel devices with the same vertical structure as the 32 nm and 22 nm templates. The mutual agreement between the different models is quite good at large inversion charges Ninv, in particular in the 32 nm FDSOI template, whereas discrepancies appear at low Ninv, especially in the 22 nm DG device.
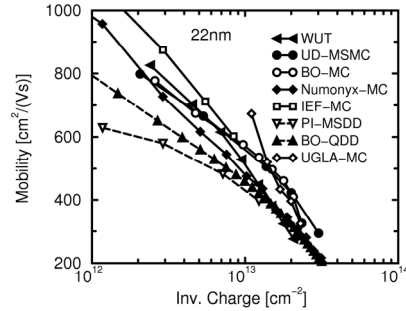
Fig. 3 Simulated low-field mobility of the 22 nm DG template. The inversion charge includes both channels (front and back interface).

B. Drain current in the 32 nm template

Figs. 4 report I/V curves of the 32 nm FDSOI template at high drain–source voltages. Considering the models of the DD family, the figures show a more than satisfactory mutual agreement, that has been observed also below threshold and at low drain–source voltages (not shown)
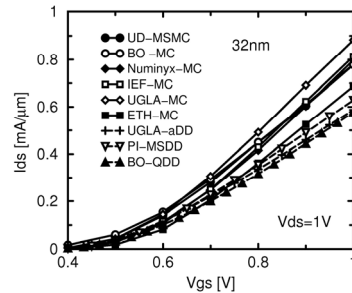
Fig. 4 Trans-characteristics of the 32 nm FDSOI template for VDS=1 V.

Considering now MC models, which take into account more accurately the quasi-ballistic nature of carrier transport in short MOSFETs, the mutual agreement is quite satisfactory, much better of what has been found in [7] mainly because in the 32 nm FDSOI device considered in this work the role of II scattering in the S/D regions is significantly reduced with respect to the devices in [7]. It is also interesting to note that different treatments of quantization (MSMC vs. quantum corrections vs. no quantization) and of different descriptions of the band-structure (full-band vs. simple non-parabolic analytical bands) only have a marginal impact on the simulated current of this device. As expected, the current provided by the MC models is larger than the one given by the DD ones at high VDS, where non-equilibrium effects become significant. At low VDS, instead, the two approaches give essentially the same current, as it is expected since the device works close to equilibrium (not shown). In some cases (ETH-MC vs. UGLA-aDD) current from DD is larger than from MC, consistently with the failure of DD models also near equilibrium.

C. Drain current in the 222 nm template

Comparisons between the MC, DD and NEGF results for the 22 nm DG template are reported in Fig. 5. Concerning the DD models, the overall agreement is essentially as good as for the 32 nm device.

Concerning the MC models, the agreement between IDS predictions is not very satisfactory. Possible explanations can be traced back to the different modeling of SR and phonon scattering in thin film Double-Gate SOI structures, since we have seen that also the differences in low-field mobility in this device are significant and the device works at lower effective field compared to the 32 nm FDSOI template.

Since in the 22 nm DG device the impact of II in the S/D regions is large (the series resistances extracted from DD simulations are 90 $\Omega\mu$), we have performed MC simulations without II scattering to isolate the effect of the various scattering mechanisms on the spread between the simulation results. As it can be seen in Fig. 6, without II the spread between the MC results is smaller than in Fig. 5, but still significant, meaning that the different treatment of II scattering implemented in the models is only one of the reasons for the spread between the MC results.
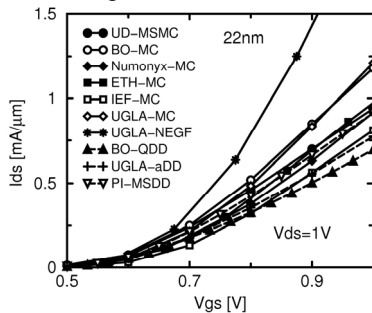

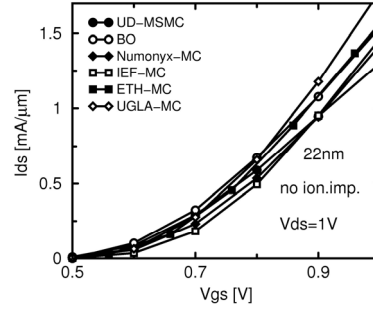Fig. 5 Trans-characteristics of the 22 nm DG template

Fig. 6 Trans-characteristic of the 22 nm DG template. Results from MC simulations without ionized impurity scattering.

VI INTERNAL QUANTITIES

To further investigate the origin of the discrepancies between the different modeling approaches, we have compared internal quantities (inversion charge and average velocity profiles) as obtained with the various simulators. Typical results are reported in Figs. 7 and 8.

The correlation between the spread of the drain currents and the spread in terms of average velocity is not so clear. The difference in the average velocity predicted by many models over a large fraction of the channel is often even larger than that in the corresponding drain currents. In fact it is the velocity near the injection point (the so-called virtual source) that essentially controls the current drive of the device. In this respect, we see that the velocity in the DD models is limited to the saturation velocity, whereas the MC models feature peak velocities that can be more than two times larger, but the differences in terms of IDS are significantly smaller. On the other hand these differences in terms of velocity have a large impact in the determination of the cut-off frequency of the devices. The differences in terms of inversion charge that can be observed in Fig. 8 in the central portion of the channel and close to the drain junction can be interpreted as differences in terms of velocity. In fact, also in the case of Ninv, it is the value at the virtual source that really controls the current. We have thus collected the inversion charge and average velocity at the virtual source for some of the modeling approaches. However also this comparison does not help too much in understanding the origin of the different model predictions. In fact, being all the approaches self-consistent, they provide different potential profiles, and thus different positions of the virtual source.
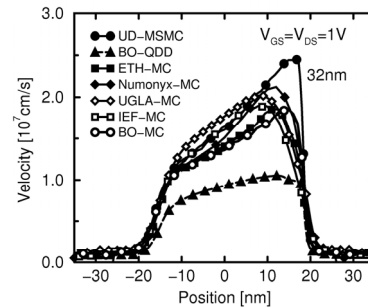

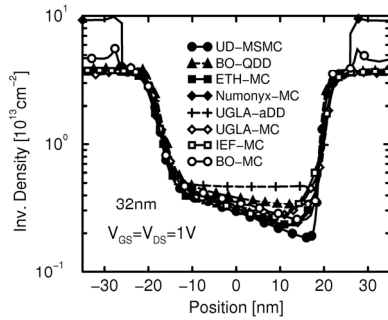Fig. 7 Velocity profiles along the channel in the 32 nm FDSOI biased at VGS=VDS=1 V.

Fig.8 Inversion charge profiles along the channel in the 32 nm FDSOI template biased at VGS=VDS= 1 V.

VII CONCLUSION

The extensive comparison presented in this work has interested four DD simulators, six MC simulators, one NEGF solver and a model for the computation of the low-field mobility. The model predictions tend to converge for the longer channel devices, whereas the predictions of the scaling trends of on-current improvement are quantitatively quite different among the models. Comparison with ballistic NEGF results, points out that even with a limited number of scattering mechanisms accounted for (II, phonons, SR) scattering still plays a remarkable role in decananometric devices. The impact of scattering on IDS becomes even larger when specific mechanisms needed to reproduce the low-field mobility of advanced devices (e.g. remote-charges in the high-j) are included in the models.

We finally emphasize that simulations of more mature technologies (e.g. 32 nm compared to 22 nm one) yields more similar predictions from the different simulators.

REFERENCES

[1]] P. Palestri et al., A comparison of advanced transport models for the computation of the drain current in nanoscale nMOSFETs Solid State Electron 2009; 53:1293-1302.

[2] International Technology Roadmap for Semiconductor; 2007. <http://public.itsr.net>.

[3] C. Jacoboni, L. Reggiani The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials. Rev Mod Phys 1983;55:645–705.

[4] S. Datta, The non-equilibrium Green's function (NEGF) formalism: an elementary introduction. In: IEDM technical digest; 2002. p. 703–6.

[5] A. Abramo et al. A comparison of numerical solutions of the Boltzmann transport equation for high-energy electron transport in silicon. IEEE Trans Electron Dev 1994; 41:1646–54.

[6] P. Palestri et al., Comparison of modeling approaches for the capacitance–voltage and current–voltage characteristics of advanced gate stacks. IEEE Trans Electron Dev, 2007;54(1):106–14..

[7] C. Fiegna et al., Comparison of Monte Carlo transport models for nanometer-size MOSFETs. In: Proceedings SISPAD; 2007. p. 57–60.

[8] S. Takagi, Toriumi A, Iwase M, Tango H. On the universality of inversion-layer mobility in Si MOSFETs. Part I – effect of substrate impurity concentration. IEEE Trans Electron Dev 1994;41(12):2357–62.

[9] L. Lucci et al., Multi-subband Monte-Carlo study of transport, quantization and electron gas degeneration in ultra-thin SOI n-MOSFETs. IEEE Trans Electron Dev 2007;54(5):1156–64.

[10] P. Palestri et al., An improved semiclassical Monte-Carlo approach for nano-scale MOSFET simulation. Solid State Electron 2005;49:727–32.

[11] F. Bufler et al,. Efficient Monte Carlo device modeling. IEEE Trans Electron Dev 2000;47(10):1891–7.

[12] A. Ghetti et al.,. 3D Monte Carlo device simulation of nanowire MOSFETs including quantum mechanical and strain effects. In: Proceedings SISPAD; 2006. p. 67.

[13] D. Querlioz et al., On the ability of the particle Monte Carlo technique to include quantum effects in nano-MOSFET simulation. IEEE Trans Electron Dev 2007;54(9):2232–42.

[14] CL Alexander et al., Increased intrinsic parameter fluctuations through ab initio Monte Carlo simulations in nano-scaled MOSFETs. In: IEDM technical digest; 2006

[15] G. Baccarani et al., Theoretical foundations of the quantum drift–diffusion and density-gradient models. Solid State Electron 2008;52(4):526–32..

[16] A.Martinez et al. A self-consistent fully 3-d real-space NEGF simulator for studying Nonpertubative effects in nano-MOSFETs. IEEE Trans Electron Dev 2007;54(9):2213–22.