

Drift-Diffusion TCAD Modelling of III-V-MOSFETs
in the Ballistic Regime

DISS. ETH NO. 26200

Drift-Diffusion TCAD Modelling of III-V-MOSFETs in the Ballistic Regime

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

PAULINA AGUIRRE

MSc. TUDelft
born on May 29th, 1986
citizen of Quito, Ecuador

accepted on the recommendation of

Prof. Dr. Andreas Schenk, examiner
Prof. Dr. Mathieu Luisier, co-examiner
Prof. Dr. Pierpaolo Palestri, co-examiner

2019

Acknowledgments

First, I would like to express my deepest gratitude to Prof. Andreas Schenk for giving me the opportunity to work in this research project and specially for his substantial commitment during this research work. I am very grateful for all the physics I learned from him, his incredible patience and vast understanding of the theoretical framework behind this work. His experience and guidance, where decisive in the success of this work. Also, I would like to express my special gratitude to Prof. Mathieu Luisier, for all the support he offered through this years with the Quantum Transport code. Besides his valuable help, his human warmth and empathy through the difficult times was always an encouragement. I am also grateful to Prof. Pierpaolo Palestri for accepting to be the co-examiner of this thesis.

I would really like to thank Oleg Penzin from Synopsys Inc. for the collaboration during the development of Chapter 3 of this thesis and its valuable discussions. As well Dr. Axel Erlebach for the help in the PMI implementations and availability. As well many thanks to Dr. Denis Dolgos for the help with the multivalley version of the tunneling model and Pawel Lenarczyk for sharing his knowledge of numerics.

As well my gratitude to Dr. Saurabh Sant for sharing his TCAD expertise and his sincere friendship through these years, many thanks for Dr. Hamilton Carrillo-Nunez for his help during the first year of my thesis. I would like to thank the friendly and cheerful people of the group. Many thanks to Christian Stieger, Dr. Aron Szabo, Dr. Mauro Caldera, Dr. Anne Ziegler, Dr. Guillermo Indalecio, Dr. Yonseug Lee and Kantawong Vuttivorakulchai. Thanks to Christoph Wicki and

Christine Haller, for creating the optimal working conditions in the institute that allowed me to focus on my research.

Finally, I would like to express my gratitude to my family, that has always encouragement me in these difficult years. My husband, for his understanding of the long hours of work that an endeavour like this implies. Especial thanks to my sister Elizabeth, for showing me with the example of her life that despite the difficulties life throws at you, it is worth to keep fighting for your dreams.

This work was funded by the European Commissions Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 619326 (III-V-MOS Project).

Abstract

This thesis presents Technology Computer Aided Design (TCAD) modelling of MOSFETs with InGaAs channels (III-V-MOSFETs) operating in the ballistic regime. It is shown how the drift-diffusion formalism can be used to simulate ultra-thin-body (UTB) and gate-all-around nanowire (GAA NW) transistors at the nanometer scale. In order to accurately cover the relevant quantum-mechanical effects as well as the characteristics of ballistic transport, several physical models must be revisited and assembled with appropriate assumptions. In the first part, drift-diffusion quantum corrections are applied to simulate geometrical confinement across the channel of two-dimensional UTB and three-dimensional GAA NW MOSFETs. Quantum drift-diffusion models available in the commercial device simulator Sentaurus-Device from Synopsys Inc., as the Density Gradient and the Modified Local Density Approximation models, are used after a careful calibration of their parameters. Quantum tunneling through the potential barrier between source and drain, as consequence of the ultra-short channels and the small electron effective mass inherent in III-V materials, becomes a severe leakage current mechanism in the off-state. The Nonlocal Tunneling model is used to simulate source-to-drain tunneling in combination with the models for geometrical confinement. The inclusion of models for the density of states of a one- and two-dimensional electron gas proves necessary to assure the accurate simulation of the electrostatics in the investigated devices. This also ensures a similar Fermi energy level of the source as the one obtained from a quantum-mechanical reference tool. An important consideration for the modelling and simulation of current-voltage characteristics is the almost negligible electron-phonon scattering in

transistors with ultra-short III-V-channels. Thus, quantum-ballistic simulations are used as reference. TCAD models of the ballistic mobility are developed and implemented in the simulation framework. They result in the same limitation of the on-current as found in the quantum-ballistic reference simulations. In these TCAD models, a local variant of the proposed ballistic velocity depends either on the quasi-Fermi potential or on the electron density. The inclusion of a ballistic mobility affects the overall shape of the quasi-Fermi potential as well as the quantum tunneling rates. The latter effect occurs because, in Sentaurus-Device, source-to-drain and band-to-band tunneling are implemented as electron generation rates with local quasi-Fermi energies. Modifications of the tunneling models are suggested that would circumvent the problem. Finally, the Shockley-Read-Hall (SRH) recombination and the band-to-band tunneling models are applied simultaneously to correctly describe the floating-body effect in UTB and GAA NW transistors. An appropriate quantum lifetime which includes the effect of geometrical confinement is coupled into the SRH recombination model.

Zusammenfassung

Die vorliegende Dissertation behandelt die TCAD-Modellierung von MOSFETs mit InGaAs-Kanälen (III-V-MOSFETs) im ballistischen Bereich. Es wird gezeigt, wie der Drift-Diffusions-Formalismus verwendet werden kann, um Ultradünnschicht- (UTB) und Gate-All-Around-Nanodraht Transistoren (GAA-NW) im Nanometerbereich zu simulieren. Um die relevanten quanten-mechanischen Effekte und die Charakteristika des ballistischen Transports zu berücksichtigen, müssen diverse physikalische Modelle überdacht und mit entsprechenden Annahmen in eine Simulationsumgebung integriert werden.

Im ersten Teil der Arbeit werden Drift-Diffusions-Quantenkorrekturen angewendet, um das geometrische Confinement senkrecht zum Kanal im zweidimensionalen UTB-MOSFET und im drei-dimensionalen GAA-NW-MOSFET zu simulieren. Dazu werden - nach sorgfältiger Kalibrierung der Modellparameter - die Quanten-Drift-Diffusions-Modelle des kommerziellen Bauelemente-Simulators Sentaurus-Device von Synopsys Inc., wie das Dichtegradient-Modell und das Modell der Modifizierten Lokal-Dichte-Approximation, verwendet. Tunneln durch die Potentialbarriere zwischen Source und Drain infolge der ultrakurzen Kanäle und der kleinen effektiven Elektronenmasse, die III-V-Materialien eigen ist, ist ein empfindlicher Leckstrom-Mechanismus im Sperrzustand. Das Modell des Nichtlokalen Tunnelns wird benutzt, um Source-Drain-Tunneln in Kombination mit den Modellen für das geometrische Confinement zu simulieren.

Die Einbeziehung von Modellen für die Zustandsdichte des ein- bzw. zweidimensionalen Elektronengases erweist sich als erforderlich, um eine genaue Simulation der Elektrostatik in den untersuchten Bauelementen zu gewährleisten. Damit hat auch die Fermi-Energie

der Source einen ähnlichen Wert wie jene in den Referenzsimulationen mit einem Quantentransport-Simulator.

Ein wichtiger Gesichtspunkt für die Modellierung und Simulation von Strom-Spannungs-Kennlinien ist die nahezu vernachlässigbare Elektron-Phonon-Streuung in Transistoren mit ultrakurzen III-V-Kanälen, weshalb quanten-ballistische Simulationen als Referenz herangezogen werden. TCAD-Modelle der ballistischen Beweglichkeit werden entwickelt und in die Simulationsumgebung implementiert. Sie ergeben dieselbe Begrenzung des Einschaltstroms wie man sie in den quanten-ballistischen Simulationen findet. Vorgeschlagene Varianten der ballistischen Geschwindigkeit in diesen TCAD-Modellen hängen entweder vom Quasi-Fermi-Potential oder von der Elektrodendichte ab. Die Verwendung einer ballistischen Beweglichkeit beeinflusst sowohl den Verlauf des Quasi-Fermi-Potentials als auch die Tunnelraten. Letzteres tritt auf, weil Source-Drain- und Interband-Tunneln in Sentaurus-Device als Elektronen-Generationsraten mit lokalen Quasi-Fermi-Energien implementiert sind. Modifikationen der Tunnel-Modelle werden vorgeschlagen, die das Problem beheben würden.

Abschliessend werden die Modelle der Shockley-Read-Hall-Rekombination und des Interband-Tunnelns gleichzeitig verwendet, um den Floating-Body-Effekt in UTB- und GAA-NW-Transistoren zu modellieren. Eine geeignete Quantenlebensdauer, die den Effekt des geometrischen Confinements berücksichtigt, wird dazu in das SRH-Modell eingeführt.

Contents

Acknowledgments	v
Abstract	vii
Zusammenfassung	ix
1 Introduction	1
2 Modelling of Channel Quantization	7
2.1 Introduction	7
2.2 Theory	8
2.2.1 Density Gradient Approach	8
2.2.2 Electric Field Dependency of γ	9
2.2.3 Anisotropic Density Gradient Model	10
2.2.4 Modified Local Density Approximation	12
2.3 Simulation Method	12
2.3.1 2D Simulations of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FETs	12
2.3.2 3D Simulations $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ GAA Nanowires	13
2.4 Simulation Results	14
2.4.1 2D DG UTB FETs	14
2.4.2 3D GAA Nanowires	17
2.5 Discussion	21
2.6 Summary	23

3	Density of States	25
3.1	Introduction	25
3.2	Theory	26
3.2.1	2D-DOS	27
3.2.2	1D-DOS	28
3.3	Simulation Results	31
3.4	Discussion	33
3.5	Summary	36
4	Ballistic Mobility	39
4.1	Introduction	39
4.2	Theory of Transport Regimes and Ballistic Mobility	40
4.3	Ballistic Velocity as Function of QFP	44
4.3.1	Velocity Comparisons between QTx and DD	47
4.4	Ballistic Mobility as Function of Density	51
4.5	Implementation	55
4.6	Results	56
4.6.1	Application of Fermi Statistics	57
4.6.2	Application of Boltzmann Statistics	61
4.7	Discussion	62
4.8	Summary	65
5	Leakage Currents in the Ballistic Regime	67
5.1	Introduction	67
5.2	Theory of Source-to-Drain Tunneling	67
5.3	Source-to-Drain Tunneling and Ballistic Mobility	68
5.3.1	Behaviour of the Nonlocal Tunneling Model	68
5.3.2	Nonlocal Tunneling using Contact Fermi Levels	72
5.4	Band-to-Band Tunneling	72
5.5	Simulation Results	73
5.5.1	Influence of the Ballistic Mobility on the STDT Rate	73
5.5.2	Influence of the Ballistic Mobility on the BTBT Rate	75
5.6	Discussion	76
5.7	Summary	80

6 Floating Body Effect	81
6.1 Introduction	81
6.2 Theoretical Aspects	81
6.2.1 The Floating Body Effect	81
6.2.2 Band gap Variation due to Geometrical Con- finement	83
6.2.3 Effect of Quantization on the SRH Lifetimes	84
6.3 Simulation Method	85
6.4 Results	86
6.4.1 2D DG UTB FETs	86
6.4.2 3D GAA NWs	89
6.5 Discussion	89
6.6 Summary	90
A Ballistic Mobility: Derivation of the TOB Density	91
Bibliography	95
Curriculum Vitae	103

Chapter 1

Introduction

The expiration of Moore's law in the past 10 years has been slowly reaching its critical point [1]. During this time, aggressive scaling of Silicon transistors has enabled a huge increase in the number of transistors per chip that has boosted an impressive growth of the microelectronics industry. Nowadays, portable devices like laptops, smart-phones and smart-watches are ingrained in the collective of our civilization. The demands of processing speed of these wide-spread personal gadgets has pushed the microelectronics industry to design every year more energy-efficient and faster integrated circuits than the previous year.

However, nowadays shrinking the size of the silicon MOSFETs is no longer making them more energy-efficient or increasing their switching speed, so they can cope with the consumer demands. Thus, resulting operating speed of high-end chips has levelled off since the mid-2000s [1] despite the increased demands for switching speed.

For confronting this stagnation of the operation speed, Intel has projected by early 2020s to produce 5 nm chips. However, this could rise their production costs to over 16 billion USD, which is a whopping third of its current annual revenue [2]. Due to the large economical disadvantages involved by Moore's law, the microelectronics industry in the last years has started to look for ways to overcome the unfavorable economic standpoint of further shrinkage of traditional Silicon MOSFETs.

New materials besides Silicon that offer a better performance in terms of switching speed and energy efficiency, have been explored as viable option for chip fabrication in the last years. Materials such as SiGe and III-V compound semiconductors on Silicon platforms have been identified as the next major step of advanced CMOS technology nodes [3]. The reason behind this prediction, is given by the intrinsic properties of these materials, as their high carrier mobilities. To date, most of these transistors have been fabricated in planar structures and their short channels fabricated with III-V materials induce source-to-drain tunnelling earlier than Silicon counterparts and the quantization effects increase the threshold voltage of their transfer characteristics.

This unwanted increment in the leakage current has demanded the industry and academia solutions to overcome this phenomenon that is inherent in nanometre-scale devices. The architecture of these devices has been the next logical point where the community has sought solutions and improvements. For example, devices like the non-planar Gate-All-Around (GAA) nanowires have been recently suggested. The tunneling currents present in these devices ($L_g \approx 10\text{nm}$) are less severe, than in their planar counterparts. The disruptive 3D architecture of the GAA nanowires provides a higher electrostatic control of the gate over the channel, which helps to reduce unwanted leakage currents [4].

The development of new device architectures is just one way to seek for improvements in energy consumption. Another way is to develop devices that rely on operation mechanisms different from the traditional Si MOSFET. In 1987 S. Banerjee et al. [5] proposed the Tunneling field-effect transistors (TFET), that in contrasts to the carrier emission principle of the MOSFET relies on band-to-band Tunneling (BTBT) emission principle that potentially allows this devices to achieve a sub-threshold swing below 60mV/decade at room temperature [6].

TCAD in the Semiconductor Industry

The development of new nano devices has created the challenge of speeding up the transfer from prototypes to mass production while keeping research and development costs under control. For responding to the demands of microelectronics industry, Technology Computer Aided Design (TCAD) is a viable tool that can save an estimated

40% to 50% of semiconductor technology development costs and is able to reduce development time by 40% [7]. These reductions in research and development costs allow the exploration of new device architectures in a fast and flexible way. With TCAD, the integration of state-of-the-art materials in Silicon processes is faster and cheaper and the modelling of new physical phenomena allows to predict output characteristics.

Another benefit of the widespread use of TCAD in the semiconductor industry is the fact that it allows scientists and engineers to understand and analyze experimental data from new fabrication trials. In these new fabrication trials, the complexity of the fabricated structures and the presence of large parasitic defects often interferes with the task of easily estimating critical design parameters (geometry, doping, etc) directly from measurements without the guidance of TCAD simulations [8], [9]. For instance, the comparison of extracted experimental data to TCAD models can be very useful for the correct interpretation of resultant device characteristics of novel fabricated structures. Concrete cases like [10] and [11] reinforce the idea that TCAD simulations can be used as a guidance for further design improvements. An example are common defects like impurities in the insulator-channel interface that come from the fabrication process and impact the static current of the devices [12]. Using TCAD to model them can explain their impact on the channel mobility and its repercussions on the device characteristics [13].

A further meaningful argument for calibrated TCAD is that it can serve as starting point for the exploration of new device designs in semiconductor industry and academic environments. One case worthwhile to mention is the TFET. Simple TCAD models of band-to-band tunneling were used for understanding the new TFET structures. Unfortunately, none of the fabricated devices ever met the predictions when implemented [14], because the BTBT models used to predict their output characteristics were physically inaccurate and lacked the inclusion of non-idealities [15]. These inaccurate models raised over-optimistic expectations on the potential of TFETs for ultra-low power energy efficiency applications [16] that were progressively reversed when it was evident that physically imprecise models were used [17]. The cost and time invested in these fabrication trials

must have been largely superior to the cost of developing experimentally verified TCAD models of non-deal BTBT.

In the specific case of MOSFETs at the nanometer scale, quantum transport simulators have been developed to predict their output characteristics. To date, QTx, a quantum transport simulator [18] has been implemented into the Sentaurus Package. This is truly an important step to allow industry develop new structures and understand the output characteristics of new nanometre-scale MOSFETs. However, these computations still require expensive computational resources (GPUs and cluster) and long CPU times. Besides these pitfalls, not all relevant and physical processes are covered by QTx.

In contrast, S-Device [19] has had almost 20 years of continuous development of diverse physical models, and the drift-diffusion formalism is computationally cheap compared to a quantum transport code. The drift-diffusion formalism has the advantage of its flexibility. With the right physical models this simulator can be useful to model the transfer characteristics of switches reaching from 3D GAA NWs to power Si-C MOSFETs [20].

Goal of this work

The four year effort of this thesis has been focused on providing ways to use existing quantum correction models and novel ballistic mobility models for simulating III-V n-MOSFETs at the nanometre scale. The problem till now was that most of the models available in the commercial S-Device were developed for structures that are larger than the current sizes that the semiconductor industry technology road map requires [21]. Geometrical confinement and source-to-drain tunneling leakage currents were not calibrated for these new FET sizes. Furthermore, most of the attempts of using quantum corrections for simulating ballistic characteristics using the drift diffusion formalism to simulate III-V n-MOSFETs, in this work are carefully tested against quantum transport. Works like [22] have used full quantum transport, semi-classical multi-valley / multi-sub band transport and TCAD models to simulate short-channel III-V n-MOSFETs.

Outline of this work

This thesis focuses on drift-diffusion TCAD modelling of III-V MOSFETs in the ballistic regime. For achieving this goal, the following steps had to be reached. First, in Chapter 2, quantum drift-diffusion (QDD) tools, e.g. S-Device from Synopsys [19], were used to mimic the described quantum effects by careful calibration of the implemented quantum correction models. Through the whole chapter, the Density Gradient (*DG*) and the Modified Local Density Approximation (MLDA) models were used to simulate geometrical confinement in the channel of these devices. Since the source-to-drain tunnelling current (STDT) is a consequence of the short channel and the light electron effective masses (m_e) inherent in III-V materials, the Nonlocal Tunneling (NLT) model [19] was used to simulate STDT. The latter was applied in combination with either MLDA or (anisotropic) DG. All these phenomena were demonstrated for technologically relevant $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ double-gate ultra-thin-body (UTB) FETs and $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ gate-all-around (GAA) NWs.

Chapter 3 explains how a 2DEG density of states (DOS) and a 1DEG DOS were developed in the drift-diffusion formalism. To date, [19] used the default DOS model for a 3DEG, which generated a wrong Fermi energy E_F , thus the electrostatics in the sub-threshold regime and in the on-state are in most of the cases wrong. In this chapter, the QV_{GS} -curves of QTx and S-Device in the inversion regime were matched. For doing this, the sub-band energies and the non-parabolicity parameter from QTx were used, and the only parameter to fit was a mass parameter which is the same for all sub-bands. It represents an approximate average over all quantization masses extracted from QTx.

After preparing the right basis, we jump to the most important part of this thesis. Chapter 3 details the development and testing of different types of ballistic mobility (μ_b) models. This is a very important effort, because if we assume an excessive diffusive mobility of $10^4 \text{ cm}^2/\text{Vs}$ as in Chapter 2 there will be a pronounced current overshoot after the onset of inversion. As in these devices the electron-phonon scattering is almost negligible, the QTx ballistic simulations were taken as reference. For this reason, TCAD models of the "ballistic mobility" were developed to limit the current to the

correct QT values. A distinctive feature of these models compared to current literature [23]- [24] is that a local variant of the proposed ballistic velocity depends either on the quasi-Fermi potential or on the density. These choices are "TCAD-friendly" and preferred to a dependence on the electric field. Since the mean ballistic velocity is the key quantity for the mobility model, the physical consequences of the different expressions are analyzed and the resulting transfer characteristics are compared with each other. This chapter concludes that the best expression conserves the ballistic current.

Making a logical link to Chapter 4, Chapter 5 examines the consequences of the introduction of a ballistic mobility on the overall shape of the quasi-Fermi potential (QFP) and the quantum tunneling rates. The models used for STDT and for BTBT rely on the local QFP for calculating its electron generation rates (G_e), thus certain unexpected behaviors appear with the introduction of μ_b . These deviations from the expected behaviors are examined, and this chapter proposes solutions to mitigate the impact of μ_b on the tunnel currents.

Finally, after the exploration of all these topics in Chapter 5 a simulation scheme for simulating the so-called floating body effect is proposed. It includes the treatment of the SRH recombination model for the cases of geometrical confinement and μ_b , by implementing a quantum lifetime τ_n previously derived in [25].

It was adapted from a previous work to 2D and 3D structures. The 2DEG and 1DEG DOS models were tested in collaboration with Synopsys were tested and quantitatively assessed with a full band quantum transport code.

Chapter 2

Modelling of Channel Quantization

Part of this Chapter is based on the analyses published in [26] and [27].

2.1 Introduction

As devices are scaled down into the nanometre range, quantum effects start to play a major role and strongly impact their performance. For instance, geometrical quantum confinement in the body of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ FETs with a body thickness below 12 nm leads to a shift of the threshold voltage, and a change of gate capacitance, therefore altering significantly the transfer characteristics of the device [28]. Moreover, the thin potential barrier between source and drain in transistors with gate lengths shorter than 25 nm causes a strong leakage current known as source-to-drain tunneling (STDT) which deteriorates the sub-threshold swing (SS) [29].

Quantum transport (QT) simulators are still not mature for industrial environments and are computationally too expensive for larger and more complex 2D and 3D devices. Furthermore, not all features available in commercial TCAD packages are ready for use in QT simulators, e.g. band-to-band tunneling and certain scattering models

(impurity, surface roughness, etc.). Calibration capabilities are also limited in QT solvers.

The aim of this chapter is, therefore, to study how so-called quantum drift-diffusion (QDD) tools, e.g. the commercial simulator Sentaurus-Device (S-Device) from Synopsys [19], can be used to mimic the effects of quantum confinement by careful calibration of the implemented quantum correction models. This is demonstrated for technologically relevant $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ double-gate ultra-thin-body (UTB) FETs. and $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ gate-all-around Nanowires (GAA) NWs.

2.2 Theory

2.2.1 Density Gradient Approach

For the simulation of geometrical confinement perpendicular to the transport direction we used the Density Gradient (DG) approach that adds a quantum potential Λ to the classical potential in the computation of the electron density (n) [30]. This quantum potential depends on the carrier density n , the effective mass m_e , and a fitting parameter γ introduced in the model of Ref. [31] :

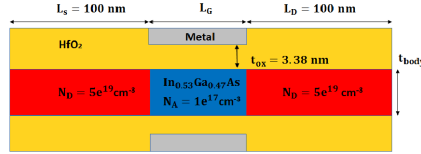
$$\Lambda_n = -\frac{\gamma \hbar^2}{6m_n} \frac{\nabla^2 \sqrt{n}}{\sqrt{n}}. \quad (2.1)$$

where $\gamma = \gamma_0 \cdot \gamma_{\text{pmi}}$. The parameter γ_0 can be used for calibration, but in this work it will be set to the default value of 1. The solution-dependent γ_{pmi} can be computed by a Physical Model Interface (PMI) to couple it with S-Device. The density n with Boltzmann statistics is defined as:

$$n = N_c^{3D} \exp\left(\frac{E_F - E_C}{k_B T}\right). \quad (2.2)$$

where N_c^{3D} is the effective density of states of a three-dimensional electron gas(3DEG) [32].

In order to calibrate γ , we simulated the 1D electron density profile along a vertical cut in the middle of the double-gate FET presented in Fig. 2.1 by using the Schrödinger-Poisson solver S-Band [32] as reference tool. Fig. 2.2(a) shows a comparison of the density profiles for a body thickness of 6 nm after calibration of γ . The corresponding

Figure 2.1: Schematic of an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ double-gate UTB FET.

QV-curves (inversion charge versus gate voltage) are presented in Fig. 2.2(b). Next, the same kind of simulations and fittings were done for a number of body thicknesses ranging from 4 nm to 10 nm in the gate voltage interval $V_G = [0, 1]$ V. The required effective masses m_e and non-parabolicity parameters α depend on the body thickness, which were taken from Ref. [33].

2.2.2 Electric Field Dependency of γ

We observed that the combination of geometrical and channel confinement makes γ a strong function of the normal electric field. The value of γ in the oxide and at the semiconductor-oxide interface (γ_{ox}) was maintained as a constant parameter and γ was modelled as a function to fit the entire V_G -range.

Based on the extracted QV-data the following form for the dependence on normal electric field at the interface (E_{norm}) was found

$$\gamma_{\text{pmi}}(E_{\text{norm}}) = (1/c) + a \cdot \tan(b E_{\text{norm}}) \quad (2.3)$$

with E_{norm} measured in [V/cm]. This function is continuous and smooth over the entire V_G -range which assures convergence when implemented in a PMI of S-Device. The fitting parameters a , b , and c depend on the body thickness t_{body} and are modelled as a quadratic polynomial, e.g.

$$a = a_1 t_{\text{body}}^2 + a_2 t_{\text{body}} + a_3, \quad (2.4)$$

where t_{body} is measured in nm. The coefficients are $a_1 = -0.0053$, $a_2 = 0.07824$, $a_3 = 0.07824$, $b_1 = -6.566 \times 10^{-7}$, $b_2 = 8.78 \times 10^{-6}$,

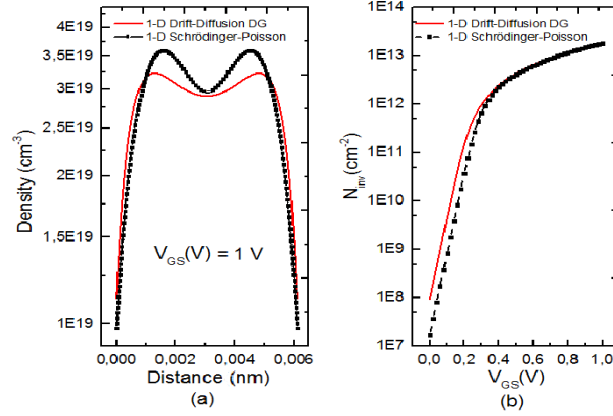


Figure 2.2: (a) Comparison of the electron density and (b) inversion charge vs. gate voltage computed with the isotropic DG model (red curves) and the Schrödinger-Poisson solution from S-Band (black dashed curves). In (b) $\gamma = 0.6$ which gives the best fit for $V_{GS} = 1 V$.

$b_3 = -2.007 \times 10^{-5}$, $c_1 = -0.006163$, $c_2 = 0.1065$, $c_3 = 0.5677$. For the simulation results presented in Fig.2.3, $t_{body} = 7 \text{ nm}$, $a = 0.499$, $b = 9.216 \times 10^{-6}$ and $c = 1.011$ were used.

Fig. 2.3(a) shows the field-dependence of γ for various body thicknesses. Fig. 2.3(b) demonstrates how the $I_D V_{GS}$ -curve in the case $t_{body} = 4 \text{ nm}$, $L_G = 10 \text{ nm}$ changes using the extreme values of γ extracted from Fig. 2.3(a) and the field-dependent γ , respectively.

2.2.3 Anisotropic Density Gradient Model

One important feature of the DG model in S-Device is its anisotropy [19]. The anisotropic DG model uses an attenuation matrix with diagonal elements α_l , α_v in order to scale A_n in longitudinal and vertical direction, respectively. The element α_v serves to reproduce the effect of geometrical confinement perpendicular to the transport direction, whereas the α_l -value lowers the height of the energy barrier between source and drain, which mimics the effect of STDT.

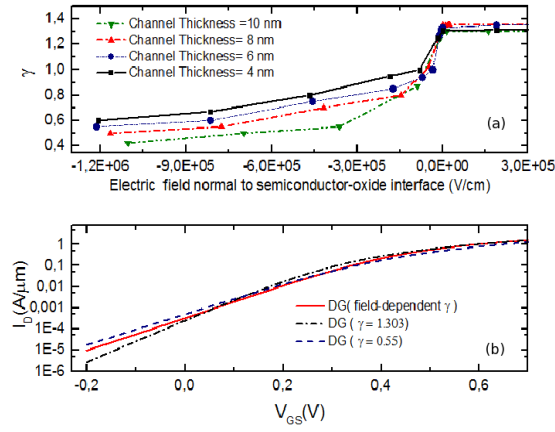


Figure 2.3: (a) DG fitting parameter γ as function of normal electric field at the interface for different body thicknesses. (b) Comparison of the $I_D V_{GS}$ characteristics ($V_{DS} = 0.05$ V) of an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ double-gate UTB FET ($t_{\text{body}} = 4$ nm, $L_G = 10$ nm) computed for $\gamma = 0.55$, $\gamma = 1.303$, and using a field-dependent γ (PMI model).

2.2.4 Modified Local Density Approximation

Another available model from [19] to simulate geometrical confinement is the Modified Local Density Approximation (MLDA). This model has no fitting parameter [34] and can be used to simulate geometrical confinement in the ultra-thin body perpendicular to the transport direction. The confined n at a distance y from the oxide/channel interface [34] with Boltzmann statistics, is defined by

$$n_{\text{MLDA}}(\eta_n) = N_c^{3\text{D}} \exp(\eta_n) \left[1 - \exp\left(-\left(\frac{y}{\lambda_n}\right)^2\right) \right] \quad (2.5)$$

where $\eta_n = \frac{E_{\text{F},n} - E_{\text{C}}}{k_{\text{B}}T}$, $\lambda_n = \left(\frac{\hbar}{2m_{\text{qm}}k_{\text{B}}T}\right)^{\frac{1}{2}}$ and m_{qm} is the quantization mass. One drawback of this model is that it does not allow the simulation of wave penetration in the oxide.

2.3 Simulation Method

2.3.1 2D Simulations of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FETs

For the simulation of STDT in S-Device in combination with geometrical confinement, the quantum transport code QTx [18] was used as reference tool. Transfer characteristics for gate lengths ranging from 10 nm to 25 nm were computed and the SS values were extracted for comparison. We extracted two sets of transfer characteristics, one for $V_{\text{DS}} = 0.05$ V (linear regime) and one for $V_{\text{DS}} = 0.63$ V (saturation regime). We applied three S-Device models to simulate STDT: (i) the anisotropic DG model with an optimal α_1 -value for each gate length that can reproduce the SS obtained from QTx (ii) the MLDA [34], and (iii) the Nonlocal Tunneling (NLT) model [19]. To simulate STDT, NLT can be applied in combination with either MLDA or (anisotropic) DG. In this combination, geometrical confinement in the ultra-thin body is enforced by the MLDA or (anisotropic) DG model. The tunneling mass (m_c) of the NLT model could of course also be used to better match SS and drain-induced barrier lowering (DIBL). The NLT model computes the semi-classical barrier tunneling current using distribution functions with local quasi-Fermi potentials (QFPs)

at the classical turning points [19] and the formula has been derived for a 3D electron gas (3DEG).

2.3.2 3D Simulations $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ GAA Nanowires

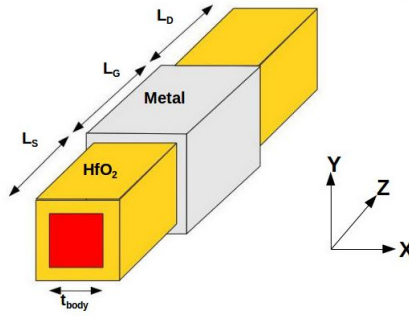


Figure 2.4: Schematic of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ gate-all-around nanowire (GAA NW) FET.

In a subsequent step in this work 3D simulations were done for gate-all-around nanowire (GAA NW) FETs depicted in Fig. 2.4. Their design parameters are the same as for the DG UTB FETs. The dimensions of the devices under study are given in Table 2.1.

The $I_D V_{GS}$ -characteristics simulated with QTx were used as a reference to calibrate the QDD simulation set-up of S-Device [32]. Effective masses and non-parabolicity parameters were calculated from a full-band version of the QTx code.

In the case of the 3D simulations of the GAA NWs we applied the anisotropic DG model. In the 3D case the α has three elements and correspond to the directions x,y and z. The α_v scales A in the y and x direction and serves to reproduce the effect of geometrical confinement perpendicular to the transport direction (both α_{vy} , α_{vx} are set to 1), whereas the α_1 (z direction)-value lowers the height of the energy barrier between source and drain, which mimics the effect of STD. Since the latter was simulated directly by the NLT model of S-Device, $\alpha_1 = 0$ was chosen. The fitting parameter γ for both transverse

Table 2.1: Dimensions of the GAA NW FETs according to future technology nodes as described in [4].

Node	L_G [nm]	t_{ox} [nm]	t_{body} [nm]	m_e/m_0	α_{np}
A	15	3.75	7	0.0642	1.2
B	10.4	3.25	5.5	0.0674	1.1

directions (x and y) was calibrated using the 1D density profile in the center of the GAA NW along the y-direction in the middle yz-plane (compare Fig. 2.4). The QTx electron density profile along the same line was taken as reference, see Fig 2.5 (a). By matching the densities in the sub-threshold regime (see Fig. 2.5(b)), the electrostatics (i.e. the threshold voltage) can be fitted very well. Next, as in the case of the 2D DG UTB FETs simulations, we also used the MLDA in the 3D GAA aNWs, without any fitting parameter [34].

Finally to include STDT, the NLT model [19] was used in combination with the MLDA and the anisotropic DG model, respectively. The tunneling mass (m_c) was set to the value of m_e .

2.4 Simulation Results

2.4.1 2D DG UTB FETs

The small difference in the threshold voltage between QTx and S-Device (≈ 60 mV) is due to different energy zeros and was not removed in Figs. 2.6-2.8 for better visibility. The first model we used to simulate STDT was the anisotropic DG model turning off NLT. Fig. 2.6(a) presents the fitted $I_D V_{GS}$ -curves for three gate lengths computed with this model for $V_{DS} = 0.05$ V and Fig. 2.6(b) presents the fitted $I_D V_{GS}$ -curves for $V_{DS} = 0.63$ V. It is important to note that these curves were obtained with $\alpha_v = 1$. This means that the quantum potential A was not further modified in vertical direction, i.e. we are able to accurately reproduce the confinement-induced shift of the threshold voltage. However, at the high $V_{DS} = 0.63$ V the

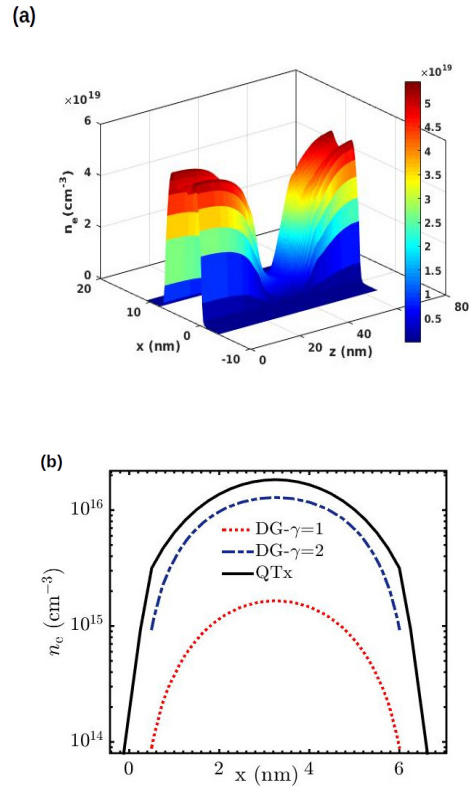


Figure 2.5: (a) Quantum-mechanical n contour plot from QTx of an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ GAA NW. (b) Vertical 1D electron density profiles at the center of the channel of the GAA NW with $t_{\text{body}} = 5.5 \text{ nm}$ at $V_{\text{GS}} = 0.1615 \text{ V}$ and $V_{\text{DS}} = 50 \text{ mV}$. Note that $\gamma = 2$ in the channel region and $\gamma = 30$ in the oxide region, fairly matches the quantum mechanical n from QTx. The γ field dependence was neglected.

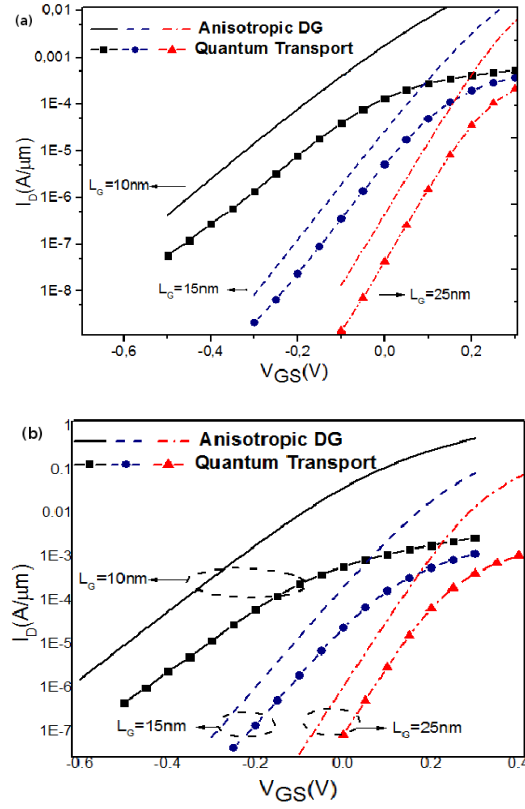


Figure 2.6: $I_D V_{GS}$ -characteristics from the anisotropic DG model for a 7 nm UTB FET with different gate lengths (a) $V_{DS} = 0.05$ V, parameters: $\gamma = 0.6$, $\alpha_v = 1$, $\alpha_1 = 0.7$ ($L_G = 10$ nm), $\alpha_1 = 0.4$ ($L_G = 15$ nm and $L_G = 25$ nm).

anisotropic DG model requires a larger α_1 to reproduce STD T for all gate lengths.

The fitted transfer characteristics for different gate lengths computed with the combination MLDA + NLT ($m_c = 0.0516 m_0$) are presented in Fig. 2.7(a) for $V_{DS} = 0.05$ V and in Fig. 2.7(b) for $V_{DS} = 0.63$ V. At high source-drain voltage ($V_{DS} = 0.63$ V), MLDA+NLT can reproduce the leakage current due to STD T better than at low V_{DS} , in particular for the shortest gate length.

Finally, the fitted characteristics obtained by the combination of DG (with field-dependent γ) + NLT are presented in Fig. 2.8(a) for $V_{DS} = 0.05$ V and in Fig. 2.8(b) for $V_{DS} = 0.63$ V. For this combination, the STD T currents a low and high V_{DS} are of comparable quality. Note that $\alpha_1 = 0$ by default when NLT is used [19].

For all the "ballistic" simulations with S-Device, a constant channel mobility of $2.26 \times 10^4 \text{ cm}^2/\text{Vs}$ was used which leads to arbitrary ON-currents. Note that in the drift-diffusion transport model the current diverges for $L_G \rightarrow 0$ and the empirical high-field saturation was not included. An important fact to point out is that the inclusion of a ballistic correction in the total mobility would reduce the ON-current by orders of magnitude. TCAD models of reduced dimensionality and their impact on the $I_D V_{GS}$ -curves will be discussed in Chapter 3, and the ON-current will be subject of Chapter 4. As the ON-current was not a subject of this Chapter, no attempt was made to fit it to the values obtained with QTx.

2.4.2 3D GAA Nanowires

This section shows the results from the simulations of GAA NW FETs with the parameters descried in Tab. 2.1.

Figs. 2.9 and 2.10 present the transfer curves of the GAA NW FETs from node A and B, respectively, in the linear regime ($V_{DS} = 50$ mV) and in the saturation regime ($V_{DS} = 0.63$ V). Again, the combinations DG + NLT and MLDA + NLT with a constant high mobility were used ($2.26 \times 10^4 \text{ cm}^2/\text{Vs}$).

The following remarks have to be made: (i) The apparent misfits in the threshold voltages between the DG and the QTx curves in Fig. 2.9 are not caused by the electrostatics, but due to the uncorrected high constant mobility. (ii) The corresponding misfits in Fig. 2.10 are also

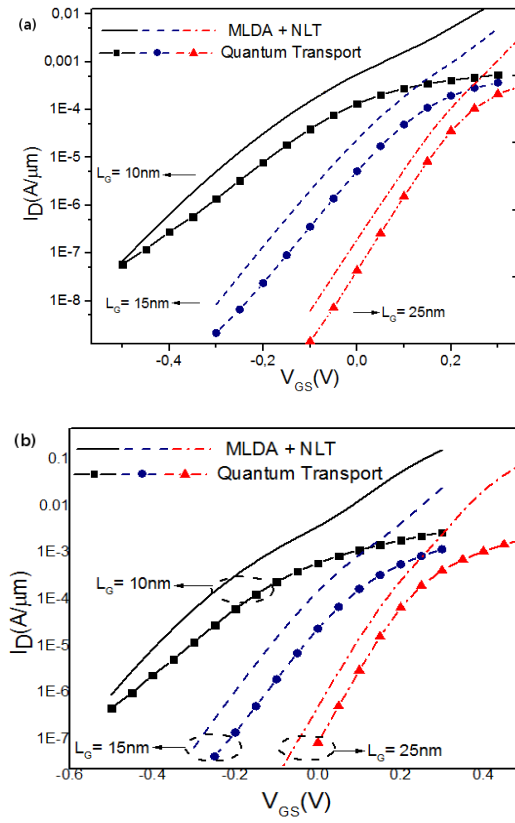


Figure 2.7: (a) $I_D V_{GS}$ -characteristics obtained from the combination MLDA + NLT for a 7 nm UTB FET with different gate lengths. (b) The same but for $V_{DS} = 0.63$ V

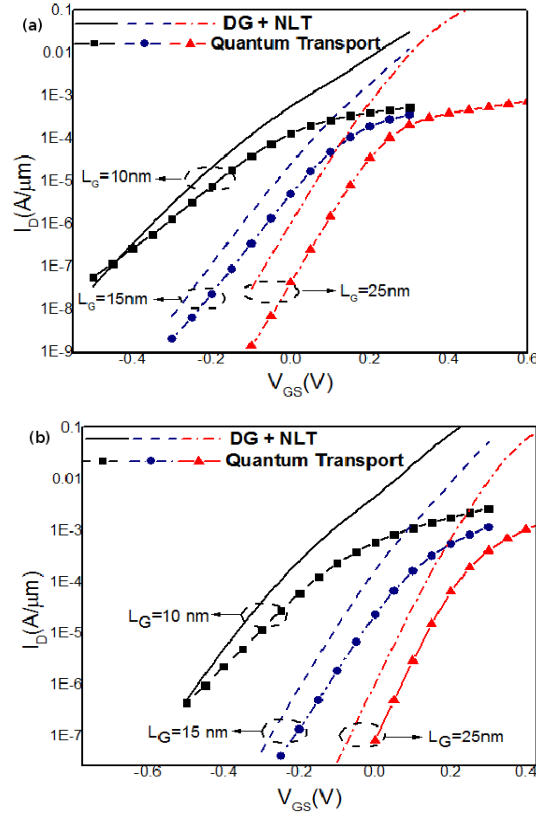


Figure 2.8: (a) $I_D V_{GS}$ -characteristics obtained from the combination of DG (with field-dependent γ) + NLT for a 7 nm UTB FET with different gate lengths. $V_{DS} = 0.05$ V, $m_c = 0.0516 m_0$. (b) The same but for $V_{DS} = 0.63$ V

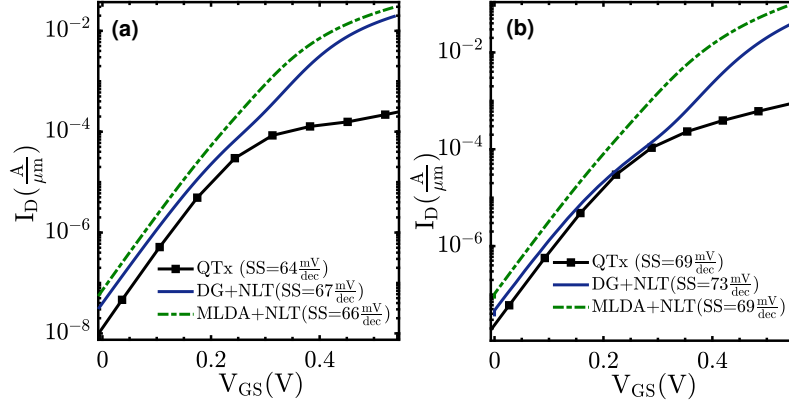


Figure 2.9: $I_D V_{GS}$ -characteristics obtained from the combination of DG + NLT and MLDA + NLT for a GAA NW FET with $t_{\text{body}} = 7$ nm at (a) $V_{DS} = 0.05$ V and (b) $V_{DS} = 0.63$ V. Parameters: $m_c = 0.0642 m_0$, $\gamma = 1.0$, $(\alpha_1, \alpha_v) = (0, 1)$, and $WF = 4.93$ eV.

due to a real difference between the threshold voltages of the DG and QTx curves, since the adjusted γ -value for $t_{\text{body}} = 5.5$ nm ($\gamma = 2$) prevented convergence, and $\gamma = 1$ was used instead.

From the extracted SS values in Figs. 2.9 - 2.10 one observes that despite the superior electrostatic control in the GAA NWs, a significant leakage current caused by STDT persists. The best way to illustrate its effect is by showing the spectral current distribution in comparison to the shape of the lowest CB edge. This is done in Fig. 2.11 for the off-state and the on-state, respectively, comparing the gate lengths 15 nm and 10.5 nm. The ON-current is basically thermionic current in both cases, i.e. STDT has no effect here. The OFF-current is dominated by STDT - almost completely in the case of $L_G = 10.5$ nm. Therefore, a gate length of 10 nm can be considered as the end of scaling for III-V-channel FETs. Further scaling will also not significantly improve the ON-current.

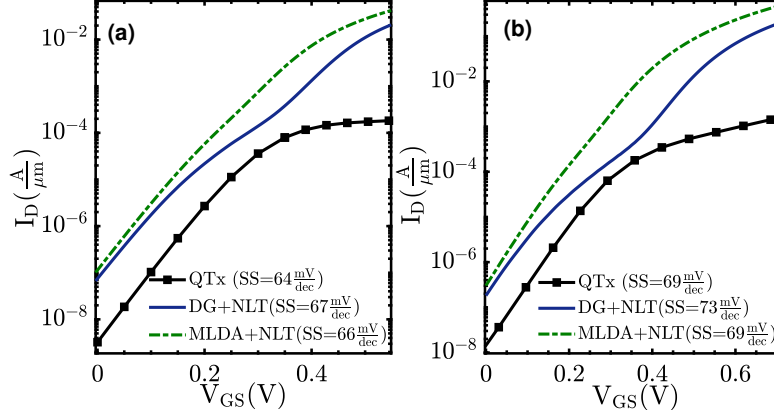


Figure 2.10: $I_D V_{GS}$ -characteristics obtained from the combination of DG + NLT, MLDA + NLT and μ_b for a GAA NW FET with $t_{\text{body}} = 5.5$ nm at (a) $V_{DS} = 0.05$ V and (b) $V_{DS} = 0.63$ V. Parameters: $V_{DS} = 0.05$ V, $m_c = 0.0674 m_0$, $\gamma = 1.0$, $(\alpha_1, \alpha_v) = (0, 1)$, and $WF = 4.98$ eV.

2.5 Discussion

In the case of the 2D simulations, the anisotropic DG model with constant γ can fairly reproduce the SS of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ UTB FETs even in case of strong STDT, but has the disadvantage that the parameter α_1 has to be fitted for each gate length. The NLT model, when used together with DG and field-dependent γ , needs less calibration of the tunneling mass m_c than in combination with MLDA to match the SS reference values. A drawback of this combination is that MLDA does not include the effect of wave penetration into the oxide [34]. As viable choice one can recommend the combination of NLT and DG with field-dependent γ , as just the same tunneling mass m_c (with physically correct value) can be used for all gate lengths and no further α -fitting is necessary.

In the case of the 3D simulations, the DG curves in Figs. 2.9 - 2.10 exhibit a pronounced bump around the onset of inversion. The strength of this effect increases with decreasing cross section

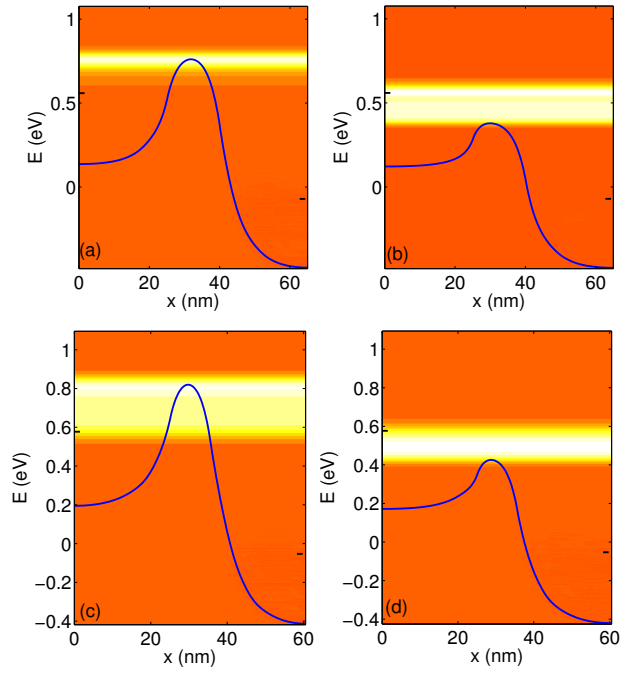


Figure 2.11: Distribution of spectral currents and lowest conduction band edge for GAA NW FETs from node (A) and (B) at $V_{DS} = 0.63$ V. (a) Off-state, $L_G = 15$ nm at $V_{GS} = 0.026$ V. (b) On-state, $L_G = 15$ nm at $V_{GS} = 0.48$ V. (c) Off-state, $L_G = 10.4$ nm at $V_{GS} = 0.03$ V. (d) On-state, $L_G = 10.4$ nm at $V_{GS} = 0.48$ V.

of the GAA NW and increasing source-drain bias. It is attributed to the breakdown of the DG method when the channel quantization fades out and only geometrical confinement remains. The DG model had been developed for the former case, but is unable to cover the latter correctly. One can even observe the artifact that in the case of extremely thin bodies a huge quantum potential is computed in the flat-band regions of the semiconductor (also when $\alpha_1 = 0$). An empirical workaround to mitigate this effect is to remove the gate oxide in small parts near the source/drain contacts which eliminates the "density gradient" between insulator and semiconductor there. However, this was not done in the present study.

2.6 Summary

In this Chapter, geometrical quantization in the channels of 2D and 3D devices was simulated with the DG, anisotropic DG and MLDA. The STDt was simulated using the NLT model or the anisotropic DG. The NLT model used in combination with the anisotropic DG model can reproduce the reference SS of the DG UTB FETs in the case of strong STDt, which is an improvement over the anisotropic DG. In GAA NWs the excellent gate electrostatic control reduces SS significantly compared to the DG UTB FETs with the same L_G [4]. Here better results are found with the combination of DG and NLT.

Chapter 3

Density of States

The content of this chapter was published in [36].

3.1 Introduction

In this Chapter TCAD models of a 2D/1D DOS will be presented. These models can reproduce the electrostatics of devices with strong geometrical confinement, such as 2D DG UTB FETs and 3D GAA NWs. To the date [19], the default DOS model in S-Device is that of a 3DEG

$$N_c^{3D} = \frac{1}{\sqrt{2}} \left(\frac{m_c k_B T}{\pi \hbar^2} \right)^{3/2} \quad (3.1)$$

where m_c is the DOS effective mass and k_B the Boltzmann constant. As seen in the previous Chapter if it is possible to reproduce the threshold voltage shift by shifting the WF in accordance with the energy barriers of QTx or introducing a model that can reproduce the quantum mechanical electron density profile.

However, as seen in the previous results (see Fig. 2.9-Fig. 2.10) from the previous Chapter the DG method breaks when the channel quantization fades out and only geometrical confinement remains. The DG model had been developed for the former case, but is unable to cover the latter correctly. Thus the simulation of low-dimensional

devices also 2DEG/1DEG DOS models are available in [19], that can be used in combination with Fermi statistics.

3.2 Theory

In the III-V devices under study, side valleys are negligible [33], only the Γ -valley is populated. Size quantization generates sub-bands (subscript i), and electron densities (in units cm^{-3}) are computed with the expressions

$$n_{2D} = \sum_i N_c^{3D} (m_i^{2D}) F_0(\eta, \epsilon_i, \alpha_{np,i}) , \quad (3.2)$$

$$n_{1D} = \sum_i N_c^{3D} (m_i^{1D}) F_{-\frac{1}{2}}(\eta, \epsilon_i, \alpha_{np,i}) . \quad (3.3)$$

Here, $F_p(\eta, \epsilon_i, \alpha_{np,i})$ is the generalized Fermi integral

$$F_p(\eta, \epsilon_i, \alpha_{np,i}) = \int_0^\infty d\epsilon \frac{[\epsilon(1 + k_B T \alpha_{np,i} \epsilon)]^p (1 + 2k_B T \alpha_{np,i} \epsilon)}{1 + \exp(\epsilon + \epsilon_i - \eta)} \quad (3.4)$$

with $\eta = (E_{F,n} - E_c)/k_B T$ and $\epsilon_i = \Delta E_{n,i}/k_B T$. The latter is the normalized sub-band energy referenced to the bulk band edge E_c . The mass parameters $m_i^{2D,1D}$ are derived from the actual quantization masses $m_{qu,i}^{2D,1D}$ by the transformations

$$m_i^{2D} = \left(\frac{m_{qu,i}^{2D}}{t_{body}} \right)^{2/3} \left(\frac{2\pi\hbar^2}{k_B T} \right)^{1/3} , \quad (3.5)$$

$$m_i^{1D} = \left(\frac{m_{qu,i}^{1D}}{t_{body}} \right)^{1/3} \frac{1}{t_{body}} \left(\frac{\pi\hbar^2}{k_B T} \right)^{2/3} , \quad (3.6)$$

and ensure that the prefactors in Eqs. (3.2) and (3.3) are the correct effective densities of states when N_c^{3D} from Eq. (3.1) is used. The parameter $\alpha_{np,i}$ accounts for the non-parabolicity of each sub-band.

The condition that the drift-diffusion current must vanish at equilibrium also when Fermi statistics is required leads to a Fermi correction term in the equation for the current density [19]:

$$j_{FS} = -\mu_n n k_B T \nabla \ln(\gamma_n) \quad \gamma_n = n/n_B , \quad (3.7)$$

where n is one of the densities Eq. (3.2) or Eq. (3.3) and n_B their Boltzmann limit. The correction term is negative, hence at high V_{GS} the current density is reduced when the quasi-Fermi level moves into the first sub-band. The term has, therefore, the same effect as a ballistic mobility correction. Before one can make any statements about the latter, a careful calibration of the DOS is necessary.

There are two possibilities to fit the 2DEG/1DEG DOS in S-Device. The first one is to extract the sub-band energies $\Delta E_{n,i}$, the quantization masses $m_{qu,i}^{1D,2D}$, and the non-parabolicity parameters $\alpha_{np,i}$ of the populated sub-bands from the QTx dispersion curves. From the quantization masses needed for the input file the mass parameters $m_i^{1D,2D}$ will be computed internally. The second possibility is to fit the DOS by matching the QV_{GS} -curves of QTx and S-Device in the inversion regime. Naturally, one can only fit *one* parameter. The best way is to take the sub-band energies $\Delta E_{n,i}$ and the non-parabolicity parameter α_{np} from QTx and to fit a mass parameter $\langle m_{qu}^{1D,2D} \rangle$ which is the same for all sub-bands. It represents an approximate average over all $m_{qu,i}^{1D,2D}$ extracted from QTx.

3.2.1 2D-DOS

To determine the work function and the input parameter $\langle m_{qu}^{2D} \rangle$ for the 2DEG DOS, the DG UTB FET of Fig. 2.1 was simulated at $V_{DS} = 0$ V both with QTx and S-Device, including 4 sub-bands.

First, to minimize the effect of quantum-mechanical charge penetration into the source-to-drain potential barrier and to make the fitting of the WF feasible, a long gate with $L_G = 40$ nm was chosen. The good match of the two QV_{GS} -curves in the sub-threshold and inversion regimes is shown in Fig. 3.1.

Next, we used the values of the WF and $\langle m_{qu}^{2D} \rangle$ obtained from the previous calibration and simulated QV_{GS} -curves for shorter gate lengths. Because STDT determines the SS at those gate length, the corresponding charge penetration into the barrier at equilibrium ($V_{DS} = 0$) also changes the slope of the QV_{GS} -curves in the sub-threshold range. Fig. 3.2 presents the case of $L_G = 15$ nm. It shows that the slope of the QTx-curve cannot be reproduced without a quantum correction in S-Device. Since the NLT model is only active when

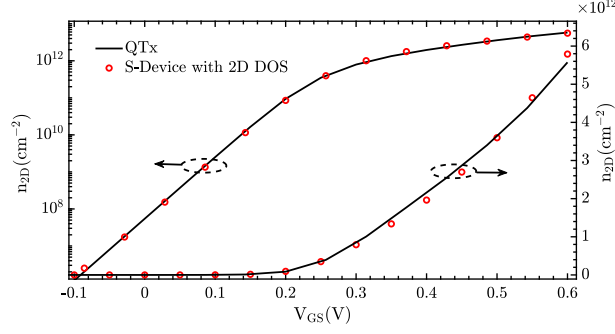


Figure 3.1: 2D electron density at the center of the channel vs. gate voltage in the DG UTB FET with $t_{\text{body}} = 7$ nm and $L_G = 40$ nm at $V_{\text{DS}} = 0$ mV. The threshold voltage was fitted including 4 sub-bands with $\langle m_{\text{qu}}^{2\text{D}} \rangle = 0.048 m_0$ and $\alpha_{\text{np}} = 1.1697 \text{ eV}^{-1}$ which gives a value of 4.78 eV for the WF.

$V_{\text{DS}} > 0$, the only way is to apply the anisotropic DG model in transport direction and to fit the value of α_1 .

3.2.2 1D-DOS

To determine the work function and the input parameter $\langle m_{\text{qu}}^{1\text{D}} \rangle$ for the 1DEG DOS, the GAA NW FET of Fig. 2.4 was simulated at $V_{\text{DS}} = 0$ V both with QTx and S-Device including 10 sub-bands. In order to minimize the effect of quantum-mechanical charge penetration into the source-to-drain potential barrier and to make the fitting of the WF feasible, a gate length of $L_G = 30$ nm is sufficient here (stronger source-to-drain barrier due to GAA architecture). A good agreement between the two QV_{GS} -curves is achieved as shown in Fig. 3.4.

Fig. 3.3 shows the extracted QTx band diagram at a point in the heavily doped source of the 3D GAA NW used for calibration. In the simulations with S-Device 10 sub-bands were included, with $\Delta E_{\text{n,i}}$ set to the value of the minimum of the i th sub-band.

Fig. 3.5 presents the results for $L_G = 15$ nm. As in the 2DEG case, the slope of the QTx-curve cannot be reproduced without a quantum

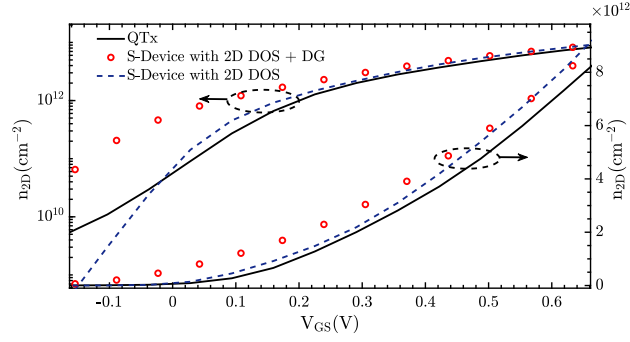


Figure 3.2: 2D electron density at the center of the channel vs. gate voltage in the DG UTB FET with $t_{\text{body}} = 7$ nm and $L_G = 15$ nm at $V_{\text{DS}} = 0$ mV. Parameters: $\langle m_{\text{qu}}^{2\text{D}} \rangle = 0.042 m_0$, $\alpha_{\text{np}} = 1.1697 \text{ eV}^{-1}$, $\text{WF} = 4.83 \text{ eV}$, and $(\alpha_l, \alpha_v) = (0.6, 0)$.

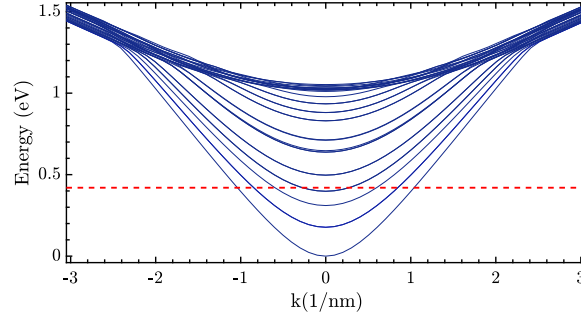


Figure 3.3: Band diagram calculated at a point in the heavily doped source of the 3D GAA NW FET with $t_{\text{body}} = 7$ nm. The dashed red line shows the position of the Fermi level.

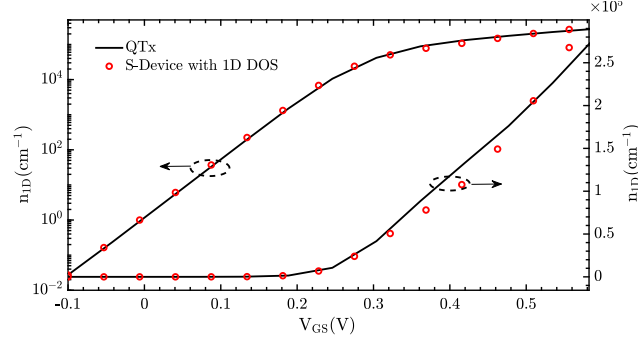


Figure 3.4: 1D electron density at the center of the channel vs. gate voltage in the GAA NW FET with $t_{\text{body}} = 7$ nm and $L_G = 30$ nm at $V_{\text{DS}} = 0$ mV. Parameters: $\langle m_{\text{qu}}^{1\text{D}} \rangle = 0.07 m_0$, $\text{WF} = 4.77$ eV, and $\alpha_{\text{np}} = 1.2 \text{ eV}^{-1}$.

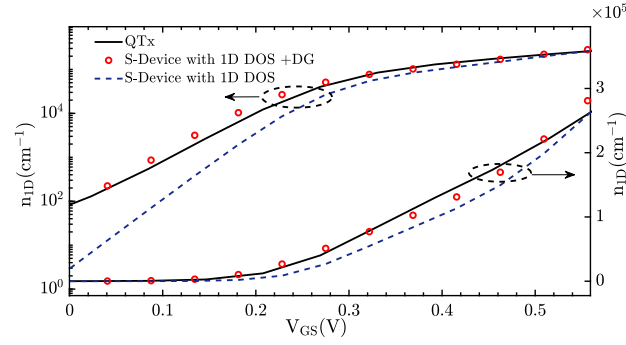


Figure 3.5: 1D electron density at the center of the channel vs. gate voltage in the GAA NW FET with $t_{\text{body}} = 7$ nm and $L_G = 15$ nm at $V_{\text{DS}} = 0$ mV. Parameters: $\langle m_{\text{qu}}^{1\text{D}} \rangle = 0.09 m_0$, $(\alpha_l, \alpha_v) = (1.1, 0)$, $\text{WF} = 4.8$ eV, and $\alpha_{\text{np}} = 1.2 \text{ eV}^{-1}$.

correction in S-Device. The anisotropic DG model in transport direction was applied with a fitted α_1 value.

3.3 Simulation Results

In the previous Chapter the default S-Device DOS model of a 3DEG and Boltzmann statistics were used. In this subsection, Fermi statistics and the DOS models for a 2DEG/1DEG are applied. As discussed above, Fermi statistics is needed since the quasi-Fermi level is way higher than the lowest sub-band energy in the studied transistors. The Fermi correction term in the current Eq.(3.7) reduces the current in the inversion regime significantly. This effect increases with decreasing dimensionality, i.e. it is stronger for a 1DEG than for a 2DEG [19].

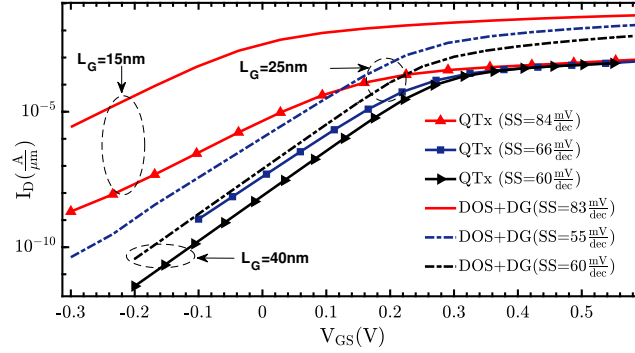


Figure 3.6: $I_D V_{GS}$ -characteristics obtained from the combination of 2DEG DOS + the anisotropic DG for a DG UTB FET with $t_{\text{body}} = 7$ nm and different gate lengths. Parameters: $V_{DS} = 0.05$ V, $\mu_d = 2.3 \times 10^4$ cm²/Vs, $\gamma = 0.7$, and $(\alpha_l, \alpha_v) = (0.8, 0)$. The fitted mass parameter $\langle m_{\text{qu}}^{2D} \rangle$ reduces from $0.038 m_0$ for $L_G = 40$ nm to $0.027 m_0$ for $L_G = 15$ nm.

Fig. 3.6 shows the fitted $I_D V_{GS}$ -characteristics of DG UTB FETs for three gate lengths computed with the combination of anisotropic DG model and a 2DEG DOS with the fitting parameters obtained from Figs. 3.1 and 3.2. As in the previous subsection, the mobility is

constant and large here, without any ballistic correction. Anisotropic

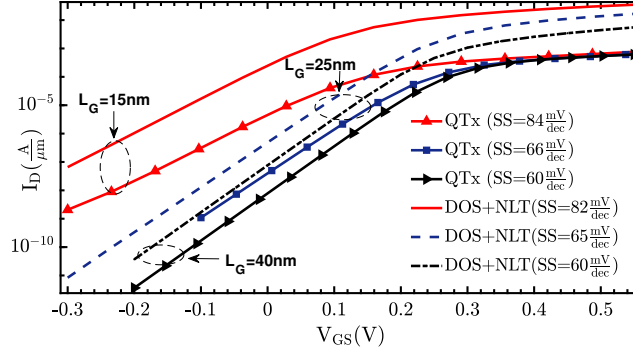


Figure 3.7: $I_D V_{GS}$ -characteristics obtained from the combination of 2DEG DOS + NLT for a DG UTB FET with $t_{\text{body}} = 7$ nm and different gate lengths. All other parameters are the same as in Fig. 3.6.

DG with fitted α_1 is able to reproduce the slopes of the QTx-characteristics very well. The same can be achieved with the NLT model with the statement "Multivalley". The model involves one tunnelling path for each sub-band and the tunneling mass is given by the corresponding $\langle m_{\text{qu},i}^{2D} \rangle$. In this particular case, the sub-bands closest to the quasi-Fermi level contribute most and their effective masses are slightly smaller than the one of the first sub-band ($\langle m_{\text{qu}}^{2D} \rangle = 0.042 m_0$ and $m_c = 0.0516 m_0$). This results in a good agreement with the sub-threshold region of the QTx reference. Fig. 3.8 presents the transfer curves in the linear regime for a GAA NW from node A with a sufficient gate length ($L_G = 30$ nm) to avoid STD T.

Fig. 3.9 shows the transfer curves of GAA NWs from node A in the linear and saturation regime, respectively. Fig. 3.10 presents the corresponding curves for node B. Again, the combinations 1DEG DOS, 1DEG DOS + DG and 1DEG DOS + NLT with a constant high mobility were used.

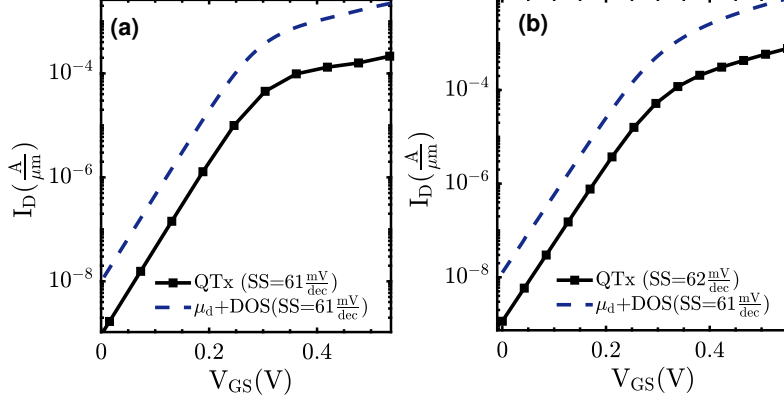


Figure 3.8: $I_D V_{GS}$ -characteristics obtained from the combination of DG + 1DEG DOS, and μ_d + 1D DOS for a GAA NW FET with $t_{\text{body}} = 7$ nm and $L_G = 30$ nm. Parameters: $\langle m_{\text{qu}}^{1D} \rangle = 0.07 m_0$, $WF = 4.77$ eV, and $\alpha_{\text{np}} = 1.2 \text{ eV}^{-1}$.

3.4 Discussion

From the above results the following remarks can be made:

(i) When using the NLT model instead of the anisotropic DG to simulate STDT, its combination with the 1DEG DOS results in a saddle point of the current from the onset of inversion to high V_{GS} . The artifact is most pronounced for node B where STDT cannot be neglected. However, as can be seen from the spectral currents in Fig. 2.11, the effect of STDT vanishes at high V_{GS} . The reason is not clear. It could be related to the fact that the STDT formula of the NLT model belongs to a 3DEG. Finally the poor SS obtained by the 1DEG + NLT combination is caused by the fact that the sub-bands closest to the quasi-Fermi level contribute the most and in this case their effective mass is much higher than that of the first sub-band ($\langle m_{\text{qu},1}^{2D} \rangle = 0.09 m_0$ and $m_c = 0.0642 m_0$).

(ii) The inclusion of a ballistic mobility correction would further scale down all currents. This will be shown in Chapter 4.

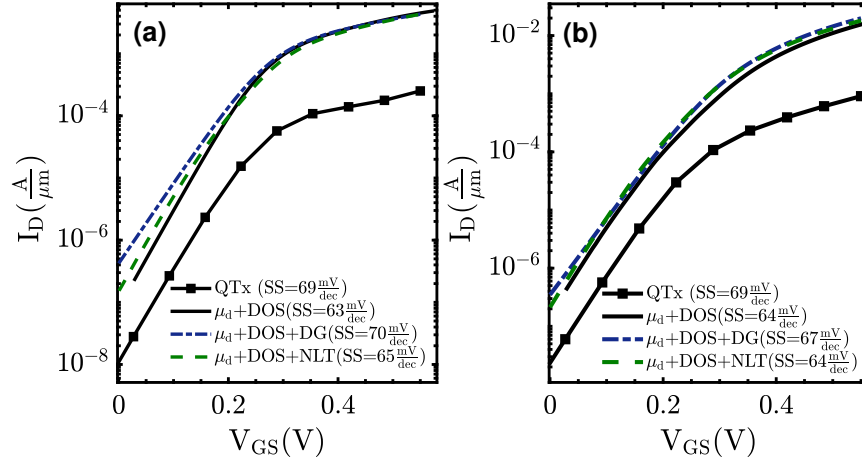


Figure 3.9: $I_D V_{GS}$ -characteristics obtained from the combinations 1DEG DOS, 1DEG DOS + NLT + μ_d for a GAA NW FET with $t_{\text{body}} = 7 \text{ nm}$ (node A) at (a) $V_{DS} = 0.05 \text{ V}$ and (b) $V_{DS} = 0.63 \text{ V}$. Parameters: $\alpha_{np} = 1.2 \text{ eV}^{-1}$, $\gamma = 0.7$, $(\alpha_l, \alpha_v) = (1.1, 0)$, and $WF = 4.8 \text{ eV}$.

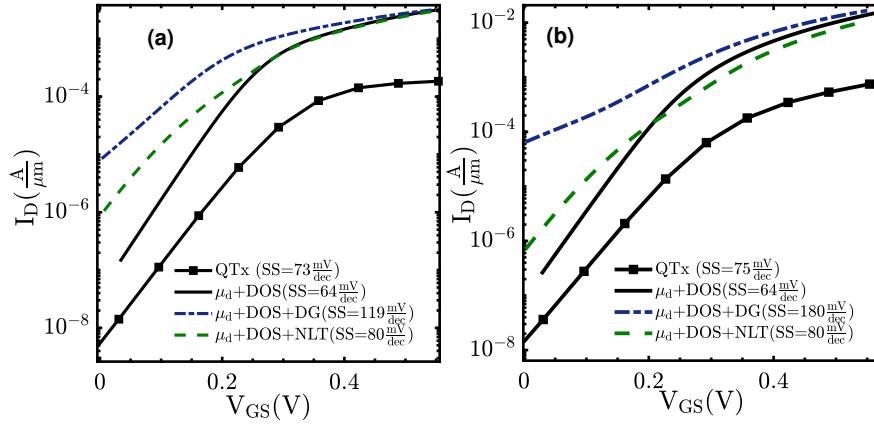


Figure 3.10: $I_D V_{GS}$ -characteristics obtained from the combinations 1DEG DOS, 1DEG DOS + NLT + μ_d for a GAA NW FET with $t_{\text{body}} = 5.5$ nm (node B) at (a) $V_{DS} = 0.05$ V and (b) $V_{DS} = 0.63$ V. Parameters: $\alpha_{\text{np}} = 1.1 \text{ eV}^{-1}$, $\gamma = 1.0$, $(\alpha_1, \alpha_v) = (0.7, 0)$, and $\text{WF} = 4.77 \text{ eV}$.

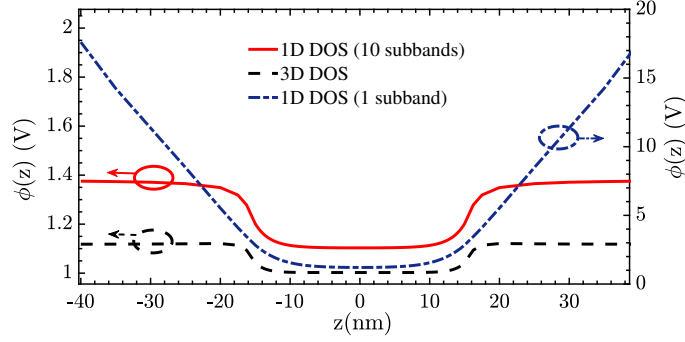


Figure 3.11: Electrostatic potential $\phi(z)$ along a longitudinal cut in a 3D GAA NW with $t_{\text{body}} = 7 \text{ nm}$ and $L_G = 30 \text{ nm}$ at $V_{\text{DS}} = 0 \text{ mV}$.

The value of $E_F = 0.92 \text{ eV}$ obtained from the QDD simulation by the inclusion of a 2DEG DOS was compared to the $E_F = 0.86 \text{ eV}$ measured in the drain contact from QTx. There is an overall agreement and an error of 6.5% only. In contrast, for the 1DEG DOS the error between QDD ($E_F = 1.06 \text{ eV}$) and QTx ($E_F = 0.42 \text{ eV}$) is 60%. This disagreement can be traced back to the form of the 1DEG DOS. The characteristic delta like shape of a 1DEG DOS cannot accommodate so much electrons as a 3DEG DOS. For having a flat electrostatic potential $\phi(z)$ as in Fig. 3.11 it is necessary to define more than one sub-band in the parameter file. That assures that $\phi(z)$ tends to be flat in the contacts, although not completely. The problem might be also related to the implementation of the algorithm inside the simulator, were the number of energy points for solving the Fermi integral to compute the carrier density might not be sufficient. In this work we used a high number of integration points in the Gauss-Laguerre quadratures when solving F_p Eq.(3.4) [19].

3.5 Summary

In this Chapter, models for a 2DEG and a 1DEG DOS were presented. The mass parameters ($m_i^{2D,1D}$) were derived from quantization masses

($m_{\text{qu},i}^{1\text{D},2\text{D}}$) and used to fit the QV_{GS} -curves to a QTx reference to reproduce the electrostatics of 2D simulations of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET and 3D simulations $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ GAA NWs from Chapter 2.

Chapter 4

Ballistic Mobility

The content of this chapter was accepted by Solid-States Electronics, A. Schenk and P. Aguirre, TCAD Models of the Ballistic Mobility in the Source-to-Drain Tunneling Regime [37]. Parts of the Results section in this Chapter were presented in [38].

4.1 Introduction

In the previous Chapters Sentaurus-Device from Synopsys [19] was used to simulate STDt in $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FETs with gate lengths L_G ranging from 10 nm to 25 nm (see Figs. 2.6-2.8). Their $I_D V_{GS}$ -characteristics exhibit a pronounced current overshoot after the onset of inversion due to the exclusive usage of a diffusive mobility (μ_d). Since electron-phonon scattering is weak in these devices, the so-called ballistic QT solution can be taken as reference. The aim of this work is to develop a TCAD model of the ballistic mobility capable to limit the current to the correct QT values. As distinct from the various models in the literature [23, 24, 39–41], the ballistic velocity is modelled explicitly either as a function of the quasi-Fermi potential (QFP) or the density (n). Though conceptually questionable in general, such choices are TCAD-friendly and hence preferred to other dependencies e.g. on the electric field. Since the mean ballistic

velocity is the key quantity for the mobility model, the physical consequences of the different expressions are analysed and the resulting transfer characteristics are compared with each other.

4.2 Theory of Transport Regimes and Ballistic Mobility

The first moment of the Boltzmann transport equation in relaxation time approximation reads in the stationary, iso-thermal, and non-degenerate case [42]:

$$\left(\mu_d^{-1} + \frac{m_e}{q} \vec{v} \cdot \nabla \right) \vec{v} = \nabla \psi_n, \quad (4.1)$$

where μ_d is the diffusive mobility, \vec{v} the mean velocity, m_e the effective mass, and q the elementary charge.

The function $\psi_n = \Phi - \frac{kT}{q} \ln(n/n_i)$ on the right-hand side contains the electrostatic potential Φ , the density n , the intrinsic density n_i , and the thermal voltage $V_T = k_B T$. The second term on the lhs, typical for hydrodynamic equations, is generally skipped in 2D/3D device simulators to ensure convergence. The function ψ_n has the meaning of the QFP, provided the concept of local equilibrium is (at least approximately) applicable. Simplifying Eq. (4.1) to the 1D case (channel direction x) one obtains

$$\frac{v}{\mu_d} + \frac{m_e}{q} v v' = \psi_n'. \quad (4.2)$$

Where $(')$ stands for the derivate over space. The second term becomes larger than the first one if $\mu_d > q/(m_e |v'|)$. For an estimate one can replace v' by v_{th}/L_G with the root mean square (r.m.s.) 1D thermal velocity $v_{th} = \sqrt{k_B T/m_e}$ which gives approximately $\mu_d > L_G \sqrt{\frac{q}{(V_T m_e)}} \equiv \mu_{lim} (V_T = k_B T/q)$. For example, $L_G = 15$ nm and $m_e = 0.0516 m_0$ yield $\mu_{lim} = 1719 \text{ cm}^2/\text{Vs}$.

The proportionality to v in the second term of Eq. (4.2) suggests to introduce a quantity called *ballistic mobility* in the form

$$\mu_b = \frac{q}{m_e} \frac{1}{v'_b} \quad (4.3)$$

4.2. THEORY OF TRANSPORT REGIMES AND BALLISTIC MOBILITY 41

emission	QFP in contacts, no QFP defined in channel
kinetic	QFP in contacts, constant QFP in channel, $\mu_d = \infty$
ballistic	non-constant QFP $\psi_n(x)$ in channel can be assumed and used, $\mu_d \gg \mu_{lim}$
quasi-ballistic	intermediate regime, $\mu_d \approx \mu_b$
diffusive	dissipative regime, $\mu_d \ll \mu_b$

Table 4.1: Nomenclature for different transport regimes in relation to QFP and mobilities.

which requires to define a local model for the mean ballistic velocity $v_b(x)$ (for nomenclature of the different transport regimes see Table 1).

The *ballistic* regime is defined by $\mu_d \gg \mu_{lim}$. The quotation marks mean that μ_d still has to be small enough such that a local, non-constant QFP $\psi_n(x)$, i.e. the concept of local thermodynamic equilibrium, can be assumed and used. In the simulation samples sketched in Fig. 2.1 and Fig. 2.4, the measured bulk mobility of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ ($\sim 10^4 \text{cm}^2/\text{Vs}$) could be seen as such a limit.

There is no analytical solution of the differential equation (4.2) even for constant mobility μ_d . Setting $\mu_d = \infty$ makes Eq. (4.2) to an Euler equation and defines the *kinetic* transport regime (see Table 1) where in analogy to Eq. 4.3 a "kinetic mobility" could be introduced in the form

$$\mu_k = \frac{v_k}{\psi'_n} \left(= \frac{q}{m_e} \frac{1}{v'_k} \right) \quad (4.4)$$

with a mean kinetic velocity v_k . This expression for μ_k can also be motivated from Heisenberg's uncertainty principle (a method often used in the past for the development of TCAD mobility models, e.g. [43]). In the limit $L_G \rightarrow 0$ one can write $\Delta t \Delta E \approx \tau_k \Delta E_k \approx \hbar/2 \approx \Delta x m_e v_k \approx \Delta x \Delta p$, where τ_k is a 'kinetic relaxation time' related to the kinetic mobility by $\mu_k = q\tau_k/m_e$. By division, $\tau_k = (m_e v_k \Delta x / \Delta k)(\Delta E_k / \Delta k)^{-1}$, and with the parabolic band approximation $\Delta E_k / \Delta v_k = m_e v_k$ one obtains (4.4).

Since $\mu_d \gg \mu_b$ holds in the ballistic transport regime, the left-hand side of Eq. (4.2) can be approximated by v_b/μ_b (neglecting the first term) which yields another definition of the ballistic mobility:

$$\mu_b(x) = \frac{v_b(x)}{\psi'_n(x)}. \quad (4.5)$$

The nomenclature for the different transport regimes in relation to QFP and mobilities is summarized in Table 4.1. In the emission regime, the electron is decoupled from any thermodynamic bath, i.e. short-range electron-electron interaction is not efficient to ensure particle and energy exchange, hence a QFP is not defined. In textbooks, the course of the QFP in emission regions is often symbolized by crosses. Diodes (pn-junction and Schottky) have been treated in the emission theory with the result that the slope of the current-voltage characteristics is the same as in the diffusion theory, only the pre-factors differ (by a moderate factor).

In the kinetic transport regime, electrons keep the QFP of the contact from which they are injected until they hit the opposite contact. Short-range electron-electron interaction is efficient to ensure particle and energy exchange in the ensemble of moving electrons. The QFP is constant in the channel because of the absence of scatterings that can relax the total momentum of the system. At the channel-drain junction, due to the vast number of electrons in the drain reservoir and due to the very quick thermalization by short-range electron-electron interaction, the incoming electrons rapidly become members of the drain ensemble, i.e. they attain the Fermi level of the drain. Thus, the change of the QFP is step-like. In most of state-of-the-art NEGF QT simulators, short-range electron-electron interaction is not included in the Hamiltonian, and the position of the step remains undetermined.

In the ballistic regime, a weak relaxation of the total momentum disturbs the kinetic motion (similar to the trajectory of a cannonball which is not parabolic but ballistic due to the air drag). In the community the name “ballistic” is commonly used instead of “kinetic/emission”. DD TCAD simulators are based on the concept of local thermodynamic equilibrium. The value of the diffusive mobility μ_d must not be much larger than the value of μ_{lim} . The latter corresponds to the case where the momentum relaxation time is approximately equal to the transit time. As shown by Frensley in his pioneering

paper [44], the shape of the QFP along the channel is essentially given by an error function if the barrier is treated with the saddle-point method. This analytical solution for $\psi_n(x)$ does not depend at all on the value of the diffusive mobility μ_d , only $\mu_d = \infty$ is excluded. However, the existence of a continuous function $\psi_n(x)$ warrants the limitation of the diffusive mobility to μ_{lim} . When $L_G \rightarrow 0$ (kinetic case), the error function is squeezed to a step function.

A step-like change of the QFP can also occur in DD transport as the result of sharp density gradients, as visible in Fig. 4.1b). The figure shows a TCAD simulation of the QFP along the channel of the test transistor of Fig. 2.1. A gate length of 40 nm was used which is long enough to safely neglect STDt. The softened step function is symmetrical in the linear regime (50 mV) with an almost linear potential drop typical for an Ohmic resistor, but strongly asymmetrical in the saturation regime (0.63 V) with a sharp edge at the drain-side pn-junction ($x = x_j = 80$ nm). To the left of x_j , the QFP belongs to the moving electrons coming from the source which have a very small density (that nevertheless determines the total value of n in this region), but to the right of x_j the QFP belongs to the thermalized high-density electrons of the drain.

A heavy S/D doping of $N_D = 5 \times 10^{19} \text{ cm}^{-3}$ was chosen to demonstrate the degeneracy effect. It shows up as a slight difference between $\psi_n(x)$ and $\Phi(x) - V_T \ln(n(x)/n_i)$ in the pn-junctions (see the insert of Fig. 4.1(a)). Whereas degeneracy is negligible for the shape of the QFP, it impacts the on-current of a FET due to a Fermi correction term $-\mu_n n k_B T \nabla \ln(\gamma_n)$ in the equation for the current density, where $\gamma_n = n/n_B$ is the ratio between the actual density and its Boltzmann form. This term tends to reduce the on-current (negative sign), thus it has the same effect as the ballistic mobility. For a better decoupling of both effects, the S/D doping is kept at $N_D = 6 \times 10^{18} \text{ cm}^{-3}$ in Sections 4.3 and 4.4.

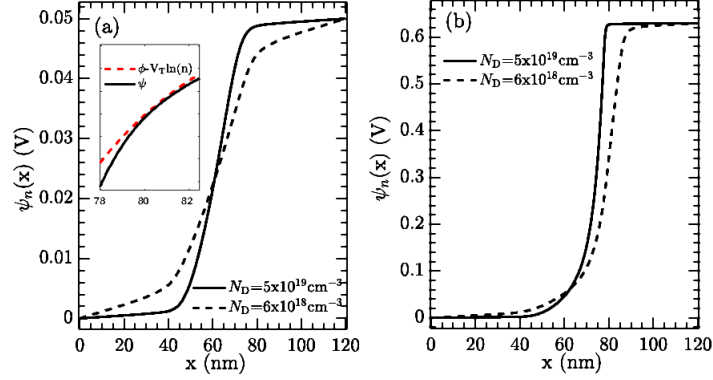


Figure 4.1: Quasi-Fermi potential $\psi_n(x)$ along the channel of the transistor shown in Fig. 2.1 at (a) $V_{DS} = 50$ mV, (b) $V_{DS} = 0.63$ V obtained from TCAD simulations with constant mobility. Parameters: $L_G = 40$ nm, $V_{GS} = 0.31$ V, $\mu_d = 10^4$ cm²/Vs. The small difference between $\psi_n(x)$ and $\Phi(x) - V_T \ln(n(x))$ in the case $N_D = 5 \times 10^{19}$ cm⁻³ is shown in the insert of (a).

4.3 Ballistic Velocity as Function of QFP

Two expressions for the ballistic mobility were suggested in Section 4.2: Eq. (4.3) and Eq. (4.5). Equating them leads to the balance equation

$$q\psi'_n(x) = \frac{m_e}{2} (v_b^2)' , \quad (4.6)$$

stating that the loss of electro-chemical energy of the ballistic electrons is compensated by the gain of their kinetic energy. Integrating this equation from a starting point x_S near the source-channel junction to a point x in the channel results in a solution for the mean ballistic velocity:

$$v_b(x) = \sqrt{v_b^2(x_S) + \frac{2q}{m_e} [\psi_n(x) - \psi_n(x_S)]} \quad (x > x_S) . \quad (4.7)$$

This solution contains two non-localities: the QFP and the velocity at the starting point x_S . In order to obtain a *local* model, one has to

(i) define the starting point as the boundary of ballistic motion, (ii) to find a reasonable expression for $v_b(x_S)$, and (iii) to replace $\psi_n(x_S)$ by a reasonable value. For the choice of x_S one can proceed as in the compact modeling of the ballistic transistor and set $x_S = x_{\text{TOB}}$, where x_{TOB} is the so-called virtual source, i.e. the top-of-the-barrier (TOB) point. The only way to remove the non-locality $v_b(x_{\text{TOB}})$ is to consider the kinetic limit ($L_G \rightarrow 0$), where only thermionic electrons with energies higher than the TOB contribute and are injected with the mean thermal velocity $v_{\text{th},i}$. Then,

$$v_k(x_{\text{TOB}}) = v_{\text{th},i} \frac{1 - e^{-V_{\text{DS}}/V_T}}{1 + e^{-V_{\text{DS}}/V_T}} = v_{\text{th},i} \tanh\left(\frac{V_{\text{DS}}}{2V_T}\right) \equiv \langle v_{\text{inj}} \rangle. \quad (4.8)$$

Note that $v_k(x_{\text{TOB}})$ vanishes at thermodynamic equilibrium ($V_{\text{DS}} = 0$ V, $\psi_n(x) = \text{const}$). The mean thermal velocity $v_{\text{th},i}$ at the injection point could be set to the r.m.s. value in one dimension v_{th} or to the mean of the value in one dimension $v_{\text{th}}\sqrt{2/\pi} \approx 0.8v_{\text{th}}$. The latter option is used in the following simulations. The actual value needed for 2D/3D device simulation might be slightly different. This will be further discussed in Section 4.7.

Finally, the value of $\psi_n(x_{\text{TOB}})$ is approximated by zero (grounded source contact). This removes the second non-locality, however it neglects the voltage drop up to the virtual source caused by the finite series resistance between contact and x_{TOB} . Inserting the above simplifications ($v_b(x_S) \rightarrow \langle v_{\text{inj}} \rangle$, $\psi_n(x_S = x_{\text{TOB}}) \rightarrow 0$) into Eq. (4.7) one obtains

$$v_b(x) = \sqrt{\langle v_{\text{inj}} \rangle^2 + 2v_{\text{th}}^2 \psi_n(x)/V_T}, \quad (4.9)$$

which is a local model up to the V_{DS} -dependence of the mean injection velocity $\langle v_{\text{inj}} \rangle$. The same form of the ballistic velocity was suggested in Ref. [45], with a fitting parameter instead of the first term under the square root, and called Kinetic Velocity Model (KVM). The non-local factor $\tanh(V_{\text{DS}}/2V_T)$ becomes 0.75 at 300 K and $V_{\text{DS}} = 50$ mV, the standard voltage for the linear regime. It quickly approaches unity, when the source-drain bias exceeds a few V_T .

Fig. 4.2 shows the velocity profiles from Eq. (4.7) and Eq. (4.9). In the linear regime ($V_{\text{DS}} = 50$ mV) Eq. (4.7) measured at x_{TOB} is 34% lower than that from Eq. (4.9) and in the saturation regime (V_{DS}

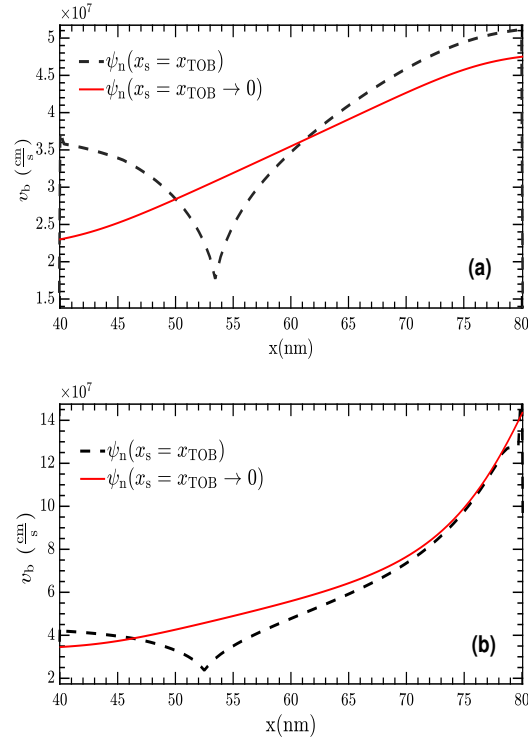


Figure 4.2: Profiles of ballistic velocities along the channel from Eq.(4.7) (dashed curve) and Eq.(4.9) (red curve) extracted from self-consistent TCAD simulations. (a) $V_{DS} = 50 \text{ mV}$, $x_{TOB} = 56 \text{ nm}$ and (b) $V_{DS} = V_{D,sat} = 0.63 \text{ V}$, $x_{TOB} = 54 \text{ nm}$. Parameters: $N_D = 6 \times 10^{18} \text{ cm}^{-3}$. STDT was turned off in all simulations.

= 0.63V) at x_{TOB} the reduction is 23%. The reason behind these important differences is due to the large source resistance ($N_{\text{D}} = 6 \times 10^{18} \text{ cm}^{-3}$). The mean ballistic velocity Eq. (4.9) was implemented in the PMI of S-Device (for details see Section 4.5).

4.3.1 Velocity Comparisons between QTx and DD

The extraction of velocity profiles from QTx is not trivial. To compare with the ballistic velocity of thermionic electrons (4.9), the QTx kinetic velocity is computed by

$$v_{\text{k}} = \frac{I_{\text{T}}}{q n_{1\text{D-T}}} . \quad (4.10)$$

To obtain v_{k} , the lower integration (energy starting at TOB) must be ϵ_{TOB} , otherwise the electrons are not thermionic. I_{T} is the thermionic current which is obtained by integration over at ϵ_{TOB} as given in (4.11) where $F_{\text{L}}(\epsilon)$ is the Fermi distribution.

$$I_{\text{T}} = -\frac{e}{\hbar} \int_{\epsilon_{\text{TOB}}}^{\epsilon_{\text{END}}} \frac{d\epsilon}{2\pi} T(\epsilon) F_{\text{L}}(\epsilon) . \quad (4.11)$$

$n_{1\text{D-T}}$, the 1D density of thermionic electrons flowing from source to drain, is obtained by integration of the thermally weighted LDOS (Z_{L}) over energies above the TOB as follows:

$$n_{1\text{D-T}} = \frac{e}{\hbar} \int_{\epsilon_{\text{TOB}}}^{\epsilon_{\text{END}}} Z_{\text{L}}(\epsilon, r) F_{\text{L}}(\epsilon) d\epsilon . \quad (4.12)$$

For ensuring an accurate comparison between the velocity profiles obtained from QTx and QDD it is very important to check the E_{F} in the contacts of both simulators. If the source/drain contacts are degenerate, the two speed will strongly diverge in the drain region.

The extracted QFP is inserted in Eq. (4.9) and the resulting mean ballistic velocity is compared with the mean kinetic velocity from the effective-mass QT solver QTx [46] in Fig. 4.5. For high doped S/D contacts (ie. $N_{\text{D}} = 5 \times 10^{19} \text{ cm}^{-3}$), a pronounced discrepancy of the speed in the drain region come from the high E_{F} in the source that of QTx compared to S-Device.

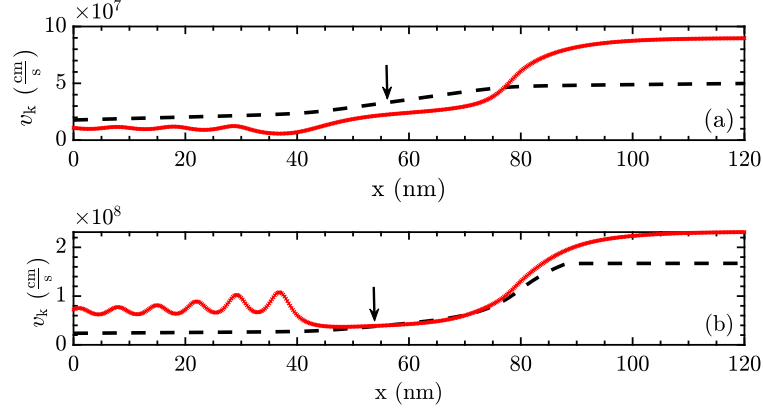


Figure 4.3: Comparison of the mean ballistic velocity Eq. (4.9) (black dashed) with the mean kinetic velocity extracted from QTx (red symbols). (a) $V_{\text{DS}} = 50$ mV, (b) $V_{\text{DS}} = 0.63$ V. Parameters: $N_{\text{D}} = 6 \times 10^{18} \text{ cm}^{-3}$, $L_{\text{G}} = 40$ nm, $V_{\text{GS}} = 0.31$ V. The arrow points to x_{TOB} , the position of the TOB.

The thermionic (kinetic) electrons are assumed to possess the QFP of the source up to the drain contact. The velocity profiles agree well in the channel region (from 40 nm to 80 nm), in particular for the high source-drain bias. The oscillatory behaviour in the source found with QTx is caused by quantum reflection. In the drain, the velocity extracted from QTx becomes much larger than the one computed by Eq. (4.9) which is attributed to the 2D DOS model of QTx. Note, that the total increase of the mean ballistic velocity is approximately given by $\sqrt{2 V_{\text{DS}}/V_{\text{T}}}$. At x_{TOB} (the point of the TOB indicated by an arrow in Fig. 4.5), the mean ballistic velocity has only little increased from its initial value $\langle v_{\text{th}} \rangle = 2.37 \times 10^7$ cm/s . In the linear regime ($V_{\text{DS}} = 50$ mV) the increase is $\approx 85\%$ and in the saturation regime ($V_{\text{DS}} = 0.63$ V) $\approx 60\%$. The idea behind choosing a channel length of 40 nm is to avoid any influence of source-to-drain tunneling. In the case of shorter channels ($L_{\text{G}} = 15$ nm) the behaviour is similar (see Fig. 4.4). However, the shorter the channel the more voltage drops over the S/D

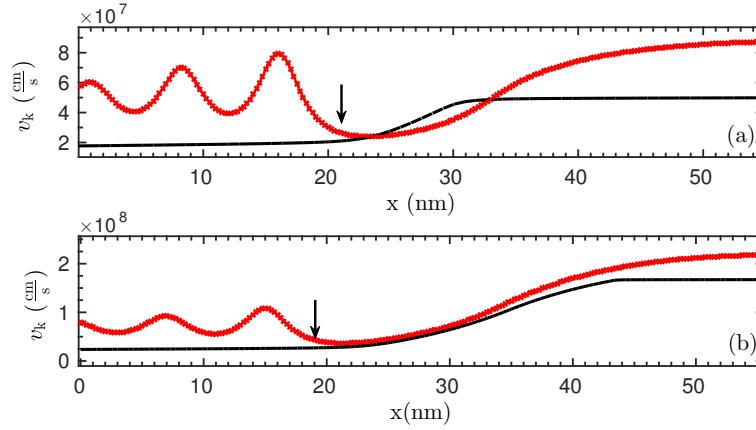


Figure 4.4: Comparison of the mean ballistic velocity Eq. (4.9) (black dashed) with the mean kinetic velocity extracted from QTx (red symbols). (a) $V_{\text{DS}} = 50 \text{ mV}$, (b) $V_{\text{DS}} = 0.63 \text{ V}$. Parameters: $N_{\text{D}} = 6 \times 10^{18} \text{ cm}^{-3}$, $L_{\text{G}} = 15 \text{ nm}$, $V_{\text{GS}} = 0.30 \text{ V}$.

regions and the more the mean velocity at the virtual source deviates from the mean thermal velocity. In the linear regime ($V_{\text{DS}} = 50 \text{ mV}$) this leads to an increasing discrepancy as suggested by the Fig. 4.4. The shorter the channel, the more difficult is the comparison with quantum transport. An increasing part of the discrepancy comes from quantum reflection and the transmitting boundary conditions (short S/D regions).

The overestimation of v_b at x_{TOB} when calculated with Eq. 4.9, as clearly visible in Fig. 4.1, is due to the non-zero value $\psi_n(x_{\text{TOB}})$, i.e. the price for the negligence of the last term in Eq. (4.7). It not only depends on source doping and source extension, but also on the gate bias V_{GS} . The latter dependence is uncritical because $\psi_n(x_{\text{TOB}})$ changes by less than $V_{\text{DS}}/10$ over the whole gate voltage range. The three assumptions that lead to a local model will be further discussed in Section 4.7. In the next section, another model of the ballistic velocity will be proposed which requires the numerical determination

of x_{TOB} . Introducing this non-locality in the above model would allow to keep the term $\psi_n(x_{\text{TOB}})$ with the result that $v_b(x_{\text{TOB}}) = \langle v_{\text{inj}} \rangle$.

Frenslley [44] suggested to make use of $\nabla \cdot \vec{j}_n = 0$ in the Euler equation and to replace $(v^2)'$ by $-2v^2(\ln n)'$. Based on this he derived the kinetic (actually: ballistic) current density of a model problem with the essential feature that the current depends on the QFP at x_{TOB} , in contrast to the emission theory, where no QFP exists in the barrier region. Applying the same replacement to the more general equation (4.2) it can be cast into the form [45]

$$\frac{v}{v_d} + \frac{v^2}{v_b^2} = 1 \quad (4.13)$$

with

$$v_b^2 = -\frac{q\psi'_n}{m_e(\ln n)'} = v_{\text{th}}^2 \left(1 + \frac{\Phi'}{\psi'_n - \Phi'} \right) = v_{\text{th}}^2 \frac{\psi'_n}{\psi'_n - \Phi'}. \quad (4.14)$$

Where $v_b = \mu_d(x)\psi'_n$ and $\mu_d(x)$ is the diffusive mobility (without the kinetic term). In this model for v_b could be used for $x \geq x_{\text{TOB}}$ because $\psi'_n(x_{\text{TOB}}) \neq 0$. However, it depends on the electric field, and gives 0/0 in the heavily doped drain region.

Fig. 4.5 compares the driving forces ψ'_n , E_x , and $V_T(\ln n)'$. As can be seen, in the source-side part of the channel ψ'_n is dominated by the density gradient whereas in the drain-side part it is dominated by the electric field. Therefore, ψ'_n cannot be replaced by one of the two. This becomes clear if one plots the local inverse mobility $\mu_b^{-1}(x)$ and the local resistivity $\rho(x) = (q\mu_b(x)n(x))^{-1}$ using the ballistic velocity (4.9). Fig. 4.6 shows that the whole channel contributes to both the inverse ballistic mobility and the ballistic resistivity. Furthermore, a sharp drop is observed towards the gate edges so that TCAD artifacts from model simplifications can be tolerated here. The QDD current after the onset of inversion is still overestimated compared to QTx. The relative deviation is quite strong in the linear regime, but much smaller in the saturation regime. This behaviour can be attributed to the shape of the QFP $\psi_n(x)$ (see Fig. 4.1). At high drain bias, it is much closer to a step function, i.e. to the kinetic case where the electrons keep their QFP from the contact. Note, that the latter had been one of the initial assumptions to construct the local model

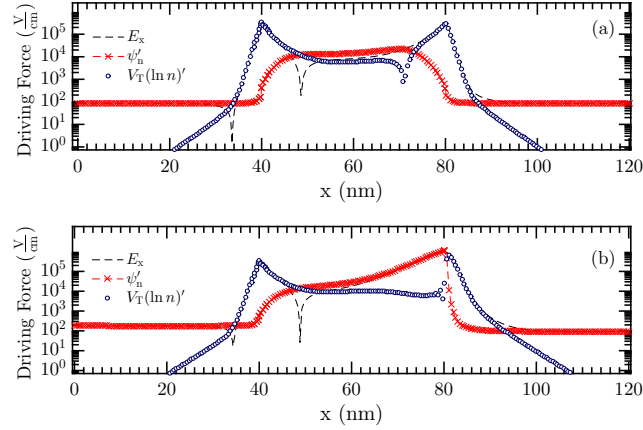


Figure 4.5: Comparison of driving forces along the channel obtained from TCAD simulations with a constant mobility of $\mu_d = 10^4 \text{ cm}^2/\text{Vs}$. (a) $V_{DS} = 50 \text{ mV}$, (b) $V_{DS} = 0.63 \text{ V}$. Parameters: $L_G = 40 \text{ nm}$, $V_{GS} = 0.31 \text{ V}$.

(4.9). In contrast, the channel behaves like a low-Ohmic resistor with an almost linear $\psi_n(x)$ in the linear regime, a situation furthest from the kinetic case. Second, due to the large series resistance, the injection velocity at the virtual source is overestimated which contributes to the overestimation of the on-current at low V_{GS} . This calls for improvements in modeling the ballistic velocity. Penzin et al. [45] added an empirical function of ψ_n in their KVM model to achieve the necessary drop of the on-current in the linear regime.

4.4 Ballistic Mobility as Function of Density

An obvious feature of the ballistic velocity (4.9) is, that the ballistic current density j_b is not conserved when used in the mobility μ_b . Evaluating the continuity equation with a density that depends

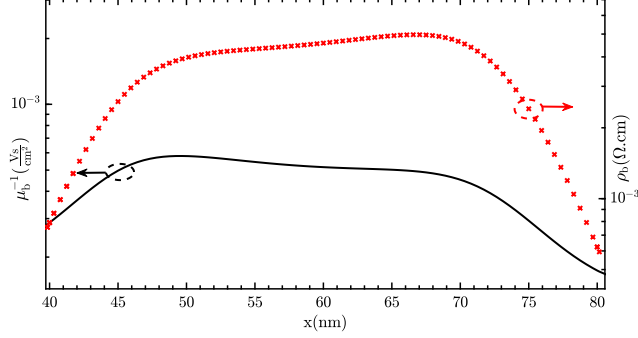


Figure 4.6: Inverse ballistic mobility (using Eq. (4.5) with the velocity model Eq. (4.9) in the TCAD simulations) and the corresponding resistivity along the channel. Parameters: $V_{DS} = 50 \text{ mV}$, $N_D = 6 \times 10^{18} \text{ cm}^{-3}$, $L_G = 40 \text{ nm}$, $V_{GS} = 0.31 \text{ V}$.

on a *local* QFP $\psi_n(x)$ results in a position-dependent $j_b(x)$. For $j_b(x_{\text{TOB}}) = j_b(x_D)$ to hold, the densities of thermionic electrons would have to fulfill $n_{\text{TOB}}(x_{\text{TOB}})/n_{\text{TOB}}(x_D) = v_b(x_D)/v_b(x_{\text{TOB}}) \approx \sqrt{2 V_{DS}/V_T}$ which is not the case. Here, x_D denotes a point near the channel-drain junction. Densities, whether TOB or total, depend exponentially on V_{DS} . A ballistic velocity that conserves the current is:

$$v_b(x) = \langle v_{\text{inj}} \rangle \frac{n_{\text{TOB}}(x_{\text{TOB}})}{n_{\text{TOB}}(x)} = \frac{j_k}{q n_{\text{TOB}}(x)}, \quad (4.15)$$

where the mean injection velocity is divided by the normalized density of thermionic electrons in channel direction. Because of current conservation,

$$\begin{aligned} j_b(x_{\text{TOB}}) &= q v_b(x_{\text{TOB}}) n_{\text{TOB}}(x_{\text{TOB}}) = \\ &= q \langle v_{\text{inj}} \rangle n_{\text{TOB}}(x_{\text{TOB}}) = q \langle v_{\text{inj}} \rangle n(x_{\text{TOB}}). \end{aligned}$$

The last identity is due to the fact that at x_{TOB} the TOB density is exactly equal to the total density. The replacement $v_b(x_{\text{TOB}}) \rightarrow \langle v_{\text{inj}} \rangle$ is the same local approximation as in Section 4.3. Inside the channel

region, using parabolic bands, a relation between $n_{\text{TOB}}(x)$ and the total density $n(x)$ can be found:

$$n_{\text{TOB}}(x) = n(x) \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}, \frac{\Phi(x) - \Phi(x_{\text{TOB}})}{V_T}\right). \quad (4.16)$$

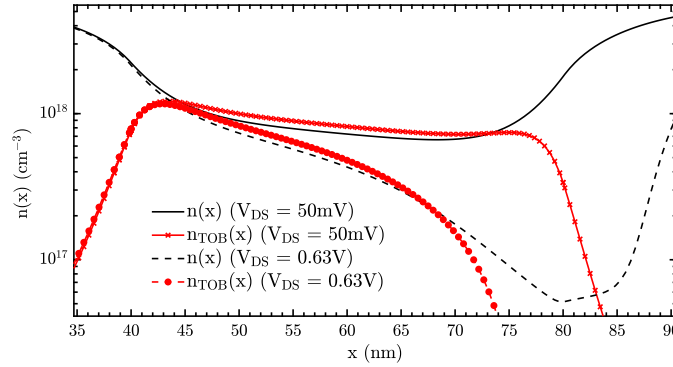


Figure 4.7: Comparison between $n_{\text{TOB}}(x)$ and $n(x)$ along the channel in the middle of the body. Parameters: $N_D = 6 \times 10^{18} \text{ cm}^{-3}$, $L_G = 40 \text{ nm}$, $V_{GS} = 0.31 \text{ V}$. The $n(x)$ -profiles are direct extractions from the self-consistent TCAD simulation, whereas the $n_{\text{TOB}}(x)$ -profiles are a post-processing evaluation of Eq. (4.16) with all densities imported from TCAD.

Here, $\Gamma(\nu, b)$ denotes the incomplete Gamma function. The derivation of Eq. (4.16) is given in Appendix A. Fig. 4.7 compares $n_{\text{TOB}}(x)$ and $n(x)$ in the channel region.

As one can see, both practically coincide in a large part of the channel. Deviations only occur where the local resistivity quickly drops to zero (compare Fig. 4.6). Hence, the model (4.15) can be simplified in the following way: the density $n_{\text{TOB}}(x)$ in the denominator can be replaced by the actual density $n(x)$ which avoids to evaluate Eq. (4.16) and results in

$$v_b(x) = \langle v_{\text{inj}} \rangle \frac{n(x_{\text{TOB}})}{n(x)}. \quad (4.17)$$

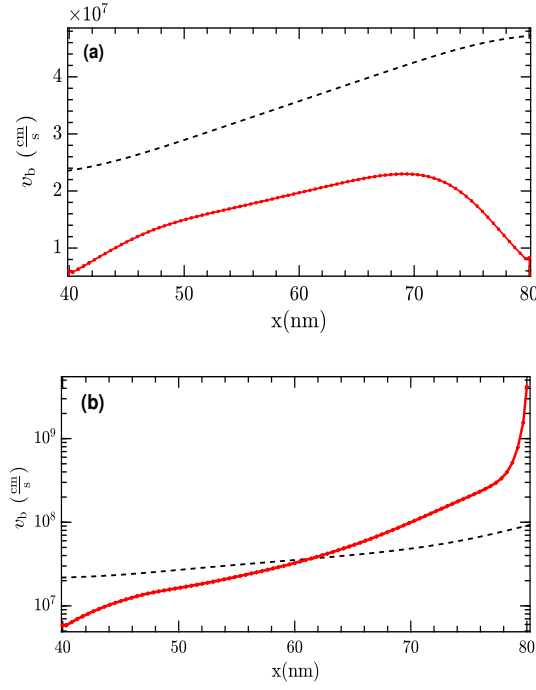


Figure 4.8: Profiles of the ballistic velocities along the channel with model (4.9) (black dashed) and with model (4.17) (red symbols) extracted from self-consistent TCAD simulations. Parameters: (a) $V_{DS} = 50$ mV, $N_D = 6 \times 10^{18} \text{ cm}^{-3}$, $L_G = 40$ nm, $V_{GS} = 0.31$ V, $v_{th,i} = 2.37 \times 10^7$ cm/s and (b) $V_{DS} = 0.63$ mV with the same parameters as in (a).

Fig. (4.8) (a) and (b) compares the ballistic velocities (4.9) and (4.17). Where needed, i.e. in the channel region, the n -dependent model Eq. (4.17) yields a slower increase of the mean velocity towards the drain in the linear regime, which results in a lower ballistic mobility. Consequently, the on-current in the linear regime is reduced. In the saturation regime shown in of Fig. (4.8) (b) the reduction provided by the n -dependent model (Eq. (4.17)) in comparison to the $\psi_n(x)$ -dependent model (Eq. (4.9)) is quite clear at x_{TOB} , thus the current in the saturation regime is reduced, too. This is also demonstrated by the comparison transfer characteristics of the same $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FETs with gate lengths of 40 nm in Fig. 4.10 (see subsection 4.6.1), where STDT is absent. The agreement with the QTx-curves is reasonable considering the fact that the DOS models are different. In contrast to the ψ_n -dependent model in the previous section, the n -dependent model (4.17) makes explicit use of the injection point with the obvious advantage that $v_b(x_{\text{TOB}}) = \langle v_{\text{inj}} \rangle$. The non-locality $n(x_{\text{TOB}})$ cannot be removed, since the point x_{TOB} has to be under full gate control. Though replacing $n(x_{\text{TOB}})$ by the source doping N_D would make the model local (analogous to setting $\psi_n(x_{\text{TOB}}) = 0$ in the previous section), this would lead to bias-dependent velocities with extreme values. The replacement of the TOB density by the density overestimates the TOB density in the region between the source-channel junction and x_{TOB} . This artificially drops the velocity and underestimates the ballistic mobility. Thus, a possible refinement of the model could be to replace v_b by $\langle v_{\text{inj}} \rangle$ for all $x < x_{\text{TOB}}$, leading to $v_b(x) = \langle v_{\text{inj}} \rangle [\Theta(x_{\text{TOB}} - x) + \Theta(x - x_{\text{TOB}})n(x_{\text{TOB}})/n(x)]$.

4.5 Implementation

All above considerations apply to the ballistic transport regime. Of practical interest is the quasi-ballistic regime, defined by $\mu_d \approx \mu_b$ (see Table 4.1). Eq. (4.2) and Eq. (4.3) yield $v = v_{\text{qb}} = \mu_n \psi'_n$, and μ_n is given by the Matthiessen rule in its common form $\mu_n^{-1} = \mu_d^{-1} + \mu_b^{-1}$. Weighting factors for diffusive and ballistic sub-populations [40, 47] are not considered here. For μ_b it would mean an additional fitting factor, and the PMI of S-Device internally combines mobility models with the common Matthiessen rule. The models (4.5) with (4.9) and

with (4.17) were implemented in the PMI “HighFieldMobility” which combines the ballistic mobility model with the user-defined diffusive mobility μ_d by the Matthiessen rule. Generalization of Eq. (4.4) to the 2D/3D case is done by $\mu_b = v_b/(|\nabla\psi_n| + \epsilon)$, where ϵ is an appropriate cut-off. The computation of an average 1D density $n(x)$ is prohibitive, hence $n(\vec{r})$ is used. Density minima perpendicular to the channel yield high local values of μ_b and do not contribute. The obviously better approach would be to compute n_{TOB} as average over a line/slice perpendicular to the transport direction. However, this would require the availability of a tensor-product grid in the PMI of S-Device which is not the case. To extract $n(\vec{r}_{\text{TOB}})$, a search algorithm was implemented to find the point where the conduction band (CB) energy is maximum. In each step of the Newton iteration, a variable $E_{c,\text{ref}}$ is first set to a large negative value. Then, the local value of the CB energy is compared with $E_{c,\text{ref}}$ on every vertex \vec{r}_n of the search path. Only if $E_c(\vec{r}_n) > E_{c,\text{ref}}$, $E_{c,\text{ref}}$ is updated and set to the value of $E_c(\vec{r}_n)$. The vertex \vec{r}_{TOB} is found for $E_c(\vec{r}_{\text{TOB}}) = E_{c,\text{ref}}$. The electron density is extracted at this vertex and used for the mobility model in the next Newton step. A sufficient mesh refinement, in particular at the source-channel junction, is required to ensure convergence.

4.6 Results

In this section two sets of simulations were done. The first one, used Fermi statistics, and the second one Boltzmann statistics. The test structure is depicted in Fig. 2.1 and the parameters used are shown Table 4.2.

Table 4.2: Summary of geometry parameters, effective masses, and $V_{\text{DS,sat}}$ of the simulated devices.

L_G	t_{body}	t_{ox}	m_e/m_0	$V_{\text{DS,sat}}$
7 nm	2.8 nm	2.6 nm	0.080	0.56 V
11.5 nm	4.6 nm	3.2 nm	0.0678	0.61 V
15 nm	7 nm	3.7 nm	0.0516	0.63 V

4.6.1 Application of Fermi Statistics

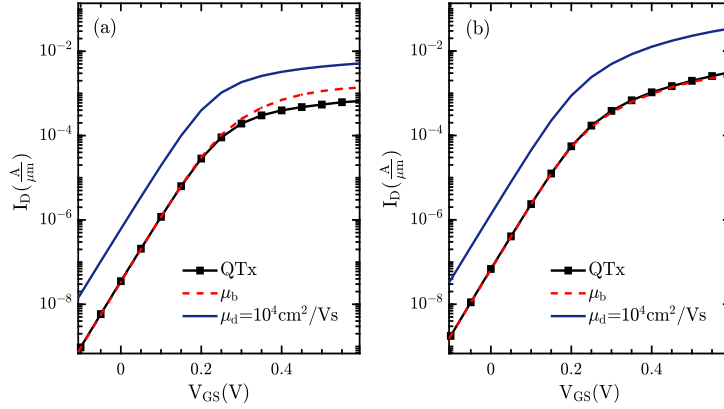


Figure 4.9: $I_D V_{GS}$ -characteristics of an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET ($L_G=15\text{ nm}$) computed with μ_b from Eq. (4.5) with the velocity model Eq. (4.9) (red-dashed curves) and with constant mobility $\mu_d = 10^4\text{ cm}^2/\text{Vs}$ (solid blue curves) for (a) $V_{DS} = 50\text{ mV}$ and (b) $V_{DS} = V_{D,\text{sat}} = 0.63\text{ V}$. Parameters: $N_D = 6 \times 10^{18}\text{ cm}^{-3}$, $v_{\text{th},i} = 2.37 \times 10^7\text{ cm/s}$. The self-consistent TCAD simulations are compared with the QTx reference characteristics (black lines with squares). STDT was turned-off in all simulations.

Fig. 4.9 presents the simulated $I_D V_{GS}$ -curves for $L_G = 15\text{ nm}$, computed with $\mu_b(x)$ (Eq. (4.5) with Eq. (4.9)) and $\mu_d = 10^4\text{ cm}^2/\text{Vs}$, respectively, for $V_{DS} = 50\text{ mV}$ (a) and for $V_{DS} = V_{D,\text{sat}} = 0.63\text{ V}$ (b). The QTx reference characteristics are plotted for comparison. The shift of the threshold voltage V_{th} due to geometrical confinement [48] found with QTx was modelled by a work function fit in S-Device. Geometry parameters [21], effective masses, and applied $V_{D,\text{sat}}$ are summarized in Table 4.2.

Devices with shorter gate were at first simulated without STDT. For this, the tunneling part of the spectral current in QTx was filtered out, and no model for STDT was used in the simulations with S-Device. In this section Fig. 4.11 presents transfer characteristics for

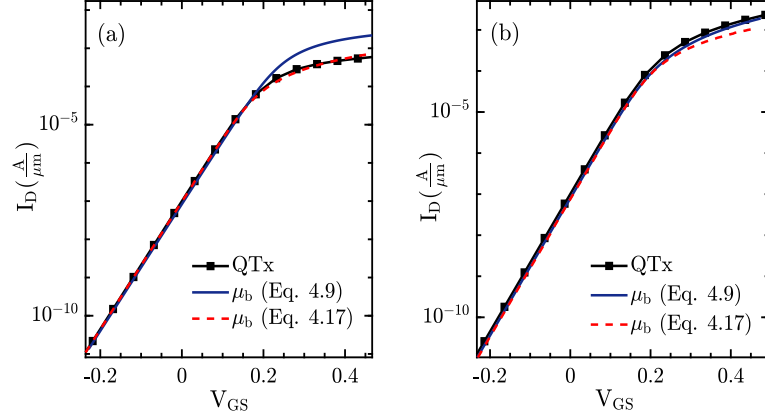


Figure 4.10: $I_D V_{GS}$ -characteristics of an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET with $L_G = 40$ nm at (a) $V_{DS} = 50$ mV and (b) $V_{DS} = 0.63$ V computed with μ_b (Eq. (4.5)). Blue solid curve: v_b -model as function of ψ_n (Eq. (4.9)), red dashed curve: v_b -model as function of n (Eq. (4.17)). Parameters: $N_D = 6 \times 10^{18} \text{ cm}^{-3}$, $v_{th,i} = 2.37 \times 10^7 \text{ cm/s}$. STDT was suppressed by the sufficiently long gate. Other parameters are given in Table 4.2.

DG FETs with $L_G = 11.5$ nm and $L_G = 7$ nm with the mean ballistic velocity Eq. (4.17). Note, that there is no explicit fitting except the electrostatics in the sub-threshold range at $V_{DS} = 0$ V (no influence of the mobility model). The agreement with QTx is very good up to an overdrive voltage of 0.2 V. In deep inversion, the deviation increases with decreasing gate length and becomes more significant in the saturation regime. This behaviour is attributed to the different DOS models which will be discussed in more detail in Section 4.7. The shorter the gate, the thinner the body (see Table 4.2) and the stronger the impact of the true 2D DOS of QTx. The dashed curves in Fig. 4.12 represent the corresponding profiles $\psi_n(x)$ and $\mu_b^{-1}(x)$ along a cut line in the middle of the body. The solid curves in this figure are obtained with the ψ_n -dependent model of the ballistic velocity.

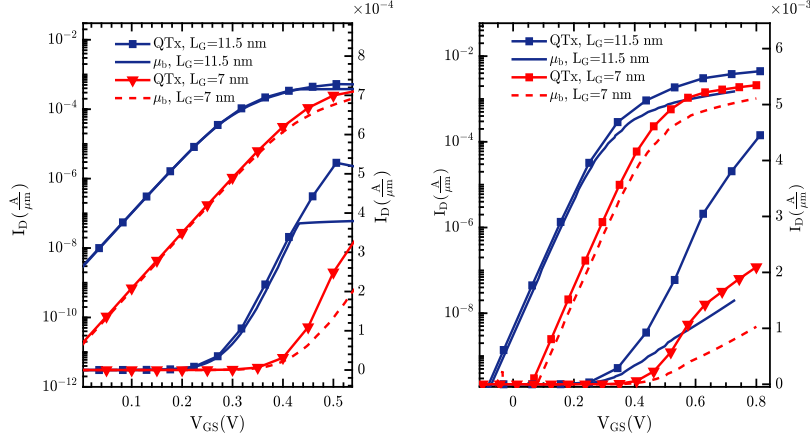


Figure 4.11: $I_D V_{GS}$ -characteristics of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FETs with $L_G = 11.5$ nm and $L_G = 7$ nm at $V_{DS} = 50$ mV (left) and $V_{DS} = V_{D,\text{sat}}$ (right) computed with μ_b (Eq. (4.5)) and $v_b(n)$ (Eq. (4.17)). Parameters: $N_D = 6 \times 10^{18} \text{ cm}^{-3}$, $v_{\text{th},i} = 2.37 \times 10^7 \text{ cm/s}$. STDT was turned off. Other parameters are given in Table 4.2.

In addition to the transfer characteristics presented in Fig. 2.1, μ_b was tested for the 3D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ GAA NWs from Chapter 3. In this case the 1DEG DOS model was applied. Fig. 4.13 presents the transfer curves in the linear regime for a GAA NW from node A with a sufficient gate length ($L_G = 30$ nm) to avoid STDT. The combination 1DEG DOS + μ_b accounts both for the Fermi and mobility correction and reproduces the QTx curve reasonably well. The remaining difference in the ON-current has its origin in the KVM-form Eq. (4.9) [45] as in the case of the 2D simulations. In this section the use of $v_b(n)$ (Eq. (4.17)) was avoided, since the implementation of the V_{GS} switch in the searching algorithm for a 3D structure was particularly challenging and numerically unstable, due to the increased number of vertices in a 3D structure.

Fig. 4.14 presents the simulated $I_D V_{GS}$ -curves for $L_G = 10$ nm, computed with $\mu_b(x)$ (Eq. (4.5) with Eq. (4.9)) and $\mu_d = 10^4 \text{ cm}^2/\text{Vs}$, respectively, for $V_{DS} = 50$ mV (a) and for $V_{DS} = V_{D,\text{sat}} = 0.63$ V

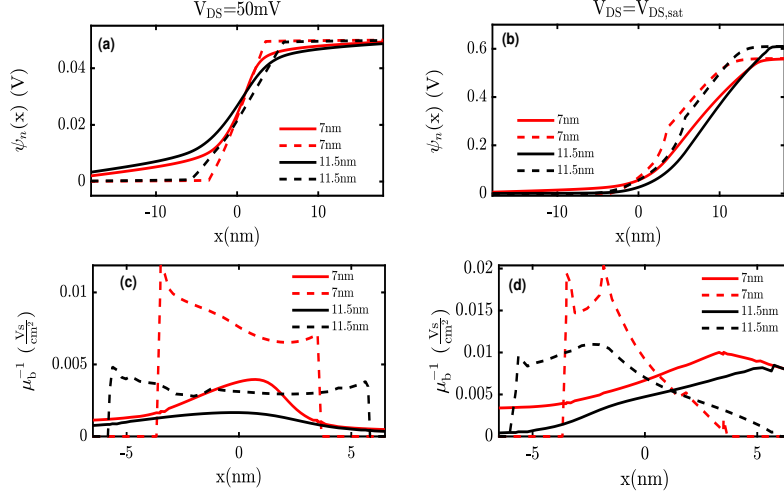


Figure 4.12: Profiles of $\psi_n(x)$ and $\mu_b^{-1}(x)$ along a cut line in the middle of the body of the $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FETs with $L_G = 11.5$ nm (black) and $L_G = 7$ nm (red) at $V_{DS} = 50$ mV (left) and $V_{DS} = V_{DS,sat}$ (right). Solid curves: $v_b(\psi_n)$ (Eq. (4.9)), dashed curves: $v_b(n)$ (Eq. (4.17)). Parameters: $V_{GS} = 0.45$ V ($L_G = 7$ nm), $V_{GS} = 0.35$ V ($L_G = 11.5$ nm, $V_{DS} = 50$ mV), $V_{GS} = 0.27$ V ($L_G = 11.5$ nm, $V_{DS} = V_{DS,sat}$), $N_D = 6 \times 10^{18} \text{ cm}^{-3}$, $v_{th,i} = 2.37 \times 10^7 \text{ cm/s}$. Other parameters are given in Table 4.2.

(b). The application of a ballistic mobility model yields an overall good agreement with the QTx transfer curves and decreases I_{ON} by one order of magnitude in comparison to the simulation with the diffusive mobility μ_d . However, it is still larger than the QTx value, in particular in the linear regime. This can be due to the fact that the true 1D DOS can hardly be mimicked in S-Device or that it is the model μ_b itself, which overestimates the current in the linear regime as seen in for the previous 2D case.

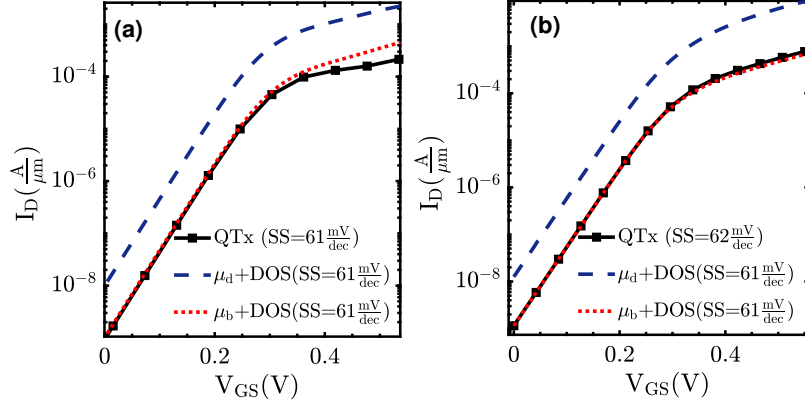


Figure 4.13: $I_D V_{GS}$ -characteristics obtained from the combination of DG + 1DEG DOS, and μ_b + 1D DOS for a GAA NW FET with $t_{\text{body}} = 7$ nm and $L_G = 30$ nm. Parameters: $\langle m_{\text{qu}}^{1D} \rangle = 0.07 m_0$, $WF = 4.77$ eV, and $\alpha_{\text{np}} = 1.2 \text{ eV}^{-1}$

4.6.2 Application of Boltzmann Statistics

In this section a heavy S/D doping of $N_D = 5 \times 10^{19} \text{ cm}^{-3}$ is used to demonstrate the behaviour of $\mu_b(x)$ under degeneracy due to the high-density electrons of the drain. Although the degeneracy is negligible for the shape of the QFP, it will impact the ON-current. Due to the Fermi correction term in the current equation. Fig. 4.15 shows simulations where this term was omitted in order to highlight the impact of the mobility model. If we compare Fig. 4.15 (b) to Fig. 4.11 (a), the ON-current mismatch between QTx and DD is higher, despite the same $v_b(n)$ (Eq. (4.17)) is used.

An important remark from this section is that in Fig. 4.15 STDT is included in the $I_D V_{GS}$ -characteristics to demonstrate the strong effect of STDT in a degenerate transistor with such a short gate length. Fig. 4.15 (a) shows a fair match in the sub-threshold regime when μ_b is included (Eq. (4.5)). On the other hand, Fig. 4.15 (b) reveals a stronger mismatch in the sub-threshold range. This is traced back to

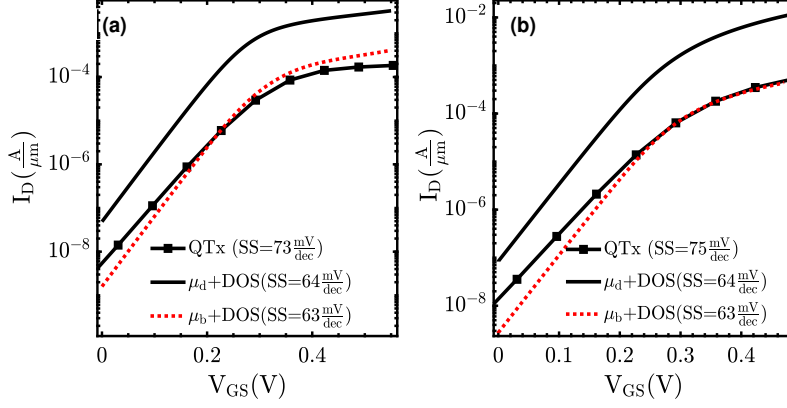


Figure 4.14: $I_D V_{GS}$ -characteristics obtained from the combinations 1DEG DOS, 1DEG DOS + NLT + μ_b for a GAA NW FET with $t_{\text{body}} = 5.5 \text{ nm}$ (node B) at (a) $V_{DS} = 0.05 \text{ V}$ and (b) $V_{DS} = 0.63 \text{ V}$. Parameters: $\alpha_{np} = 1.1 \text{ eV}^{-1}$, $\gamma = 1.0$, $(\alpha_l, \alpha_v) = (0.7, 0)$, and $WF = 4.77 \text{ eV}$.

the deformation of the QFP in the presence of $v_b(n)$ (Eq. (4.17)) and will be explained and discussed in depth in Chapter 5.

4.7 Discussion

Introducing a ballistic mobility can be seen as the re-introduction of the hydrodynamic term in the balance equation for the mean velocity. In FETs with ultra-short channels this term is crucial to prevent the divergence of the DD current when $L_G \rightarrow 0$ and no series resistance is present. The expression of the ballistic mobility contains the mean ballistic velocity. Its determination would necessitate the iterative solution of the whole equation system in the ballistic regime. In order to avoid this, two explicit models of the mean ballistic velocity as function of solution variables were suggested, $v_b(\psi_n)$ and $v_b(n)$. Utilizing the kinetic limit, locality can be approximately achieved in the case of $v_b(\psi_n)$. In the $v_b(n)$ -model, the TOB density $n(x_{\text{TOB}})$

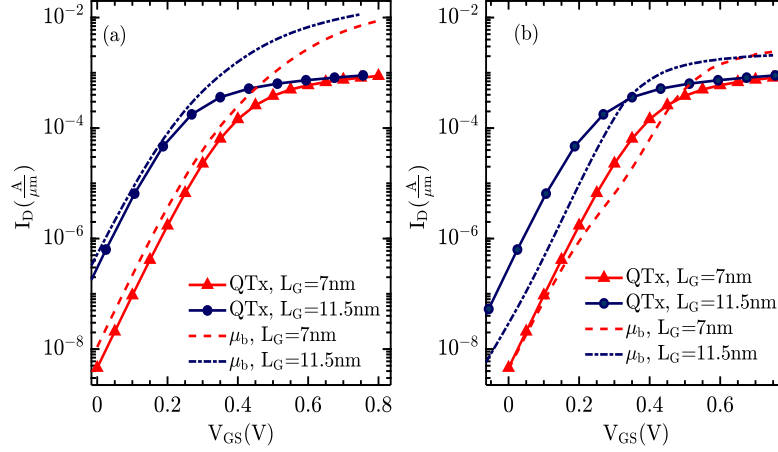


Figure 4.15: $I_D V_{GS}$ -characteristics of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FETs with $L_G = 11.5 \text{ nm}$ and $L_G = 7 \text{ nm}$ at $V_{DS} = 50 \text{ mV}$ (left). Subfigure (a) $I_D V_{GS}$ -characteristics are computed with μ_b (Eq. (4.5)) and (b) are computed with $v_b(n)$ (Eq. (4.17)). Parameters: $N_D = 5 \times 10^{19} \text{ cm}^{-3}$, STDT is included by the NLT model and the Fermi correction term in the equation for the current density is omitted. Other parameters are given in Table 4.2.

has to be extracted which requires to find the virtual source x_{TOB} numerically. This is the only non-local remnant of the hydrodynamic term. The models have no free parameters, but the mean injection velocity is not unique and can serve as TCAD parameter with the mean thermal velocity as default.

The models were implemented in a PMI of S-Device, and transfer characteristics of InGaAs DG FETs with L_G ranging from 7 nm to 40 nm were simulated. In InGaAs both the small transport and tunnel mass highlight the ballisticity effect and the interplay with STDT. Benchmarking against real devices would require a fully characterized (electrical and physical) ultra-short-channel FET which was not available. Instead, the simulations were benchmarked against quantum-transport results obtained with QTx, but also here a precise one-to-one

comparison is impossible. This is mainly due to the different DOS models. The DOS in S-Device used in this chapter is the common 3D DOS of a 3DEG populating one valley. QTx features the correct 2D DOS of a 2DEG, and the sub-band dispersions are self-consistent Poisson-Schrödinger solutions. For any comparison between S-Device and QTx the first step is always to match the electrostatics. This is done at zero current to avoid any influence of the mobility model. Furthermore, a longer gate is used in order to suppress the effect of charge penetration into the S-D potential barrier. However, it is not possible to match the electrostatics over the entire V_{GS} -range. In deep inversion, not only the relative contribution of different sub-bands has changed, but also their dispersion. In other words, the DOS in QTx self-consistently depends on the gate voltage, but is a constant in S-Device.

The developed models of the ballistic velocity are oversimplified in the following sense. First, the starting point of ballistic motion is not well defined. The virtual source has been chosen, but a better option could be a point between the source-channel junction and x_{TOB} . The practical advantage of x_{TOB} is that it can be easily found numerically because of $\nabla\Phi(x = x_{TOB}) = 0$. In 2D/3D devices, under flat-band conditions, the point can degenerate to a contour line/surface. Second, the injection velocity at the virtual source is, in general, not a mean thermal velocity. Even the mean *kinetic* velocity has already increased here from the latter due to the steep density gradient (continuity equation holds in any transport regime). The mean ballistic velocity differs from the mean kinetic velocity as result of the weak scattering related to the finite value of μ_d . The corresponding small voltage drop adds to the possibly much greater voltage drop due to the source resistance. Third, a local model requires to replace the QFP at x_{TOB} by zero. This artificially increases the mean ballistic velocity at the virtual source. Fourth, the locality in the range $x < x_{TOB}$ has the price that the ballistic velocity is overestimated in the ψ_n -dependent model, but underestimated in the n -dependent model.

The models of the ballistic velocity were derived for a homogeneous band structure (BS). However, alloy- or strain-induced gradients can be present in the channel of modern FETs. In the TCAD simulator S-Device those gradients are treated by the driving force they induce

in the current equation. A similar term, the position-dependent Fermi statistics correction, was included in the above simulations of the IV -characteristics. All those terms self-consistently change n , ψ_n , ψ'_n , and x_{TOB} . As the models directly depend on them, they capture the influence of an inhomogeneous BS to some extent. However, BS gradients, or even band edge discontinuities, could change the injection velocity at x_{TOB} significantly. A deeper investigation of BS inhomogeneities is beyond the scope of this thesis.

Under conditions where STDT dominates the current (high S/D doping, ultra-short channels, deep sub-threshold regime) the ballistic mobility models cannot be used. The sharp maximum of the STDT rate leads to a high local electron density exceeding the density of thermionic electrons. This high local density has a destructive impact on the ballistic mobility which depends on ψ'_n in the denominator and either on the inverse density or on $\sqrt{\psi_n}$ in the numerator. In turn, the deformed QFPs shrink or even extinguish the tunnel current. This is an artifact, because the true tunnel current is a spectral current that flows *below* the TOB. The generated density which mimics this current *has nothing to do* with the thermionic density above the TOB in the model of the ballistic velocity. One faces a fundamental TCAD problem here. The off-current becomes corrupted by the mobility model which is needed for the on-current, where, on the other hand, the influence of STDT fades away. Thus, one cannot simulate the entire transfer characteristics with the same model set. This explains that not only the concept of a local QFP breaks down in the STDT regime but also the concept of a mean ballistic velocity. In other words, the situation in the deep sub-threshold region of an ultra-short FET becomes comparable to a MIM (metal-insulator-metal) structure where a $\psi_n(x)$ does not exist in the insulator.

4.8 Summary

The concept of a local ballistic mobility $\mu_b(x)$ was developed from the first moment of the Boltzman equation. Two models were proposed exploiting the kinetic transport limit and using quantum transport results as reference. These models were implemented in a PMI coupled to S-Device and tested on 2D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FETs and 3D

$\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ GAA NWs. The agreement between QTx and the DD $I_D V_{GS}$ -curves proved to be good independent of the gate length of the devices. Devices with gate lengths from 7 to 40 nm were simulated. As their body thickness was scaled with the gate, small disagreements due to the 2DEG started to be evident in thinner devices. The inclusion of $\mu_b(x)$ in the simulation of ultra short devices has a moderate to strong effect on the generation rate in the NLT model that is used to simulate STDt. When STDt dominates in the sub-threshold regime (high S/D doping and ultra-short gates) the deformation of the QFP affects the electron generation from the NLT model.

Chapter 5

Leakage Currents in the Ballistic Regime

Part of this chapter is based on the representations in Solid-States Electronics, A. Schenk and P. Aguirre, TCAD Models of the Ballistic Mobility in the Source-to-Drain Tunneling Regime [37]. The second part of this Chapter has been presented in IWCN, May 2019.

5.1 Introduction

In this section we will focus on showing how the ballistic mobility models from the previous Chapter can affect the leakage currents due to Source-to-Drain Tunneling (STDT) and Band-to-Band Tunneling (BTBT) present in the 2D simulation of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FETs.

5.2 Theory of Source-to-Drain Tunneling

The Nonlocal Tunneling (NLT) model in S-Device [19] is an adaptation of the Schottky barrier model originally proposed in Ref. [49] to a general barrier that carriers encounter in a device. The actual spectral tunnel current through the source-to-drain potential barrier

up to E_{TOB} is mimicked by a generation-recombination current with the net rate

$$\begin{aligned}
 R_n(u, l, \epsilon) - G_n(u, l, \epsilon) = & \frac{4\pi q^2 m_0 k_B T}{h^3} \delta[\epsilon - E_c(u)] \delta[\epsilon - E_c(l)] |F(u)| |F(l)| \Theta[-F(u)] \Theta[F(l)] \\
 & \times \mathcal{T}_n(u, l, \epsilon) \left\{ \ln \left(1 + \exp \left[\frac{E_{F,n}(u) - \epsilon}{k_B T} \right] \right) - \ln \left(1 + \exp \left[\frac{E_{F,n}(l) - \epsilon}{k_B T} \right] \right) \right\}
 \end{aligned} \tag{5.1}$$

where u and l are positions on a tunnel path with energy ϵ , $F(x) = d\Phi(x)/dx$ is the local field strength, and

$$\mathcal{T}_n(u, l, \epsilon) = \exp \left(-\frac{2}{\hbar} \int_l^u dx \sqrt{2m_e [E_c(x) - \epsilon]} \Theta[E_c(x) - \epsilon] \right) \tag{5.2}$$

the WKB transmission probability. The contribution to the STDT current density by electrons that tunnel from the CB edge at points ahead u , to the CB edge at the point u is obtained by a double integral over the net rate:

$$\frac{dj_{\text{tun},n}}{du}(u) = -q \int_{-\infty}^u dl \int_{-\infty}^{\infty} d\epsilon [R_n(u, l, \epsilon) - G_n(u, l, \epsilon)] . \tag{5.3}$$

5.3 Source-to-Drain Tunneling and Ballistic Mobility

5.3.1 Behaviour of the Nonlocal Tunneling Model

In the previous chapter, we examined two TCAD-friendly mobility models for the simulation of ballistic transport in $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FETs and 3D GAA NWs. As the gate length gets shorter, in the case below 20 nm, STDT is not negligible and notably increases the leakage current in the sub-threshold regime [50, 51]. In the case of a FET with $L_G = 7$ nm and $N_D = 5 \times 10^{19} \text{ cm}^{-3}$, at $V_{DS} = 50$ mV and $V_{GS} = 0$ V, over 95% of the spectral current is tunnel current. STDT can be simulated in S-Device with the Nonlocal Tunneling (NLT) model [19] (Eq. 5.1) using the effective transport mass extracted from

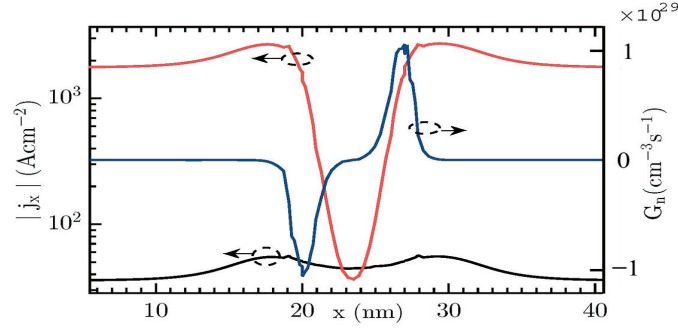


Figure 5.1: Profiles of the STDT rate (blue curve) and the current densities with STDT (red curve) and without STDT (black curve) along a cut in the middle of the channel of the $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET with $L_G = 7$ nm. Parameters: $N_D = 5 \times 10^{19} \text{ cm}^{-3}$, $V_{DS} = 50$ mV. The v_b -model as function of ψ_n (Eq. (4.9)) was used.

QTx as the tunneling mass (m_t). This value changes with body thickness and is given in Table. 4.2.

The choice of the model for the ballistic mobility has a strong effect on the sub-threshold current, which can be explained as follows. When using the NLT model in S-Device, electrons “recombine” at the beginning of a tunnel path, whereas they are “generated” at its end. The generation/recombination rates, shown in Fig. (5.1), are computed with the local QFPs at the classical turning points. These G/R rates are the source/sink of an additional DD current which adds to the thermionic current and determines the shape of the QFP in the self-consistent solution. As a result, the DD current is driven by the gradient of this QFP, and the current level is the integral over the G/R rates with the corresponding QFP values at the classical turning points. The profile of the current density in Fig. (5.1) reveals a strong compensation of the oppositely flowing partial currents of recombining and generated electrons in the channel region which leads to a sharp drop of the total current under the gate. The small thermionic current (shown by the black curve in Fig. (5.1)) is not perfectly constant due to a x -dependent y -component that still exists even in the narrow slab of only 2.8 nm width. As the tunnel-generated density is much

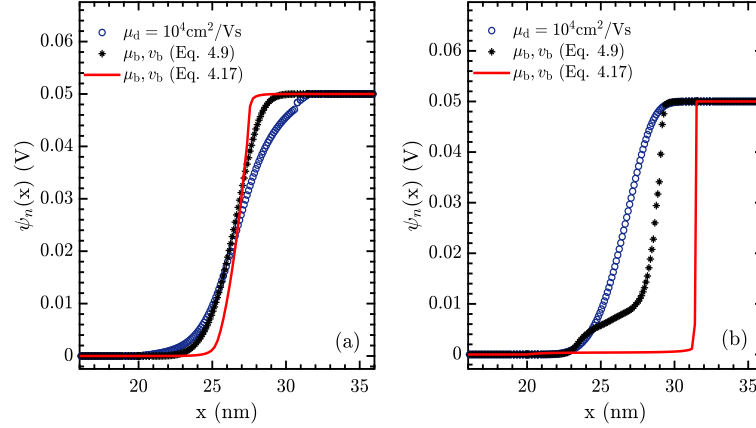


Figure 5.2: Profiles of the quasi-Fermi potential along a cut in the middle of the channel for different mobility models. a) STDT turned off, b) with STDT. Parameters: $N_D = 5 \times 10^{19} \text{ cm}^{-3}$, $V_{DS} = 0.05 \text{ V}$, $V_{GS} = -0.23 \text{ V}$.

smaller than the channel doping in the simulated transistors, it has no electrostatic effect, i.e. the tunnel barrier remains unchanged. Hence, the strong dependence of the off-current on the mobility model originates from the changed values of $\psi_n(x)$ at the classical turning points.

If the ballistic mobility, modelled by Eq. (4.5), becomes dominant, it strongly changes the shape of the QFP $\psi_n(x)$ ($\sim \ln n(x)$) in the sub-threshold region. This is a consequence of current conservation enforced by the continuity equation and the small electron concentration in the channel. Only a constant mobility has no effect (see Fig. 4.1). The high sensitivity of $\psi_n(x)$ to the ballistic mobility is demonstrated in Fig. 5.2 for $L_G = 11.5 \text{ nm}$. If STDT is absent (Fig. 5.2 a)), the edge in the error-function-like profile becomes steeper with the ballistic mobilities - more pronounced when the n -dependent model of the ballistic velocity is used. The position of the edge is not changed and remains at the center of the device. If the strong

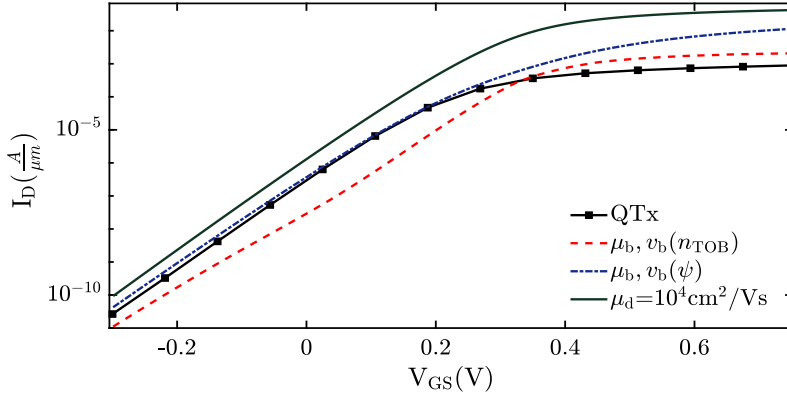


Figure 5.3: $I_D V_{GS}$ -characteristics of an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET with $L_G = 11.5 \text{ nm}$ for different mobility models (Eq. (4.17) and Eq. (4.9)). STDT is included by the NLT model. Parameters: $N_D = 5 \times 10^{19} \text{ cm}^{-3}$, $V_{DS} = 50 \text{ mV}$.

STDT is now turned on (Fig. 5.2 b)), the slope further increases and the edge is shifted towards the drain. The electron quasi-Fermi potentials at the drain side further decrease towards zero (the value at the source) which reduces the STDT rate and the off-current. In the case of the n -dependent model of the ballistic velocity, the QFP shape degenerates into a step function located at the channel-drain junction. The constant QFP under the gate suppresses the STDT rate and, therefore, the sub-threshold current as shown in Fig. 5.3. This behaviour can also be understood analytically inserting the models of the ballistic mobility into the continuity equation and using a simplistic x -dependence of the STDT rate.

If in the implemented NLT model of S-Device just the Fermi levels of the S/D contact regions would be used (emission approach, $\psi_n(x) \rightarrow \psi_n(x_S)\Theta(x_{TOB} - x) + \psi_n(x_D)\Theta(x - x_{TOB})$), the mobility would not change the STDT rate at all. The implementation of such a model variant has been suggested to Synopsis, since no PMI is available for testing it.

5.3.2 Nonlocal Tunneling using Contact Fermi Levels

In the previous section the reasons behind the corrupted STDT rate in the sub-threshold regime with $\mu_b[v_b(n_{\text{TOB}})]$ (see Fig. 5.3) were discussed. Therefore, the purpose of this section is to describe how a NLT model that uses the Fermi levels of the S/D regions becomes independent of the choice of the mobility model.

In order to demonstrate this, Eq. (5.1) was implemented as a post-processing step, after the self-consistent simulations without STDT. The required conduction band energies ($\epsilon(x)$), local field strengths ($F(x)$) and quasi-Fermi energies ($E_F(x)$) were extracted from the same self-consistent simulations that included $\mu_b[v_b(n_{\text{TOB}})]$. For each gate bias the data for 9 spatially equal slabs from a horizontal cut across the structure were extracted from the previous simulations. In the implementation of Eq. (5.1), $E_F(u)$ was replaced by the maximum of $E_F(x)$, and $E_F(l)$ by the minimum of $E_F(x)$, which corresponds to the Fermi levels of the contacts.

Next, Eq. (5.3) was implemented in the code to calculate j_{tun} and its output was normalized to the area of the device. Finally, it was added to the simulated j_n that does not include STDT. As long as STDT dominates over j_n , the result represents the total current correctly.

5.4 Band-to-Band Tunneling

The dynamic non-local path Band-to-Band Tunneling (BTBT) from [19] determines dynamically the tunnel path that starts from the VB edge, continues along the direction of the electric field at the vertex and ends at the CB edge. For a given tunneling path of length L that starts at $x=0$ and ends at $x=L$, holes are generated at $x=0$ and electrons are generated at $x=L$. Kane's formalism [52] defines the generation rate along a tunnel path by the following equation:

$$G_{\text{btb}} = \frac{g}{72\hbar} \frac{1 - \exp\left(-\kappa_m^2 \int_0^L \frac{dx}{\kappa(x)}\right)}{\int_0^L \frac{dx}{\kappa(x)}} \exp\left(-2 \int_0^L \kappa(x) dx\right)$$

$$\times \left[\left(\exp \left[\frac{\epsilon - E_{F,n}(L)}{k_B T(L)} \right] + 1 \right)^{-1} - \left(\exp \left[\frac{\epsilon - E_{F,p}(0)}{k_B T(0)} \right] + 1 \right)^{-1} \right]. \quad (5.4)$$

there, x is the position along the tunneling path, g is the degeneracy of the joint DOS, L is the length of the tunnel path, $\kappa(x)$ denotes the imaginary wave vector, $E_{F,n}(L)$ and $E_{F,p}(0)$ are the Fermi energies of electrons at point L and holes at point 0 , respectively. $\kappa(x)$ is calculated using Kane's two band dispersion model [52]. It is important to clarify that the previous equation is set for direct BTBT, where phonon assisted tunneling is not considered here.

5.5 Simulation Results

In this section the $I_D V_{GS}$ -characteristics of the device of Fig. 2.1. STDT and BTBT were simulated together with $\mu_b[v(\psi_n)]$ and $\mu_b[v(n_{TOB})]$. Ballistic $I_D V_{GS}$ -characteristics obtained from the quantum-transport tool QTx [53] were used as reference.

For studying the influence of $\mu_b[v(n_{TOB})]$ in the BTBT rate, another leakage mechanism inherent in DG transistors and FinFETs was considered, the floating-body effect (FBE) [54] caused by BTBT and affected by Shockley-Read-Hall (SRH) recombination. Models available in [19] were used in the TCAD simulations.

5.5.1 Influence of the Ballistic Mobility on the STDT Rate

Fig. 5.4(a) shows that the $v_b(n_{TOB})$ -model can well reproduce the quantum-ballistic I_{ON} , but the sub-threshold current becomes corrupted. As previously explained, this can be traced back to the deformation of $\psi_n(x)$ (see Fig. 5.4(b)). As the STDT rate of the NLT model is computed with the local QFPs at the classical turning points (u,l) for each tunnel path, the deformed $\psi_n(x)$ artificially suppresses the tunnel current.

The red curve in Fig. 5.5 was obtained by a post-processing calculation of the STDT current using the contact Fermi levels in the NLT model instead of $\psi_n(x)$. This removes the artifact in the deep

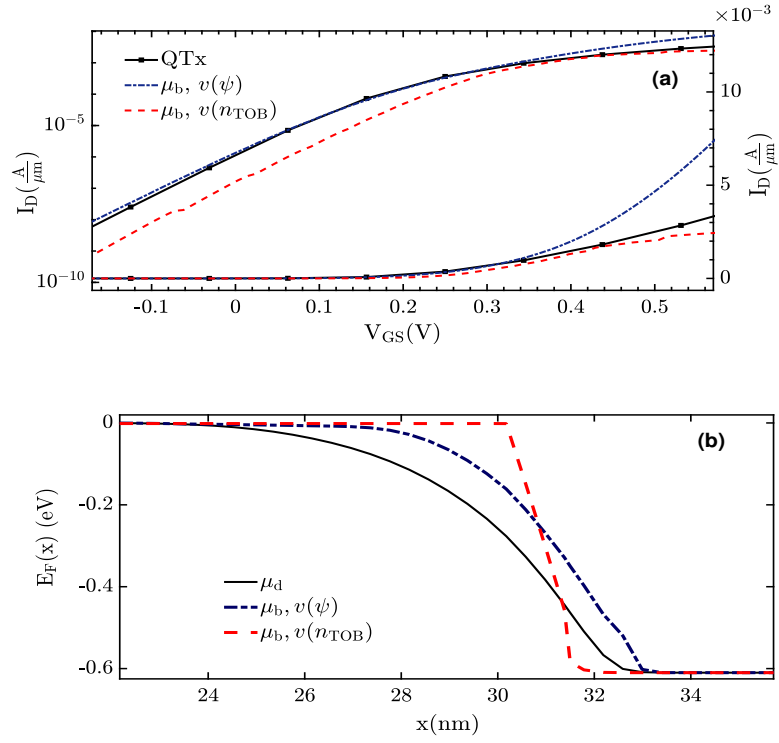


Figure 5.4: (a) $I_D V_{GS}$ -characteristics for $\mu_d = 1 \times 10^3 \text{ cm}^2/\text{Vs}$, $\mu_b[v(\psi_n)]$, and $\mu_b[v(n_{TOB})]$ extracted at $V_{DS} = 1 \text{ V}$. Parameters: $N_D = 5 \times 10^{19} \text{ cm}^{-3}$, $L_D = 20 \text{ nm}$, $L_G = 11.5 \text{ nm}$, $t_{\text{body}} = 4.2 \text{ nm}$ and $m_e = 0.0678 m_0$. (b) Corresponding profiles of the electron Fermi energy $-e\psi_n(x)$, extracted at $V_{GS} = 0 \text{ V}$ and $V_{DS} = 1 \text{ V}$

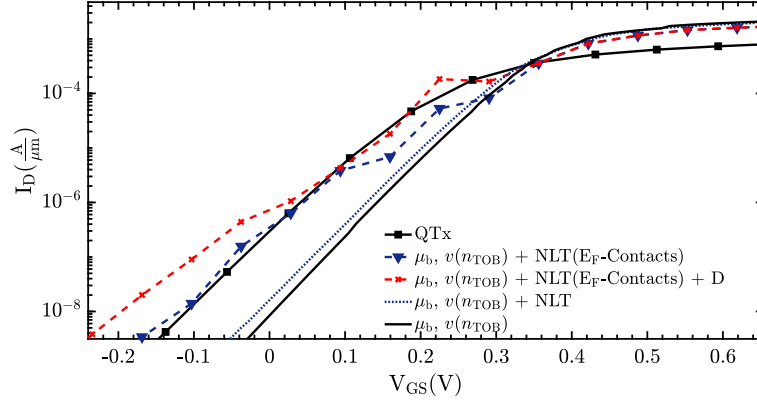


Figure 5.5: $I_D V_{GS}$ -characteristics of an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET with $L_G = 11.5$ with a tunnel mass $m_t = m_e$. The red curve was obtained by a post-processing calculation of the STD T current using the contact Fermi energies in the NLT model instead of local QFPs. In the legend, +D, accounts for adding extra discretization to the meshing points. Parameters: $N_D = 5 \times 10^{19} \text{ cm}^{-3}$, $m_e = 0.0678 m_0$, $V_{DS} = 50 \text{ mV}$.

sub-threshold range. However, close to the onset of inversion, it underestimates the QTx current. One way to better match the solution to the QTx reference in this region is increasing the number of discretization points in $\epsilon(x)$ used for the calculation of $j_{\text{tun},n}$. The number of tunneling paths and the generation-recombination current net rate will be higher then. Fig. 5.5 (blue curve) shows that the region before the onset of the resultant curve with discretization slowly converges to the QTx reference.

5.5.2 Influence of the Ballistic Mobility on the BTBT Rate

The ballistic velocity models also impact the BTBT rates which locally depend on the QFPs (see Fig. 5.7(a)).

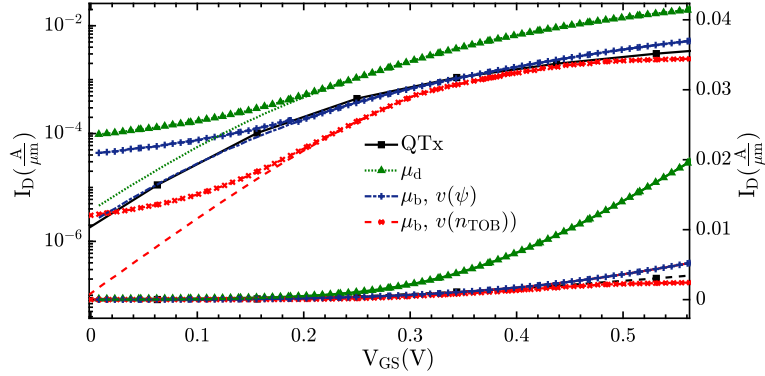


Figure 5.6: $I_D - V_{GS}$ characteristics at $V_{DS} = 1$ V of the device depicted in Fig. 2.1 for different electron mobilities ($\mu_d = 1 \times 10^4 \text{ cm}^2/\text{Vs}$, $\mu_b[v(\psi_n)]$, and $\mu_b[v(n_{TOB})]$) with $m_e = 0.0678 m_0$, $m_h = 0.446 m_0$, and SRH lifetimes $\tau_e = \tau_h = 1$ ns. The actual band gap of $E_G = 1.7$ eV caused by the strong confinement was lowered to $E_G = 0.7$ eV in order to enhance the effect of BTBT. Curves without BTBT are shown for comparison (dashed and dot-dashed lines).

The transfer curves in Fig. 5.6 (with BTBT + SRH + STDT) exhibit the additional FBE-induced leakage current. The stronger sensitivity of the BTBT rate to the $v_b(n_{TOB})$ -model as compared to the $v_b(\psi_n)$ -model can be traced back to a stronger deformation of $\psi_n(x)$ in the channel-drain junction where the electron BTBT rate is maximal (Fig. 5.7 (b)). The relative effect is not much changed even at an extreme rate of $10^{32} \text{ cm}^{-3}\text{s}^{-1}$ (see Fig. 5.8).

5.6 Discussion

Fig. 5.5 shows the $I_D - V_{GS}$ -characteristics obtained by the NLT model that uses Fermi levels of the contacts. The dark blue dashed curve reproduces the current in the sub-threshold regime effortlessly, however in the region close to the onset the number of tunneling paths

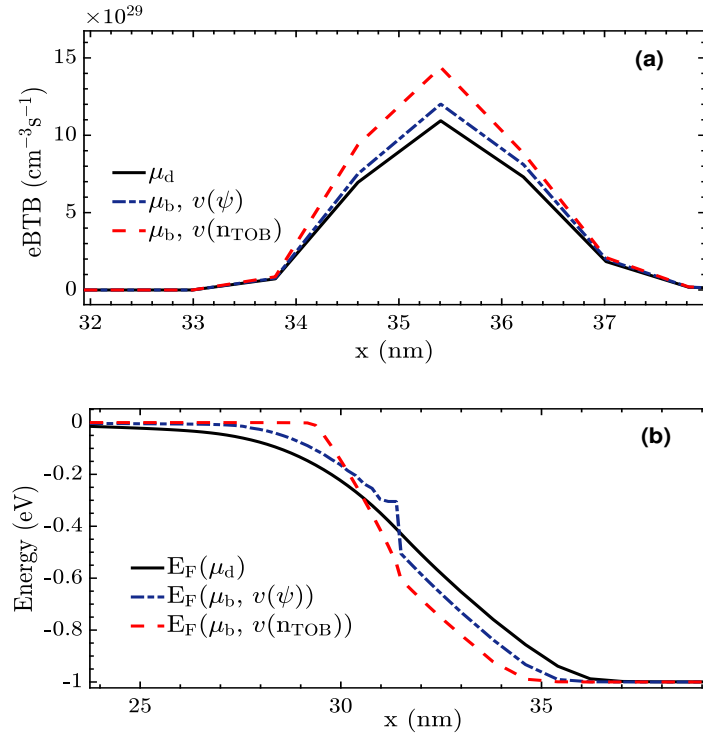


Figure 5.7: (a) Electron BTBT rates for different electron mobilities $\mu_d = 1 \times 10^3 \text{ cm}^2/\text{Vs}$, $\mu_b[v(\psi_n)]$, and $\mu_b[v(n_{\text{TOB}})]$ extracted at $V_{\text{GS}} = 0 \text{ V}$ and $V_{\text{DS}} = 1 \text{ V}$. Parameters are the same as in Fig. 4.5. (b) Corresponding profiles of the electron Fermi energy $-q\psi_n(x)$.

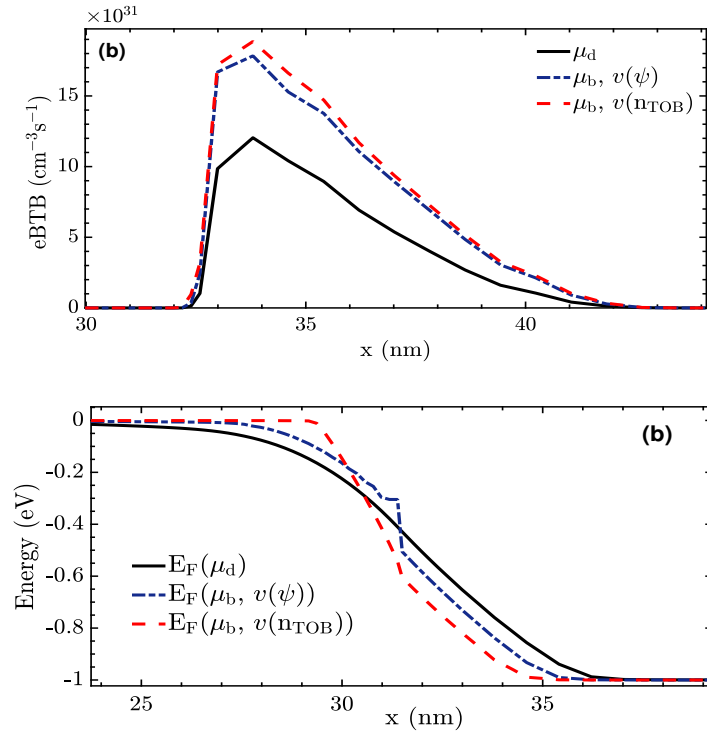


Figure 5.8: (a) Electron BTBT rates for different electron mobilities $\mu_d = 1 \times 10^3 \text{ cm}^2/\text{Vs}$, $\mu_b[v(\psi_n)]$, and $\mu_b[v(n_{\text{TOB}})]$ extracted at $V_{\text{GS}} = 0 \text{ V}$ and $V_{\text{DS}} = 2.5 \text{ V}$. Parameters are the same as in Fig. 4.5. (b) Corresponding profiles of the electron Fermi energy $-q\psi_n(x)$.

seems insufficient, because the generation current is much lower than the QTx reference.

This issue was addressed by increasing the space discretization. For a given energy discretization, a higher number of discretization points in x-direction leads to a larger number of aligned turning points, i.e. to more tunneling paths. The red curve shows a better agreement with the QTx reference curve in the voltage range from 0.1 to 0.3 V. However, since the idea was to increase the electron generation-recombination rate the sub-threshold region, increases further than the QTx reference as expected. This experiment was only performed for $\mu_b[v_b(n_{\text{TOB}})]$, the $\mu_b[v_b(\psi)]$ -model did not show any significant deformation of the QFP.

Fig. 5.4 (a) shows that the deformation of the QFP due to the inclusion of μ_b clearly underestimates the leakage current that is caused by STDt. When $\mu_b[v_b(n_{\text{TOB}})]$ is used together with the BTBT and NLT models as in Fig. 4.5, an enhanced electron BTBT rate (see Fig. 5.7) will have the same effect (see Fig. 5.7(b)) in Eq. (5.4). In Fig. 4.5, the leakage current with $\mu_b[v_b(\psi)]$ is still much higher than the one simulated with $\mu_b[v_b(n_{\text{TOB}})]$. This is a consequence of the NLT electron generation rate ($10^{33} \text{ cm}^{-3}\text{s}^{-1}$), that is greater than the electron BTBT rate. Thus, for a drain bias of $V_{\text{DS}}=1 \text{ V}$ STDt prevails as the dominant mechanism for the leakage current.

When V_{DS} increases to 2.5V, as presented in Fig. 5.8, the size of the electron BTBT rate starts approaching that of the NLT electron generation rate thus the total leakage current will be higher. As explained before, the cause of this change in the electron BTBT rate comes from the deformation of E_F at the end of the gate edge close to the source contact Eq. (5.4). A way to remediate this issue is to redefine beginning and end of the tunneling paths, so that the E_F used to calculate Eq. (5.4) is located before and after the conduction and valence band edges. It can be implemented in S-Device using the nonlocal recombination PMI [55].

Finally, it is important to point out that SRH recombination that is necessary for simulating the FBE does not affect the eBTBT rate, but a change of the lifetimes might have an effect on the NLT rate. This will be examined in detail in the next chapter.

5.7 Summary

In this chapter we discussed the effect of the ballistic mobility $\mu_b[v_b(n_{\text{TOB}})]$ on the leakage current in the sub-threshold regime. STD T was simulated by the NLT model and BTBT by the dynamical non-local path BTBT model. For correcting the issues that the NLT model has when used together with a ballistic mobility model, a version dependent on the Fermi level of the contacts was developed as a post-processing step.

Chapter 6

Floating Body Effect

6.1 Introduction

This chapter presents a study of the so called floating body effect (FBE) in InGaAs DG UTB FETs and GAA NWs, taking into account the quantum-mechanical effects due to the short and thin channels.

6.2 Theoretical Aspects

Certain theoretical considerations are quite important to address in order to be able to simulate the FBE in the ballistic regime. In this section we will present them in a consistent way relating them to quantum-mechanical effects due to geometrical confinement and the short-channel.

6.2.1 The Floating Body Effect

The FBE refers to the turn-on of a parasitic bipolar transistor (BJT) [54]- [56] in the channel of a FET. The parasitic BJT operation has its origin in electron-hole pairs generated in the channel-drain junction by BTBT. Electrons flow to the drain, but holes accumulate in the body under the gate and lower the potential barrier of the forward-biased source-channel junction. This induces an additional electron current

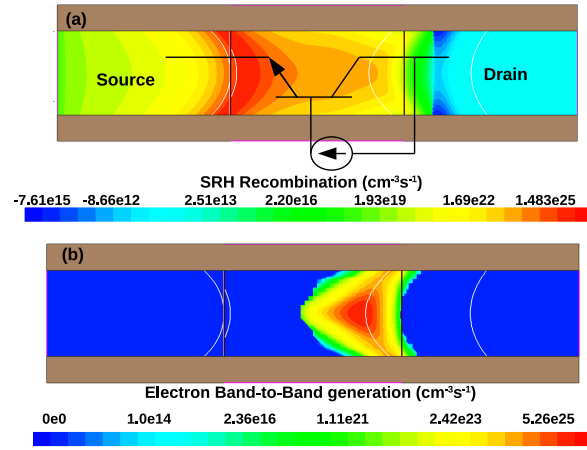


Figure 6.1: (a) SRH recombination rate in the channel of an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET and (b) the distribution of the electron BTBT generation rate.

form the "emitter" to the base. Part of these electrons recombine in the "base" by SHR recombination, the others reach the drain. The total current increases and can dominate the "leakage" current in the sub-threshold range. SRH rate and BTBT rate are depicted in Fig .6.1.

6.2.2 Band gap Variation due to Geometrical Confinement

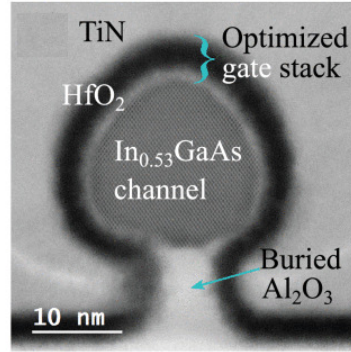


Figure 6.2: Transmission Electron Microscopy image of the cross section of the NW MOSFET normal to the transport direction [54].

For the 2D DG UTB transistor shown in Fig. 2.1 and the 3D GAA NW FET, the effective masses of electrons and holes (m_e , m_h), non-parabolicities (α_{np}) and band gaps (E_G) were extracted from the band structure simulations using the full-band version QTx.

The effect of geometrical confinement on the m_e and E_G was analysed by varying the body thickness t_{body} of the 2D DG UTB like in Fig. 6.3(a). Strong geometrical confinement E_G increases, as shown in Ref. [33]. In the case of a 3D GAA NW, the same trend is visible (Fig. 6.3 (b)).

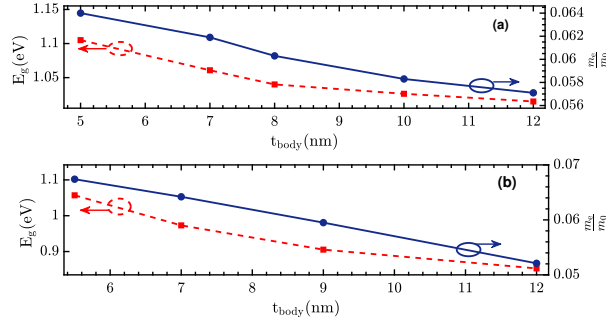


Figure 6.3: Variation of band gap and effective masses due to geometrical confinement in the channel of (a) 2D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET and (b) 3D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ GAA NW FET.

6.2.3 Effect of Quantization on the SRH Lifetimes

As shown in Chapter 2, quantum effects due to geometrical confinement in the channel can be modelled by the DG method and the quantized electron densities will be used as input for the Shockley-Read-Hall (SRH) rate. To be consistent with the E_G , m_e and m_h computations in the previous sections, it is assumed that the one-dimensional confinement causes an increased energetic separation between the electronic states in the conduction and valence bands and the deep trap levels as consequence of subband formation. Therefore, to appropriately capture this behaviour, the SRH model requires that the hole and electron lifetimes (τ) are modelled to account the presence of quantization [25].

In [25] an analytical approximation of a quantum lifetime (τ_n) was developed. The complete derivation details can be found in [25]. First, we assume Boltzmann statistics and the case of strong geometrical confinement where the contribution of a certain subband i at a position y is the stronger the closer it is to the trap energy level and the larger $|\Psi_i(y)|^2$. Furthermore, all subbands except the lowest one can be neglected if the confinement is sufficiently strong to provide a large subband separation ($\Delta\varepsilon_i \gg k_B T$). Then one obtains

$$\tau_n^{-1}(z) = \frac{c_n^0 N_t C \sqrt{\varepsilon_r}}{\sqrt{\pi k_B T} (\Delta + \varepsilon_r)} \exp\left(\frac{E_{\text{act}}^{\text{QM}}(\Gamma(y))}{k_B T}\right), \quad (6.1)$$

where the activation energy is defined as:

$$E_{\text{act}}^{\text{QM}} = \frac{(E_t + \Gamma(y) - \varepsilon_r)^2}{4\varepsilon_r} \quad (6.2)$$

where $\Gamma(y) = E_0 - E_c(z)$ is the additional offset between the conduction band energy and the lowest eigenenergy. N_t is the trap density, $\varepsilon_r = S\hbar\omega_0$ where S is the Huang-Rhys factor [57] and $\hbar\omega_0$ the effective phonon energy, which for InGaAs is typically 30 meV.

Finally, C is defined as $C = \exp(-(\varepsilon_r(2f_b + 1))/(\hbar\omega_0) + \zeta + a/(4k_B T))$, where $f_b = \frac{1}{\exp(\hbar\omega_0/k_B T) - 1}$, $\zeta = 2S\sqrt{f_b(f_b + 1)}$ and $a = (\hbar\omega_0\hbar\omega_0\zeta)/(2k_B T)$.

First, Eq.(6.1) was implemented in the PMI and a 1D profile across the structure of Fig. 2.1 was simulated. Then the model was adapted to the 2D structure of Fig. 2.1 and E_0 was obtained from the first energy level of a QTx band structure obtained from QTx. The traps energy level was set to $E_t = E_C - E_g/2$ and Δ from Eq.(6.1) is defined as $\Delta = E_0 - E_C + E_t$.

6.3 Simulation Method

For the simulation of the leakage current caused by STDT, the NLT model was used. The BTBT generation in the reversed-biased channel-drain junction was simulated by the dynamic nonlocal path BTBT model [52] as in Chapter 4. To describe the FBE correctly, not only the BTBT generation is needed, but also SRH and Auger recombination rates. For the SRH lifetimes a constant τ of 10^{-9} s was used. The values of the Auger coefficients for both electrons and holes were set to $2.5 \times 10^{-28} \text{cm}^6 \text{eV}^{-1}$ based on experimental values taken from Refs. [58] and [59].

The electron mobility model used was $\mu_b[v(\psi_n)]$ because for simulating STDT and BTBT, their respective generation rates are less sensitive to the QFP-dependent model. Furthermore, the shape of the QFP at $V_{\text{DS}} = 1 \text{V}$ is a step-like function, thus successfully reducing the on-current to the QTx reference.

6.4 Results

Both structures from Fig. 2.1 were simulated. The addition of spacers in these structures is aimed to reduce the electric field as origin of BTBT. Thus, both structures were simulated with additional spacers of $t_S = 3$ nm length on both sides of the channel and without spacers.

6.4.1 2D DG UTB FETs

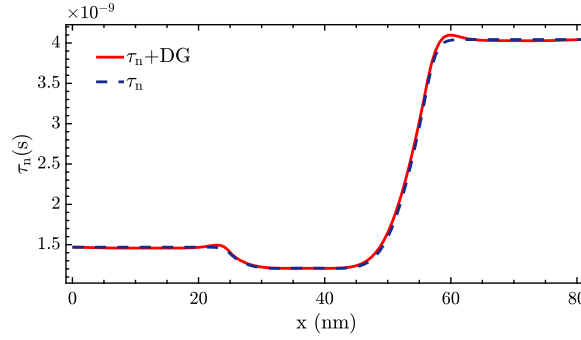


Figure 6.4: Electron lifetime (τ_n) profile in the channel of a 2D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET measured at $V_{DS} = 1$ V and $V_{GS} = -0.1$ V. A constant mobility $\mu_d = 10^4$ cm^2/Vs was used and $m_e = 0.0571m_0$, $m_h = 0.0558m_0$, $\hbar\omega_0 = 30$ meV. Parameters: $t_{\text{body}} = 12$ nm, $L_G = 25$ nm, $L_D = 25$ nm, $t_S = 3$ nm and $N_D = 1 \times 10^{19} \text{ cm}^{-3}$.

As explained in the previous section, Eq.(6.1) for τ_n was implemented in to assess its effect on the overall SRH recombination rate.

Fig. 6.4 shows the profile of τ_n along the channel. If the DG is applied the electron distribution, there is a slight increase in the gate-drain junction. The origin of the τ_n - profile along the channel comes from a 1-D cut across the channel.

Fig. 6.5 shows the SRH recombination rates that were obtained by different values of a constant τ and a quantum τ_n from Eq.(6.1). The inclusion of τ_n does not have a big effect on the SRH rate if compared to $\tau = 10^{-9}$ s. Therefore, the overall leakage current is almost the same (see Fig. 6.6).

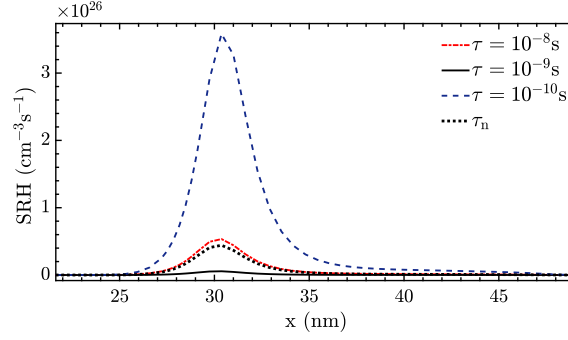


Figure 6.5: SRH recombination rates in a 2D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UT FET channel measured at $V_{\text{DS}} = 1 \text{ V}$ and $V_{\text{GS}} = -0.1 \text{ V}$. A constant $\mu_{\text{d}} = 10^4 \text{ cm}^2/\text{Vs}$ was used. Parameters: $t_{\text{body}} = 12 \text{ nm}$, $L_{\text{G}} = 25 \text{ nm}$, $L_{\text{D}} = 25 \text{ nm}$, $t_{\text{S}} = 3 \text{ nm}$ and $N_{\text{D}} = 1 \times 10^{19} \text{ cm}^{-3}$.

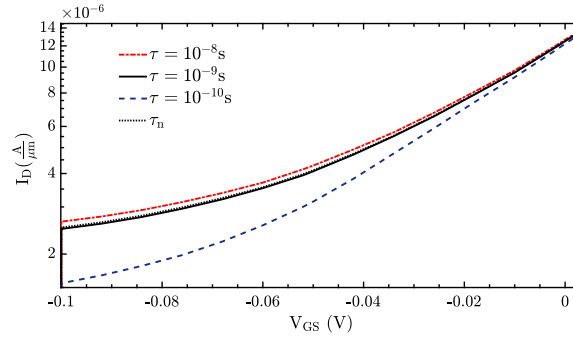


Figure 6.6: $I_{\text{D}}V_{\text{GS}}$ -characteristics of a 2D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET measured at $V_{\text{DS}} = 1 \text{ V}$ and $V_{\text{GS}} = -0.1 \text{ V}$. A constant $\mu_{\text{d}} = 10^4 \text{ cm}^2/\text{Vs}$ was used. Parameters: $t_{\text{body}} = 12 \text{ nm}$, $L_{\text{G}} = 25 \text{ nm}$, $L_{\text{D}} = 25 \text{ nm}$, $t_{\text{S}} = 3 \text{ nm}$ and $N_{\text{D}} = 1 \times 10^{19} \text{ cm}^{-3}$.

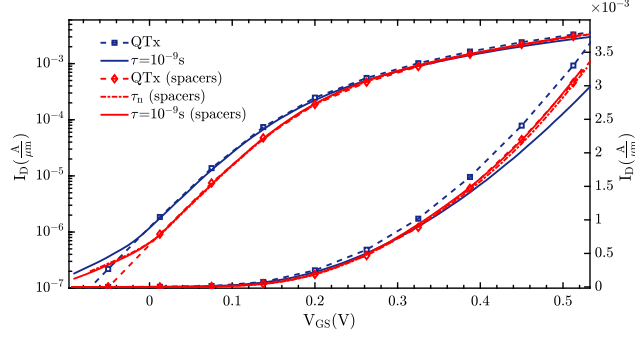


Figure 6.7: $I_D V_{GS}$ -characteristics a 2D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET measured at for a $V_{DS}=1$ V. The red curves include spacers and blue curves have no spacers. A ballistic mobility $\mu_b[v(\psi_n)]$ was used, $\tau = 10^{-9}\text{s}$ and $m_e = 0.0571m_0$, $m_h = 0.0558m_0$ where used. It is important to point out that the QTx references dont include SRH Recombination and the FBE is not present. Parameters: $t_{\text{body}} = 12$ nm, $L_G = 25$ nm, $L_D = 25$ nm, $t_S = 3$ nm and $N_D = 1 \times 10^{19} \text{cm}^{-3}$

Fig. 6.7 presents the $I_D V_{GS}$ -characteristics of a 2D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET. The inclusion of spacers reduces the effect of the strong electric field at the channel-drain interface [60] and thus, reduces the off-state leakage current due to BTBT. However, these spacers increase the parasitic capacitance between channel and S/D contacts. In certain applications like Radio Frequency (RF) this might be undesirable. Therefore, the same 2D DG UTB FET was simulated for both cases. The dashed curves, are the QTx ballistic references in the case of the presence and lack of spacers. No BTBT was included in this simulations. Finally, the inclusion of μ_b and Fermi statistics assures that the on-current will match to the QTx reference. The DG method was avoided, since it was very difficult to achieve convergence with the Fermi correction on, so the WF was fitted to the barrier height of the QTx reference at $V_{DS} = 0$ V and $V_G = -0.3$ V.

Fig. 6.8 shows the SRH recombination rates when the NLT model to simulate STDT is used and when STDT is absent respectively. The arrows point to the gate-drain junctions and the inclusion of

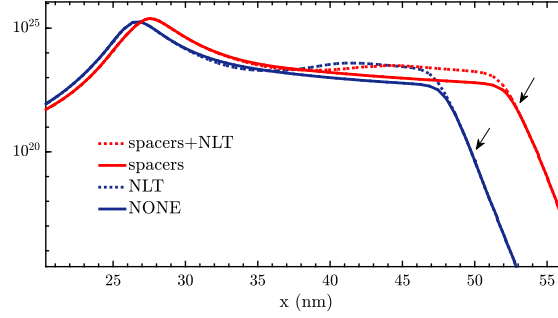


Figure 6.8: SRH recombination rate in a 2D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET measured at $V_{\text{DS}} = 1 \text{ V}$ and $V_{\text{G}} = -0.1 \text{ V}$.

NLT slightly increases the recombination rate due to the electron recombination peak that happens in the right part of the channel.

6.4.2 3D GAA NWs

Fig. 6.9 presents the $I_{\text{D}}V_{\text{GS}}$ -characteristics of a 3D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ GAA NW with spacers on the side of the channel.

6.5 Discussion

The reason that τ_n affects slightly the $I_{\text{D}}V_{\text{GS}}$ -characteristics of Fig. 6.6 compared to the ones with a constant $\tau = 10^{-9} \text{ s}$ could be attributed to the fact that we used a Huang Riss factor(S) in the PMI implementation that assures that τ_n remains in a range close to $\tau = 10^{-9} \text{ s}$. The pre factor c_n^0 is from Eq.(6.1) was set to 1 and S was fitted to 1 fit τ_n .

The introduction on the DG and Boltzmann statistics in the simulations assured a quantum density and wave penetration in the oxide follow the consider assumptions in the model [25]. However, in these simulations from Fig. 6.7 and Fig. 6.9 Fermi statistics were included to lower the on-state current to match the QTx reference and the DG omitted for ease of converge.

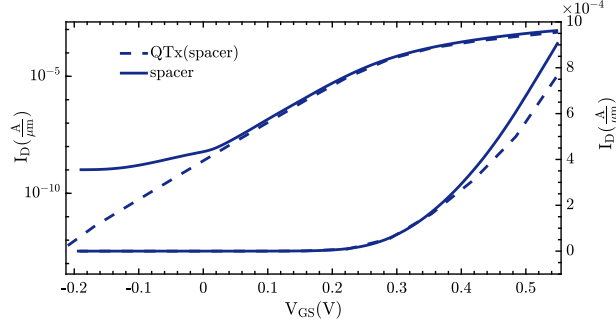


Figure 6.9: $I_D V_{GS}$ -characteristics of a 3D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ GAA NW at $V_{DS} = 1$ V. Red curves include spacers $t_S = 3$ nm and blue curves are without spacers. A ballistic mobility $\mu_b[v(\psi_n)]$, $\tau = 10^{-9}$ s and $m_e = 0.0642m_0$, $m_h = 0.059m_0$ where used. Parameters: $t_{\text{body}} = 7$ nm, $L_G = 15$ nm, $L_D = 25$ nm, $t_S = 3$ nm and $N_D = 1 \times 10^{19} \text{ cm}^{-3}$.

6.6 Summary

In this chapter the FBE in a 2D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ DG UTB FET and a 3D $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ GAA NW was simulated. The inclusion of μ_b and Fermi correction reproduce the ON-current from the QTx reference and the models of BTBT and NLT reproduced the leakage current. The inclusion of a quantum lifetimes has a negligible effect on the SRH recombination rate as compared to a constant value. Adding spacers at the sides of the channel reduces the leakage current.

Appendix A

Ballistic Mobility: Derivation of the TOB Density

This appendix presents the derivation of Eq. (4.16) which relates $n_{\text{TOB}}(x)$ to $n(x)$ in the channel region. Assuming Maxwell-Boltzmann statistics and a parabolic conduction band (CB), the density is given by

$$\begin{aligned} n(x) &= N_c \frac{2}{\sqrt{\pi}} \int_{\frac{E_c(x)}{k_B T}}^{\infty} d\epsilon \sqrt{\epsilon - \frac{E_c(x)}{k_B T}} \exp\left(-\frac{\epsilon - E_{F,n}(x)}{k_B T}\right) \\ &= N_c \exp\left(-\frac{\psi_n(x) + E_c(x)/q}{V_T}\right) \quad (\text{A.1}) \end{aligned}$$

where $E_c(x)$ is the energy of the CB edge, N_c the effective CB DOS, and $E_{F,n}(x)$ the quasi-Fermi energy $E_{F,n}(x) = -q\psi_n(x)$.

The TOB density is computed with the same integrand, but the lower integration limit has to be replaced by the energy of the TOB $E_{\text{TOB}} = E_c(x_{\text{TOB}})$:

$$\begin{aligned}
n_{\text{TOB}}(x) &= N_c \frac{2}{\sqrt{\pi}} \int_{\frac{E_{\text{TOB}}}{k_B T}}^{\infty} d\epsilon \sqrt{\epsilon - \frac{E_c(x)}{k_B T}} \exp\left(-\frac{\epsilon - E_{F,n}(x)}{k_B T}\right) \\
&= N_c \frac{2}{\sqrt{\pi}} \int_0^{\infty} d\epsilon \sqrt{\epsilon + \frac{E_{\text{TOB}} - E_c(x)}{k_B T}} e^{-\epsilon} \exp\left(-\frac{\psi_n(x) + E_{\text{TOB}}/q}{V_T}\right). \tag{A.2}
\end{aligned}$$

The last integral is proportional to the incomplete Gamma function [61]

$$\Gamma\left(\frac{3}{2}, b\right) = e^{-b} \int_0^{\infty} d\epsilon \sqrt{\epsilon + b} e^{-\epsilon}. \tag{A.3}$$

Expressing energies by the electrostatic potential, i.e. $E_{\text{TOB}} - E_c(x) = q\Phi(x) - q\Phi(x_{\text{TOB}})$, the TOB density can be written with (A.3) as

$$\begin{aligned}
n_{\text{TOB}}(x) &= \exp\left(\frac{E_{\text{TOB}} - E_c(x)}{qV_T}\right) N_c \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}, \frac{E_{\text{TOB}} - E_c(x)}{qV_T}\right) \\
&\quad \times \exp\left(-\frac{q\psi_n(x) + E_{\text{TOB}}}{qV_T}\right) \\
&= N_c \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}, \frac{\Phi(x) - \Phi(x_{\text{TOB}})}{V_T}\right) \exp\left(-\frac{\psi_n(x) - E_c(x)/q}{V_T}\right) \\
&= \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}, \frac{\Phi(x) - \Phi(x_{\text{TOB}})}{V_T}\right) n(x) \tag{A.4}
\end{aligned}$$

which is Eq. (4.16). The assumption of Maxwell-Boltzmann statistics holds in a large part of the channel, but breaks down where $E_c(x) - E_{F,n}(x) < k_B T$, i.e. near the pn-junctions (see Fig. 4.7).

Since $\Gamma(\frac{3}{2}, 0) = \sqrt{\pi}/2$ it follows that $n(x_{\text{TOB}}) = n_{\text{TOB}}(x_{\text{TOB}})$ as required.

Bibliography

- [1] C. Gartenberg, “On its 50th anniversary, intel is forced to do less with moore,” 2018. [Online]. Available: <https://www.theverge.com/2018/7/19/17590242/intel-50th-anniversary-moores-law-history-chips-processors-future>.
- [2] T. Quarterly, “After Moore’s law,” 2016. [Online]. Available: <https://www.economist.com/technology-quarterly/2016-03-12/after-moores-law>.
- [3] J. A. del Alamo, “Nanometre-scale electronics with III-V compound semiconductors,” *Nature*, vol. 479, p. 317, 2011.
- [4] M. Rau, E. Caruso, D. Lizzit, P. Palestri, D. Esseni, A. Schenk, L. Selmi, and M. Luisier, “Performance projection of III-V ultra-thin-body, FinFET, and nanowire mosfets for two next-generation technology nodes,” in *International Electron Device Meeting (IEDM) Tech. Digest*, 2016, pp. 758–761.
- [5] S. Banerjee, W. Richardson, J. Coleman, and A. Chatterjee, “A new three-terminal tunnel device,” *IEEE Electron Device Letters*, vol. 8, no. 8, pp. 347–349, Aug 1987.
- [6] H. Lu and A. Seabaugh, “Tunnel Field-Effect Transistors: State-of-the-Art,” *IEEE Journal of the Electron Devices Society*, vol. 2, no. 4, pp. 44–49, July 2014.
- [7] “The international technology roadmap for semiconductors (ITRS),” Tech. Rep., 2012.

- [8] P. De Wolf, T. Clarysse, W. Vandervorst, L. Hellemans, and P. Niedermann, "Cross-sectional nano-spreading resistance profiling," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, vol. 16, no. 1, pp. 355–361, 1998. [Online]. Available: <https://avs.scitation.org/doi/abs/10.1116/1.589810>
- [9] N. Khalil, J. Faricelli, C. Huang, and S. Selberherr, "Two-Dimensional Dopant Profiling of Submicron Metal-Oxide-Semiconductor Field-Effect Transistor Using Nonlinear Least Squares Inverse Modeling," *Journal of Vacuum Science Technology B: Microelectronics and Nanometer Structures*, vol. 14, pp. 224 – 230, 02 1996.
- [10] A. Bandiziol, P. Palestri, F. Pittino, D. Esseni, and L. Selmi, "A TCAD-Based Methodology to Model the Site-Binding Charge at ISFET/Electrolyte Interfaces," *IEEE Transactions on Electron Devices*, vol. 62, no. 10, pp. 3379–3386, Oct 2015.
- [11] S. Sant, K. Moselund, D. Cutaia, H. Schmid, M. Borg, H. Riel, and A. Schenk, "Lateral InAs/Si p-Type Tunnel FETs Integrated on Si Part 2: Simulation Study of the Impact of Interface Traps," *IEEE Transactions on Electron Devices*, vol. 63, no. 11, pp. 4240–4247, Nov 2016.
- [12] P. Osgnach, E. Caruso, D. Lizzit, P. Palestri, D. Esseni, and L. Selmi, "The impact of interface states on the mobility and drive current of In_{0.53}Ga_{0.47}As semiconductor n-MOSFETs," *Solid-State Electronics*, vol. 108, 01 2015.
- [13] O. Badami, E. Caruso, D. Lizzit, P. Osgnach, D. Esseni, P. Palestri, and L. Selmi, "An Improved Surface Roughness Scattering Model for Bulk, Thin-Body, and Quantum-Well MOSFETs," *IEEE Transactions on Electron Devices*, vol. 63, no. 6, pp. 2306–2312, June 2016.
- [14] A. C. Seabaugh and Q. Zhang, "Low-Voltage Tunnel Transistors for Beyond CMOS Logic," *Proceedings of the IEEE*, vol. 98, no. 12, pp. 2095–2110, Dec 2010.

- [15] S. Sant and A. Schenk, “The effect of density-of-state tails on band-to-band tunneling: Theory and application to tunnel field effect transistors,” *Journal of Applied Physics*, vol. 122, no. 13, p. 135702, 2017. [Online]. Available: <https://doi.org/10.1063/1.4994112>
- [16] D. Esseni, M. Pala, P. Palestri, C. Alper, and T. R., “A review of selected topics in physics based modeling for tunnel field-effect transistors,” *Semiconductor Science and Technology*, vol. 32, p. 083005, 08 2017.
- [17] W. G. Vandenberghe, B. Soree, W. Magnus, M. V. Fischetti, A. S. Verhulst, and G. Groeseneken, “Two-dimensional quantum mechanical modeling of band-to-band tunneling in indirect semiconductors,” in *2011 International Electron Devices Meeting*, Dec 2011, pp. 5.3.1–5.3.4.
- [18] M. Luisier, A. Schenk, and W. Fichtner, “Quantum transport in two- and three-dimensional nanoscale transistors: Coupled mode effects in the nonequilibrium Green’s function formalism,” *Journal of Applied Physics*, vol. 100, no. 4, p. 043713, 2006.
- [19] *Sentaurus-Device User Guide*, Synopsys Inc, Mountain View, California, 2018.
- [20] J. Mütting, B. Kakarla, and U. Grossner, “Comprehensive and detailed study on the modeling of commercial SiC power mosfet devices using TCAD,” in *Silicon Carbide and Related Materials 2016*, ser. Materials Science Forum, vol. 897. Trans Tech Publications, 6 2017, pp. 553–556.
- [21] IMEC, “Extended intel technology road map,” IMEC, Report, www.iii-v-mos-project.eu.
- [22] L. Selmi, E. Caruso, S. Carapezzi, M. Visciarelli, E. Gnani, N. Zagni, P. Pavan, P. Palestri, D. Esseni, A. Gnudi *et al.*, “Modelling nanoscale n-MOSFETs with III-V compound semiconductor channels: From advanced models for band structures, electrostatics and transport to TCAD,” in *2017 IEEE International Electron Devices Meeting (IEDM)*, Dec 2017, pp. 13.4.1–13.4.4.

- [23] M. S. Shur, “Low ballistic mobility in submicron HEMTs,” *IEEE Electron Device Letters*, vol. 23, no. 9, pp. 511–513, Sep. 2002.
- [24] E. Gnani, A. Gnudi, S. Reggiani, and G. Bacarani, “Effective mobility in nanowire fets under quasi-ballistic conditions,” *IEEE Transactions on Electron Devices*, vol. 57, no. 1, pp. 336–344, Jan 2010.
- [25] T. Höhr, A. Schenk, and W. Fichtner, “Revised Shockley Read Hall lifetimes for quantum transport modeling,” *Journal of Applied Physics*, vol. 95, pp. 4875 – 4882, 06 2004.
- [26] P. Aguirre, H. Carrillo, A. Ziegler, M. Luisier, and A. Schenk, “Drift-diffusion quantum corrections for In_{0.53}Ga_{0.47}As double gate ultra-thin-body FETs,” in *2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Sep. 2016, pp. 53–56.
- [27] P. Aguirre, M. Rau, and A. Schenk, “2D and 3D TCAD simulation of III-V channel FETs at the end of scaling,” in *2018 Joint International EUROSIOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSIOI-ULIS)*, March 2018, pp. 1–4.
- [28] A. Wettstein, A. Schenk, and W. Fichtner, “Quantum device-simulation with the density-gradient model on unstructured grids,” *IEEE Transactions on Electron Devices*, vol. 48, no. 2, pp. 279–284, Feb 2001.
- [29] F. Heinz and A. Schenk, “Self-consistent modeling of longitudinal quantum effects in nanoscale double-gate metal oxide semiconductor field effect transistors,” *Journal of Applied Physics*, vol. 100, 10 2006.
- [30] A. Schenk and A. Wettstein, “Simulation of DGSOI MOSFETs with a Schrödinger-Poisson based mobility model,” in *International Conference on Simulation of Semiconductor Processes and Devices*, Sep. 2002, pp. 21–24.
- [31] M. G. Ancona, “Density-gradient theory: A macroscopic approach to quantum confinement and tunneling in semiconductor

- devices,” *Journal of Computational Electronics*, vol. 10, pp. 65–97, 06 2011.
- [32] *Sentaurus Monte Carlo User Guide*, Synopsys Inc, Mountain View, California, 2015.
- [33] G. Zerveas, E. Caruso, G. Bacarani, L. Czornomaz, N. Daix, D. Esseni, E. Gnani, A. Gnudi, R. Grassi, M. Luisier *et al.*, “Comprehensive comparison and experimental validation of band-structure calculation methods in III-V semiconductor quantum wells,” *Solid-State Electronics*, vol. 115, pp. 92 – 102, 2016, selected papers from the EUROSIO-ULIS conference. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038110115002580>
- [34] O. Penzin, G. Paasch, and L. Smith, “Nonparabolic multivalley quantum correction model for InGaAs double-gate structures,” *IEEE Transactions on Electron Devices*, vol. 60, no. 7, pp. 2246–2250, July 2013.
- [35] M. Luisier, A. Schenk, and W. Fichtner, “Quantum transport in two- and three-dimensional nanoscale transistors: Coupled mode effects in the nonequilibrium Green’s function formalism,” *Journal of Applied Physics*, vol. 100, no. 4, p. 043713, 2006. [Online]. Available: <https://doi.org/10.1063/1.2244522>.
- [36] P. Aguirre, M. Rau, and A. Schenk, “2D and 3D TCAD simulation of III-V channel FETs at the end of scaling,” *Solid-State Electronics*, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038110119301546>
- [37] A. Schenk and P. Aguirre, “TCAD models of the ballistic mobility in the source-to-drain tunneling regime,” *Solid-State Electronics*, vol. 157, pp. 1 – 11, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038110118306439>.
- [38] P. Aguirre and A. Schenk, “Ballistic Mobility Model for QDD Simulation of Ultra-short Transistors,” in *2018 IEEE 2nd Electron Devices Technology and Manufacturing Conference (EDTM)*, March 2018, pp. 77–79.

- [39] M. Lundstrom and X. Sun, “Some useful relations for analyzing nanoscale MOSFETs operating in the linear region,” March. 2016.
- [40] R. Kotlyar, R. Rios, C. E. Weber, T. D. Linton, M. Armstrong, and K. Kuhn, “Distributive quasi-ballistic drift diffusion model including effects of stress and high driving field,” *IEEE Transactions on Electron Devices*, vol. 62, no. 3, pp. 743–750, March 2015.
- [41] S. Martinie, G. Le Carval, D. Munteanu, S. Soliveres, and J. Autran, “Impact of ballistic and quasi-ballistic transport on performances of double-gate mosfet-based circuits,” *IEEE Transactions on Electron Devices*, vol. 55, no. 9, pp. 2443–2453, Sep. 2008.
- [42] K. Blotekjaer, “Transport equations for electrons in two-valley semiconductors,” *IEEE Transactions on Electron Devices*, vol. 17, no. 1, pp. 38–47, Jan 1970.
- [43] S. A. Schwarz and S. E. Russek, “Semi-empirical equations for electron velocity in silicon: Part III-MOS inversion layer,” *IEEE Transactions on Electron Devices*, vol. 30, no. 12, pp. 1634–1639, Dec 1983.
- [44] W. R. Frensley, “Barrier-limited transport in semiconductor devices,” *IEEE Transactions on Electron Devices*, vol. 30, no. 12, pp. 1619–1623, Dec 1983.
- [45] O. Penzin, L. Smith, A. Erlebach, M. Choi, and K. Lee, “Kinetic velocity model to account for ballistic effects in the drift-diffusion transport approach,” *IEEE Transactions on Electron Devices*, vol. 64, no. 11, pp. 4599–4606, Nov 2017.
- [46] A. Ziegler, M. Frey, L. Smith, and M. Luisier, “A nonparabolic bandstructure model for computationally efficient quantum transport simulations,” *IEEE Transactions on Electron Devices*, vol. 63, no. 5, pp. 2050–2056, May 2016.
- [47] A. Erlebach, K. H. Lee, and F. M. Bufler, “Empirical ballistic mobility model for drift-diffusion simulation,” in *2016 46th*

European Solid-State Device Research Conference (ESSDERC), Sep. 2016, pp. 420–423.

- [48] A. Schenk and A. Wettstein, “Simulation of DGSOI MOSFETs with a Schrödinger-Poisson based mobility model,” in *International Conference on Simulation of Semiconductor Processes and Devices*, Sep. 2002, pp. 21–24.
- [49] M. Jeong, P. M. Solomon, S. E. Laux, H. . P. Wong, and D. Chidambarrao, “Comparison of raised and schottky source/drain mosfets using a novel tunneling contact model,” in *International Electron Devices Meeting 1998. Technical Digest (Cat. No.98CH36217)*, Dec 1998, pp. 733–736.
- [50] C. Convertino, C. Zota, S. Sant, F. Eltes, M. Sousa, D. Caimi, A. Schenk, and L. Czornomaz, “InGaAs-on-insulator FinFETs with reduced Off-Current and Record Performance,” 12 2018, pp. 39.2.1–39.2.4.
- [51] M. B. et. al, “Multisubband ensemble Monte Carlo analysis of tunneling leakage mechanisms in ultrascaled FDSOI, DGSOI, and finFET Devices,” *IEEE Transactions on Electron Devices*, vol. PP, pp. 1–8, 01 2019.
- [52] E. O. Kane, “Theory of tunneling,” *Journal of Applied Physics*, vol. 32, no. 1, pp. 83–91, 1961. [Online]. Available: <https://doi.org/10.1063/1.1735965>.
- [53] M. Luisier, A. Schenk, and W. Fichtner, “Quantum transport in two- and three-dimensional nanoscale transistors: Coupled mode effects in the nonequilibrium Green function formalism,” *Journal of Applied Physics*, vol. 100, no. 4, p. 043713, 2006. [Online]. Available: <https://doi.org/10.1063/1.2244522>
- [54] S. Sant, P. Aguirre, H. Hahn, V. Deshpande, L. Czornomaz, and A. Schenk, “Impact of Floating Body Effect, Back-Gate Traps, and Trap-Assisted Tunneling on Scaled In_{0.53}Ga_{0.47}As Ultrathin-Body MOSFETs and Mitigation Measures,” *IEEE Transactions on Electron Devices*, vol. 65, no. 6, pp. 2578–2584, June 2018.

- [55] S. Sant, “Role of non-idealities in III-V/Si and A11 III-V Tunnel Field Effect Transistors,” Ph.D. dissertation, ETH Zurich, 2018.
- [56] M. Yoshimi, M. Terauchi, A. Nishiyama, O. Arisumi, A. Murakoshi, K. Matsuzawa, N. Shigyo, S. Takeno, M. Tomita, K. Suzuki *et al.*, “Suppression of the floating-body effect in SOI MOSFET’s by the bandgap engineering method using a si/sub 1-x/ge/sub x/ source structure,” *IEEE Transactions on Electron Devices*, vol. 44, no. 3, pp. 423–430, March 1997.
- [57] A. Schenk, “An improved approach to the Shockley-Read-Hall recombination in inhomogeneous fields of space-charge regions,” *Journal of Applied Physics*, vol. 71, no. 7, pp. 3339–3349, 1992. [Online]. Available: <https://doi.org/10.1063/1.350929>.
- [58] S. Hausser, G. Fuchs, A. Hangleiter, K. Streubel, and W. T. Tsang, “Auger recombination in bulk and quantum well InGaAs,” *Applied Physics Letters*, vol. 56, no. 10, pp. 913–915, 1990. [Online]. Available: <https://doi.org/10.1063/1.103175>
- [59] M. E. Prise, M. R. Taghizadeh, S. D. Smith, and B. S. Wherrett, “Picosecond measurement of Auger recombination rates in InGaAs,” *Applied Physics Letters*, vol. 45, no. 6, pp. 652–654, 1984. [Online]. Available: <https://doi.org/10.1063/1.95344>.
- [60] C. Huang, S. Lee, V. Chobpattana, S. Stemmer, A. C. Gossard, and M. J. W. Rodwell, “S5-H6: Leakage current suppression in InGaAs-channel MOSFETs: Recessed InP source/drain spacers and InP channel caps,” in *2014 Lester Eastman Conference on High Performance Devices (LEC)*, Aug 2014, pp. 1–4.
- [61] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover Publications Inc., p. 260.

Curriculum Vitae

Personal Details

Name: Delia Paulina Aguirre Fernandez
Birth: 29th May 1986, Quito, Ecuador
Citizenship: Ecuador

Education

Swiss Federal Institute of Technology: **PhD, April 2019**
Delft University of Technology/Swiss Federal Institute of Technology:
MSc. in Biomedical Engineering, 2013
Ecole Nationale Supérieure d'Ingenieurs du Sud Alsace, France:
Master in Control, Signal and Image Processing, 2011
Universidad San Francisco de Quito:
Electrical Engineer, 2010

Research Experience

Guest Researcher: TU Delft
09.2014 to 12.2014 .
Master Thesis: Electronics Laboratory, ETH Zurich
09.2012 to 03.2013
Research Assistant: Politecnico Institute of Grenoble
GIPSA Lab, 12.2009 to 06.2010
Research Assistant: Bern University of Applied Sciences
Radiofrequency Laboratory, 05.2009 to 08.2009

Industrial Experience**Development Engineer:**

Roche Diagnostics, Rotkreutz, Switzerland, 11.2013 to 09.2014

Research and Development Intern:

Philips Research Center, Eindhoven, The Netherlands, 05.2013 to 10.2013

Development Engineering Intern:

Albis Technologies, Electromagnetic Compatibility Certification Laboratory,
Zurich, Switzerland, 03.2011 to 09.2011