electronics

MDPI

Article

Modeling Nanoscale III–V Channel MOSFETs with the Self-Consistent Multi-Valley/Multi-Subband Monte Carlo Approach

Enrico Caruso [1,2], David Esseni [1], Elena Gnani [3], Daniel Lizzit [1], Pierpaolo Palestri [1,*], Alessandro Pin [1], Francesco Maria Puglisi [4], Luca Selmi [4] and Nicolò Zagni [4]

1 Polytechnic Department of Engineering and Architecture, University of Udine, 33100 Udine, Italy; Enrico.Caruso@infineon.com (E.C.); david.esseni@uniud.it (D.E.); daniel.lizzit@uniud.it (D.L.); alessandro.pin.86@gmail.com (A.P.)
2 Infineon Technologies Austria AG Siemensstrasse 2, 9500 Villach, Austria
3 ARCES Research Center and Department of Electrical Engineering (DEI), University of Bologna, 40136 Bologna, Italy; elena.gnani@unibo.it
4 Department of Engineering "Enzo Ferrari", University of Modena and Reggio Emilia, via P. Vivarelli 10, 41125 Modena, Italy; francescomaria.puglisi@unimore.it (F.M.P.); luca.selmi@unimore.it (L.S.); nicolo.zagni@unimore.it (N.Z.)
* Correspondence: pierpaolo.palestri@uniud.it; Tel.: +39-0432-55-82-49

Abstract: We describe the multi-valley/multi-subband Monte Carlo (MV–MSMC) approach to model nanoscale MOSFETs featuring III–V semiconductors as channel material. This approach describes carrier quantization normal to the channel direction, solving the Schrödinger equation while off-equilibrium transport is captured by the multi-valley/multi-subband Boltzmann transport equation. In this paper, we outline a methodology to include quantum effects along the transport direction (namely, source-to-drain tunneling) and provide model verification by comparison with Non-Equilibrium Green's Function results for nanoscale MOSFETs with InAs and InGaAs channels. It is then shown how to use the MV–MSMC to calibrate a Technology Computer Aided Design (TCAD) simulation deck based on the drift–diffusion model that allows much faster simulations and opens the doors to variability studies in III–V channel MOSFETs.

Keywords: III–V semiconductors; modeling and simulation; Monte Carlo

Citation: Caruso, E.; Esseni, D.; Gnani, E.; Lizzit, D.; Palestri, P.; Pin, A.; Puglisi, F.M.; Selmi, L.; Zagni, N. Modeling Nanoscale III–V Channel MOSFETs with the Self-Consistent Multi-Valley/Multi-Subband Monte Carlo Approach. *Electronics* **2021**, *10*, 2472. https://doi.org/10.3390/electronics10202472

Academic Editor: Yuning Li

Received: 8 September 2021
Accepted: 6 October 2021
Published: 12 October 2021

1. Introduction

Despite being investigated and employed for a long time as semiconducting materials for radio-frequency/microwave applications and for opto-electronics devices, only recently III–V compounds and alloys have attracted interest as possible replacements for silicon as channel material for nanoscale MOSFETs [1–5].

Device modeling and simulation plays a key role in technology development and optimization. This task becomes particularly complicated in nanoscale MOSFETs with III-V materials in the channel. In fact, selecting the right compound (or alloy), devising a layered structure to help carrier confinement and reduce short-channel effects (SCE), selecting the right gate stack, and considering the presence of interface and border traps in the dielectrics require specific models that significantly deviate from the TCAD models developed for silicon devices.

Full-quantum simulators based on Non-Equilibrium Green's Function (NEGF) formalism [6–8] are suited to fully capture the quantization effects normal to transport (carrier confinement at the semiconductor/dielectric interfaces as well as confinement in quantum wells) and along the transport direction (i.e., source-to-drain tunneling, SDT, and band-to-band tunneling, BTBT). These are mandatory model ingredients in order to optimize the device in terms of short-channel effects and low off-current. The use of tight binding

(TB) or density functional theory (DFT) Hamiltonians allows for a full-band description of the dispersion relationship of most of the III–V compounds and alloys, while the k·p approach [9] provides an accurate description of the warped nature of the valence band and of the non-parabolic nature of the Γ valley of the conduction band. Simpler effective mass (EMA) approaches can accurately describe the non-parabolic/multi-valley features of the band structure once calibrated against the advanced models mentioned above [10]. On the other hand, the computational burden of full-quantum simulations is significant and precludes the inclusion of a complete set of scattering mechanisms. Furthermore, the approach is usually applicable only to devices with short channel length and very thin quantum wells. It is thus hard to calibrate the model on experimental data that are usually available for devices with fairly long channel. The full quantum NEGF approaches are thus at the top of the model hierarchy outlined in Figure 1, since they provide the most accurate description of quantum effects.

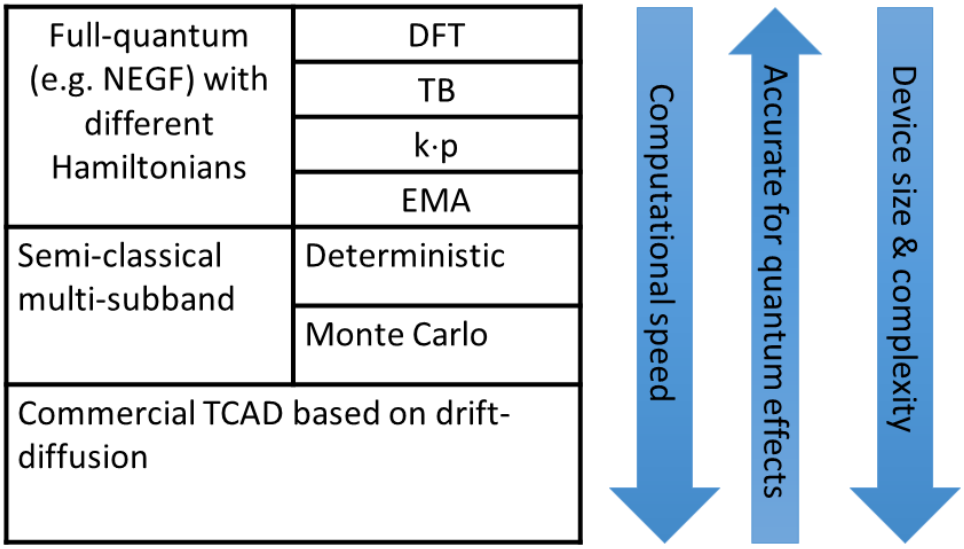


Figure 1. Model hierarchy for III–V channel nanoscale MOSFETs.

On the opposite side of the model hierarchy of Figure 1, we find the commercial TCAD tools based on the drift–diffusion approach (e.g., [11]). They can handle very complex device geometries going from planar to 3D architectures with non-planar access regions and non-negligible contact resistance. Accurate models for traps in the gate stack can be activated [12]. The drift–diffusion model, however, does not naturally account for carrier quantization and quasi-ballistic transport. In recent years, models have been proposed to include quantization effects [13] as well as high-field transport in short-channel MOSFETs [14]. These models open the door to accurate device design and optimization spanning a large set of parameters not possible with NEGF simulators, as well as variability studies [15]. However, the various model ingredients require calibration against more accurate (although time consuming) simulation approaches.

In the middle of the transport model hierarchy of Figure 1, we have the semi-classical models based on the multi-subband framework. In this approach, quantization normal to the transport direction is handled by solving the Schrödinger equation in each device slice and then solving the Boltzmann Transport Equation (BTE) along the channel direction using the derivative of the subband energy as the driving force [16,17]. The problem can be tackled either with the Monte Carlo method [18,19] or with deterministic approaches [20,21]. In the former case, we obtain what is known as Multi-valley/Multi-subband Monte Carlo (MV–MSMC) [17]. The label "Multi-valley" indicates that subbands can originate from quantization from different minima (valleys) of the conduction band. It is understood that all these models should be solved self-consistently with the Poisson equation.

In this paper, we review the MV–MSMC approach originally developed for silicon-channel MOSFETs and show the main model upgrades needed to describe phenomena relevant in MOSFETs with III–V compounds and alloys in the channel from the standpoint of the intrinsic material's transport properties. Compared to other papers by our group [22–25], here we show how to empirically include source-to-drain tunneling in the MV–MSMC and calibrate the model parameters vs. NEGF simulations. In addition, we use the MV–MSMC to calibrate the quantization and transport modeling parameters of a TCAD simulator to allow variability studies in III–V channel MOSFETs. Although unpassivated interface and border traps may have a significant impact on the electrical performance of III–V MOSFETs [24], they are not considered here, to highlight the maximum upper performance limits achievable thanks to the intrinsic transport properties of the material.

2. Materials and Methods

The Multi-Subband Monte Carlo method has been extensively described elsewhere [16–18]. Although the approach can be used also for hole transport [26], here we review the model for electron transport, considering only n-type MOSFETs. Since the conduction band of many semiconductors features many minima relevant for electron transport, each of these minima generates its own set of subbands in the presence of confining fields as in MOS structures. For this reason, we will refer to the model as "multi-valley/multi-subband Monte Carlo" (MV-MSMC). Although the methodology may be used for devices of the nanowire type [27], here we consider planar structures featuring an essentially 2D electron gas (i.e., carriers are free to move in the transport plane and quantized in the direction normal to the transport plane). We also assume that the device is uniform in the width direction, so that the electrostatic problem is 2D.

A simplified flowchart of the MV–MSMC code is shown in Figure 2. Starting from a 2D potential profile (i.e., that varies along the channel direction x and along the quantization direction z), the 1D Schrödinger equation (SE) is solved along z in the different sections along the channels for each set of valleys. Penetration of the wave function into the gate dielectric can be accounted for. The derivative of the subband energy (for the given valley and subband index) is then used as the driving force in the set of BTEs (one for each subband) along the channel direction x. The BTE also requires computing the scattering rates for the electrons in the inversion layer [17]: they depend on the subband energy and corresponding wave functions. Inter-subband scattering couples the BTEs for the different valleys/subbands. The solution of the BTE provides the occupation of the subbands in the sections along the channel. The squared modulus of the wave function associated with the subbands describes how the charge is distributed along z. It is thus possible to derive the 2D charge concentration profile. This is input into the 2D Poisson equation to derive a new guess for the potential profile and iterate the whole procedure until convergence is reached.

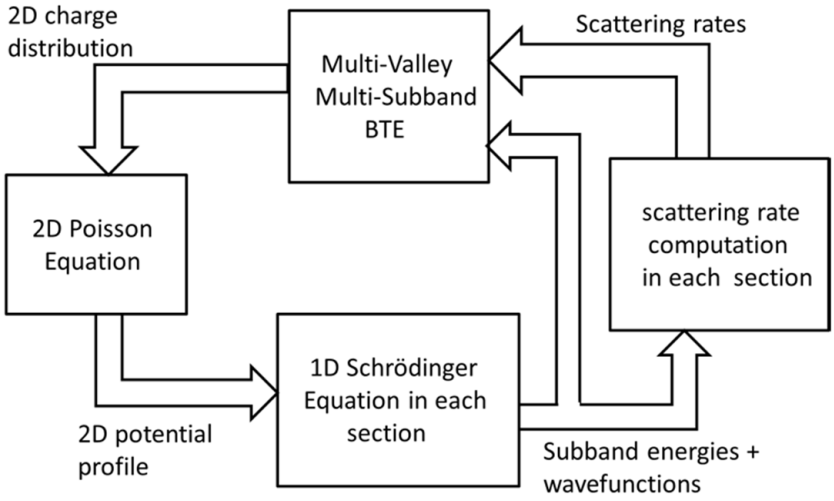


Figure 2. Simplified flowchart of the multi-valley/multi-subband Monte Carlo method in the spatially two-dimensional implementation used in this work.

The MV–MSMC code used in this work was originally conceived for silicon devices [16,18]. Here, we list the main add-ons included to simulate MOSFETs with III–V channel.

First of all, the Γ-valley of most III–V materials is strongly non-parabolic (see, for example, Figure 3 for InAs). While in silicon the non-parabolicity of the valleys can be added as a simple correction of the in-plane energy dispersion that does not depend on the subband energy [17], in III–V materials one should apply parabolic corrections to the subband value and use an in-plane energy dispersion that depends on the subband energy itself [10,28].

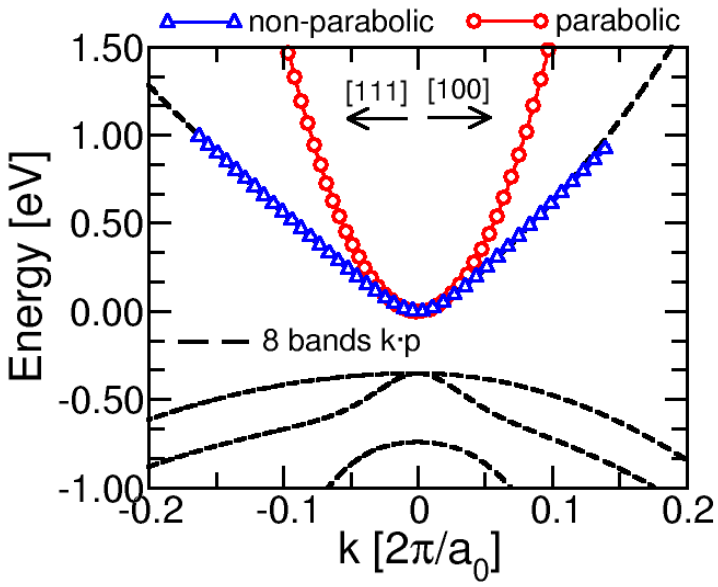


Figure 3. Dispersion relationship of the Γ valley of InAs. The 8-band k·p of [9] (with the parameters for InAs from [29]) is compared with the parabolic and non-parabolic analytical expressions.

In addition, the density of states of the Γ valley is very small, so that the Pauli's exclusion principle has to be enforced in order to recover Fermi statistics at equilibrium. This means rejecting scattering events based on the occupation of the final state [17]. This is needed also in Si devices, but in III–V MOSFETs this has an additional effect on the device electrostatics, so that the effective channel length differs from its geometrical value [30].

The offset between the Γ valley and the satellite valleys is small, so that the latter have a significant occupation, and this demands for a multi-valley approach. This is

of fundamental importance when considering high-field transport [31,32] as well as in quantum wells, where the higher quantization mass of satellite valleys makes the respective subbands very close in energy to the subbands originating from the Γ valley (e.g., in (111)-GaAs [33]).

The small energy barrier between the III–V channel and the high-k dielectric in the gate stack makes the wave function penetration into the gate stack relevant. This has an effect on the electrostatics, as we will see later in Section 3.2. Furthermore, the model for surface roughness scattering should account for that, and the linear model [34] should be replaced by non-linear expressions [35]. Overall, the calculation of the scattering rates in III–V MOSFETs has a significant computational burden due to the multi-valley nature of the material, the need for non-linear surface roughness, and the need to include polar phonons where the computation of the matrix element requires multiple integrals in the quantization direction [17].

Due to the small energy gap of most III–V compounds and alloys proposed as channel materials for nanoscale MOSFET (i.e., In-rich InGaAs, GaSb, InSb) tunneling (direct and band-to-band) has a stronger impact on the device characteristics than in silicon devices, in particular, in the sub-threshold region of operation. These effects are not naturally included in the MV–MSMC, where quantum effects are restricted to carrier quantization normal to the transport direction. However, many workarounds have been proposed. In [36], direct source-to-drain-tunneling (SDT) has been included by accounting for tunneling (via WKB formalism) of electrons that hit the classical turning point. Our group instead has proposed to convolve the subband energy profiles with a Gaussian function mimicking the finite size of the wave packet [37], similar to what proposed in [38] for semiclassical Monte Carlo. In this way, the charge associated with the tunneling carriers is implicitly included in the device electrostatics. The approach has been verified against full-quantum simulations for silicon devices [37]. In this paper (Section 3.1), we will show that the approach can be used also in III–V MOSFETs where tunneling phenomena across the source barrier are more severe. Concerning band-to-band tunneling (BTBT), it can be added as a generation mechanism. The reader is referred to [39]. In this work, we neither consider tunnel FETs nor apply drain-to-source voltages high enough to induce BTBT in MOSFETs. Consequently, BTBT is ignored.

The interface between III–V semiconductors and high-k gate dielectrics features a large concentration of traps with energy inside the conduction band of the semiconductor. This may result in significant Fermi-level pinning [40]; in other words, above a given point, increasing the gate voltage does not efficiently increase the inversion charge but only the population of the interface states that do not contribute to transport. These interface states can be included in the MV–MSMC, as described in [41], but are not activated in the simulations reported below.

As in nanoscale silicon MOSFETs, in III–V channel MOSFETs the source and drain series resistances have values comparable to those of the channel resistance and should be part of the device model. How to compute the contact resistance itself is not the aim of this work. An example of a model based on WKB tunneling and thermionic emission is reported in [25]. Once the value of the source (R_S) and drain (R_D) series resistances is known, they can be introduced in the MV–MSMC loop by modifying the source and drain Fermi levels as $E_{FS} = R_S I_{DS}$ and $E_{FD} = V_{DS} - R_D I_{DS}$.

3. Results

It is clear from the description in the previous section that the MV–MSMC used in this work is a very powerful and complete tool for the modeling and optimization of nanoscale MOSFETs, including the main physical mechanisms that control the device performance. However, there are two main limitations; in this section, we will analyze and propose ways to overcome them.

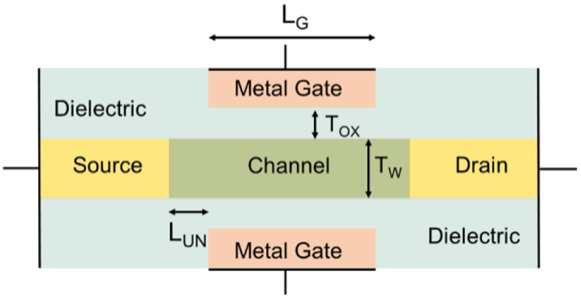
First of all, although quantum effects in the direction normal to transport are naturally accounted for by the solution of the SE, tunneling along the transport direction, that

may play a significant role in determining the off-current [42], needs to be included. In Section 3.1, we apply to III–V MOSFETs the methodology proposed in [37] and calibrate the model parameters against NEGF simulations.

Secondly, the MV–MSMC approach is based on a separation between transport direction (where the set of BTEs is solved) and quantization direction (where the SE is solved) that limits its applicability to simple device architectures. Access regions of complex shape, e.g., raised source/drain, need to be included as lumped series resistance (see above), losing the ability to describe momentum redirection and accurately model all aspects of source starvation phenomena [43]. Furthermore, the computational burden of the MV–MSMC prevents its use for extensive sensitivity/variability studies. Following our previous work in [14,15], we show in Section 3.2 how to calibrate the channel transport models in a commercial drift–diffusion TCAD based on the MV–MSMC simulation results. This step then allows us to exploit the available TCAD models to analyze more complex device architectures, consider tunneling from/to border traps, and perform AC/noise as well as variability analysis.

3.1. Inclusion of Quantum Effects along the Channel Direction in the MV–MSMC

This section reports the calibration of the empirical model for quantum effects along the channel direction implemented in the MV–MSMC [37] against NEGF simulations of a set of nanoscale MOSFETs with InAs and $In_{0.53}Ga_{0.47}As$ channels and a double-gate architecture. This structure can represent either a planar double-gate structure or the 2D cross section of a tall FinFET with a thick oxide on the top and negligible fringing fields at the bottom, such that only the lateral gates control the channel. The geometrical parameters for these devices were chosen starting from the semiconductor roadmap (ITRS) indications for Ge/III–V semiconductors [44], carefully adjusted to preserve electrostatic integrity in terms of subthreshold slope (SS) and drain-induced barrier lowering (DIBL). The devices, sketched in Figure 4, have gate lengths of 8.3, 10.4, and 15 nm. For those devices, the gate dielectric is scaled accordingly to 3.3, 3.3, and 3.84 nm. Having a relative permittivity of 22, the corresponding EOT is 0.595, 0.595, and 0.68 nm, respectively. The channel doping is 10^{17} cm^{-3}, and the source and drain doping is 5×10^{19} cm^{-3}.



Param.	Devices		
L_G [nm]	8.3	10.4	15
T_W [nm]	4	4	7
T_{OX} [nm]	3.3	3.3	3.84
L_{UN} [nm]	0	2	2

Figure 4. Sketch of the simulated devices and table with geometrical parameters.

The NEGF solver is described in [45]. It includes only the Γ valley of the conduction band, differently from the MV–MSMC that models also satellite valleys. We however verified that the occupation of satellite valleys is negligible in the simulations reported here. Since scattering is not included in the NEGF simulator, scattering was turned off in the MV–MSMC as well, for a fair comparison.

The model employed in the MV–MSMC convolves the subband energy profiles with a Gaussian function with variance σ [37]. The same function is also used to smooth the charge profile before entering it into the Poisson's equation, in order to mimic the effect of the finite size of the electron wave packet. The goal of the following figures is to show that a unique value of σ provides a reasonably good match between MV–MSMC and NEGF. To this end, the gate work functions of the devices in the two models were adjusted to match

the off-current at high V_{DS}. In this comparison we turned off the wavefunction penetration in the MV–MSMC for consistency with the NEGF solver.

The simulated I_{DS}-V_{GS} for the $In_{0.53}Ga_{0.47}As$ devices with length of 15 nm, 10.4 nm, and 8.3 nm are shown in Figures 5–7, respectively. The MV–MSMC without quantum correction (i.e., with $\sigma = 0$) shows a much steeper current increase with gate voltage, since source-to-drain tunneling is not accounted for. The use of $\sigma = 4.25$ nm for the three devices provides a good match between MV–MSMC and NEGF at high V_{DS}. The agreement is a little less satisfactory at low V_{DS}. The good agreement at high V_{DS} can be better appreciated by looking at Figure 8 that shows the trans-conductance (left) and the subthreshold slope (right) for the 8.3 nm device. It should be said that the value of σ is definitely large, considering that the corresponding value for silicon devices is 1.1 nm [37]. This observation suggests that Gaussian smoothening and the σ parameter should be seen as an effective way to mimic what happens in NEGF simulations, and it is hard to relate it to possible definitions of the size of the electron wave packet.

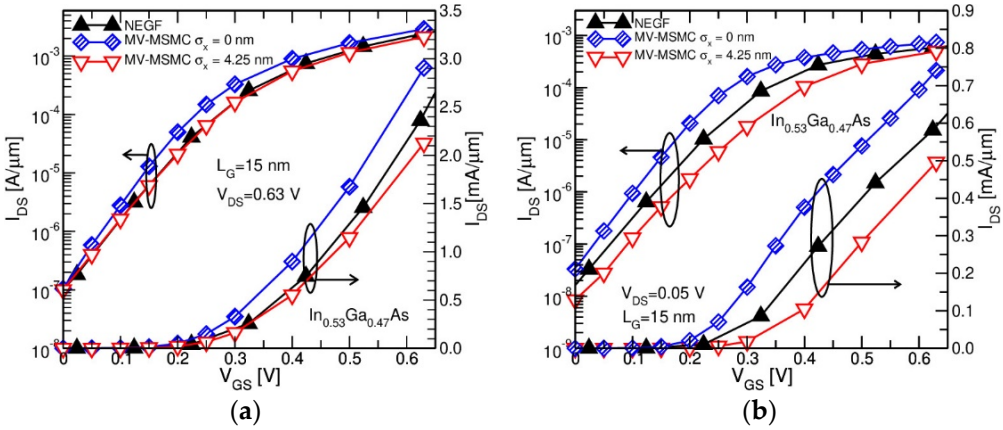


Figure 5. I_{DS}-V_{GS} comparison between NEGF (Non-Equilibrium Green's Function) (ascending triangles) and MV–MSMC (multi-valley/multi-subband Monte Carlo) simulations using σ = 0 nm (diamonds) and σ = 4.25 nm (descending triangles) for an InGaAs n-MOSFET with L_G of 15 nm. V_{DS} is 0.63 V for (**a**) and 50 mV for (**b**).

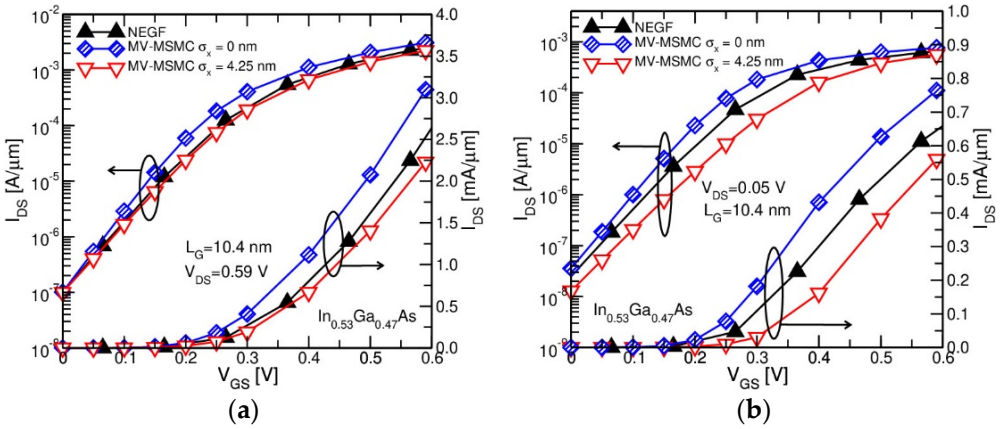


Figure 6. Same as Figure 4 for an InGaAs n-MOSFET with L_G of 10.4 nm. V_{DS} is 0.59 V for (**a**) and 50 mV for (**b**).

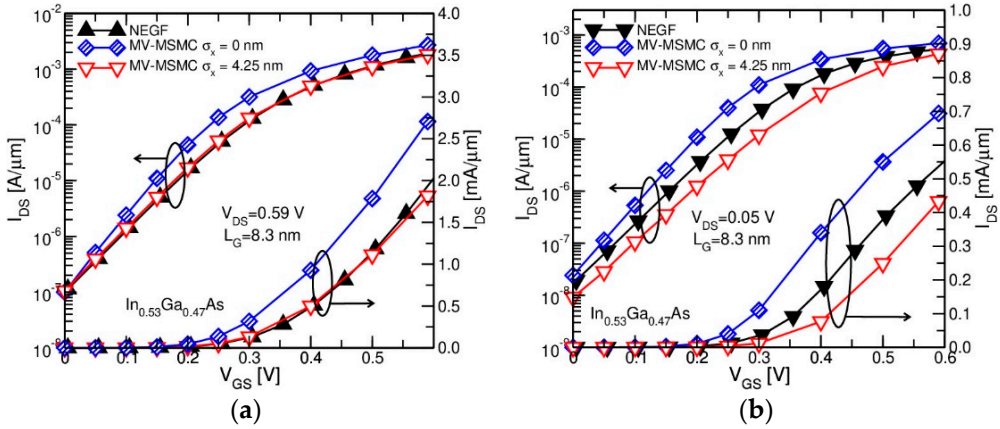


Figure 7. Same as Figure 4 for an InGaAs n-MOSFET with L_G of 8.3 nm. V_{DS} is 0.59 V for (**a**) and 50 mV for (**b**).

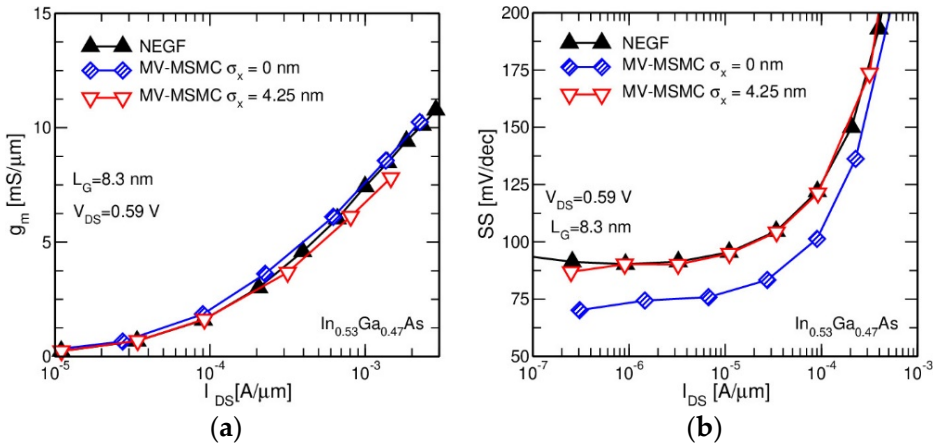


Figure 8. Comparison of (**a**) g_m and (**b**) SS between NEGF (ascending triangles) and MV–MSMC simulations using σ = 0 nm (diamonds) and σ = 4.25 nm (descending triangles) for an InGaAs n-MOSFET with L_G of 8.3 nm. V_{DS} is 0.59 V. MV–MSMC and NEGF were run over the same voltage range, resulting in different current ranges in the *x*-axis of the two plots.

Luckily, the same value works fine for InAs as well, as demonstrated by the match between MV–MSMC and NEGF for the 10.4 nm device in Figure 9. In this latter case, the off-current was not matched to highlight that the NEGF shows a contribution of band-to-band tunneling (due to the smaller bandgap of InAs compared to $In_{0.53}Ga_{0.47}As$) that is not included in the MV–MSMC. The values of the on-current from NEGF and MV–MSMC are summarized in Table 1.

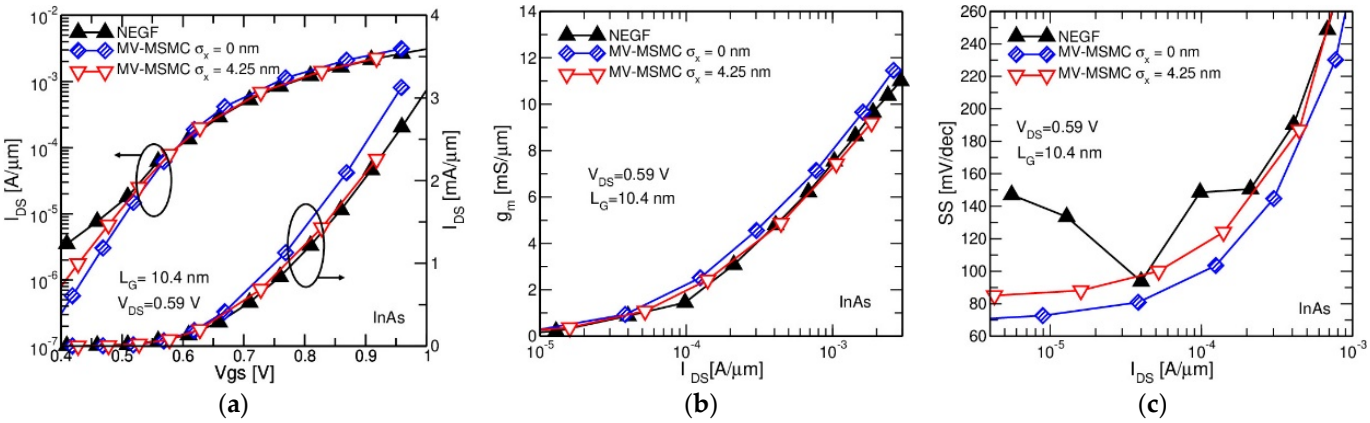


Figure 9. Comparison of (**a**) I_{DS}-V_{GS}, (**b**) g_m, and (**c**) SS between NEGF (ascending triangles) and MV–MSMC simulations using σ = 0 nm (diamonds) and σ = 4.25 nm (descending triangles) for an InAs n-MOSFET with L_G of 10.4 nm. V_{DS} is 0.59 V.

Table 1. On-current (drain current and $V_{GS} = V_{DS} = V_{DD}$) computed from NEGF and MV–MSMC for the devices in Figure 4. Results extracted from Figures 5–9.

	I_{ON} [mA/μm]		
L_G [nm]	**NEGF**	**MV-MSMC σ_x = 0 nm**	**MV-MSMC σ_x = 4.25 nm**
15 (InGaAs)	2.43	2.91	2.13
10.4 (InGaAs)	2.53	3.10	2.23
10.4 (InAs)	2.04	2.45	2.09
8.3 (InGaAs)	1.94	2.71	1.82

It is important to remark that in order to match the MV–MSMC and NEGF results, the former was run after turning off scattering. However, scattering is relevant even at small gate lengths in high-mobility materials. Figure 10 shows that activating scattering in the 10.4 nm device of Figure 4 reduces the on-current by roughly a factor of two. The scattering parameters were set to reproduce mobility data for long-channel MOSFET, as detailed in [42]. External and contact series resistances are not considered here, although they may further reduce the on-current.

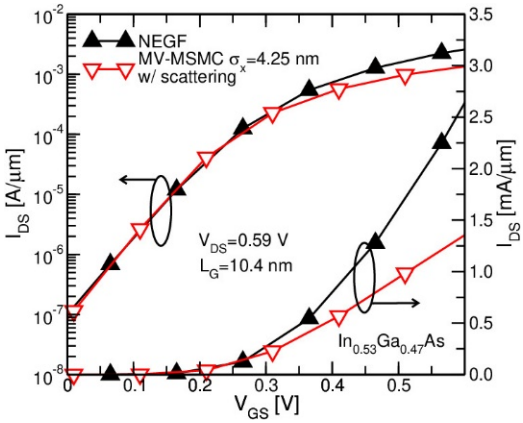


Figure 10. Comparison of I_{DS}-V_{GS} between NEGF and MV–MSMC including scattering. $In_{0.53}Ga_{0.47}As$ device with 10.4 nm gate length. V_{DS} = 0.59 V. Simulation with scattering includes also wave function penetration.

3.2. Use of MV–MSMC Simulations to Calibrate a Commercial TCAD Modeling Tool

As already discussed in Section 2, the MV–MSMC approach features simulation times per bias point that are not suited for sensibility/variability analysis. Nevertheless, to carry out this kind of simulations while retaining the accuracy of the full MV–MSMC model, we calibrated the transport and quantization models of a commercial TCAD simulator (based on the drift–diffusion formalism [14]) to reproduce the MV–MSMC results.

We consider in the following the 10.4 nm device of Figure 4. To reproduce the MV–MSMC calculations, we resort to the quantum-corrected drift–diffusion model (QDD), that considers both the increased confinement of carriers in the channel and the quasi-ballistic transport at a very short channel length. Geometric confinement is accounted for by considering the bandgap widening effect due to splitting of conduction/valence bands into discrete levels in the computation of carrier densities. A further quantum correction is implemented by the Modified Local Density Approximation (MLDA) model to properly account for carrier confinement near the semiconductor/insulator interfaces. The same metal gate work function is used in QDD and MV–MSMC, and the curves are compared "as they are" without performing V_{GS} shifts to align the thresholds. At this stage, quantum effects along the transport direction are absent from both the QDD and the MV–MSMC calculations. We will return to this point later.

Since the QDD, differently from the MSMC, does not consider the electron wave function penetration (WFP) in the gate oxide, an effective T_W is employed in order to reproduce the N_{INV}-vs.-V_{GS} curve (see Figure 11a). This approach is different from the one used in [14,15,23], where T_W was kept equal to the geometrical value, and the parameters of the MV–MSMC were recalibrated disabling the WFP. Although the value T_W = 4.9 nm appears to fit better the N_{INV}-vs.-V_{GS} curve in Figure 11a, a value of 4.7 nm was chosen because it improved the overall agreement between the QDD and MV–MSMC models when considering the complete IV curves, as shown in Figure 12.

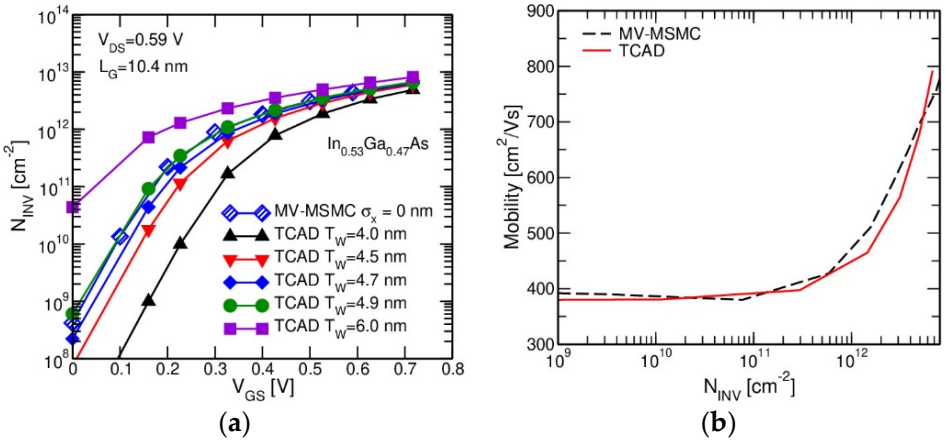


Figure 11. (**a**) N_{INV} vs. V_{GS} for V_{DS} = 0.59 V comparing TCAD (Technology Computer Aided Design) (closed symbols) with different T_W calibrated against MV–MSMC (diamonds) simulations. (**b**) Comparison of the low-field mobility between MV–MSMC (dashed line) and calibrated TCAD (solid line).

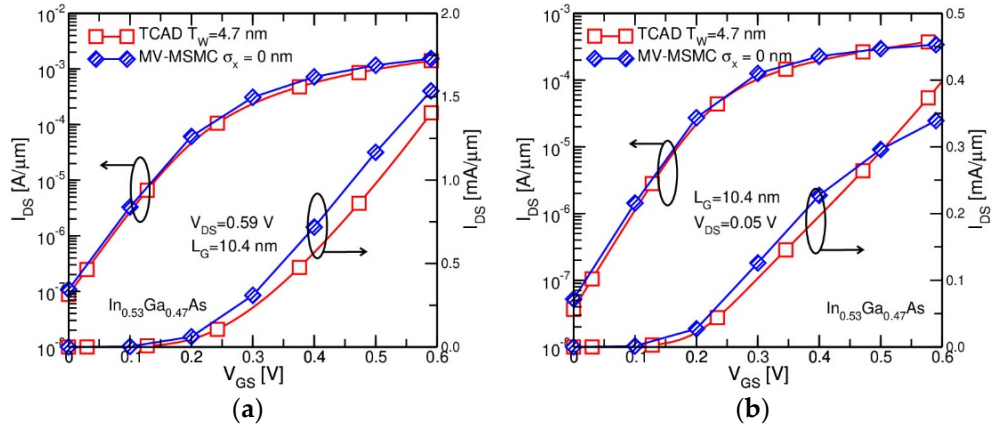


Figure 12. I_{DS}-V_{GS} comparison between TCAD (squares) and MV–MSMC (diamonds) for V_{DS} of (**a**) 0.59 V and (**b**) 50 mV. Device with L_G = 10.4 nm, sketched in Figure 4.

The calibration of the mobility model of the QDD vs. MV–MSMC was performed starting from a long-channel device (L_G = 100 nm) biased at low V_{DS} (25 mV) to avoid the influence of short-channel effects, quasi-ballistic transport, and velocity saturation effects. The parameters of the reference MV–MSMC simulation setup are summarized in Table 2. As indicated therein, satellite valleys are included in these simulations. The results of the comparison between QDD and MSMC are reported in Figure 11b. The increase of mobility with inversion density (N_{INV}) may look surprising, and it is actually due to the very thin well: when N_{INV} increases, due to the low density of states, the electron gas becomes more and more degenerate; the Fermi wave vector (i.e., the radius of the set of k-states having an energy close to the Fermi level) becomes larger and larger, so that surface roughness scattering (the dominant mechanism in a thin quantum well) is less effective in deflecting the velocity. In fact, the surface roughness matrix element is reduced as the exchanged wave vector increases, and this for a large Fermi wave-vector, results in a small change in the angle between the velocity vector and the channel direction [46]. Experimental mobility data for such a thin film are not available, to the best of our knowledge. On the other hand, the data in [47] show that at 8 nm thickness, the mobility roll-off at high inversion density, usually observed in bulk devices, disappears. This is an indication that the interplay between scattering and quantized carrier gas changes from bulk to thin film devices. Notice that the parameters for surface roughness in the MV–MSMC

were calibrated in [42] by comparison with experiments for bulk devices that showed the usual roll-off at high inversion density. Interestingly, when applied to thin films, the same parameter set gives the trend shown in Figure 11b.

Table 2. Main model parameters of the MV–MSMC.

Band structure	Γ valley	M = 0.043 (m_0) α = 1.5 eV^{-1}
	Satellite valleys	From [48]
Scattering parameters	Phonons	From [49]
	Alloy scattering	From [16]
	Roughness: r.m.s. value	0.25 nm [42]
	Roughness: correlation length	1.5 nm [42]
	Interface charge	5×10^{12} cm^{-2}

Besides the calibration of the low-field mobility model and the effective T_W, the good agreement between the MV–MSMC and the QDD (shown in Figure 12 for the device with a 10.4 nm gate length, but a similar agreement was found for the 15 nm device) required the inclusion of a ballistic mobility in the QDD with value $\mu_{bal} = K \cdot L_G$ with K = 30 $cm^2/(Vs \cdot nm)$, as well as high-field mobility with a saturation velocity $v_{sat} = 4 \times 10^7$ cm/s. The main model parameters of the QDD are summarized in Table 3.

Table 3. Main model parameters of the QDD.

Physical Mechanism	Model	Parameters	Value
Electron Density	Fermi-Dirac Statistics	Default from [11]	
Band Structure	Multi-valley, Non-parabolic Bands	Default from [11]	
Geometric Confinement	Bandgap Widening	Default from [11]	
Quantum Confinement	MLDA	Default from [11]	
Low-Field Mobility	Constant	μ_n	380 $cm^2/(Vs)$
Ballistic Mobility	$\mu_{bal} = K \times L_G$	K	30 $cm^2/(Vs\ nm)$
High-field Saturation	Canali Model	v_{sat}	4×10^7 cm/s
		β	2

Inclusion of quantum effects along the transport direction in the QDD is not straightforward and in our opinion still an open issue with the modeling options available at present. A systematic analysis of possible ways to include size-induced quantization (vertical direction) and source-to-drain tunneling (transport direction) has been carried out in [50].

We analyze two options in Figure 13. We start by using the MLDA for quantization in the vertical direction (to be consistent with the analysis in Figures 11a and 12). MLDA only handles quantization in the vertical direction; therefore, non-local tunneling (NLT) from source-to-drain tunneling is activated in the TCAD with a tunneling mass $m_{tun} = 0.05m_0$. Plot (a) in Figure 13 reports the results for the 10.4 nm device. The IV curves from the NEGF are added only to compare the subthreshold behavior, although a match was not targeted, since forcing ballistic transport in the QDD model is not possible. We observe that activating NLT in the QDD increases the inverse subthreshold slope (as expected) but also significantly increases the current in a strong inversion regime. This is an artificial and unphysical effect: the electrons that tunnel below the potential barrier at the source are not self-consistently included in the device electrostatics, so that the height of the barrier is related only to the carriers that surmount the barrier classically. As a result, tunneling adds a current, but the associated charge is not consistent with the device electrostatics. This is not the case of NEGF, where the evanescent part of the wavefunctions describing

the tunneling electrons contributes to device electrostatics. This point was discussed in detail [37] and led us to include quantum phenomena along the transport in the MV–MSMC via subband smoothening, i.e., by lowering the barrier rather than making electrons tunnel below it. Unfortunately, this option is not available in TCAD. The option that more closely resembles it is the Density Gradient (DG) model, where the potential barrier is modified based on the gradient of the electron concentration in order to account for quantum effects. In TCAD, a density gradient cannot be activated only along the transport direction but should be active in all simulated coordinates. As a result, we had to turn off MLDA and use DG also for size-induced quantization in the vertical direction. A plot similar to Figure 11a (not shown here) pointed out that a good match between the inversion density of MV–MSMC and QDD with DG was possible by using an effective T_W = 4.7 nm (as for MLDA) and setting to 0.1 the empirical γ parameter of the DG model (i.e., the parameter that set the strength of the potential correction [11]). Plot (b) in Figure 13 reports QDD simulations employing the DG model. The curves refer to different values of the α parameter [11] that sets the strength of the potential corrections imposed by the DG in the transport direction compared to the one in the vertical direction (i.e., $\alpha = 10^{-4}$ means essentially zero DG corrections along the transport direction). We observe that increasing α increases the drain current in the sub-threshold region (leading to a larger inverse sub-threshold slope) but has a minor effect in strong inversion, consistently with the fact that tunneling is mimicked by barrier lowering, so that all electrons contribute to determining the electrostatics. While far from being conclusive, this analysis outlines possible strategies to include SDT in QDD to more accurately reproduce NEGF results.

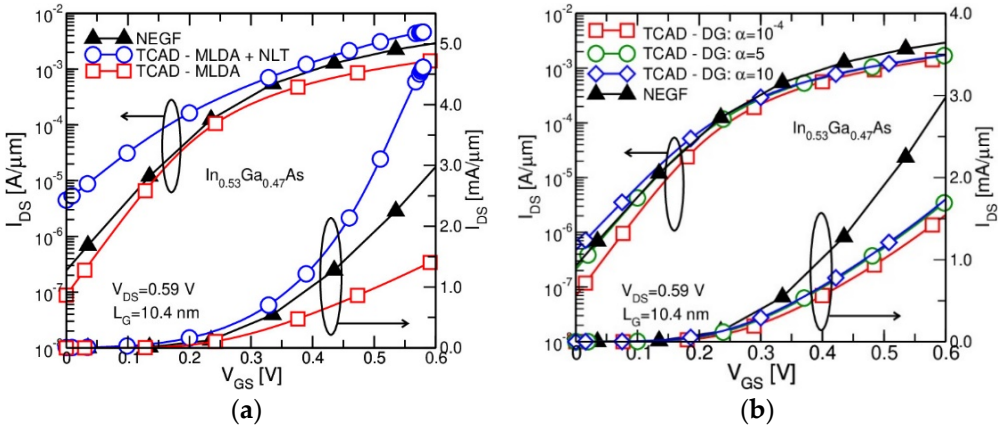


Figure 13. Comparison between TCAD and NEGF simulation results for the 10.4 nm device of Figure 4. In plot (**a**), the MLDA (Modified Local Density Approximation) is used in TCAD to describe quantization in the vertical direction and non-local-tunneling (NLT) for source-to-drain tunneling. In plot (**b**), instead, the Density Gradient (DG) model is used with different values of the parameter α described in the text.

4. Discussion and Conclusions

We have presented a hierarchy of modeling approaches for nanoscale MOSFETs and applied it to devices based on III–V compounds. It is evident that each approach has its own pros and cons. It is thus not possible to identify the generally best approach. Instead, the whole suite of methods should be employed to get accurate and trustworthy results. The full-quantum tools (e.g., the NEGF simulator employed here) are very powerful in providing information about the band structure in the presence of strong size- and bias-induced quantization, as well as in predicting the influence on the device behavior of quantum phenomena, such as the impact of source-to-drain tunneling on the off-current and inverse sub-threshold slope in nanoscale MOSFETs. On the other hand, inclusion of scattering (that is relevant also in decananometric MOSFETs, as seen in Figure 10) is very demanding from a computational point of view, so that an accurate estimate of the device

on-current requires approaches such as the MV–MSMC. The calibration of the MV–MSMC requires a mix of comparisons against full-quantum models (to set the band structure of the quantized electron gas as well as the empirical parameter describing tunneling along the channel) and against low-field mobility data (to calibrate the scattering parameters). Experimental velocity-field curves for thin III–V structures would be very useful as a further calibration/verification of the scattering rates at high energy, but no data are presently available in the literature, at least to our knowledge. Another ingredient that is missing in the MV–MSMC is band-to-band tunneling, which may have an impact on the off-current. Again, this mechanism is well described by full-quantum tools, so that one may use those models to calibrate and validate more empirical approaches suitable to be integrated in the MV–MSMC (e.g., approaches similar to the one implemented in the semi-classical MC described in [39]).

The NEGF and the MV–MSMC simulators used here are research tools suited for expert users. Extensive device design and optimization are typically carried out by using commercial TCAD tools, which can also deal with the critical role of the statistical variability of device parameters. The NEGF and the MV–MSMC simulators can be also very useful for the calibration of TCAD models, although such calibration must be repeated when considering new materials and different device architectures. Furthermore, calibration is almost universally carried out in terms of DC current–voltage characteristics, which does guarantee the trustworthiness when dealing with AC simulations or when performing sensitivity analysis as needed for variability studies.

In this work, we restricted the analysis to 2D structures (i.e., planar double-gate MOSFETs or tall FinFETs). TCAD can handle more complicated 3D devices such as tri-gate FinFETs, gate-all-around MOSFETs, and stacked nanowires that are candidates for aggressively scaled technology nodes. The TCAD calibration procedure provided here vs. 2D MV–MSMC can be applied to these structures as well, although verification against 3D full-quantum approaches and solutions of the BTE for the 1D carrier gas [20,27,51] may be needed.

Author Contributions: Conceptualization all authors; methodology E.C., P.P., N.Z.; software D.E., D.L., E.C., E.G., L.S., P.P.; validation E.C., N.Z., A.P.; writing—original draft preparation P.P.; writing—review and editing: all authors; funding acquisition L.S. All authors have read and agreed to the published version of the manuscript.

Funding: The research reported in the paper received partial support from the European Commission III–V-MOS project (g.a. 619326) via the IUNET consortium.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The MV–MSMC code is subject to licensing restrictions.

Acknowledgments: We would like to acknowledge the work of many people who contributed to the development of the MV–MSMC, among them, Luca Lucci, Patrik Osgnach, Paolo Toniutti, Alberto Revelant, Fabiano Zaninotto.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Radosavljevic, M.; Dewey, G.; Basu, D.; Boardman, J.; Chu-Kung, B.; Fastenau, J.M.; Kabehie, S.; Kavalieros, J.; Le, V.; Liu, W.K.; et al. Electrostatics improvement in 3-D tri-gate over ultra-thin body planar InGaAs quantum well field effect transistors with high-K gate dielectric and scaled gate-to-drain/gate-to-source separation. In Proceedings of the 2011 International Electron Devices Meeting, Washington, DC, USA, 5–7 December 2011; pp. 33.1.1–33.1.4. [CrossRef]
2. Convertino, C.; Zota, C.; Sant, S.; Eltes, F.; Sousa, M.; Caimi, D.; Schenk, A.; Czornomaz, L. InGaAs-on-Insulator FinFETs with Reduced Off-Current and Record Performance. In Proceedings of the 2018 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 1–5 December 2018; pp. 39.2.1–39.2.4. [CrossRef]
3. Riel, H.; Wernersson, L.-E.; Hong, M.; del Alamo, J.A. III–V compound semiconductor transistors—From planar to nanowire structures. *MRS Bull.* **2014**, *39*, 668–677. [CrossRef]

4. Lee, S.; Chobpattana, V.; Huang, C.-Y.; Thibeault, B.J.; Mitchell, W.; Stemmer, S.; Gossard, A.C.; Rodwell, M.J.W. Record Ion (0.50 mA/μm at V_{DD} = 0.5 V and I_{off} = 100 nA/μm) 25 nm-gate-length ZrO_2/InAs/InAlAs MOSFETs. In Proceedings of the 2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers, Honolulu, HI, USA, 9–12 June 2014; pp. 1–2. [CrossRef]
5. Zota, C.B.; Lindelow, F.; Wernersson, L.-E.; Lind, E. InGaAs tri-gate MOSFETs with record on-current. In Proceedings of the 2016 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 3–7 December 2016; pp. 3.2.1–3.2.4. [CrossRef]
6. Visciarelli, M.; Gnudi, A.; Gnani, E.; Reggiani, S. A full-quantum simulation study of InGaAs NW MOSFETs including interface traps. In Proceedings of the 2016 46th European Solid-State Device Research Conference (ESSDERC), Lausanne, Switzerland, 12–15 September 2016; pp. 180–183. [CrossRef]
7. Park, S.H.; Liu, Y.; Kharche, N.; Jelodar, M.S.; Klimeck, G.; Lundstrom, M.S.; Luisier, M. Performance Comparisons of III–V and Strained-Si in Planar FETs and Nonplanar FinFETs at Ultrashort Gate Length (12 nm). *IEEE Trans. Electron Devices* **2012**, *59*, 2107–2114. [CrossRef]
8. Luisier, M.; Klimeck, G. Atomistic full-band simulations of silicon nanowire transistors: Effects of electron-phonon scattering. *Phys. Rev. B* **2009**, *80*, 155430. [CrossRef]
9. Bahder, T.B. Eight-bandk·pmodel of strained zinc-blende crystals. *Phys. Rev. B* **1990**, *41*, 11992. [CrossRef] [PubMed]
10. Zerveas, G.; Caruso, E.; Baccarani, G.; Czornomaz, L.; Daix, N.; Esseni, D.; Gnani, E.; Gnudi, A.; Grassi, R.; Luisier, M.; et al. Comprehensive comparison and experimental validation of band-structure calculation methods in III–V semiconductor quantum wells. *Solid-State Electron.* **2016**, *115*, 92–102. [CrossRef]
11. *Sentaurus Device User Guide*, version Q-2019; Synopsys, Inc.: Mountain View, CA, USA, 2019.
12. Caruso, E.; Lin, J.; Monaghan, S.; Cherkaoui, K.; Gity, F.; Palestri, P.; Esseni, D.; Selmi, L.; Hurley, P.K. The Role of Oxide Traps Aligned with the Semiconductor Energy Gap in MOS Systems. *IEEE Trans. Electron Devices* **2020**, *67*, 4372–4378. [CrossRef]
13. Penzin, O.; Paasch, G.; Smith, L. Nonparabolic Multivalley Quantum Correction Model for InGaAs Double-Gate Structures. *IEEE Trans. Electron Devices* **2013**, *60*, 2246–2250. [CrossRef]
14. Carapezzi, S.; Caruso, E.; Gnudi, A.; Palestri, P.; Reggiani, S.; Gnani, E. TCAD Mobility Model of III-V Short-Channel Double-Gate FETs Including Ballistic Corrections. *IEEE Trans. Electron Devices* **2017**, *64*, 4882–4888. [CrossRef]
15. Zagni, N.; Caruso, E.; Puglisi, F.M.; Pavan, P.; Palestri, P.; Verzellesi, G. Systematic Modeling of Electrostatics, Transport, and Statistical Variability Effects of Interface Traps in End-of-the-Roadmap III–V MOSFETs. *IEEE Trans. Electron Devices* **2020**, *67*, 1560–1566. [CrossRef]
16. Fischetti, M. Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures. I. Homogeneous transport. *IEEE Trans. Electron Devices* **1991**, *38*, 634–649. [CrossRef]
17. Esseni, D.; Palestri, P.; Selmi, L. *Nanoscale MOS Transistors: Semi-Classical Transport and Applications*; Cambridge University Press: Cambridge, UK, 2011.
18. Lucci, L.; Palestri, P.; Esseni, D.; Bergagnini, L.; Selmi, L. Multisubband Monte Carlo Study of Transport, Quantization, and Electron-Gas Degeneration in Ultrathin SOI n-MOSFETs. *IEEE Trans. Electron Devices* **2007**, *54*, 1156–1164. [CrossRef]
19. Sampedro, C.; Gamiz, F.; Godoy, A.; Valín, R.; Garcia-Loureiro, A.; Ruiz, F.J.G. Multi-Subband Monte Carlo study of device orientation effects in ultra-short channel DGSOI. *Solid-state Electron.* **2010**, *54*, 131–136. [CrossRef]
20. Jin, S.; Pham, A.-T.; Choi, W.; Nishizawa, Y.; Kim, Y.-T.; Lee, K.-H.; Park, Y.; Jung, E.S. Performance evaluation of InGaAs, Si, and Ge nFinFETs based on coupled 3D drift-diffusion/multisubband boltzmann transport equations solver. In Proceedings of the 2014 IEEE International Electron Devices Meeting, San Francisco, CA, USA, 15–17 December 2014; pp. 7.5.1–7.5.4. [CrossRef]
21. Badami, O.; Lizzit, D.; Specogna, R.; Esseni, D. Surface roughness limited mobility in multi-gate FETs with arbitrary cross-section. In Proceedings of the 2016 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 3–7 December 2016; pp. 36.1.1–36.1.4. [CrossRef]
22. Palestri, P.; Caruso, E.; Driussi, F.; Esseni, D.; Lizzit, D.; Osgnach, P.; Venica, S.; Selmi, L. State-of-the-art semi-classical Monte Carlo method for carrier transport in nanoscale transistors. In Proceedings of the 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 25–29 May 2015; pp. 1–8. [CrossRef]
23. Selmi, L.; Caruso, E.; Carapezzi, S.; Visciarelli, M.; Gnani, E.; Zagni, N.; Pavan, P.; Palestri, P.; Esseni, D.; Gnudi, A.; et al. Modelling nanoscale n-MOSFETs with III-V compound semiconductor channels: From advanced models for band structures, electrostatics and transport to TCAD. In Proceedings of the 2017 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2–6 December 2017; pp. 13.4.1–13.4.4. [CrossRef]
24. Chen, S.-H.; Liao, W.-S.; Yang, H.-C.; Wang, S.-J.; Liaw, Y.-G.; Wang, H.; Gu, H.; Wang, M.-C. High-performance III-V MOSFET with nano-stacked high-k gate dielectric and 3D fin-shaped structure. *Nanoscale Res. Lett.* **2012**, *7*, 431. [CrossRef]
25. Palestri, P.; Caruso, E.; Badami, O.; Driussi, F.; Esseni, D.; Selmi, L. Semi-classical modeling of nanoscale nMOSFETs with III-V channel. In Proceedings of the 2019 Electron Devices Technology and Manufacturing Conference (EDTM), Singapore, 12–15 March 2019; pp. 234–236. [CrossRef]
26. De Michielis, M.; Esseni, D.; Palestri, P.; Selmi, L. Semiclassical Modeling of Quasi-Ballistic Hole Transport in Nanoscale pMOSFETs Based on a Multi-Subband Monte Carlo Approach. *IEEE Trans. Electron Devices* **2009**, *56*, 2081–2091. [CrossRef]
27. Lenzi, M.; Palestri, P.; Gnani, E.; Reggiani, S.; Gnudi, A.; Esseni, D.; Selmi, L.; Baccarani, G. Investigation of the Transport Properties of Silicon Nanowires Using Deterministic and Monte Carlo Approaches to the Solution of the Boltzmann Transport Equation. *IEEE Trans. Electron Devices* **2008**, *55*, 2086–2096. [CrossRef]

28. Jin, S.; Fischetti, M.; Tang, T.-W. Modeling of electron mobility in gated silicon nanowires at room temperature: Surface roughness scattering, dielectric screening, and band nonparabolicity. *J. Appl. Phys.* **2007**, *102*, 083715. [CrossRef]
29. Vurgaftman, I.; Meyer, J.R.; Ram-Mohan, L.R. Band parameters for III–V compound semiconductors and their alloys. *J. Appl. Phys.* **2001**, *89*, 5815–5875. [CrossRef]
30. Lizzit, D.; Esseni, D.; Palestri, P.; Osgnach, P.; Selmi, L. Performance Benchmarking and Effective Channel Length for Nanoscale InAs, $In_{0.53}Ga_{0.47}As$, and sSi n-MOSFETs. *IEEE Trans. Electron Devices* **2014**, *61*, 2027–2034. [CrossRef]
31. Yokoyama, K.; Hess, K. Monte Carlo study of electronic transport inAl1−xGaxAs/GaAs single-well heterostructures. *Phys. Rev. B* **1986**, *33*, 5595–5606. [CrossRef]
32. Caruso, E.; Pin, A.; Palestri, P.; Selmi, L. On the electron velocity-field relation in ultra-thin films of III–V compound semiconductors for advanced CMOS technology nodes. In Proceedings of the 2017 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS), Athens, Greece, 3–5 April 2017; pp. 152–155.
33. Caruso, E.; Palestri, P.; Lizzit, D.; Osgnach, P.; Esseni, D.; Selmi, L. Quasi-Ballistic Γ- and L-Valleys Transport in Ultrathin Body Strained (111) GaAs nMOSFETs. *IEEE Trans. Electron Devices* **2016**, *63*, 4685–4692. [CrossRef]
34. Prange, R.E.; Nee, T.-W. Quantum Spectroscopy of the Low-Field Oscillations in the Surface Impedance. *Phys. Rev.* **1968**, *168*, 779. [CrossRef]
35. Lizzit, D.; Esseni, D.; Palestri, P.; Selmi, L. A new formulation for surface roughness limited mobility in bulk and ultra-thin-body metal–oxide–semiconductor transistors. *J. Appl. Phys.* **2014**, *116*, 223702. [CrossRef]
36. Medina, A.G.; Padilla, J.L.; Sampedro, C.; Godoy, A.; Donetti, L.; Gamiz, F. Source-to-Drain Tunneling Analysis in FDSOI, DGSOI, and FinFET Devices by Means of Multisubband Ensemble Monte Carlo. *IEEE Trans. Electron Devices* **2018**, *65*, 4740–4746. [CrossRef]
37. Palestri, P.; Lucci, L.; Tos, S.D.; Esseni, D.; Selmi, L. An improved empirical approach to introduce quantization effects in the transport direction in multi-subband Monte Carlo simulations. *Semicond. Sci. Technol.* **2010**, *25*, 055011. [CrossRef]
38. Ferry, D.K. Effective potentials and the onset of quantization in ultrasmall MOSFETs. *Superlattices Microstruct.* **2000**, *28*, 419–423. [CrossRef]
39. Revelant, A.; Palestri, P.; Osgnach, P.; Selmi, L. Calibrated multi-subband Monte Carlo modeling of tunnel-FETs in silicon and III–V channel materials. *Solid-state Electron.* **2013**, *88*, 54–60. [CrossRef]
40. Taoka, N.; Yokoyama, M.; Kim, S.H.; Suzuki, R.; Iida, R.; Lee, S.; Hoshii, T.; Jevasuwan, W.; Maeda, T.; Yasuda, T.; et al. Impact of Fermi level pinning inside conduction band on electron mobility of InxGa1xAs MOSFETs and mobility enhancement by pinning modulation. In Proceedings of the 2011 International Electron Devices Meeting, Washington, DC, USA, 5–7 December 2011; pp. 27.2.1–27.2.4. [CrossRef]
41. Osgnach, P.; Caruso, E.; Lizzit, D.; Palestri, P.; Esseni, D.; Selmi, L. The impact of interface states on the mobility and drive current of $In_{0.53}Ga_{0.47}As$ semiconductor n-MOSFETs. *Solid-state Electron.* **2015**, *108*, 90–96. [CrossRef]
42. Rau, M.; Caruso, E.; Lizzit, D.; Palestri, P.; Esseni, D.; Schenk, A.; Selmi, L.; Luisier, M. Performance projection of III-V ultra-thin-body, FinFET, and nanowire MOSFETs for two next-generation technology nodes. In Proceedings of the 2016 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 5–7 December 2016; pp. 30.6.1–30.6.4. [CrossRef]
43. Rau, M.; Lin, J.; Antoniadis, D.A.; Del Alamo, J.A.; Luisier, M. Investigation of Source Starvation in High-Transconductance III–V Quantum-Well MOSFETs. *IEEE Trans. Electron Devices* **2019**, *66*, 4698–4705. [CrossRef]
44. ITRS. International Technology Roadmap for Semiconductors. 2013. Available online: http://www.itrs2.net/2013-itrs.html. (accessed on 8 August 2021).
45. Visciarelli, M.; Gnani, E.; Gnudi, A.; Reggiani, S.; Baccarani, G. Impact of Strain on Tunneling Current and Threshold Voltage in III–V Nanowire TFETs. *IEEE Electron Device Lett.* **2016**, *37*, 560–563. [CrossRef]
46. Badami, O.; Caruso, E.; Lizzit, D.; Osgnach, P.; Esseni, D.; Palestri, P.; Selmi, L. An Improved Surface Roughness Scattering Model for Bulk, Thin-Body, and Quantum-Well MOSFETs. *IEEE Trans. Electron Devices* **2016**, *63*, 2306–2312. [CrossRef]
47. Takei, K.; Fang, H.; Kumar, S.B.; Kapadia, R.; Gao, Q.; Madsen, M.; Kim, H.S.; Liu, C.-H.; Chueh, Y.-L.; Plis, E.; et al. Quantum Confinement Effects in Nanoscale-Thickness InAs Membranes. *Nano Lett.* **2011**, *11*, 5008–5012. [CrossRef] [PubMed]
48. O'Regan, T.P.; Fischetti, M.; Soree, B.; Jin, S.; Magnus, W.; Meuris, M. Calculation of the electron mobility in III-V inversion layers with high-κ dielectrics. *J. Appl. Phys.* **2010**, *108*, 103705. [CrossRef]
49. Mateos, J.; Gonzalez, T.; Pardo, D.; Hoel, V.; Happy, H.; Cappy, A. Improved Monte Carlo algorithm for the simulation of δ-doped AlInAs/GaInAs HEMTs. *IEEE Trans. Electron Devices* **2000**, *47*, 250–253. [CrossRef]
50. Aguirre, P.; Carrillo-Nunez, H.; Ziegler, A.; Luisier, M.; Schenk, A. Drift-diffusion quantum corrections for $In_{0.53}Ga_{0.47}As$ double gate ultra-thin-body FETs. In Proceedings of the 2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Nuremberg, Germany, 6–8 September 2016; pp. 53–56. [CrossRef]
51. Badami, O.; Lizzit, D.; Driussi, F.; Palestri, P.; Esseni, D. Benchmarking of 3-D MOSFET Architectures: Focus on the Impact of Surface Roughness and Self-Heating. *IEEE Trans. Electron Devices* **2018**, *65*, 3646–3653. [CrossRef]