Physics-based compact modeling of quasi-ballistic transistors

Ajaykumar, Arjun

2017

Ajaykumar, A. (2017). Physics-based compact modeling of quasi-ballistic transistors.
Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/72954>

<https://doi.org/10.32657/10356/72954>

PHYSICS-BASED COMPACT MODELING OF QUASI-BALLISTIC TRANSISTORS

ARJUN AJAYKUMAR

SCHOOL OF ELECTRICAL AND ELECTRONIC ENGINEERING

2017

PHYSICS-BASED COMPACT MODELING OF QUASI-BALLISTIC TRANSISTORS

ARJUN AJAYKUMAR

School of Electrical & Electronic Engineering

A Thesis Submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy

2017

Acknowledgment

During my Ph.D. study, I have received guidance and support from many people. I would like to express my gratitude to all of them.

First of all, I would like to express my deepest gratitude to my supervisor A/P Zhou Xing for guidance during the past 4-years of my Ph.D. study in NTU. He has guided my research with much patience.

I am also thankful to Dr. Chiah Siau Ben, Binit Syamal, Liu Xu, and Zhou Hongtao for support and valuable discussions. I would also like to thank Prof. Dimitri Antoniadis for many insightful discussions and the SOI measurement data.

I would also like to thank all my friends at NTU. I'm really thankful to Ajith, Jishnu, Aamani, and Rijil for their support and encouragement during the past few years.

Last but not least, I would like to express my gratitude to my wife Harsha and my parents Ajaykumar and Sathee Devi for their understanding and encouragement.

Abstract

This Ph.D. study deals with the development of physics-based compact models for advanced semiconductor devices with a focus on quasi-ballistic (QB) transistors. The initial part of the thesis deals with Fermi potential modeling of III-V (InGaAs)-based High Electron Mobility Transistors with rectangular quantum well. The III-V material based devices usually have higher mobility due to their large mean free path (λ). When compared to silicon-based devices, the large λ devices show a faster onset of ballistic effects with channel length scaling. The drift-diffusion (DD) based devices are scattering limited whereas the QB devices are limited by the injection velocity. The mobility (μ), which is defined in a scattering dominant environment, loses its validity in QB regime. Hence, the DD based models cannot be used to study the physics of devices operating in QB regime. Similarly, the QB models in literature fail to capture the DD device physics. The current literature lacks models that exclusively study the DD to QB transition even though DD and QB devices have been well modeled individually.

In the major part of this thesis, critical length (δ), defined based on the thermal energy drop with respect to the virtual source, is presented as a better parameter for ballisticity. The δ is derived from the Quasi-2D solution of the Poisson equation and captures the device parameter (oxide thickness,

body thickness, doping, and channel length) dependences. The proposed model is validated with extracted Monte-Carlo simulation results from the literature. The model can be used to predict the onset of ballistic behavior based on the relative magnitude of the δ to λ . The δ model is further used to calculate the apparent μ , which could be used similarly to diffusive μ to predict the electrical characteristics in QB regime. To exclusively study the DD to QB transition, the velocity vs. electric field characteristics of the DD devices is extended to the QB regime. This new approach is used for an in-depth analysis of the transition region in conceptual MOS devices. The impact of various material parameters like λ , thermal velocity, and saturation velocity are well studied at different channel length scales. The proposed model tallies well with the DD and QB regime device physics. The model is also tested for its temperature dependence, which is, to the best of our knowledge, the first in literature. The model is validated with SOI measurement data for channel lengths spanning from DD to QB regimes.

The last part of the thesis focuses on the study of QB transport on the low-frequency noise (LFN). A modified $(g_m/I_d)^2$ method for the sub-threshold regime is proposed to aid in characterizing the LFN origin. Also, the recent measurement results on devices operating in ballistic regime show LFN scaling trend that cannot be explained using the well-established carrier number (CNF) and correlated mobility fluctuation (CMF) theories developed for DD devices. The CNF/CMF theory is extended to the QB regime to give a physical explanation for the anomalous LFN scaling behavior. The proposed models are validated with measurement data extracted from the literature.

Contents

List of Figures	x
List of Tables	xxiv
1 Introduction	1
1.1 Motivation	1
1.2 Objective	4
1.3 Major Contributions of the Thesis	4
1.4 Organization of the Thesis	6
List of Abbreviations	1
2 Surface-Potential-Based MOSFETs and Fermi-Potential-Based HEMTs	8
2.1 Surface Potential Based Compact Models	9
2.1.1 Pao–Sah Solution	9
2.1.2 Piecewise Regional Approach	12
2.2 Fermi-potential Model for High Electron Mobility Transistors	16

2.2.1	HEMT Structure Overview	16
2.2.2	Coupled Poisson-Schrödinger Solution	19
2.2.3	Unified Regional Triangular Potential Well Solution	21
2.3	Double Heterostructure InGaAs Based HEMTs	22
2.3.1	Rectangular Quantum Well	22
2.3.2	Comparison with Triangular Well Model	25
2.4	Model Formulation	26
2.4.1	Unified Surface Potential Model	34
2.4.2	Model Validation and Scalability	35
2.4.3	Accumulation	39
2.5	Drain Current Modeling	41
2.6	Conclusion and Summary	48
3	Critical Length Based DD and Ballistic Transition Model	50
3.1	Drift-Diffusion to Quasi-Ballistic Regime	51
3.1.1	Shur Model	57
3.1.2	Arora Model	58
3.1.3	Landauer Transport and Scattering Theory	60
3.2	Critical-Length Model from Q2D Poisson for Calculating the Onset of Ballisticity	64

3.2.1	Model Description	65
3.2.2	Results and Discussion	68
3.3	Conclusion	74
4	Drift-Diffusion to Ballistic Regime Transition	76
4.1	Study of DD to QB Transition	76
4.2	Model Description	80
4.3	Results and Discussion	85
4.4	Conclusion	100
5	Low Frequency Noise in Deep Submicron Devices	101
5.1	Low Frequency Noise, $1/f$ Noise	101
5.1.1	Hooge model	102
5.1.2	Charge number fluctuation (CNF) model	103
5.1.3	Number fluctuation and correlated mobility fluctuation in MOS-FET	110
5.2	Noise in Subthreshold Regime	117
5.3	Results and Discussion	123
5.4	Conclusion	131
6	Low Frequency Noise in Drift-Diffusion to Ballistic Devices	132
6.1	Introduction	132

6.2	Model Formulation	134
6.3	Results and Discussion	136
6.4	Conclusion	143
7	Conclusion and future work	144
7.1	Conclusion	144
7.2	Future Work	146
	Publication	147
	Bibliography	149

List of Figures

2.1	MOSFET schematic for the surface potential solution. Poisson equation is solved in the Gauss box marked. Adapted from [4].	10
2.2	The unified regional surface potential solution, along with the regional pieces, as indicated, for a s-DG FinFET with $T_{ox} = 3nm$, $T_{Si} = 50nm$, and a heavily-doped body with $N_A = 10^{18}cm^{-3}$. Adapted from [4].	15
2.3	GaN HEMT with triangular quantum well carrier confinement.	17
2.4	Triangular low-energy notch in conventional MOSFETs for carrier confinement.	18
2.5	Energy band diagram of a typical MIS-HEMT with a triangular quantum well. Fermi potential and energy levels referenced to the bottom of quantum well are shown. Adapted from [8].	21
2.6	The schematic of InGaAs HEMT with double heterostructure resulting in a rectangular QW confinement.	22
2.7	Energy band diagram of HEMT with a rectangular quantum well. Fermi potential is in volts and referenced to the bottom of quantum well as shown.	24

2.8	The InGaAs HEMT structure with single heterojunction, resulting in a triangular quantum well for confinement.	24
2.9	Comparison between two devices with same gate stack till the channel region but with different quantum wells. One uses a rectangular quantum well and the other with triangular quantum well.	25
2.10	Comparison between the two strong inversion regional pieces derived by including and not including the electric field effect on the QW.	33
2.11	Unified surface potential model and the regional models matches well in subthreshold, weak inversion and strong inversion regions.	36
2.12	Unified surface potential model and the extracted data from coupled Poisson Schrödinger simulation. The PS simulation and the model matches well in subthreshold, weak inversion and strong inversion regions.	36
2.13	Absolute difference between the unified surface potential model and coupled Poisson Schrödinger extraction for InGaAs HEMT with physical parameters as given in Table 2.3.	37
2.14	Model (lines) and extracted Coupled Poisson Schrödinger data (symbols) for different oxide thicknesses. The proposed model matches the extracted data very well.	37

2.15 Model (lines) and extracted coupled Poisson Schrödinger data (symbols) for different quantum well dimensions. The proposed model matches the extracted data very well.	38
2.16 Model (lines) and extracted coupled Poisson Schrödinger data (symbols) for different δ -doping. The proposed model matches the extracted data very well.	38
2.17 Model accurately matches the charge variation with gate bias extracted from numerical simulations.	40
2.18 The full unified model with the accumulation region due to the low band gap InAs layer. The E_f matches the numerical solution obtained from the coupled PS solution accurately	41
2.19 The first derivative of E_f shows no abrupt discontinuities indicating a smooth transition between different operating regimes.	41
2.20 The model matches the $I_d - V_g$ characteristics of an InGaAs HEMT with channel length (L) and width (W) of 500nm and 1 μ m accurately. The drain voltages applied are 0.01V, 0.1V, 0.2V, and 0.3V.	47
2.21 The model matches $I_d - V_d$ characteristics of an InGaAs HEMT with channel length (L) and width (W) of 500nm and 1 μ m accurately. The gate voltage is varied from $-0.2V$ to $0.5V$ with a step size of 0.1V. The short-channel effects like DIBL and CLM are incorporated into the E_f model as discussed. . . .	47
3.1 The schematic of a MOSFET, the operating regime of which is determined by the source, drain and gate biases.	51

3.2	The <i>Electric – field Vs. Velocity</i> schematic for a material. The velocity saturates for electric fields higher than the critical field, E_c . For silicon $v_{sat} \sim 10^7 cm/s$	52
3.3	Maxwellian distribution functions at different Fermi energies corresponding to the E_f (Solid) and a quasi-Fermi level F_n (dashed). Adapted from [23].	53
3.4	Maxwellian distribution functions where the carrier temperature is equal to the lattice temperature (Solid), and the carrier temperature that exceeds that of the lattice (dashed). Adapted from [23].	54
3.5	The schematic of drift-diffusive and ballistic regime carrier movement in a material at different length scales as indicated.	56
3.6	Shur ballistic mobility vs. L for GaAs at various temperatures, as indicated. Adapted from [34].	58
3.7	The schematic of the linearized carrier flow based Arora model. For $\vec{E} < E_c$, the random thermal motions prevail and for $\vec{E} > E_c$, the linearized flow starts.	59
3.8	The scattering matrix representation of a nano-MOSFET. The in and out fluxes at the source S and drain D are governed by the electrostatics of the channel region.	60
3.9	Mobility extracted from montecarlo simulation (symbols) of SOI devices. Ballistic model (Dashed line) along with the μ_{eff} (Solid line) are indicated. Dotted line: Field corrected stationary mobility. Adapted from [40].	62

3.10 The model comparison between the Arora, Shur and Ballistic mobility model from Lundstrom. Adapted from [31].	63
3.11 Potential variation from source to drain for SOI devices with varying channel length at $V_{ds} = 1mV$ and $V_{gs} = 0.3V$. The dot denotes the point of kT drop (y_1) from the virtual source (y_0). The inset shows schematic of a MOSFET and its potential energy diagram from the source to drain.	67
3.12 Critical-length variation against channel length with plus (dashed) and minus (dotted) variations in oxide (square) and body (diamond) thicknesses around the nominal value (solid, circle). The scaling rates for long and short devices are different with a transition region around $0.2 - 0.4\mu m$. The inset shows a zoomed-in region, with measurement data from [45] and MC simulations from [46]. The relative invariant δ ($< 2\%$) for $L < 60$ nm is predicted by our model with ideal source/drain junctions, which is consistent with MC results for SOI 90/65/45-nm technology nodes.	69
3.13 Potential variations at $V_{ds}=1mV$, $0.1V$, $0.2V$, and $0.3V$ for $L=500nm$ SOI device. For the device, y_0 and y_1 show weak V_{ds} dependency.	69
3.14 Potential variations at $V_{ds}=1mV$, $0.1V$, $0.2V$, and $0.3V$ for $L=50nm$ SOI device. For the device, the y_0 and y_1 shift towards the source side at increasing V_{ds}	71

3.15	Potential diagrams of two $500nm$ SOI devices with different gate oxide thickness. The quasi-2D Poisson solution predicts increased QB behavior for thicker oxide due to δ being comparable to λ	71
3.16	δ contour plots for different gate oxide thicknesses. Dashed and dash-dotted lines show the boundaries between DD (right-side) and QB (left-side) transport in InGaAs and Si devices, respectively. The λ 's correspond to low gate bias.	72
3.17	δ contour plots for different silicon body thicknesses against the channel length. Dashed and dash-dotted lines show the boundaries between DD (right-side) and QB (left-side) transport in InGaAs and Si devices, respectively. The λ 's correspond to low gate bias.	73
3.18	Inverse mobility vs. inverse channel length calculated from (dashed-square) $\mu_B = v_T L / (2kT/q)$ and (dotted-square) $\mu_{B,Q2D} = v_T \delta / (2kT/q)$, and the corresponding (dashed line) μ_{eff} and (dotted line) $\mu_{eff,Q2D}$, with (solid line) $\mu_D = v_T \lambda / (2kT/q)$, $\lambda = 13nm$	73
4.1	Schematic of the empirical model used to connect the linear at saturation regimes in QB based MITVS model to simulated the DC characteristic. Adapted from [30].	78
4.2	The schematic of conduction band energy along the channel of a MOSFET at near equilibrium. Electrons are injected with velocity v_T from the S and D sides. Critical lengths of δ_s and δ_d are seen at the S and D sides, respectively.	80

4.3	Diffusive velocity (v_D) plots for varying channel lengths for λ of 15 nm and 40 nm. L_B and L_S are defined as the lengths corresponding to $v_D = v_T$ and $v_D = v_{sat}$, respectively.	83
4.4	Velocity–field characteristics modeled by (4.10) for varying channel lengths and with different $v_{sat}=2 \times 10^7$ cm/s and $v_T=1 \times 10^7$ cm/s. v_{avg} is limited by v_{sat} and v_T in DD and QB devices, respectively.	84
4.5	Mobility vs. channel length for $\lambda = 10nm$ and its impact on μ_1 studied by increasing (+) or decreasing (–) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively).	85
4.6	%change in mobility vs. channel length for $\lambda = 10nm$ and its impact on μ_1 studied by increasing (+) or decreasing (–) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively). A small transition region results in low t_{Si} and t_{ox} dependence (about 1–2%).	86
4.7	Mobility vs. channel length for $\lambda = 35nm$ and its impact on μ_1 studied by increasing (+) or decreasing (–) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively).	87
4.8	%change in mobility vs. channel length for $\lambda = 35nm$ and its impact on μ_1 studied by increasing (+) or decreasing (–) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively). The transition region widens, resulting in increased ballistic composition and t_{Si} and t_{ox} dependence (about 2–4%).	87

4.9	Mobility vs. channel length for $\lambda = 100nm$ and its impact on μ_1 studied by increasing (+) or decreasing (–) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively).	88
4.10	%change in mobility vs. channel length for $\lambda = 100nm$ and its impact on μ_1 studied by increasing (+) or decreasing (–) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively). The transition region widens, resulting in increased ballistic composition and t_{Si} and t_{ox} dependence (about 2–4%). The transition region further increases, resulting in increased ballistic composition and t_{Si} and t_{ox} dependence (about 5–10%). . .	89
4.11	G'_m vs. V_{gs} for varying channel lengths, $\lambda=30nm$ and $V_{ds}=0.5V$	89
4.12	G'_m vs. L for various λ extracted from G'_m vs. V_{gs} by fixing the extraction voltages ($V_{ds}=0.5V$, $V_{gs}=0.2V$) to ensure the device to be in current saturation. G'_m vs. L varies from unity (quadratic) to lower values (linear) whose transition region is significantly impacted by λ . With channel-length scaling, the devices with larger λ show a faster onset of ballisticity when compared to devices with smaller λ	90
4.13	$I_d - V_d$ characteristics at $V_{gs}=0V$, 0.1V, 0.2V, 0.3V, and 0.4V for an SOI device with $L=500nm$ and λ of 600nm. The device operates in QB regime and shows linear V_{gs} dependence. . .	91
4.14	$I_d - V_d$ characteristics at $V_{gs}=0V$, 0.1V, 0.2V, 0.3V, and 0.4V for an SOI device with $L=500nm$ and $\lambda = 5nm$. The device operates in DD regime and shows the typical quadratic V_{gs} dependence.	92

4.15	The extracted I_{ON} saturation current (at $V_{gs}=0.2V$ and $V_{ds}=0.5V$) for varying channel lengths and fixed $\lambda=50nm$ for conceptual devices. The model can predict the physical behavior for any source/drain/channel material combinations that could have different v_T and v_{sat} . I_{ON} in ballistic regime is determined by v_T , where as in DD, it is determined by v_{sat}	93
4.16	v_T and v_{sat} variation with temperature determined by (4.16) and (4.17), respectively.	94
4.17	v_{avg} vs. L at various temperatures, as indicated.	94
4.18	$I_{ON} - L$ plot for varying temperature. The slope in the transition region changes according to the changing v_{sat} and v_T . At QB and DD limits, the I_{ON} is still limited by v_T and v_{sat} , respectively, the magnitude of which is fixed by corresponding temperature dependent v_T , v_{sat} , and μ_1	95
4.19	Model comparison (lines) with SOI measurement data [45] (symbols) for $V_{ds} = 0.01V$ and various channel lengths, as indicated.	96
4.20	Model comparison (lines) with SOI measurement data [45] (symbols) for $V_{ds} = 1V$ and various channel lengths, as indicated.	97
4.21	Model comparison (lines) with SOI measurement data [45] (symbols) for $L = 30nm$ and various V_{ds} , as indicated.	98
4.22	Model comparison (lines) with SOI measurement data [45] (symbols) for $L = 980nm$ and various V_{ds} , as indicated.	98

4.23	$I_d - V_{ds}$ model comparison (lines) with SOI measurement data [45] (symbols) for $L = 30nm$ and various V_{gs} , as indicated. . .	99
4.24	$I_d - V_{ds}$ model comparison (lines) with SOI measurement data [45] (symbols) for $L = 980nm$ and various V_{gs} , as indicated. .	99
5.1	Typical $1/f$ noise behavior in an n channel MOSFET. Adapted from [69].	102
5.2	Typical gate bias dependence of the $1/f$ noise. Adapted from [69].	103
5.3	Typical drain bias dependence of the $1/f$ noise. Adapted from [69].	104
5.4	Traps distributed in space and in energy within the gate oxide. Adapted from [76].	106
5.5	Schematic of the carrier capture and emission process. Adapted from [77].	107
5.6	RTN associated with fixed capture and emission time con- stants. The power spectrum of such a signal is a Lorentzian, as shown. Adapted from [79].	108
5.7	Multiple traps with varying time constants of capture and emis- sion results in the LFN as illustrated. Adapted from [79]. . . .	109

5.8	The possible trapping detrapping mechanisms involved in the LFN are illustrated. Tunneling and capture mechanism is shown by (i) and (ii). Capture and tunneling mechanism is shown by (a–f). Electron transition is indicated by the arrows. Adapted from [78].	110
5.9	Tunneling involved in the capture and emission process is illustrated. The tunneling probability decreases exponentially with distance, into the oxide. Adapted from [75].	111
5.10	Typical volume element that is considered in the LFN model derivation. The noise in this volume element is integrated over volume and energy to calculate the full LFN spectrum. Adapted from [69].	112
5.11	S_{Id}/I_d^2 for a MOSFET at various gate bias. It can be noted that in subthreshold regime, effect of mobility fluctuation is negligible. Adapted from [83].	116
5.12	S_{Id}/I_d^2 vs. I_d model behavior described by (5.30) for varying Λ . The inset shows the schematic of source side of the channel with the $f(E) \times \text{DOS}$ and its dependency on V_{gs} . For low V_{gs} , the carrier injection into the channel is governed by the exponential tail of the Fermi–Dirac distribution.	118
5.13	S_{Id}/I_d^2 vs. I_d and the slope m variation for varying Λ . m varies from $m \sim 0$ to -2 depending on the C_{ox} and μ_{eff} that determine the Λ . The temperature is fixed to 300K resulting in fixed Λ	122

5.14 S_{Id}/I_d^2 vs. I_d for HfO ₂ and ZrO ₂ based devices [93]. The extracted Λ 's of 0.68×10^7 and 0.66×10^7 for HfO ₂ and ZrO ₂ , respectively, result in a constant S_{Id}/I_d^2 plateau.	123
5.15 S_{Id}/I_d^2 vs. I_d for InGaAs based devices B and C [94].	124
5.16 The blown-up of weak-inversion region for InGaAs device B , showing increased slope due to the increasing contribution from parabolic nature of the $(1 + \Lambda I_d)^2$ term.	124
5.17 The S_{Id}/I_d^2 vs. I_d for InAs NW in the subthreshold regime [88]. The dashed line with symbols fitted with the old $(g_m/I_d)^2$ model does not follow the measurement data [16] (dash-star line). The new model (solid line with symbols) with the additional factor described in (5.30) follows the measurement data.	125
5.18 The S_{Id}/I_d^2 vs. I_d for a silicon NW biochemical FET at 100K and 300K [95].	126
5.19 The $\Lambda(T)$ predicts the m to change by ~ 0.33 , as indicated.	127
5.20 S_{Id}/I_d^2 vs. I_d for PMOS and NMOS SNWTs and their comparison with models given by (5.17) and (5.35). The values of N_t (in $\text{eV}^{-1}\text{cm}^{-3}$) for PMOS and NMOS are 5.6×10^{17} and 3.7×10^{17} , respectively.	128
5.21 Measured and modeled $\langle S_{Id}/I_d \rangle$ vs. V_{ds} at $f = 10\text{Hz}$ for PMOS and NMOS SNWTs of $L = 90\text{nm}$ and $R = 40\text{nm}$ operating in the subthreshold region with $V_{gs} = -0.22\text{V}$ and $V_{gs} = 0.1\text{V}$, respectively. The respective values of N_t (in $\text{eV}^{-1}\text{cm}^{-3}$) and C_{ox} (in F/cm^2) are indicated.	129

5.22	The measured (symbols for different devices on the wafer) and modeled [dashed line, by (5.35)] drain current noise power spectral density S_{Id} for a PMOS SNWT with $L = 90nm$ and biased in the subthreshold regime at $I_d = 3.1nA$. The inset shows $I_d - V_{gs}$ characteristics of the p-type and n-type SNWTs biased at $V_{ds} = 50mV$	130
6.1	Normalized noise power for Ge nanowire MOSFETs operating in ballistic regime. Adapted from [99].	134
6.2	Model behavior and comparison of the proposed CNF/CMF model with the new λ_{app} model for two conceptual MOS devices with $\lambda = 90nm$ and $180nm$, at varying channel lengths. .	137
6.3	S_{Id}/I_d^2 vs. channel length. The model parameters used are indicated.	138
6.4	$\sqrt{S_{Id}}$ for 20nm QB device. The model parameters used are indicated.	138
6.5	$\sqrt{S_{Id}}$ for $1\mu m$ DD device. The model parameters used are indicated.	139
6.6	Measurement data and model comparison for InGaAs nanowire devices. The model parameters used are indicated.	140
6.7	$I_d - V_{gs}$ fitted to the DC model described in [4].	141
6.8	$I_d - V_{ds}$ fitted to the DC model described in [4].	141

6.9	L_{pk} vs. λ with varying V_g at $T = 300K$. For large λ devices, $L_{pk} \approx \lambda$ with weak V_g dependence. For small λ devices, L_{pk} is dependent on both V_g and λ , as shown in the inset.	142
-----	---	-----

List of Tables

2.1	Symbols Used	12
2.2	Complete ϕ_s for DG devices	16
2.3	Materials and Physical Parameters	39
3.1	Symbols Used	68
3.2	Model equations	68
4.1	Key extracted parameters	97

Chapter 1

Introduction

1.1 Motivation

Integrated circuits are found everywhere and form the basic unit in all computing devices like mobile phones, computers etc. Technology and technological demands are growing at a much faster rate to meet our improved standards. Integrated circuits are comprised of many MOS transistors and most of us individually own billions of transistors through our phone, tablets, laptops and other gadgets. There is a growing demand for circuits operating at low-power and faster speed. From the semiconductor device side, this is achieved by device-scaling and introduction of new materials with higher mobility for the carriers. Another important technology that could potentially revolutionize human life is Internet of Things (IoT) and sensors. IoTs and sensors rely on low power consuming circuitries that have the required speed and low noise behavior. They could also be based on silicon or other materials (2D materials, III-V, etc.).

To better design the circuits, the physics of scaled and new material based devices need to be well studied. Technology Computer-Aided Design (TCAD) is used to study the physics of devices. These could be based on Monte Carlo (MC) simulations or numerical solution of various transport equations coupled with quantum mechanics. Both MC and numerical device simulation approaches are time consuming and cannot be used in circuit simulators since a typical IC consists of many transistors. Circuit designers rely on analytical compact models for simulating the electrical characteristics in the circuit simulators. Many studies, recent and past, have shown the impact of quasi-ballistic transport in nano-MOSFETs. In the QB regime, scattering is a low probability event and hence the conventionally used DD equations based on well defined mobilities lose their validity. To accurately simulate a device working in ballistic regime, new physics based formalisms are needed. Most of the current simulators still rely on compact models based on drift-diffusion formalism that employ many fitting parameters to match the actual measurement results. The compact model for a particular device type can be used to fit the measurement data of another device type provided there are sufficient fitting parameters. Such models are not scalable and require tuning of the fitting parameters with varying device dimensions. To achieve a scalable model, physics-based modeling approach is used. From the device physics perspective, such empirically based models are not suitable tools for studying and improving the device characteristics.

Also new material based devices like high electron mobility transistors (HEMTs) are also gaining importance as potential candidates for devices. HEMTs are heterostructure devices that are used in high-power and high-frequency applications due to their high-mobility and low-noise characteristics. HEMTs localize the carriers in a material with higher mobility and low defect and doping densities, unlike the conventional MOSFETs. This

confinement of carriers result in quantization of energy levels that the carriers occupy. The conventional MOSFET compact model uses electrostatics and Poisson solutions to calculate the surface potential. Such models do not apply to HEMTs due to the pronounced quantum mechanical effects that the conventional model ignores. So, to get a scalable model, the surface potential solution needs to be re-derived including the QM effects. It is interesting to note that the high mobility material based devices show a faster onset of ballisticity than conventional silicon based MOSFETs. It is also interesting to note that irrespective of the material used, the fundamental transport mechanism (DD or QB) is the same for devices and it is just that the onset of transition between various transport regimes vary depending on certain material parameters. A systematic study of the DD to QB transition is missing in the literature. Another important performance metric of devices, especially sensors and IoTs, are the low-frequency noise (LFN). LFN is also used as a tool to characterize the nature of transport in various new material based devices. Since scaling affects the transport mechanisms, an impact is expected in the noise behavior also. The existing compact noise models are based on DD theory. Some empirical models are also used to study the LFN behavior.

The main motivation of this thesis is to systematically study the device scaling with a focus on the transition from DD to QB using analytical models. The developed models are also used to study the LFN behavior in devices. These models are physics-based explicit analytical mathematical solutions and can be used in compact model implementations.

1.2 Objective

The objectives of this research are to

- Study the physics of MOSFETs and III-V based HEMTs using existing models. Develop Fermi potential based compact models for III-V HEMTs with rectangular quantum well.
- Develop physics-based models to study the onset of ballistic transport.
- Develop physics based models that can be used to systematically study the DD to QB transition in devices with varying mean free paths.
- Extend the developed DD to QB models to study the low frequency noise behavior in devices in various transport regimes.
- Validate the models using TCAD and measurement data.

1.3 Major Contributions of the Thesis

The main contributions are

- (i) **Compact Fermi potential model for InGaAs/InAs/InGaAs HEMTs** It was shown that the compact model developed for InGaN/GaN HEMTs that uses a triangular potential well for carrier confinement cannot be used for InGaAs based devices where the electrons are confined in a rectangular quantum well. A new scalable physics based model which can capture the Fermi potential variation was developed. The model was extended using the short channel DD model to simulate the DC

characteristics of InGaAs based HEMTs. The model is validated using TCAD simulation results.

- (ii) **Quasi-2D surface potential based critical length model** A new critical-length based model was derived from the quasi-2D Poisson solution. It was shown that near equilibrium critical-length would be a better estimate for ballisticity than channel length. The model was validated using MC and measurement data extracted from literature.
- (iii) **Average velocity model to study DD to QB transition extensively** Mobility derived from the Quasi-2D surface potential based critical length model and a new average velocity model were proposed to extend the velocity–field characteristics for the DD based devices in the QB regime. The proposed model was used to study the effect of device geometries and material parameters on DD to QB transition. The model was validated with measurement data for SOI devices with channel lengths varying from DD to QB regimes.
- (iv) **Modified subthreshold LFN formalism** In literature, LFN is not only used to study the material quality but also to infer the nature of carrier transport. LFN is used to study carrier transport in III-V and in new material based devices. It was shown that the existing transconductance based formalism used to model LFN needed reformulation in the subthreshold regimes. This comes directly from the exponential tail of the Fermi–Dirac statistics, based on which the carriers populate the channel region in subthreshold MOSFET operating regimes. Based on this new formalism, LFN for III-V and nanowire devices were studied and validated.
- (v) **Impact of DD to QB transition on the LFN** It was also shown that the changing transport mechanism from DD to QB has a direct impact on

the LFN behavior. The existing LFN models in literature are derived based on the DD transport mechanism. These were extended to the QB regime in accordance. The scattering limited noise model in literature was also extended to the QB regime. It was shown that the LFN in QB regime is influenced greatly by scattering as opposed to the carrier number and mobility fluctuation seen in DD devices. This model was used to explain the anomalous LFN scaling trend seen in the recent measurement results.

1.4 Organization of the Thesis

- Chapter 2 briefly reviews the surface potential models used for modeling the conventional MOSFETs. An overview of triangular quantum well based HEMTs is also given. The major portion of this chapter discusses the new Fermi potential model derived for double heterostructure based InGaAs HEMTs.
- Chapter 3 briefly describes the existing mobility models in the quasi-ballistic regime. Major portion of this chapter discusses the new quasi-2D surface potential based critical length model used to study the onset of ballistic transport with channel length scaling. The model behavior is studied for various device parameters and validated with data extracted from literature.
- Chapter 4 briefly describes the existing quasi-ballistic transport models in terms of the DC characteristics. Major portion of this chapter discusses the newly proposed critical length based mobility and the average velocity model. The model behavior is extensively studied for

various device and material parameters. The proposed model is also validated with measurement data.

- Chapter 5 gives an overview of LFN in MOSFETs. It also discusses the existing models in literature used to model LFN. Major portion of this chapter discusses the new subthreshold regime noise model, which is a reformulation of the existing transconductance based noise model. The model behavior is extensively studied and validated with measurement data.
- Chapter 6 studies the impact of DD and QB transport mechanism on the LFN in MOSFETs. The existing LFN model is extended to QB regime. The model behavior is extensively studied and used to explain the anomalous LFN scaling trends.
- Chapter 7 presents the conclusion and summary. Possible extensions of the proposed models are also discussed in this chapter.

Chapter 2

Surface-Potential-Based MOSFETs and Fermi-Potential-Based HEMTs

TCAD based device simulators solve Boltzmann and hydrodynamic transport equations to simulate the electrical characteristics of a semiconductor device. These equations have many parameters and need to be solved iteratively using numerical methods to predict the electrical characteristics of a single device. Solving such complex differential equations is time consuming and hence not suitable in circuit simulators where an IC designer has to work with many devices. The circuit simulators rely on explicit analytical solutions that can be solved faster, to simulate the device behavior. Such equations or models are known as compact device models. Compact models can be charge based [1] , threshold voltage based [2] or surface potential based. Surface potential based model is advantageous over charge and threshold voltage based models, as it has fewer approximations and is also the current

trend. In this work, surface potential based compact models are developed to study the MOS field effect transistor physics and DC electrical characteristics.

2.1 Surface Potential Based Compact Models

Surface potential based models have been in development since a long time. For conventional silicon based MOSFETs, such models rely on the solution of Pao–Sah differential equation which is derived by solving the electrostatics of the MOS gate stack [3]. The solution to Pao–Sah equation gives the surface potential dependence to the gate bias of the MOSFET. Once the surface potential is known, various electrostatic dependent parameters like the depletion and inversion charge concentrations could be easily obtained. This, together with the charge sheet approximation, could be used to obtain the DC characteristics of a long channel MOSFET [3], [4].

2.1.1 Pao–Sah Solution

Fig.2.1 shows the schematic of a MOS field effect transistor. By solving Poisson equation and by assuming the Boltzmann statistics for the carrier concentration in the rectangular Gaussian box and by considering the electrostatics of the gate stack, as illustrated in Fig. 2.1, we get [4]

$$\begin{aligned} \frac{d^2\phi}{dx^2} &= -\frac{q(p - n + N_D - N_A)}{\epsilon_{Si}} \\ &= \frac{qp_0}{\epsilon_{Si}} \left[e^{(\phi=2\phi_F-V_c)/v_{th}} - e^{-(\phi-V_r)/v_{th}} + 1 - e^{-(2\phi_F+V_{cr})/v_{th}} \right] \end{aligned} \quad (2.1)$$

Figure 2.1: MOSFET schematic for the surface potential solution. Poisson equation is solved in the Gauss box marked. Adapted from [4].

where V_c , $\phi_F = -E_F/q$, and V_r are the channel potential, Fermi potential and reference potential at flatband ($V_{gr} = V_{FB}$) and near equilibrium ($V_{ds} = 0$). Solving (2.1) together with the Gauss law given as [4]

$$E_s = \frac{C_{ox}(V_{gf} - \phi_s)}{\epsilon_{Si}} \quad (2.2)$$

results in the Pao–Sah surface potential ϕ_s solution given as [3],[4]

$$V_g - V_{FB} - \phi_s = \text{sgn}(\phi_s - \phi_0) \gamma \sqrt{f_\phi(\phi_s, \phi_0, V_c, V_r)} \quad (2.3)$$

This equation, known as Pao–Sah equation, forms the basis for most of the surface potential based compact MOSFET models available in the literature.

$$f_{\phi} = e^{-\frac{(2\phi_f + V_{cr})}{\nu_{th}}} \left[\nu_{th} e^{-\frac{V_r}{\nu_{th}}} \left(e^{\frac{\phi_s}{\nu_{th}}} - e^{\frac{\phi_0}{\nu_{th}}} \right) - (\phi_s - \phi_0) \right] + \nu_{th} e^{\frac{V_r}{\nu_{th}}} \left(e^{-\frac{\phi_s}{\nu_{th}}} - e^{-\frac{\phi_0}{\nu_{th}}} \right) + (\phi_s - \phi_0) \quad (2.4)$$

γ is the *body factor* given by [4]

$$\gamma = \frac{\sqrt{2q\epsilon_s p_0}}{C_{ox}} \quad (2.5)$$

$$p_0 = n_i \exp \left[\sinh^{-1} \left(\frac{N_A - N_D}{2n_i} \right) \right] \quad (2.6)$$

$$\phi_f = \nu_{th} \sinh^{-1} \left(\frac{N_A - N_D}{2n_i} \right) \quad (2.7)$$

The solution of Pao–Sah equation gives the variation in surface potential ϕ_s with gate bias V_g . It can be noted that equation (2.3), due to its implicit relationship between ϕ_s and V_g , cannot be solved analytically. It can be solved using numerical methods like Newton-Raphson. As mentioned before, numerical methods are not preferred for compact models. Hence, to get an explicit analytic expression for ϕ_s , a piecewise regional approach is commonly used. In piecewise regional approach, the Pao–Sah equation is simplified in each regions of MOSFET operation (accumulation, depletion, and inversion). This results in simplified regional models, one for each operating regime and valid only in the respective regime. These regional models are then joined using appropriate smoothing functions to get a unified regional model (URM) [4]. A brief overview of this approach for conventional MOSFETs is given in the next section.

Table 2.1: Symbols Used

Symbol	Unit	Description
V_g	V	Gate bias
V_{FB}	V	Flat band voltage
N_A	m^{-3}	Acceptor concentration
N_D	m^{-3}	Donor concentration
n_i	m^{-3}	Intrinsic carrier concentration
ν_{th}	V	Thermal voltage
ϕ_s	V	Surface potential
C_{ox}	F/m^2	Oxide capacitance
ϵ_{si}	F/m	Silicon dielectric constant
V_c	V	Channel potential $\left(\frac{-E_i(0,y)}{q}\right)$
V_r	V	Reference potential
V_t	V	Threshold Voltage
V_{gf}	V	$V_g - V_{FB}$
δ, σ		Smoothing parameters

2.1.2 Piecewise Regional Approach

The Pao–Sah equation is reduced to simpler forms and solved for the different n-channel MOSFET operation regimes identified as follows.

2.1.2.1 Accumulation ($V_{gr} < V_{FB}$)

In the accumulation region, holes accumulate at the oxide/silicon interface. In this regime, the negative charge comprising of fixed depletion charges and electrons could be safely neglected from Pao–Sah equation. The simplified equation (2.3) when solved results in the surface potential in accumulation

region and is given as [4]

$$\phi_{cc} = V_{gf} + 2\nu_{th}W \left[\frac{\gamma}{2\sqrt{\nu_{th}}} e^{\left(\frac{-(V_{gf}-V_r)}{2\nu_{th}} \right)} \right] \quad (2.8)$$

where W is the Lambert W function.

2.1.2.2 Depletion $V_{FB} < V_{gr} < V_t$

With increasing positive gate bias, from simple electrostatics, the channel region becomes more negative or the concentration of holes from accumulation earlier, decreases. A regime is reached where the negative charges are comprised purely of the fixed depletion charges in the bulk. The Pao–Sah equation (2.3) is solved in this regime by considering only these depletion charges. The ϕ_s in the depletion mode is obtained as [4]

$$\phi_{dd} = \phi_0 + \left(-\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + (V_{gf} - \phi_0)^2} \right) \quad (2.9)$$

2.1.2.3 Inversion $V_{gr} > V_t$

Above threshold voltage V_t , inversion region is reached where the mobile inversion charges accumulate at the oxide-substrate interface. The Pao–Sah equation can again be simplified considering only the inversion charges. The surface potential in the inversion region is obtained as [4]

$$\phi_{ss} = V_{gf} - 2\nu_{th}W \left[\frac{\gamma}{2\sqrt{\nu_{th}}} e^{\left(\frac{-(V_{gf}-2\phi_f-V_c)}{2\nu_{th}} \right)} \right] \quad (2.10)$$

where W is the Lambert W function. The above regional solutions are then unified using appropriate smoothing functions to get a unified-regional-model. The smoothing functions used are [4]

$$v_f \{x; \sigma\} = 0.5 \left(x + \sqrt{x^2 + 4\sigma} \right) \quad (2.11)$$

and

$$v_r \{x; \sigma\} = 0.5 \left(x - \sqrt{x^2 + 4\sigma} \right) \quad (2.12)$$

and

$$v_{eff} \{x, x_{sat}; \delta\} = x_{sat} - 0.5 \left(x_{sat} - x - \delta - \sqrt{(x_{sat} - x - \delta)^2 + 4\delta x_{sat}} \right) \quad (2.13)$$

The unified regional pieces are defined as

$$\phi_{acc} = v_r \{V_{gf}; \sigma_a\} + 2\nu_{th} W \left[\frac{\gamma}{2\sqrt{\nu_{th}}} e^{\left(\frac{-(V_{gf} - V_r)}{2\nu_{th}} \right)} \right] \quad (2.14)$$

$$\phi_{sub} = \phi_0 + \left(-\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + v_f(V_{gf} - \phi_0; \sigma_f)} \right) \quad (2.15)$$

$$\phi_{str} = v_f \{V_{gf}; \sigma_s\} - 2\nu_{th} W \left[\frac{\gamma}{2\sqrt{\nu_{th}}} e^{\left(\frac{-(V_{gf} - 2\phi_f - V_c)}{2\nu_{th}} \right)} \right] \quad (2.16)$$

Here ϕ_{acc} , ϕ_{sub} , and ϕ_{str} are the accumulation, subthreshold and strong inversion pieces of the surface potential. σ_a , σ_f , and σ_s are the smoothing parameters. The unified regional solution is obtained as

$$\phi_s = \phi_{acc} + \phi_{ds} \quad (2.17)$$

where

$$\phi_{ds} = v_{eff} \{\phi_{sub}, \phi_{str}; \delta_\phi\} \quad (2.18)$$

Figure 2.2: The unified regional surface potential solution, along with the regional pieces, as indicated, for a s-DG FinFET with $T_{ox} = 3nm$, $T_{Si} = 50nm$, and a heavily-doped body with $N_A = 10^{18}cm^{-3}$. Adapted from [4].

is the sub-threshold and strong inversion regional solutions stitched using the smoothing parameter δ_ϕ . These equations have been extended to model double gate and fully depleted MOS devices [4]. The complete URM ϕ_s for such devices is summarized in Table 2.2. The fully depletion potential ϕ_{FD} is explained in detail in [4]. Fig. 2.2 shows the URM solution, along with the regional pieces, as indicated, for a s-DG FinFET with $T_{ox} = 3nm$, $T_{Si} = 50nm$, and a heavily-doped body with $N_A = 10^{18}cm^{-3}$ [4].

The regional equations and the unified regional solution are compared with the numerical solution. As shown, the regional solutions respectively match the regions of operation. The URM predicts the numerical simulation accurately, thus eliminating the need for iteration. URM method has

Table 2.2: Complete ϕ_s for DG devices

$$\begin{aligned}
 \phi_{acc} &= v_r \{V_{gf}; \sigma_a\} + 2\nu_{th} W \left[\frac{\gamma}{2\sqrt{\nu_{th}}} e^{\left(\frac{-(V_{gf}-V_r)}{2\nu_{th}} \right)} \right] \\
 \phi_{sub} &= \phi_0 + \left(-\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + v_f (V_{gf} - \phi_0; \sigma_f)} \right) \\
 \phi_{dep} &= v_{eff} \{ \phi_{sub}, \phi_{FD}; \delta_{dp} \} \\
 \phi_{dv} &= V_{gf} - \Gamma \sqrt{\phi_{FD}} \\
 \phi_{str} &= v_f \{V_{gf}; \sigma_s\} - 2\nu_{th} W \left[\frac{\gamma}{2\sqrt{\nu_{th}}} e^{\left(\frac{-(V_{gf}-2\phi_f-V_c)}{2\nu_{th}} \right)} \right] \\
 \phi_{ds} &= v_{eff} \{ \phi_{sub}, \phi_{str}; \delta_{\phi} \} \\
 \phi_{seff} &= \phi_{acc} + \phi_{ds}
 \end{aligned}$$

already been confirmed as a viable solution for compact modeling [4]. In the next section, the URM solution is extended for Fermi potential solutions in high electron mobility devices (HEMTs). A detailed derivation and modeling is also discussed.

2.2 Fermi-potential Model for High Electron Mobility Transistors

2.2.1 HEMT Structure Overview

Apart from the Si-based MOSFETs, III-V material based HEMTs are also gaining importance. III-V materials have high mobility and low noise characteristics and, hence, are seen as potential candidates for new devices [5], [6] to meet growing technology demands. Apart from the use of high mobility materials, the HEMTs achieve higher carrier mobility, also by confining the

carriers in an undoped region, thus, reducing the impurity scattering. The doped regions act as supplier of carriers [7]. The spatial separation between the doped and undoped regions is achieved with the heterostructure band alignment that results in a low energy notch or potential well, as shown in Fig.2.3. A similar low energy notch is formed in MOSFETs under strong inversion, as shown in Fig.2.4. But in MOSFETs, this region is at the doped Si below the SiO_2 . Thus, the carriers formed at MOSFET inversion are in constant interaction with the interface and the ionized dopants in the channel region. This results in increased scattering and mobility reduction.

Figure 2.3: GaN HEMT with triangular quantum well carrier confinement.

Several improvements have been made to the HEMT structure to enhance its performance [7]. Nitride based HEMTs (GaN), owing to the built-in polarization field, do not need a doping region to have localized carriers. Because of the pronounced quantum mechanical effects, the electrons are commonly referred to as 2D electron-gas (2DEG). Unlike the MOSFETs, HEMTs use a Schottky gate contact to control the charges in the 2DEG layer. This can affect the ON-OFF ratio due to increased gate-carrier leakage. Thus, MIS-HEMT structures (shown in Fig.2.5) that have an additional

oxide layer is used to control the 2DEG charge [8].

Figure 2.4: Triangular low-energy notch in conventional MOSFETs for carrier confinement.

It can be noted that like Si-based MOSFETs, GaN HEMT also uses a triangular notch where the electrons accumulate. The triangular potential notch is more like a triangular quantum well with quantized subbands whose effect is significant in III-V devices. This is because, unlike silicon based devices, the Fermi level lies deep within the triangular well resulting in significant quantum mechanical effects [7]. Because of these pronounced quantum mechanical effects, ϕ_s -based models, as discussed in previous section, cannot be used to accurately capture the physics of devices even though they can be used for mere fitting of the measurement data. For use in the circuit simulators, there is a growing demand for physics-based compact de-

vice models that can capture and simulate the different device operation regimes. In the past, many compact models have been developed for the Al-GaN/GaN based heterostructure devices with undoped triangular-potential-wells for electron confinement [8], [9]. Because of the pronounced quantum mechanical confinement effect, the analytical models are based on the Coupled Poisson-Schrödinger (PS) solutions [8], [10].

2.2.2 Coupled Poisson-Schrödinger Solution

2.2.2.1 Schrödinger solution

The time independent Schrödinger equation is given as

$$\frac{-\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial x^2} + U(x)\psi(x) = E\psi \quad (2.19)$$

where $U(x)$ is the potential energy and E is the total energy. The wave function (ψ) solutions of the triangular potential well ($U(x) = q\vec{F}x$, where $\vec{F}$ is the electric field) are of the form (Fang-Howard wave functions) [11]

$$\psi(x, \alpha_x) = 2\alpha_x^{3/2} x e^{-\alpha_x x} \quad (2.20)$$

where α_x is a parameter that depends on the gradient of the potential. The energy levels in the potential well are of the form [10]

$$E_n = \gamma_n \vec{F}^{2/3}. \quad (2.21)$$

Do note that the $\vec{F}$ directly modulates the QW and, hence, the subband energies. The schematic of the triangular potential well at the interface of a metal-insulator-semiconductor structure and the wave function solution is

shown in Fig.2.3. Assuming the Fermi-Dirac statistics, the total 2DEG concentration for n sub-bands, n_s , is obtained as [10]

$$n_s = D \frac{kT}{q} \sum_n \log \left(1 + e^{q/kT(E_f - E_n)} \right) \quad (2.22)$$

where D is the 2D density of states of the QW region.

2.2.2.2 Poisson solution

In HEMTs, since the E_f lies within the triangular-QW and since the electron concentration could be directly obtained from the relative position of E_f and E_n , it could be used as a better potential reference than the ϕ_s . The E_f at the gate stack and channel interface, referenced to the bottom of the conduction band, as shown in Fig.2.5, is termed as Fermi-potential [8]. Fermi-potential in III-V devices is analogous to surface potential in conventional MOSFETs since both lie at the interface of gate stack-channel region, both are directly dependent on the electrostatics of the gate stack, and both give a direct estimate of the carrier concentration. Considering 2DEG as a charge-sheet, like in MOSFETs, at the interface of the gate stack and channel simplifies the calculations. If we solve the Poisson equation, as discussed for the MOSFET in the previous section, we could obtain a solution to 2DEG given as [8]

$$n_s = \frac{C_D}{q} (V_{gc} - E_f - V_{off}) \quad (2.23)$$

where V_{gc} , E_f , and V_{off} are the gate bias, Fermi-potential, and turn-off voltage of the device referenced to the bottom of the conduction band. C_d is the equivalent gate stack capacitance till the 2DEG region. The electric field in the Schrödinger solution (equation (2.21)) can be related to the n_s through

Poisson solution (equation (2.23)). This results in explicit solutions for E_n , given as [8]

$$E_n = \gamma_n n_s^{2/3} \quad (2.24)$$

The equations (2.21)–(2.24) can only be solved numerically. Such solutions are commonly known as coupled Poisson-Schrödinger solution.

Figure 2.5: Energy band diagram of a typical MIS-HEMT with a triangular quantum well. Fermi potential and energy levels referenced to the bottom of quantum well are shown. Adapted from [8].

2.2.3 Unified Regional Triangular Potential Well Solution

Since the coupled PS solutions can only be solved numerically, the solutions are not favored for compact models. To get explicit analytical solutions, the Poisson and Schrödinger expressions can be approximated and solved to obtain regional solutions [8] similar to the ϕ_s regional ones for conventional

MOSFETs discussed earlier. For compact modeling, usually only the first two sub-bands, E_0 and E_1 , are considered. A detailed study of the URM approach for GaN based MIS-HEMT that uses triangular-QW for confinement is discussed in [8]. In this thesis, detailed derivation and URM modeling study of the InGaAs-based HEMTs, that use a double heterostructure for carrier confinement, is discussed.

2.3 Double Heterostructure InGaAs Based HEMTs

2.3.1 Rectangular Quantum Well

Figure 2.6: The schematic of InGaAs HEMT with double heterostructure resulting in a rectangular QW confinement.

For conventional MOSFETs and GaN HEMTs discussed in the previous section, carrier are confined in the low energy notch formed at the interface of a large band gap and a small band gap (SiO_2/Si or AlGaIn/GaN)

material. To further improve the carrier confinement and mobility, rectangular QW based structures have been introduced [12], [13]. Such devices, unlike the single heterojunction HEMTs, use a rectangular quantum well for confining electrons [7]. AlGaAs/GaAs/AlGaAs based double heterostructures are examples of such devices [7]. The schematic and conduction band profile of a double heterostructure HEMT is shown in Fig.2.6 and Fig.2.7, respectively [14], [15]. Note that if the InGaAs region that forms the back heterojunction in Fig. 2.6 was extended to the bulk by removing the InAlAs back barrier, a triangular quantum well could have been obtained, as shown in Fig.2.8. The advantage of double heterostructure is not only in the improved carrier confinement but also in reducing the short channel effects due to the presence of quantized subbands, whose energy offsets could be further tuned by changing the QW thickness [7]. To further improve the electron confinement and electron mobility, an additional InAs region is used [16]. The mobility can be improved by tuning the position and width of the additional QWs. Further improvement in performance can be obtained by tuning the QW structure. To further aid this, a delta doping is sometimes used as source of carriers, instead of a uniformly doped region, like in GaN HEMT. By having a delta doped region, thickness of other layers of the gate stack could be reduced to get improved capacitance [7]. Also, the use of thin layers reduces the defect density, thereby improving the device performance [7]. Such a device is analyzed in this section of the thesis. Fig.2.6 shows the schematic of such an InGaAs HEMT device. The conduction band profile along the cut line is also shown (Fig.2.7). The channel is formed by the 2DEG in the rectangular QW, as indicated in Fig. 2.7.

Figure 2.7: Energy band diagram of HEMT with a rectangular quantum well. Fermi potential is in volts and referenced to the bottom of quantum well as shown.

Figure 2.8: The InGaAs HEMT structure with single heterojunction, resulting in a triangular quantum well for confinement.

Figure 2.9: Comparison between two devices with same gate stack till the channel region but with different quantum wells. One uses a rectangular quantum well and the other with triangular quantum well.

2.3.2 Comparison with Triangular Well Model

As discussed in the previous sections, triangular QW based devices are well modeled in literature [8]. The validity of these models on rectangular QW based devices needs to be checked [14], [15]. If the physics of these devices differs, new compact models should be derived to capture the physics accurately. Two devices with same gate stack, but with different confinements, as shown in Fig.2.6 and Fig.2.8, were simulated by numerically solving coupled Poisson-Schrödinger equation. The simulations were used to study the E_f dependence on gate bias [14]. The same gate stack ensures that any difference in the E_f between the structures results from the nature of carrier population in the respective QWs. From Fig.2.9 it is seen that the E_f shows significant difference with increasing gate bias or carrier concentration. This arises only due to the difference in confinement region [14]. The triangular QW perturbation due to the electric field is given by equation 2.21. The rectangular

QW behaves differently in the presence of electric field [14], [17]. This difference directly affects the carrier concentration since the subbands and E_f are dependent on the well perturbation [14], [15]. Hence, the triangular-well compact models cannot be used to accurately capture the operation of rectangular QW devices. The next section gives a detailed model derivation and validation for rectangular QW devices [14].

2.4 Model Formulation

The conduction band profile of the device is shown in Fig. 2.7. From electrostatic analysis, that is, by considering the voltage drop, electric field, and dopant concentration in respective layers on the conduction band profile in the Fig. 2.7, it gives

$$V_{gc} = \Phi_M - \vec{F}_{ox}T_{ox} - \Phi_S - \vec{F}_{cap}T_{cap} + \Delta E_1 - \vec{F}_{1b}T_b + \vec{F}_{2b}T_b - \left(\Delta E_2 - \vec{F}_{wb}T_{wb} + \Delta E_3 - E_f \right) \quad (2.25)$$

where V_{gc} , $\vec{F}_{ox}$, $\vec{F}_{cap}$, $\vec{F}_{1b}$, $\vec{F}_{2b}$, and $\vec{F}_{wb}$ are the gate voltage relative to the channel region and electric fields, and t_{ox} , t_{cap} , t_b , and t_{wb} are the thicknesses of oxide, cap, barrier and quantum well layers, respectively. ΔE_1 , ΔE_2 , and ΔE_3 are the conduction band offsets at the interfaces of cap, barrier, and quantum-well regions, as shown in Fig. 2.7. These are formed due to the material work function differences. E_f is the Fermi level referenced to the bottom of quantum-well at the interface between the barrier and the quantum-well. This has a similar definition as the ϕ_s in silicon based devices and its unit is in volt. It should be noted that the electric field changes

the direction once the delta doping region is encountered. This is a direct consequence of the Gauss law. From Gauss law,

$$\epsilon_{ox}\vec{F}_{ox} = \epsilon_{cap}\vec{F}_{cap} = \epsilon_b\vec{F}_{1b} \quad (2.26)$$

$$\epsilon_b\vec{F}_{2b} = \epsilon_{wb}\vec{F}_{wb} = qn_s \quad (2.27)$$

where n_s is the 2DEG concentration in the confinement region. Applying Gauss law to the rectangular region in Fig 2.7, it gives

$$\vec{F}_{1b} + \vec{F}_{2b} = \frac{q\delta_d}{\epsilon_b} \quad (2.28)$$

From the above equations, we get

$$\vec{F}_{1b} = \frac{q\delta_d - qn_s}{\epsilon_b} \quad (2.29)$$

and

$$\vec{F}_{ox} = \left(\frac{q\delta_d - qn_s}{\epsilon_b} \right) \frac{\epsilon_b}{\epsilon_{ox}} \quad (2.30)$$

and

$$\vec{F}_{cap} = \left(\frac{q\delta_d - qn_s}{\epsilon_b} \right) \frac{\epsilon_b}{\epsilon_{cap}} \quad (2.31)$$

from (2.25)–(2.31)

$$\begin{aligned} V_{gc} - E_f = & \Phi_M - \Phi_S - \Delta E_1 - \Delta E_2 - \Delta E_3 - \left(\frac{q\delta_d - qn_s}{\epsilon_b} \right) \frac{\epsilon_b}{\epsilon_{ox}} t_{ox} \\ & - \left(\frac{q\delta_d - qn_s}{\epsilon_b} \right) \frac{\epsilon_b}{\epsilon_{cap}} t_{cap} - \frac{q\delta_d - qn_s}{\epsilon_b} t_b + \frac{qn_s}{\epsilon_b} t_b + \frac{qn_s}{\epsilon_{wb}}. \end{aligned} \quad (2.32)$$

This equation can be written as

$$n_s = \frac{C_d(V_{gc} - E_f - V_{off})}{q} \quad (2.33)$$

where

$$V_{off} = \Phi_M - \Phi_S + \Delta E_1 - \Delta E_2 - \Delta E_3 - \frac{qT_{ox}\delta_d}{\epsilon_{ox}} - \frac{qT_{cap}\delta_d}{\epsilon_{cap}} - \frac{qT_b\delta_d}{\epsilon_b} \quad (2.34)$$

$$C_d = \frac{1}{\frac{T_{ox}}{\epsilon_{ox}} + \frac{T_{cap}}{\epsilon_{cap}} + \frac{2T_b}{\epsilon_b} + \frac{T_{wb}}{\epsilon_{wb}}} \quad (2.35)$$

Here V_{off} is the voltage required to turn off the device and C_d is the capacitance of the gate stack till the QW region. Note that these are analogous to the gate capacitance and threshold voltage in conventional MOSFETs. Since there are two unknowns, E_f and n_s , there is no explicit solution for (2.33). Numerically, the above could be coupled with Schrödinger equation to calculate the E_f relation to V_g . To develop a compact model, three regions for the device operation can be identified as $E_f < E_0$, $E_0 < E_f < E_1$, and $E_1 < E_f$ where E_0 and E_1 are the first two energy levels in the quantum well, as shown in the Fig 2.7. From (2.22), the carrier concentration in the QW considering the first two subbands can be expressed as [10]

$$n_s = D \frac{kT}{q} \left[\ln \left(1 + e^{q(E_f - E_0)/kT} \right) + \ln \left(1 + e^{q(E_f - E_1)/kT} \right) \right] \quad (2.36)$$

2.4.0.1 Sub-threshold $E_f < E_0$

In this regime, E_f lies below E_0 . So, almost all of the carriers are populated in E_0 itself and, hence, the effect of second subband can be neglected. The Schrödinger solution (2.36) can be approximated by using a first order Taylor series expansion to give

$$n_s = D_w \frac{kT}{q} e^{\frac{q(E_f - E_0)}{kT}} \quad (2.37)$$

where D_w is the 2D density of states in the quantum well region.

Let us denote the E_f in subthreshold regime as E_{f1} . From the Poisson solution,

$$n_s = \frac{C_d(V_{gc} - E_{f1} - V_{off})}{q} \quad (2.38)$$

From (2.33) and (2.38)

$$D_w kT e^{\frac{q(E_{f1} - E_0)}{kT}} = C_d(V_{gc} - E_{f1} - V_{off}) = qn_s \quad (2.39)$$

and gives

$$\frac{qD_w \nu_{th}}{C_d} e^{\frac{q(E_{f1} - E_0)}{kT}} + E_{f1} = V_{gc} - V_{off} \quad (2.40)$$

which can be solved analytically and is given as

$$E_{f1} = V_{gc} - V_{off} - \nu_{th} W \left[\frac{qD_w e^{-E_0/\nu_{th}}}{C_d} e^{\frac{V_{gc} - V_{off}}{\nu_{th}}} \right] \quad (2.41)$$

where W is the Lambert W function.

2.4.0.2 Weak-inversion $E_0 < E_f < E_1$

In the weak inversion regime, E_f is above E_0 and below E_1 . Significant population of electrons occur in E_0 rather than E_1 . E_1 is populated according to the exponentially decaying tail of the Fermi-Dirac distribution and hence, could be neglected when compared to the electron concentration in E_0 . Let us denote the E_f in weak inversion regime as E_{f2} . The Schrödinger solution (2.36) can be approximated by considering significant population in the E_0 ,

resulting in

$$n_s = D_w \frac{kT}{q} \left[\ln \left(1 + e^{q(E_{f2}-E_0)/kT} \right) + \ln \left(1 + e^{q(E_{f2}-E_1)/kT} \right) \right] \quad (2.42)$$

$$\sim D_w \frac{kT}{q} \ln \left(e^{q(E_{f2}-E_0)/kT} \right)$$

and gives

$$n_s = D (E_{f2} - E_0) \quad (2.43)$$

Also, from the Poisson solution, n_s is given as

$$n_s = \frac{C_D}{q} (V_{gc} - E_{f2} - V_{off}) \quad (2.44)$$

From (2.43) and (2.44)

$$qD (E_{f2} - E_0) = C_d (V_{gc} - E_{f2} - V_{off}) \quad (2.45)$$

and gives

$$E_{f2} = \frac{V_{gc} - V_{off} + \frac{qD_w E_0}{C_d}}{1 + \frac{qD_w}{C_d}} \quad (2.46)$$

2.4.0.3 Strong-inversion $E_f > E_0, E_1$

In this regime, there is significant electron concentration in both the sub-bands. The Schrödinger solution (2.36) can be approximated as

$$n_s = D_w \frac{kT}{q} \sum_2 \ln \left(1 + e^{q(E_f - E_n)/kT} \right) \quad (2.47)$$

$$n_s \sim D_w (E_{f3} - E_0) + D_w (E_{f3} - E_1) \quad (2.48)$$

Combining this with the Poisson solution, $n_s = \frac{C_D}{q} (V_{gc} - E_{f3} - V_{off})$, gives

$$D_w (E_{f3} - E_0) + D_w (E_{f3} - E_1) = \frac{C_D}{q} (V_{gc} - E_{f3} - V_{off}) \quad (2.49)$$

which could be solved for E_{f3} , the strong inversion Fermi-potential solution, as

$$E_{f3} = \frac{C_D (V_{gc} - V_{off}) + qD_w(E_0 + E_1)}{2qD_w + C_D} \quad (2.50)$$

At higher gate bias, it is seen that the electrons no longer remain localized in the QW since they have high probability for electrons to spill out and occupy other regions of the gate stack. This is known as *parallel channel* formation or parasitic HEMT behavior and results in degraded electrical performance. These spilled charges should be also be included to get an accurate working model. From charge balance,

$$Q_g + Q_{pc} + Q_{2DEG} = 0 \quad (2.51)$$

where Q_g is the total gate charge, given by

$$Q_g = C_d (V_{gc} - V_{off} - E_f) \quad (2.52)$$

$$Q_{pc} = -q[D_b (E_f - E_1) + D_{cap} (E_f - E_1) + 2(D_{wb} (E_f - E_1))] \quad (2.53)$$

is the total electron charge due to parallel-channel formation and D_b , D_{wb} , D_{cap} , are the 2D density of states in the barrier, potential well barrier, and cap regions, respectively.

$$Q_{2DEG} = -q[D_w(E_f - E'_0) + D_w(E_f - E_1)] \quad (2.54)$$

is the 2DEG concentration in the quantum well region and E'_0 is the perturbed ground state potential. Electric field perturbation on a finite quantum well introduces a potential level shift [17], as given by

$$E'_0 = E_0 + \delta E_0 \quad (2.55)$$

$$\delta E_0 = -\frac{\Omega^2 m_w^* q F^2 t_w^4}{8 \hbar^2} \quad (2.56)$$

where

$$\Omega = P \left[\frac{1}{3} + \frac{\sin k_0}{k_0} + \frac{2 \cos k_0}{k_0^2} - \frac{2 \sin k_0}{k_0^3} + \frac{2}{q_0} \left[1 + \frac{2}{q_0} + \frac{2}{q_0^2} \right] \cos^2 \frac{k_0}{2} \right] \quad (2.57)$$

$\vec{F}$ is the electric field in the channel region. This field is shielded by electrons in the second quantum-well barrier region and is given as

$$\vec{F} = \frac{q D_{wb} (E_f - E_1)}{\epsilon_{wb}} \quad (2.58)$$

$$P = \left[1 + \frac{\sin k_0}{k_0} + \frac{2}{q_0} \cos^2 \frac{k_0}{2} \right]^{-1} \quad (2.59)$$

$$k_0^2 = \frac{2 m_w^* t_w^2 q}{\hbar^2} E_0 \quad (2.60)$$

$$q_0^2 = \frac{2 m_w^* t_w^2 q}{\hbar^2} (U_0 - E_0) \quad (2.61)$$

where U_0 is the barrier height of the finite potential well. Here $U_0 = \Delta E_3$, the conduction band offset. It is important to note that the behavior of the triangular and rectangular QW is very different under electric field.

Figure 2.10: Comparison between the two strong inversion regional pieces derived by including and not including the electric field effect on the QW.

From (2.52)–(2.61),

$$\begin{aligned}
 C_d (V_{gc} - V_{off} - E_{f3}) &= q \left[2D_w E_{f3} - D_w (E'_0 + E_1) \right] \\
 &\quad + q [(E_{f3} - E_1) (D_b + D_{cap} + 2D_{wb})] \\
 &= q \left[2D_w E_{f3} - D_w \left(E_0 - \frac{\Omega^2 m_w^* q (q D_{wb} (E_{f3} - E_1))^2 t_w^4}{\epsilon_{wb}^2 \hbar^2} \right) + E_1 \right] \\
 &\quad + q [(E_{f3} - E_1) (D_b + D_{cap} + 2D_{wb})]
 \end{aligned} \tag{2.62}$$

which is a quadratic equation whose solution is given as

$$E_{f3} = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \tag{2.63}$$

$$A = D_w \frac{\Omega^2 m_w^* q^3 D_{wb}^2 t_w^4}{8 \epsilon_{wb}^2 \hbar^2} \tag{2.64}$$

$$B = 2D_w + 2D_{wb} + D_b + D_{cap} + \frac{C_d}{q} - 2E_1 A \tag{2.65}$$

$$C = AE_1^2 - (D_w + 2D_{wb} + D_{cap} + D_b)E_1 - D_w E_0 - \frac{C_d}{q}(V_{gc} - V_{off}) \quad (2.66)$$

A comparison between the regional E_{f3} solution with (2.50) and with the electric field (2.63) perturbation is shown in Fig. 2.10.

2.4.1 Unified Surface Potential Model

The three regional models can be joined using smoothing function to get the URM, like in MOSFETs, as discussed in section 2.1.

2.4.1.1 Subthreshold to weak inversion

To get a compact form, $V_{go,eff}$ is introduced to prevent any incalculable results. It goes to zero when $V_{gc} - V_{off}$ is negative. For positive $V_{gc} - V_{off}$, $V_{go,eff}$ approaches $V_{gc} - V_{off}$.

$$V_{go,eff} = 0.5 \left(V_{gc} - V_{off} + \sqrt{(V_{gc} - V_{off})^2 + 4\delta_s} \right) \quad (2.67)$$

δ_s is a smoothing parameter. In the subthreshold regime,

$\nu_{th} W \left[\frac{qD_w e^{-E_0/\nu_{th}}}{C_d} e^{\frac{V_{gc}-V_{off}}{\nu_{th}}} \right] \approx 0$. So from (2.41) E_{f1} approaches $V_{gc} - V_{off}$ and is given as

$$E'_{f1} = E_{f1}|_{V_{gc}-V_{off}=V_{go,eff}} \quad (2.68)$$

To join the subthreshold piece E_{f1} with the weak inversion piece E_{f2} , $\theta(x, y; \alpha)$ given as

$$\theta(x, y; \alpha) = 0.5 \left(x + y + \alpha - \sqrt{(x - y - \alpha)^2 + 4\alpha x} \right) \quad (2.69)$$

is used as the smoothing function, where δ_{sub} is a smoothing parameter and θ is the smoothing function (2.69). The joined piece E_{fSW} , which is valid for subthreshold to weak inversion, is given as

$$E_{fSW} = \theta(E_{f2}, E'_{f1}; \delta_{sub}) \quad (2.70)$$

2.4.1.2 Weak to strong inversion

The weak inversion and strong inversion regions are joined using the similar smoothing function. It should be noted that E_{fSW} is used to join with E_{f3} to get the URM and is given as

$$E_f = \theta(E_{f3}, E_{fSW}; \delta_{str}) \quad (2.71)$$

where δ_{str} is a smoothing parameter used. Fig.2.11 shows the regional models along with the URM. Smooth transition is observed between different operating regimes. It should be noted that the unification is similar to the MOSFET regional and URM modeling that is well studied in the literature. Hence, the model accuracy from the smoothing functions should be comparable with those for the MOSFET models.

2.4.2 Model Validation and Scalability

The physical parameters and material composition used for validating the model is given in the Table 2.3. Silvaco TCAD simulator was used to get the numerical solution of the coupled Poisson Schrödinger equations. E_f is extracted from the PS solution and compared with the proposed model. As shown Fig. 2.12, the model matches the numerical solution in the subthresh-

Figure 2.11: Unified surface potential model and the regional models matches well in subthreshold, weak inversion and strong inversion regions.

Figure 2.12: Unified surface potential model and the extracted data from coupled Poisson Schrödinger simulation. The PS simulation and the model matches well in subthreshold, weak inversion and strong inversion regions.

old, weak inversion, and strong inversion regions accurately. The error between the model and the PS solution is estimated to be in the millivolt range

Figure 2.13: Absolute difference between the unified surface potential model and coupled Poisson Schrödinger extraction for InGaAs HEMT with physical parameters as given in Table 2.3.

Figure 2.14: Model (lines) and extracted Coupled Poisson Schrödinger data (symbols) for different oxide thicknesses. The proposed model matches the extracted data very well.

as shown in Fig 2.13.

Figure 2.15: Model (lines) and extracted coupled Poisson Schrödinger data (symbols) for different quantum well dimensions. The proposed model matches the extracted data very well.

Figure 2.16: Model (lines) and extracted coupled Poisson Schrödinger data (symbols) for different δ -doping. The proposed model matches the extracted data very well.

The model is also compared with the numerical data extracted for various well and oxide thicknesses and δ -doping as shown in Fig.2.14, Fig.2.15,

Table 2.3: Materials and Physical Parameters

Oxide	HfO_2
Cap	InP
Barrier	$In_{0.52}Al_{0.48}As$
WellBarrier	$In_{0.53}Ga_{0.47}As$
Well	$InAs$
δ_d	$5 \times 10^{12} cm^{-2}$
T_{ox}	2nm
T_{cap}	2nm
T_b	3nm
T_{wb}	3nm
T_w	2nm
$E0$	0.213V
$E1$	0.49V
D_{cap}	$3.316 \times 10^{17} m^{-2} V^{-1}$
D_b	$3.57 \times 10^{17} m^{-2} V^{-1}$
D_{wb}	$1.37 \times 10^{17} m^{-2} V^{-1}$
D_w	$1.08 \times 10^{17} m^{-2} V^{-1}$

and Fig.2.16, respectively. A scalable model should be able to predict the variations accurately. The model matches the numerical data accurately, thus, making it scalable. The charge calculated using the model matches the extracted charge value from the coupled PS solution and is shown in Fig.2.17.

2.4.3 Accumulation

Due to the low band gap in InAs, an accumulation region occurs at very negative voltage bias [15], [18]. The Fermi potential in the sub-threshold

Figure 2.17: Model accurately matches the charge variation with gate bias extracted from numerical simulations.

regime, including the accumulation region, is given as [15]

$$E_{f1} = V_{gc} - V_{off} + \frac{kT}{q} W \left\{ \frac{q^2 N_v T_w e^{\frac{-qE_g}{kT}}}{kTC_d} e^{\frac{-q(V_{gc}-V_{off})}{kT}} \right\} \quad (2.72)$$

where W is the Lamberts W function, E_g and N_v are the band gap and effective density of states in the valence band, respectively. V_{off} and C_d are the off voltage and effective gate capacitance, respectively.

This completes the E_f based modeling of the potential. Fig.2.18 shows the regional and unified Fermi-potential variations with gate bias. The sub-threshold regional piece, E_{f1} , saturates at higher negative bias because of the accumulation region arising due to low band gap ($0.33eV$) of the InAs channel. The unified model, E_f , matches the numerical solution obtained from the coupled PS solution accurately. The derivative of E_f (Fig. 2.15) shows no abrupt discontinuities, indicating smooth inter-operating regime transitions.

Figure 2.18: The full unified model with the accumulation region due to the low band gap InAs layer. The E_f matches the numerical solution obtained from the coupled PS solution accurately

Figure 2.19: The first derivative of E_f shows no abrupt discontinuities indicating a smooth transition between different operating regimes.

2.5 Drain Current Modeling

For the conventional MOSFETs, assuming the drift-diffusion based transport, DC characteristics are obtained by assuming the charge-sheet approxima-

tion [4]. The drain current could be expressed as

$$I_{ds} = W \mu Q_i(y) \frac{d\phi_s(y)}{dy} \quad (2.73)$$

where μ is the carrier mobility and Q_i is the inversion charge density given as

$$Q_i = -C_{ox} \left(V_{gf} - \phi_s - \gamma \sqrt{\phi_s - \phi_0} \right) \quad (2.74)$$

The full current expression could be obtained by integrating (2.73) and is given as

$$\int_0^L I_{ds} dy = \int_{\phi_s(0)}^{\phi_s(L)} W Q_i d\phi_s(y) dy \quad (2.75)$$

Making use of the symmetric charge linearization, (2.75) results in [4]

$$I_{ds0} = \bar{\beta}_0 (\bar{q}_i + \bar{A}_b v_{th}) \Delta\phi_s \quad (2.76)$$

where

$$\bar{q}_i = \left(V_{gf} - \bar{\phi}_s - \gamma \sqrt{\bar{\phi}_s - \bar{\phi}_0} \right) \quad (2.77)$$

$$\bar{A}_b = 1 + \frac{\gamma}{2\sqrt{\bar{\phi}_s - \bar{\phi}_0}} \quad (2.78)$$

$$\bar{\beta}_0 = \bar{\mu}_{eff0} C_{ox} (W/L) \quad (2.79)$$

$$\bar{\mu}_{eff0} = \frac{(\mu_{eff0,s} + \mu_{eff0,d})}{2} \quad (2.80)$$

$$\bar{\phi}_s = \frac{(\phi_{s,s} + \phi_{s,d})}{2} \quad (2.81)$$

The longitudinal-field-dependent mobility based on the piecewise velocity–field model is given as [4]

$$\mu_{eff0,c} = \frac{\mu_0}{1 + \delta_L (V_{c,eff} - V_r) / (LE_{sat})} \quad (c = s, d) \quad (2.82)$$

where $V_{c,eff}$ is an effective channel voltage with respect to the reference voltage V_r , and δ_L is a fitting parameter [4]. The longitudinal-field-dependent mobility μ_0 is given as

$$\mu_0 = \frac{\mu_1}{1 + (\mu_1/\mu_2) E_{eff}^{1/3} + (\mu_1/\mu_3) E_{eff}^\nu} \quad (2.83)$$

where μ_1 , μ_2 and μ_3 have definite physical meaning and are associated to the coulombic, 2D phonon and surface roughness scattering, respectively. ν is a fitting parameter. The effective electric field is given as [4]

$$E_{eff} = \frac{\zeta_n C_{ox}}{\epsilon_{Si}} \left(\overline{V_{gt}} + \frac{\zeta_b}{\zeta_n} \gamma \sqrt{\phi_s - \phi_0} \right) \quad (2.84)$$

where ζ_n and ζ_b are fitting parameters and $\overline{V_{gt}} = (V_{gt,s} + V_{gt,d})/2$, where

$$V_{gt,c} = \gamma \sqrt{\phi_{s,c} - \phi_0 + v_{th} e^{(\phi_{s,c} - 2\phi_F - 2V_c)/v_{th}}} - \gamma \sqrt{\phi_{s,c} - \phi_0} \quad (c = s, d) \quad (2.85)$$

The model also need to include the velocity saturation effects that are seen in DD devices. The voltage (at drain and source given as $V_{d,sat}$ and $V_{s,sat}$, respectively) at which velocity saturates is given as

$$V_{d,sat} = \frac{V_{gt,s} L E_{sat}}{V_{gt,s} + A_{b,s} L E_{sat} + 2A_{b,s} v_{th}} + V_s \quad (2.86)$$

and

$$V_{s,sat} = \frac{V_{gt,d} L E_{sat}}{V_{gt,d} + A_{b,d} L E_{sat} + 2A_{b,d} v_{th}} + V_d \quad (2.87)$$

. The saturation field E_{sat} is given as

$$E_{sat} = \frac{2v_{sat}}{\mu_0} \quad (2.88)$$

where v_{sat} is the saturation velocity. The full transition from the linear to

saturation regime is obtained using the transition function given as

$$V_{c,eff} = V_{c,sat} - \frac{1}{2} \left(V_{c,sat} - V_c + \sqrt{(V_{c,sat} - V_c)^2 + 4\delta_{sat}V_{cc',sat} + \delta_{sat}} \right) \quad (2.89)$$

$$(c = s, d; c' = d, s)$$

where

$$V_{cc',sat} = V_{c,sat} - V_{c'} \quad (c = s, d; c' = d, s) \quad (2.90)$$

Also, $\phi_{s,c}$ and $\Delta\phi_s$ are given as

$$\phi_{s,c} = \phi_{ds}(V_{c,eff}) \quad (2.91)$$

and

$$\Delta\phi_s = \phi_{s,d} - \phi_{s,s} \quad (2.92)$$

For undoped buffer like in HEMTs and other other devices like double gate, gate all around devices and SOI, the basic expressions given need to be modified slightly and is given as [4]

$$I_{ds0} = \overline{\beta_0} (\overline{q_i} + 2v_{th}) V_{ds,eff} \quad (2.93)$$

$$V_{d,sat} = \frac{V_{gt,s}LE_{sat}}{V_{gt,s} + LE_{sat} + 4v_{th}} + V_s \quad (2.94)$$

and

$$V_{s,sat} = \frac{V_{gt,d}LE_{sat}}{V_{gt,d} + LE_{sat} + 4v_{th}} + V_d. \quad (2.95)$$

Another important factor to be considered in the drain current modeling is the drain induced barrier lowering (DIBL). The DIBL is implemented using the quasi-2D surface potential lowering given as [4]

$$\delta\phi_{s,c}(y) = (V_{bi,c} + V_c - \phi_s) \frac{\sinh[(L - y)/\lambda_c]}{\sinh(L/\lambda_c)} \quad (c = s, d) \quad (2.96)$$

where

$$\lambda_c = \frac{\epsilon_{Si} X_0}{C_{ox} \eta}; \quad X_0 = \min\{X_{dm}, T_{Si}/2\} \quad (2.97)$$

and $V_{bi,c}$ denotes the built-in potential at the S/D sides. The DIBL is used to update the V_{FB} based on $\delta\phi_s$ calculated at $y = y_{min}$ and is given as [4]

$$\begin{aligned} \delta\phi_s &= \delta\phi_{s,s}(y_{min}) + \delta\phi_{s,d}(y_{min}) \\ &= \frac{(V_{bi,s} + V_{bi,d} - 2\phi_s) + \alpha_{dibl} (V_s + V_d)}{2 \cosh(y_{min}/\lambda)} \end{aligned} \quad (2.98)$$

where

$$y_{min} = \frac{L}{2} - \frac{\gamma_{sc} \lambda_c}{2} \ln \left[\frac{V_{bi} + V_d - 2\phi_f}{V_{bi} + V_s - 2\phi_f} \right] \quad (2.99)$$

with γ_{sc} and α_{dibl} being fitting parameters.

The short channel effect modeling is not complete without the channel length modulation (CLM). The CLM results in non saturating drain current. This is modeled as velocity overshoot and is expressed in terms of an effective Early voltage (V_{Aeff}), and is given as [4]

$$V_{Aeff,c} = \frac{LE_{sat} [1 + h_c (V_c - V_{c,eff})] + \delta_L (V_{c,eff} - V_r)}{\delta_L h_c (V_{c,eff} - V_r)} \quad (2.100)$$

$$h_c = \frac{\frac{\xi_c}{l^2} \left[1 + \sqrt{1 + \left(\frac{V_c - V_{c,eff}}{lE_{sat}} \right)^2} \right]}{\sqrt{1 + \left(\frac{V_c - V_{c,eff}}{lE_{sat}} \right)^2} - \frac{\xi_c}{l^2} (V_{c,eff} - V_r)} \quad (2.101)$$

where l is the characteristic length and ξ_c is a fitting parameter.

As discussed before, the E_f is analogous to ϕ_s and can be used with the conventional DD equations to simulate the device DC characteristics as [4], [8].

$$I_{ds0} = \mu_{eff} C_d \frac{W}{L} [V_g - V_{off} - \overline{E_f} + 2v_{th}] V_{dseff} \quad (2.102)$$

For HEMTs, the E_{eff} field needs modification to include the electrostatic effects multiple material layers in the gate stack. For InGaAs HEMT, E_{eff} is given as

$$E_{eff} = \zeta_n \frac{(V_g - V_{off} - \overline{E_f})}{(t_{ox} + t_{cap} + 2t_{wb} + t_b)} \quad (2.103)$$

The complete drain-current model including all the short-channel effects and series resistances (R_{sd}) is given as [4]

$$I_{ds} = \frac{\overline{g_{\nu 0}} I_{ds0}}{1 + R_{sd} I_{ds0} / V_{ds,eff}} \quad (2.104)$$

where $\overline{g_{\nu 0}}$ represents the energy-balance-based velocity overshoot factor of the non-saturating drain current given as [4]

$$\overline{g_{\nu 0}} = \frac{g_{\nu 0,s} + g_{\nu 0,d}}{2} \quad (2.105)$$

$$g_{\nu 0,c} = 1 + \left(\frac{V_c - V_{c,eff}}{V_{Aeff,c}} \right) \quad (c = s, d) \quad (2.106)$$

Figure 2.20: The model matches the $I_d - V_g$ characteristics of an InGaAs HEMT with channel length (L) and width (W) of 500nm and $1\mu\text{m}$ accurately. The drain voltages applied are 0.01V, 0.1V, 0.2V, and 0.3V.

Figure 2.21: The model matches $I_d - V_d$ characteristics of an InGaAs HEMT with channel length (L) and width (W) of 500nm and $1\mu\text{m}$ accurately. The gate voltage is varied from -0.2V to 0.5V with a step size of 0.1V . The short-channel effects like DIBL and CLM are incorporated into the E_f model as discussed.

The change in the sub-threshold slope of $I_d - V_g$ with drain bias is captured by modifying the V_{off} according to the quasi-2D surface potential shift discussed for the conventional MOSFET given by (2.96)–(2.99). Figs. 2.20 and 2.21 show a comparison of the full DC model with numerical $I_d - V_d$ and $I_d - V_g$ of an InGaAs HEMT with channel length of 500nm and width of $1\mu\text{m}$. The complete unified model incorporated with the DIBL and CLM captures the DC characteristics accurately.

2.6 Conclusion and Summary

In this chapter, a brief overview of ϕ_s -based and E_f -based models used for conventional MOSFETs and triangular QW HEMTs were discussed. In this section, it was shown that the physics of the double heterostructure devices differs from that of single heterostructure based devices. The difference is owed to the nature of carrier confinement. In double heterostructure devices, carriers are confined in a rectangular QW as opposed to a triangular QW in single heterostructures. A new compact model, based on drift-diffusion for current transport, for such devices was developed to capture the electrostatics and DC characteristics. The model was validated with numerical simulation data. Other issues like kink effect, impact ionization, and self heating have not been considered in the model. Those have been developed for triangular well devices [19]–[22] and could be extended to rectangular QW. These act as add ons to the core surface potential model developed here.

It should be noted that the model developed in this chapter is still based on drift and diffusion. But actual studies have shown significant ballistic transport effects in III-V based devices. A detailed study and modeling of

a new critical length based model to predict ballistic onset and Drift-Diffusion to Quasi-Ballistic transition is presented in chapters 4 and 5.

Declaration

Content of this chapter has been partly published in 44th European Solid State Device Research Conference DOI: 10.1109/ESSDERC.2014.6948811, and 2016 International Conference on Simulation of Semiconductor Processes and Devices DOI: 10.1109/SISPAD.2016.7605188.

Chapter 3

Critical Length Based DD and Ballistic Transition Model

In Chapter 2, surface potential and Fermi potential based compact model derivation, that forms the basis of the DC characteristics, of Si-based and III-V based field effect devices operating in the drift-diffusion regime were discussed. The practical devices could work in drift-diffusion (DD) or quasi-ballistic (QB) regime depending on the channel length and the material used. To get accurate models, the effect of channel length scaling on the carrier transport mechanism needs to be well modeled. It should be noted that DD model could still be used to fit any measurement data even though the physics is wrongly captured. This would be just a curve fitting and the model may not be of much use to study the full device behavior. This chapter discusses the impact of channel length scaling on the transport mechanism encountered in MOSFETs.

3.1 Drift-Diffusion to Quasi-Ballistic Regime

Fig 3.1 shows schematic of a MOSFET operating under bias conditions determined by the gate, source, and the drain voltages. Using the classical

Figure 3.1: The schematic of a MOSFET, the operating regime of which is determined by the source, drain and gate biases.

mechanics, motion of electrons under an electric field ($\vec{E}$) and with momentum relaxation time, τ_m , is given as

$$m^* \frac{dv}{dt} = -q\vec{E} - \frac{m^* v_d}{\tau_m} \quad (3.1)$$

where m^* and v_d are the effective mass and drift velocity of electron, respectively. At steady state, the conventional expression for the drift velocity-electric field dependency is obtained as

$$v_d = \mu \vec{E} \quad (3.2)$$

where

$$\mu = \frac{q\tau_m}{m^*} \quad (3.3)$$

is the low-electric-field mobility. The Fig. 3.2 shows the typical velocity-field characteristic for a bulk or a long channel device. As shown in the Fig. 3.2, at electric fields higher than a critical field E_c , the velocity saturates to v_{sat} which is $\sim 10^7 \text{ cm/s}$ for silicon. This is the typical bulk silicon velocity-field behavior. It should be noted that scattering plays a very important role

Figure 3.2: The *Electric – field Vs. Velocity* schematic for a material. The velocity saturates for electric fields higher than the critical field, E_c . For silicon $v_{sat} \sim 10^7 \text{ cm/s}$

in maintaining the velocity-field relationship. It should also be noted that the mobility, governed by the slope of the velocity-field relationship, is well defined only in a scattering dominant environment.

The equilibrium Fermi-Dirac could be written as [23]

$$f_0(p) = \frac{1}{1 + e^{[E_C(r,p) - E_{F0}]/kT}} \quad (3.4)$$

where E_{f0} , k , and T are the equilibrium Fermi level, Boltzmann constant and

Figure 3.3: Maxwellian distribution functions at different Fermi energies corresponding to the E_f (Solid) and a quasi-Fermi level F_n (dashed). Adapted from [23].

temperature, respectively.

$$E_C(r, p) = E_{C0}(r) + E(p) \quad (3.5)$$

is the total energy where $E_{C0}(r)$ and $E_p = p^2/2m^*$ denote the positional potential and kinetic energies of the carrier with momentum p and effective mass m^* . For non-degenerate semiconductors, the $f_0(p)$ could be approximated by a Maxwellian distribution (Fig. 3.3) given as [23]

$$f_0(r, p) = e^{[E_{F0} - E_{C0}(r)]/kT} \times e^{-p^2/2m^*kT} \quad (3.6)$$

Under an electric field, the carrier could be described using a displaced

Figure 3.4: Maxwellian distribution functions where the carrier temperature is equal to the lattice temperature (Solid), and the carrier temperature that exceeds that of the lattice (dashed). Adapted from [23].

Maxwellian distribution (Fig. 3.4) given as [23]

$$f_0(r, p) = e^{[E_F - E_{C0}(r)]/kT} \times e^{-|p - p_d|^2 / 2m^*kT} \quad (3.7)$$

where $p_d = m^*v_d$ denotes the drift motion with velocity v_d . The randomly moving carriers thermally injected into the channel region from the source and the drain contacts follow a Maxwellian distribution and under an electric field, one could see the carrier motion to follow a displaced Maxwellian distribution given above. The energy gained from the electric field is distributed (thermal energy and kinetic energy associated with drift of carriers) through scattering processes thus resulting in the Maxwellian distribution displaced v_d .

In a combined picture, one could see that velocity-field relationship

comes from directly from the motion of the carriers governed by the displaced Maxwellian distribution along the channel, moving with v_d . Thus, scattering plays an important role in carrier transport in DD regime.

Let us denote the mean free path (MFP) for momentum relaxation as λ . With scaling, when L scales down comparable to λ , intuitively one could expect less scattering. Such operating regions are known as quasi-ballistic (QB) regimes and if $L \ll \lambda$, a purely scattering free ballistic regime is reached [24]. Since scattering is less, there is no well defined mobility in this region [25]. The best way to explain the carrier transport in this region is based on scattering probabilities since at these length scales scattering becomes a probability based event [26]. It should be noted that due to the significantly larger λ of high-electron-mobility (III-V) materials and many other 2D materials, the QB effects are noticeable even at much larger length scales as opposed to silicon-based semiconductor devices [27], [28]. If $L \gg \lambda$, the carrier undergo significant scattering and mobility is dominated by diffusive mobility and could be written in terms of λ as [23]

$$\mu_D = v_T \lambda / (2kT/q) \quad (3.8)$$

where v_T is the carrier thermal velocity. Fig. 3.5 shows the schematic of transition from scattering dominated to scatteringless transport regimes. It can be seen that in the very small length scales, the carriers are limited by injection itself [23], [29]. This schematic is a very intuitive description of scaling effect on carrier transport.

In Chapter 2, DD based derivation of DC characteristics were discussed in detail. The validity of these equations is questionable in QB regime due to the absence of a well defined mobility. It should also be noted

that the current saturation in DD devices is due to velocity saturation phenomenon which is absent in QB regime where velocity injection plays the key role [26], [29]. In DD regime, velocity saturates as illustrated in Fig. 3.2 due to the emission of optical phonons, the energy of which is material dependent. In QB regime, there is no optical phonon emission and the high electric field velocity is limited by the injection mechanism itself. This could be seen from the all right-directed flux in ballistic regime, as illustrated in Fig. 3.5.

In the literature [27], [30], [31], there are many models that discuss the mobility in QB regime. The following section gives a detailed overview of the existing models. A new model for mobility, based on the quasi-2D Poisson solution is also discussed. This is one of the key contributions of this thesis [32]. This new model could be used to study the onset of ballisticsity with L scaling that the existing models fail to capture. A more detailed study of DD-to-QB transition interms of the DC characteristics is explained in detail in the next chapter, which is another important contribution of this thesis.

Figure 3.5: The schematic of drift-diffusive and ballistic regime carrier movement in a material at different length scales as indicated.

3.1.1 Shur Model

The conventional DD equations are derived from the Boltzmann transport equation using the relaxation time approximation (RTA) [23]. The energy gained by the carriers from the field is lost by scattering events. In a small device, the electron only has a finite time to interact with the electric field. When the transit time, $L/v_T \ll \tau_m$, the relaxation time approximation solution is modified to include the time spent by an electron in the channel region before getting collected by the drain contact. Since the time spent in the electric field is limited by the channel length, the modified RTA solution given as [33]

$$f_k^{(1)} = -q\bar{v}_k F \tau_k \frac{\partial f_k^{(0)}}{\partial E} \left(1 - \exp\left(\frac{-t_k}{\tau_k}\right) \right) \quad (3.9)$$

where $f_k^{(1)}$ is relaxation-time approximation solution of the Boltzmann transport equation with a correction introduced as t_k . t_k is the transit time of carriers. $\bar{k}$, $\bar{v}_k$, τ_k , F , $f_k^{(0)}$ are the wave vector, the carrier velocity, the momentum relaxation time in terms of $\bar{k}$, the electric field, and the equilibrium Boltzmann distribution function, respectively. This results in a length-dependent ballistic-mobility given as [33], [34]

$$\mu_B = \frac{2qL}{\pi m^* v_T}. \quad (3.10)$$

Fig. 3.6 shows the model behavior for GaAs based HEMT device at varying temperature [34]. A decreasing mobility is observed and is attributed to the reduced scattering due to less time spent by the electron in the channel during transit.

Figure 3.6: Shur ballistic mobility vs. L for GaAs at various temperatures, as indicated. Adapted from [34].

3.1.2 Arora Model

From the Newtonian mechanics for electron motion provided in above equation, the relaxation time transient is given as [35] [36]

$$\tau = \tau_m \left(1 - e^{-\frac{t}{\tau_m}} \right) \quad (3.11)$$

For a device of shorter L , the time t is limited by the carrier transit time. A modified mean free path λ_t is deduced from the equation as [37]

$$\lambda_t = \lambda \left(1 - e^{-\frac{L}{v_T \tau_m}} \right) \quad (3.12)$$

and interprets the scattering as a probability based event. The expression $\left(1 - e^{-\frac{L}{v_T \tau_m}} \right)$ gives the likelihood of a scattering event during the transit from

source to drain side. This could be used to extend the μ_D to QB regime as

$$\mu_{B,arora} = \mu_D \left(1 - e^{-\frac{L}{\lambda}} \right) \quad (3.13)$$

which also results in a length dependency. A more physical explanation to this was given by Arora [35], [38], who described that in the presence of an electric field, the random thermal motion of carriers is obstructed and approaches a linearized flow. He explained that beyond a critical field the velocity saturates to a value which is determined by v_T or a multiplication factor depending on the presence (or absence) of quantum emission, thus, resulting in the velocity saturation [35], [38]. Fig. 3.7 shows the random thermal motion of carriers and the transition to linearized flow for electric fields lower and higher than the critical field, E_c , respectively.

Figure 3.7: The schematic of the linearized carrier flow based Arora model. For $\vec{E} < E_c$, the random thermal motions prevail and for $\vec{E} > E_c$, the linearized flow starts.

Arora model used v_T as the ultimate limit for the velocity saturation

and the observed saturation velocity in long-channel devices is obtained by including the effect of quantum emission (optical phonons) to modulate the v_T [35], [38].

3.1.3 Landauer Transport and Scattering Theory

Figure 3.8: The scattering matrix representation of a nano-MOSFET. The in and out fluxes at the source S and drain D are governed by the electrostatics of the channel region.

Landauer theory looks at the current transport as carrier fluxes at the left and right contacts of the device under consideration. From Landauer theory, current flow between two contacts in nanoscale regime is given as [29] [23]

$$I \propto TM(f_s - f_d) \quad (3.14)$$

where T and M are transmission and number of modes (that is related to the density of states of channel region) for current flow, respectively. f_s and f_d are the Fermi distributions at source and drain contacts respectively. The scattering matrix representation of a nano-MOSFET is shown in Fig. 3.8. The source and drain fluxes are the result of carrier injection from the contacts and scattering in the channel region. The corresponding scattering

matrix is given as [23] [26] [29]

$$S = \begin{bmatrix} (1 - R) & \left[1 - (1 - R) e^{-\frac{qV_{ds}}{k_B T}} \right] \\ R & (1 - R) e^{-\frac{qV_{ds}}{k_B T}} \end{bmatrix} \quad (3.15)$$

where

$$R = \begin{cases} \frac{L}{L+\lambda} & \text{in linear regime;} \\ \frac{\delta}{\delta+\lambda} & \text{in saturation regime.} \end{cases} \quad (3.16)$$

where λ is the mean free path, L is the channel length and δ is the critical length defined as kT drop length near the beginning of channel region. This method is inspired by the Bethe thermionic emission condition [39]. It should be noted that the current is determined by the net flux coming from the S to D side and increasing drain bias suppresses the flux from the D side. Beyond the saturation point, the flux from D is completely suppressed, thus resulting in current saturation. This is completely different from the DD theory where the optical phonon emission is the responsible factor for current saturation. From Landauer transport and scattering theory, Lundstrom derived the basic transport equations for ballistic devices and is given as [30]

$$I_{DLIN} = T_{LIN} W Q_n(V_{gs}, V_{ds}) \frac{v_T}{2 (k_B T / q)} V_{ds} \quad (3.17)$$

and

$$I_{SAT} = \frac{T_{SAT}}{2 - T_{SAT}} W Q_n(V_{gs}, V_{ds}) v_T \quad (3.18)$$

where T_{LIN} and T_{SAT} are the transmission at linear and current saturation regimes, respectively. They can be obtained as $(1 - R)$ with R values corresponding to each regime as given by (3.16). From the above equations, in linear regime, an injection velocity v_{inj} given as

$$v_{inj} = v_T \frac{L}{L + \lambda} \quad (3.19)$$

Figure 3.9: Mobility extracted from montecarlo simulation (symbols) of SOI devices. Ballistic model (Dashed line) along with the μ_{eff} (Solid line) are indicated. Dotted line: Field corrected stationary mobility. Adapted from [40].

is identified as main determinant of current in QB devices. As discussed before, the definition of μ in QB regime is not clear due to the probability based scattering events. But one could define a ballistic mobility μ_B for convenience from the equations (3.17)–(3.19) as [30]

$$\mu_B = \frac{v_T L}{2k_B T / q}. \quad (3.20)$$

Lundstrom showed that Matthiessen's rule is valid in the QB regime and introduced an effective mobility defined as [30]

$$\frac{1}{\mu_{eff}} = \frac{1}{\mu_B} + \frac{1}{\mu_D} \quad (3.21)$$

Figure 3.10: The model comparison between the Arora, Shur and Ballistic mobility model from Lundstrom. Adapted from [31].

where $\mu_D = v_T \lambda 2k_B T / q$ is the diffusive mobility. It is advantageous to write in terms of μ_{eff} since it is similar to the conventional DD based equations and is useful for easy interpretation from a compact model perspective. It is noted that this is the most widely accepted theory for modeling QB devices.

Fig. 3.9 shows the mobility extracted for SOI devices with different channel lengths spanning DD and QB regimes [40]. It can be seen that the mobility decreases towards the QB regime, which could be explained by any of the above discussed length-dependent-mobility models. In the DD regimes, the mobility is well defined and determined by μ_D . The lateral and vertical bias dependence of mobility is given by equations (2.82) and (2.83) as discussed in chapter 2. A comparison between Arora, Shur, and ballistic mobility models by Lundstrom is given in Fig. 3.10 [31].

3.2 Critical-Length Model from Q2D Poisson for Calculating the Onset of Ballisticity

In previous section, an overview of existing mobility models used for studying the ballistic transport was discussed. It can be noticed that these models typically characterize the onset on ballistic effects based on the L and λ . The experimental results from Ghiboudo *et al.* [41] showed that μ_B contributed only 15% to 20% of actual mobility reduction. It was also noticed that the onset of ballisticity happens even at incomparable L and λ . This means that the L may not be the best factor to characterize the QB effects. Most of the DC models implicitly assume the source and drain regions of MOS transistor as infinite reservoirs of carriers. Natori *et al.* [42] suggested that the finite source and drain plays important role in very small dimensions and that their effect needs to be included so as to better explain the ballistic effects. He proposed $\mu_{CB} \sim (1 - R) L$ as a modification to the existing μ_B to predict the QB behavior accurately. The parameter R was used to include the *finite* source and drain effects that could appear in the QB regime [43]. This could be a possible explanation for the missing mobility reduction factors pointed out by Ghiboudo. The real transition between μ_B and μ_D with channel length scaling is still unclear. The models described in the previous sections does not give a clear picture on the onset of ballistic transport.

In this section, a new critical-length-based model is derived by solving the quasi-2D Poisson equation [32]. The critical length, δ , like in the Landauer theory, is defined as the length over which the energy of carriers drops by kT or thermal energy relative to that at the virtual source (VS). Since the quasi-2D Poisson is solved, the derived δ is scalable with the device electrostatics. δ is presented as a better estimate for ballisticity calculations.

3.2.1 Model Description

For fast simulations, compact device models are used on explicit analytical solutions. The 2D Poisson is difficult to solve. It is usually converted to a quasi-2D form [32], [44], as shown below. This could now be solved analytically to derive the potential profile. By applying Gauss's law to the Gaussian box of length dy and height X_o , where $X_o = \min\{X_{dm}, t_{Si}\}$, and including an inversion charge concentration (per unit area) $qN_i = C_{ox} (V_{gs} - V_{t0} + \Delta V_t) / \eta_x$, we have [32]

$$\epsilon_{Si} \frac{X_o}{\eta_y} \frac{dE_s}{dy} + \epsilon_{ox} \frac{(V_{gs} - V_{fb} - 2\phi_f - V)}{t_{ox}} = qN_A X_o + qN_i \quad (3.22)$$

where ΔV_t is the shift in threshold voltage due to the drain bias and is given as [44]

$$\begin{aligned} \Delta V_t = & \frac{2(V_{bi} - 2\phi_f) + [V_{ds} + (V_{bi} - 2\phi_f)](1 - e^{-L/l})}{4 \sinh^2(L/2l)} \\ & + \frac{2\sqrt{(V_{bi} - 2\phi_f)^2 + (V_{bi} - 2\phi_f)[(V_{bi} - 2\phi_f) + V_{ds}](e^{L/l} - 1)}}{4 \sinh^2(L/2l)}. \end{aligned} \quad (3.23)$$

Above equations can be rearranged as

$$\frac{d^2V}{dy^2} - Ay = C \quad (3.24)$$

where

$$A = \frac{C_{ox}\eta_y}{\epsilon_{Si}X_o} \quad (3.25)$$

and

$$C = \frac{\eta_y C_{ox}}{x_n \epsilon_{Si}} \Delta V_t - \left(1 - \frac{1}{\eta_x}\right) \left[\frac{\eta_y C_{ox}}{\epsilon_{Si} X_o} (V_{gs} - V_{fb} - 2\phi_f) - \frac{q\eta_y N_A}{\epsilon_{Si}} \right] \quad (3.26)$$

Equation (3.24) is a second-order nonhomogeneous linear differential equation. The boundary conditions for solving the equation are $V = V_{bi}|_{y=0}$ and $V = V_{bi} + V_{ds}|_{y=L}$.

Complementary solution is given as

$$y_c = C_1 e^{\sqrt{A}y} + C_2 e^{-\sqrt{A}y} \quad (3.27)$$

where

$$C_1 = V_{bi} + \frac{C}{A} - C_2 \quad (3.28)$$

and

$$C_2 = \frac{V_{bi} e^{\sqrt{A}L} - \left(V_{bi} + V_{ds} + \frac{C}{A} e^{\sqrt{A}L} - \frac{C}{A} \right)}{2 \sinh(\sqrt{A}L)} \quad (3.29)$$

Particular solution is

$$y_p = \frac{-C}{A} \quad (3.30)$$

The general solution is $y_p + y_c$ and given as [32]

$$V = -\frac{C}{A} + C_1 e^{\sqrt{A}y} + C_2 e^{-\sqrt{A}y} \quad (3.31)$$

To find the location of minimum potential or maximum energy, equate

$$\frac{dV}{dy}|_{y=y_0} = 0 \quad (3.32)$$

This gives

$$y_0 = \frac{1}{2\sqrt{A}} \ln \left(\frac{C_2}{C_1} \right) \quad (3.33)$$

Figure 3.11: Potential variation from source to drain for SOI devices with varying channel length at $V_{ds} = 1mV$ and $V_{gs} = 0.3V$. The dot denotes the point of kT drop (y_1) from the virtual source (y_0). The inset shows schematic of a MOSFET and its potential energy diagram from the source to drain.

The position along the channel, y_1 where the energy drops by kT is

$$y_1 = \delta + y_0 \quad (3.34)$$

$$V(y_1) - V(y_0) = kT/q \quad (3.35)$$

where δ is the kT -length or the critical length, given as [32]

$$\delta = \frac{1}{\sqrt{A}} \cosh^{-1} \left(\frac{1}{2} \left[2 + \frac{kT}{2\sqrt{C_1 C_2}} \right] \right) \quad (3.36)$$

Inset of Fig. 3.11 shows a schematic of the energy profile along the channel from source to drain.

Symbols and other model equations are listed in Tables. 3.1 and 3.2.

Table 3.1: Symbols Used

Symbol	Unit	Description
X_{dm}	m	Space charge region thickness
N_A	m ⁻³	Body doping concentration
V_{gs}	V	Gate–source bias
V_{ds}	V	Drain–source bias
V_{bi}	V	Built-in voltage
η_x	-	Fitting parameter for inversion charge thickness
η_y	-	Fitting parameter for average charge depth
E_s	V/m	Surface field along the channel
V	V	Channel voltage
V_{fb}	V	Flatband voltage
L	m	Channel length
C_{ox}	F/m ²	Oxide capacitance
ϕ_f	V	Bulk Fermi potential
t_{ox}	m	Oxide thickness
ϵ_{ox}	F/m	Oxide dielectric permittivity
ϵ_{Si}	F/m	Silicon dielectric permittivity
l	m	Characteristic length
t_{Si}	m	Silicon body thickness
V_{t0}	V	Linear threshold voltage
ΔV_t	V	Shift in threshold voltage due to drain bias
q	C	Electronic charge
k	JK ⁻¹	Boltzmann constant
T	K	Temperature

Table 3.2: Model equations

$$\Delta V_t = \frac{2(V_{bi}-2\phi_f) + [V_{ds} + (V_{bi}-2\phi_f)](1-e^{-L/l})}{4\sinh^2(L/2l)} + \frac{2\sqrt{(V_{bi}-2\phi_f)^2 + (V_{bi}-2\phi_f)[(V_{bi}-2\phi_f) + V_{ds}]}(e^{L/l}-1)}{4\sinh^2(L/2l)}$$

$$l = \sqrt{\frac{\epsilon_{Si}t_{ox}X_o}{\epsilon_{ox}\eta_y}}$$

3.2.2 Results and Discussion

Conceptual SOI devices with varying physical dimensions were simulated to study the model behavior. Fig. 3.11 shows the potential profiles from source to drain side of devices with different channel lengths operating at drain and gate voltages of $1mV$ and $0.3V$, respectively. It can be noted that long-channel devices show a gradual potential variation when compared to

Figure 3.12: Critical-length variation against channel length with plus (dashed) and minus (dotted) variations in oxide (square) and body (diamond) thicknesses around the nominal value (solid, circle). The scaling rates for long and short devices are different with a transition region around $0.2 - 0.4 \mu\text{m}$. The inset shows a zoomed-in region, with measurement data from [45] and MC simulations from [46]. The relative invariant δ ($< 2\%$) for $L < 60 \text{ nm}$ is predicted by our model with ideal source/drain junctions, which is consistent with MC results for SOI 90/65/45-nm technology nodes.

Figure 3.13: Potential variations at $V_{ds} = 1\text{mV}$, 0.1V , 0.2V , and 0.3V for $L = 500\text{nm}$ SOI device. For the device, y_0 and y_1 show weak V_{ds} dependency.

shorter devices that show rapid potential variation. This is due to the proximity to the source and drain junctions in short channel devices. The δ is extracted and shown in Fig. 3.12. The different δ scaling rates arise from that the difference in variation of potential from long to short devices. The critical length scales linearly for the longer channel lengths. A transition region can be noticed at around $0.2 - 0.4\mu m$. In this region, the scaling rate changes due to the more *rounded* potential profile in deep submicron devices. The proposed critical-length model includes 2D electrostatics physically in a compact form and is consistent with measurement data [45] and Monte Carlo (MC) simulations [46], as shown in the inset of Fig. 3.12. The extraction from measurements is based on simple analytical models and calculations by MC depend on specification of scattering mechanisms. The close match of the δ model with that from MC simulations (with scattering) can be a good validation of the proposed model, while MC (without scattering) as well as experimentally-extracted data tend to under-estimate the critical length, which can be attributed to the presumption of ballistic transport. This can be used as a very good tool to study the mobility and DC characteristics of devices together with the physics-based λ modeling [28].

Fig. 3.13 and Fig. 3.14 show potential variations at different drain bias for the $500nm$ and $50nm$ devices. For the $500nm$ device, y_0 and δ does not vary much with the drain bias resulting in a weak V_d dependence. Since $\delta \gg \lambda$ ($\sim 10 - 15nm$ for electrons in silicon [45]), significant scattering effects are expected, resulting in DD-dominated transport. For the $50nm$ device, with increasing V_{ds} , y_0 and the δ -drop position shift towards the source side. Strong ballistic effects are expected in this device since $\delta \sim 10nm$ is comparable with λ .

Fig. 3.15 compares two $500nm$ conceptual SOI devices with differ-

Figure 3.14: Potential variations at $V_{ds}=1mV$, $0.1V$, $0.2V$, and $0.3V$ for $L=50nm$ SOI device. For the device, the y_0 and y_1 shift towards the source side at increasing V_{ds} .

Figure 3.15: Potential diagrams of two $500nm$ SOI devices with different gate oxide thickness. The quasi-2D Poisson solution predicts increased QB behavior for thicker oxide due to δ being comparable to λ .

ent gate oxide thickness. The δ extracted are indicated in the figure. For thin-oxide device, $\delta \sim 140nm$ and hence, significant scattering and long-channel behavior is expected. The thicker-oxide device, because of the reduced gate

Figure 3.16: δ contour plots for different gate oxide thicknesses. Dashed and dash-dotted lines show the boundaries between DD (right-side) and QB (left-side) transport in InGaAs and Si devices, respectively. The λ 's correspond to low gate bias.

control, results in a rounded potential profile, like in a short-channel device. This device has $\delta \sim 57.2\text{nm}$. Since the λ and δ are comparable, the device is expected to have elevated ballistic and short-channel behavior.

Fig. 3.16 and 3.17 shows the δ contour plots for different gate oxide and body thicknesses against the channel length. The potential profile and the critical length that characterize the dominant transport mechanism is calculated from the 2D Poisson electrostatic solution as discussed. The extracted δ values are comparable to the data in literature [45]. The typical low carrier concentration mean free path for InGaAs quantum-well MOSFETs is $\sim 160\text{nm}$ [28]. The demarcation shown in the contour plots show the transition from DD to QB transport for silicon-based MOSFETs and InGaAs quantum-well MOSFETs with varying device dimensions based on two λ 's taken from [28] and [45]. These are also dependent on gate bias and indicates the DD/QB boundaries for devices with varying L , t_{ox} , t_{Si} , or X_{dm} .

Figure 3.17: δ contour plots for different silicon body thicknesses against the channel length. Dashed and dash-dotted lines show the boundaries between DD (right-side) and QB (left-side) transport in InGaAs and Si devices, respectively. The λ 's correspond to low gate bias.

Figure 3.18: Inverse mobility vs. inverse channel length calculated from (dashed-square) $\mu_B = v_T L / (2kT/q)$ and (dotted-square) $\mu_{B,Q2D} = v_T \delta / (2kT/q)$, and the corresponding (dashed line) μ_{eff} and (dotted line) $\mu_{\text{eff,Q2D}}$, with (solid line) $\mu_D = v_T \lambda / (2kT/q)$, $\lambda = 13\text{nm}$.

Fig. 3.18 shows a comparison of the effective mobility calculated from the proposed δ [$\mu_{B,Q2D} = v_T \delta / (2kT/q)$] with that based on μ_B . The difference lies in using δ or L in μ_B . As “short-channel” effects depend on device aspect ratios (t_{ox} , t_{Si} , or X_{dm} with respect to L), it is not the absolute value of L that determines the device being long or short; instead, it should be δ in comparison with λ and L .

It should also be noted that a power-law based channel potential profile proposed by Rahman and Lundstrom [47]. This was used to derive a closed-form critical length. This is an empirical model and will not scale with device geometries and biases and requires fitting to each device condition.

3.3 Conclusion

This chapter discusses the QB transport mechanism and the existing models that are used to study the QB devices. This section also presents a new model to predict the onset of ballistic transport. Critical length (or kT -length) is presented as a better estimate for QB behavior predictions. A physics-based model for the critical length is derived from the quasi-2D solution of the Poisson equation, and its dependence on key parameters (L , t_{ox} , t_{Si} , N_A , V_{ds}) has been captured. Since the model scales with the device electrostatics, it is useful to accurately depict the DD to QB transition in silicon, III-V and other material based devices with varying mean free paths. The model shows in particular that the QB effects could arise in longer channel length devices if the respective critical length and mean free path are in comparable range. Physics-based mobility and current compact models can be further developed based on the scalable critical-length model, incorporating DD to

QB transport and 2D electrostatics in nanoscale devices. In the next chapter, a detailed study of DD to QB transition and the DC modeling of DD and QB devices is discussed.

Declaration

Content of this chapter has been partly published in IEEE Electron Device Letters DOI: 10.1109/LED.2016.25787181.

Chapter 4

Drift-Diffusion to Ballistic Regime Transition

4.1 Study of DD to QB Transition

In the section 5 of chapter 2, an overview of DD based modeling was discussed. Also, in the section 1 of chapter 3, an overview of the existing ballistic models was presented. It should be noted that there are many models in literature that individually study DD and QB transport. The models developed for QB devices fail for DD devices and vice versa. As discussed in the *delta*-based modeling in chapter 3, the transition between the long-channel to the short-channel behavior is still not well studied. The MVS (MIT virtual-source) model, which is based on the Landauer transport theory, used empirical-based δ equations for long and short devices. These equations are used to derive a compact model for studying the DD to QB transition [48]. Since this model uses empirical relations, it is not physically-based and, hence, not appropriate to examine the transition accurately. Such empirical models

require refitting between different channel lengths so as to model the DC characteristics accurately.

It is well known that in long channel devices, the electron velocity saturates by emitting optical phonons. The energy gained by the electrons from the electric field (the energy associated with the drift, as explained in the displaced Maxwellian distribution discussed in chapter 3) is lost in the emission of optical phonons thereby resulting in a constant velocity termed as saturation velocity, or v_{sat} . The models based on Landauer transport (injection-limited mechanism) are not valid in the case of phonon emission limited velocity saturation. Another point to be noted is that such models may require empirical relations to join the linear and saturation regions obtained from the Landauer transport theory [30]. The best example is the MIT virtual source model which uses an empirical function F_{sat} . The complete model is given as [30]

$$I_{ds} = WQ_n(V_{gs}, V_{ds})F_{sat}(V_{ds})v_{sat} \quad (4.1)$$

where F_{sat} is given as

$$F_{sat}(V_{ds}) = \frac{(V_{ds}/V_{dsat})}{(1 + (V_{ds}/V_{dsat})^\beta)^{1/\beta}} \quad (4.2)$$

This is illustrated in Fig. 4.1. There are not many models in literature that clearly study the DD to QB transition. The model developed by Mugnaini and Iannaccone [49], [50] considers long-channel devices as being consisted of many Landauer devices, each with v_T -limited saturation. This model is similar to the work by Arora [35], [36], discussed in section 2 of chapter 3, who used v_T as the ultimate limit for the velocity saturation. This model considered saturation velocity in long-channel devices to be obtained by including the effect of quantum emission (optical phonons) that modulates v_T . These

Figure 4.1: Schematic of the empirical model used to connect the linear at saturation regimes in QB based MITVS model to simulated the DC characteristic. Adapted from [30].

models showed that velocity-saturation-like effect naturally appears in ballistic devices with v_T as the key limit irrespective of channel length.

These models may not be suitable to get a clear understanding of the transition regime encountered during the channel length scaling. For silicon, the optical phonon energy is 0.063eV. This translates to a $v_{sat} \sim 10^7$ cm/s. This value is comparable to the v_T and hence in silicon, a clear transition between the two cannot be observed. But, for InGaAs and other III-V based devices, studies have shown that there exist different v_{sat} ($\sim 0.6 - 0.8 \times 10^7$ cm/s) [51], [52] and v_T ($\sim 2.7 \times 10^7$ cm/s) [53]. The models by Mugnaini and Arora may not be suitable to analyze the velocity limiting effects. These

models also fail to study the effect of temperature on transport at different length scales. It is well studied that v_{sat} [54] decreases while v_T increases with temperature. Hence, there could be a clear difference in the transition from a longer channel length to very short channel length at various temperatures.

Some studies have shown that scattering is dominant in devices up to a length scale of 22nm and, hence, the velocity-saturation effect is still valid [55]. These results were explained using the hydrodynamic models. The extracted v_{sat} increases for nanoscale devices and forms the limit for the velocity [55]–[57]. Khandelwal *et al.* [58] proposed a model that used both v_T (as a limit for ballistic) and DD-velocity v_{DD} for DD to QB transition. This model used a Matthiessens-like rule for v_{DD} for long-channel and a charge-based critical length for v_T for short-channel devices. The model is empirical and relies on many fitting parameters for different device dimensions [58]. Dasgupta *et al.* [59] proposed a model based on the surface-potential solution to calculate accurately v_T for different length scales. This model considers DD and ballistic transports to coexist at all length scales, the validity of which is questionable from the denumerable measurement results, for long and short devices, in the literature. The transition between DD and QB regime is not clear in this model and it wrongly predicts the transport mechanism in InGaAs device as DD dominant as opposed to ballistic dominant transport seen clearly from the experimental results [28]. So, there exists a gap in literature to explicitly study the DD to QB transition illustrating the effect of the velocity limiting mechanism (v_T or v_{sat}).

In this chapter, the bulk velocity–field model is extended to ballistic regime so as to precisely study the transition encountered during the channel-length scaling. The transition region is analyzed for different device

parameters on a conceptual MOSFET. The v_T -to- v_{sat} transition is studied by still using the DD-based conventional equations. The compact model is physically-based since the quasi-2D (Q2D) Poisson solution is employed for calculating the transition characteristics. The model is validated with SOI MOSFET data with channel lengths varying from 980nm (long) to 30nm (short) without any refitting from device to device as opposed to model refittings in the literature [59].

4.2 Model Description

Figure 4.2: The schematic of conduction band energy along the channel of a MOSFET at near equilibrium. Electrons are injected with velocity v_T from the S and D sides. Critical lengths of δ_s and δ_d are seen at the S and D sides, respectively.

Fig. 4.2 shows the schematic of conduction band energy along the channel of a MOSFET at near equilibrium. From the Quasi-2D Poisson solution, the position along the channel, y_0 where the energy is maximum, is

given as [32]

$$y_0 = \frac{1}{2\sqrt{A}} \ln \left(\frac{C_2}{C_1} \right) \quad (4.3)$$

The position along the channel near the source side, y_s where the energy drops by kT , is given as

$$y_s = y_0 - \delta_s \quad (4.4)$$

Another kT drop position can be identified near the drain side as y_d and is given as

$$y_d = y_0 + \delta_d \quad (4.5)$$

At near equilibrium, $\delta_s|_{V_{ds}=0} = \delta_d|_{V_{ds}=0} = \delta$, given as [32]

$$\delta = \frac{1}{\sqrt{A}} \cosh^{-1} \left(\frac{1}{2} \left[2 + \frac{kT}{2\sqrt{C_1 C_2}} \right] \right) \quad (4.6)$$

Symbols and other model equations are defined in Tables. 3.1 and 3.2 [32], respectively.

Considering the carrier injection from both the source (S) and drain (D) in ballistic limit, Q2D-solution-based ballistic mobility (μ_{Q2D}) is defined as

$$\mu_{Q2D} = \frac{v_T (2\delta)}{2 (kT/q)} \quad (4.7)$$

μ_{Q2D} approached μ_B in literature at very small L . The transverse-field-dependent mobility is given by [4], [60]

$$\frac{1}{\mu_0} = \frac{1}{\mu_1} + \frac{1}{\mu_2} E_{eff}^{1/3} + \frac{1}{\mu_3} E_{eff}^\nu \quad (4.8)$$

where μ_1 is the near-equilibrium mobility, defined as

$$\frac{1}{\mu_1} = \frac{1}{\mu_D} + \frac{1}{\mu_{Q2D}}. \quad (4.9)$$

μ_2 and μ_3 have definite physical meaning and are associated to the 2D phonon and surface roughness scattering, respectively. We extend the piecewise velocity–field ($v - E$) model, that is well studied for conventional MOS-FETs to get the linear to quadratic transition, to ballistic devices as

$$v = \mu_0 \frac{E}{1 + \frac{E}{E_{sat}}} \quad (4.10)$$

where the saturation electric field is given as

$$E_{sat} = 2v_{avg}/\mu_0. \quad (4.11)$$

The new average velocity, v_{avg} , is defined as

$$v_{avg} = \begin{cases} v_T; (L \leq L_{min}) \\ v_{sat}; (L \geq L_{max}) \\ \frac{(v_T L_{max} - v_{sat} L_{min})}{L_{max} - L_{min}} + \frac{(v_{sat} - v_T)L}{L_{max} - L_{min}}; (L_{min} < L < L_{max}) \end{cases} \quad (4.12)$$

where $L_{min} = \min\{L_B, L_S\}$ and $L_{max} = \max\{L_B, L_S\}$. L_B is an upper bound for QB devices and L_S is a lower bound for DD devices. The transition region limits are heuristically derived by extending the Landauer theory to DD regime.

From $\mu = \mu_B \mu_D / (\mu_B + \mu_D)$ (Lundstrom *et al.* [30]) and $v = v_T v_D / (v_T + v_D)$ (Khandelwal *et al.* [58]), a general diffusive velocity can be defined as $v_D = v_T (\lambda/L)$ (shown in Fig. 4.3). The transition region is defined to be lying between $L_B = \lambda$ and $L_S = \lambda (v_T/v_{sat})$ where L_B and L_S are defined to be the lengths at which v_D becomes v_T and v_{sat} , respectively. The v_{avg} is limited to v_{sat} in DD-based devices due to phonon emission and v_T in ballistic

Figure 4.3: Diffusive velocity (v_D) plots for varying channel lengths for λ of 15 nm and 40 nm. L_B and L_S are defined as the lengths corresponding to $v_D = v_T$ and $v_D = v_{sat}$, respectively.

devices due to the limit from thermal injection velocity. For devices in the transition region between DD and QB, the v_{avg} linearly varies from v_{sat} to v_T . The underlying basis for these equations is that the shift between v_T and v_{sat} cannot be unphysical or abrupt with channel length, but should always happen within a small length scale about the λ .

Fig. 4.4 shows the velocity–field characteristics modeled by (4.11) for varying channel lengths and with different v_{sat} and v_T . It can be seen that v_{avg} is limited by v_{sat} and v_T in DD and QB devices, respectively.

The longitudinal-field-dependent mobility based on the piecewise velocity–field model (4.11) is given as [4]

$$\mu_{eff} = \frac{\mu_0}{1 + \delta_L (V_{c,eff} - V_r) / (LE_{sat})} \quad (4.13)$$

where $V_{c,eff}$ is an effective channel voltage with respect to the reference

Figure 4.4: Velocity–field characteristics modeled by (4.10) for varying channel lengths and with different $v_{sat}=2 \times 10^7$ cm/s and $v_T=1 \times 10^7$ cm/s. v_{avg} is limited by v_{sat} and v_T in DD and QB devices, respectively.

voltage V_r , and δ_L is a fitting parameter [4]. Following the unified solution from subthreshold to strong inversion [4], the DD-to-QB drain current is given as

$$I_{ds0} = \beta_0 [\bar{q}_i + \bar{A}_b(kT/q)] V_{ds,eff} \quad (4.14)$$

in which the averaged (between S and D) quantities $\beta_0 = \mu_{eff0} C_{ox}(W/L)$, $q_i = V_{gf} - \phi_s - \gamma\sqrt{\phi_s - \phi_0}$, and $A_b = 1 + \gamma/\sqrt{\phi_s - \phi_0}$ are explained in [32]. The complete drain-current model including all the short-channel effects and series resistances (R_{sd}) is given as [4]

$$I_{ds} = \frac{\bar{g}_{\nu 0} I_{ds0}}{1 + R_{sd} I_{ds0} / V_{ds,eff}} \quad (4.15)$$

where $\bar{g}_{\nu 0}$ represents the energy-balance-based velocity overshoot factor of the non-saturating drain current. Details of this formalism is given in Chapter 2.

4.3 Results and Discussion

Fig. 4.5 shows mobility vs. channel length plots for λ of 10nm and varying t_{Si} and t_{ox} . The nominal mobility vs. channel length plot for t_{ox} and t_{Si} of 3nm and 6nm, respectively, is used as a reference to study the impact of t_{Si} and t_{ox} (increase (+) or decrease (-)) on QB and DD transition. Δt_{ox} and Δt_{Si} of 25% and 17% and its impact on the % change in mobility is shown in Fig. 4.6. As seen, the mobility transition region dependence on the t_{Si} and t_{ox} is low (about 1–2%). Such a device has predominantly DD-based transport in the transition region and the impact of δ on mobility, through (4.9), is low. This explains the low dependence of μ_1 on t_{Si} and t_{ox} .

Figs. 4.7–4.8, and Figs. 4.9–4.10 show similar plots for λ of 35nm and 100nm, respectively (2–4% to 5–10% for λ of 35nm and 100nm, re-

Figure 4.5: Mobility vs. channel length for $\lambda = 10nm$ and its impact on μ_1 studied by increasing (+) or decreasing (-) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively).

Figure 4.6: %change in mobility vs. channel length for $\lambda = 10nm$ and its impact on μ_1 studied by increasing (+) or decreasing (–) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively). A small transition region results in low t_{Si} and t_{ox} dependence (about 1–2%).

spectively). It is seen that the % change in mobility dependence on t_{Si} and t_{ox} increases with λ . The device becomes more ballistic with increasing λ and the ballistic composition in the transition region also increases. This results in increasing dependence of μ_1 on t_{Si} and t_{ox} through δ . The diminishing changes at very long and very short channel lengths also tally with DD theory for reduced Q2D effect and QB theory for 2δ approaching L , respectively.

Figure 4.7: Mobility vs. channel length for $\lambda = 35nm$ and its impact on μ_1 studied by increasing (+) or decreasing (−) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively).

Figure 4.8: %change in mobility vs. channel length for $\lambda = 35nm$ and its impact on μ_1 studied by increasing (+) or decreasing (−) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively). The transition region widens, resulting in increased ballistic composition and t_{Si} and t_{ox} dependence (about 2–4%).

Figure 4.9: Mobility vs. channel length for $\lambda = 100\text{nm}$ and its impact on μ_1 studied by increasing (+) or decreasing (−) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively).

Figure 4.10: %change in mobility vs. channel length for $\lambda = 100nm$ and its impact on μ_1 studied by increasing (+) or decreasing (−) t_{Si} and t_{ox} (Δt_{ox} and Δt_{Si} of 25% and 17%, respectively). The transition region widens, resulting in increased ballistic composition and t_{Si} and t_{ox} dependence (about 2–4%). The transition region further increases, resulting in increased ballistic composition and t_{Si} and t_{ox} dependence (about 5–10%).

Figure 4.11: G'_m vs. V_{gs} for varying channel lengths, $\lambda=30nm$ and $V_{ds}=0.5V$.

Figure 4.12: G'_m vs. L for various λ extracted from G'_m vs. V_{gs} by fixing the extraction voltages ($V_{ds}=0.5\text{V}$, $V_{gs}=0.2\text{V}$) to ensure the device to be in current saturation. G'_m vs. L varies from unity (quadratic) to lower values (linear) whose transition region is significantly impacted by λ . With channel-length scaling, the devices with larger λ show a faster onset of ballisticity when compared to devices with smaller λ .

Figure 4.13: $I_d - V_d$ characteristics at $V_{gs}=0V, 0.1V, 0.2V, 0.3V$, and $0.4V$ for an SOI device with $L=500nm$ and λ of $600nm$. The device operates in QB regime and shows linear V_{gs} dependence.

Conceptual SOI devices are modeled to study the linear to quadratic, $[I_{dsat} \sim (V_{gs} - V_{th})]$ to $[I_{dsat} \sim (V_{gs} - V_{th})^2]$, transition and the effect of v_{sat} , v_T , and λ with channel length scaling. To study these effects exclusively, the series resistances and the vertical electric field degradation to mobility given by the μ_2 and μ_3 terms in (4.8) were neglected. It was also assumed that $2\delta \sim L$ (this is true for a well-constructed MOSFET). Fig. 4.11 shows the normalized $G'_m \equiv g'_m / \max(g'_m)$ vs. L plot, where $g'_m = d^2 I_d / dV_g^2$, extracted from G'_m vs. V_{gs} (Fig. 4.12) at $V_{gs} = V_{gs0} = 0.2V$, which is the gate voltage corresponding to $G'_m = 1$ for $L = 1 \mu m$. The extraction voltages ($V_{ds} > V_{gs}$) ensure current saturation. From DD to QB, the G'_m vs. L varies

from unity (quadratic) to lower values (linear). It can also be seen that the linear-to-quadratic shift is significantly impacted by λ . Consistent to the existing literature, the devices with larger λ show a faster onset of ballistics with channel-length scaling when compared to devices with smaller λ . This is further illustrated in Figs. 4.13 and 4.14 which show linear and quadratic $I_d - V_d$ behavior for a fixed channel length ($L=500\text{nm}$) but different λ (600nm showing linear and 5nm showing quadratic behaviors, respectively).

Figure 4.14: $I_d - V_d$ characteristics at $V_{gs}=0\text{V}$, 0.1V , 0.2V , 0.3V , and 0.4V for an SOI device with $L=500\text{nm}$ and $\lambda = 5\text{nm}$. The device operates in DD regime and shows the typical quadratic V_{gs} dependence.

Fig. 4.15 shows the extracted I_{ON} saturation current (at $V_{gs}=0.2\text{V}$ and $V_{ds}=0.5\text{V}$) for varying channel lengths. The model is tested for conceptual devices with various combination of v_{sat} and v_T . A λ of 50nm is assumed

Figure 4.15: The extracted I_{ON} saturation current (at $V_{gs}=0.2V$ and $V_{ds}=0.5V$) for varying channel lengths and fixed $\lambda=50nm$ for conceptual devices. The model can predict the physical behavior for any source/drain/channel material combinations that could have different v_T and v_{sat} . I_{ON} in ballistic regime is determined by v_T , where as in DD, it is determined by v_{sat} .

for these devices. To our knowledge, there is no other model in literature that can clearly study this effect. For fixed v_T and varying v_{sat} , the I_{ON} remains unaffected in ballistic regime below ($\lambda = 50nm$) showing the injection-limited transport mechanism. For fixed v_{sat} and varying v_T , the I_{ON} is limited by v_{sat} but is not clearly seen (due to the multiplication factor from μ_1) in which the v_T also affects μ_1 since both μ_D and μ_{Q2D} are dependent on v_T .

Another test for the model is the impact of temperature (T) on the I_{ON} current. It is well known that v_{sat} [54] and v_T are temperature dependent, given as

$$v_T = \sqrt{\frac{2kT}{\pi m_e^*}} \quad (4.16)$$

and

$$v_{sat} = \frac{v^*}{1 + \zeta(T/\theta)} \quad (4.17)$$

Figure 4.16: v_T and v_{sat} variation with temperature determined by (4.16) and (4.17), respectively.

Figure 4.17: v_{avg} vs. L at various temperatures, as indicated.

where $v^* = 2.4 \times 10^7 \text{ cm/s}$, $\zeta = 0.8$, and $\theta = 300 \text{ K}$ [54]. Fig. 4.16 shows v_{sat} and v_T variation with temperature with the corresponding variation of v_{avg} vs. L from (4.12), as shown in Fig 4.17. Fig. 4.18 shows $I_{ON} - L$ plot for varying temperature. The impact of T can be seen in the varying slope in

Figure 4.18: $I_{ON} - L$ plot for varying temperature. The slope in the transition region changes according to the changing v_{sat} and v_T . At QB and DD limits, the I_{ON} is still limited by v_T and v_{sat} , respectively, the magnitude of which is fixed by corresponding temperature dependent v_T , v_{sat} , and μ_1 .

the transition region, according to the changing v_{avg} (shown in Fig. 4.17). The transition region is affected due to the larger separation between v_T and v_{sat} . It can be noted that at QB and DD limits, the I_{ON} is still limited by v_T and v_{sat} , the magnitude of which is temperature dependent. It can also be noted that the μ_1 also changes due to its dependence on v_T and T . At 300K, $v_T \sim v_{sat}$ and the change in I_{ON} is only due to physical effects of L scaling. At lower T (100K), the proposed model predicts an anomalous net lowering of I_{ON} at decreasing L around 50–100nm. This can be explained by the drop in v_{avg} (between L_B and L_S) due to $v_T \sim v_{sat}/2$ in the transition region. At higher T (400K, 500K, and 600K), the change in I_{ON} around the λ is due to $v_T > v_{sat}$. This results in a net rise in I_{ON} at decreasing L with increased slope at increasing T , as shown in Fig. 4.18. To our best knowledge, there is no other model that captures these effects clearly.

The proposed model with its full features (including the critical-length-

Figure 4.19: Model comparison (lines) with SOI measurement data [45] (symbols) for $V_{ds} = 0.01V$ and various channel lengths, as indicated.

based μ_1 , μ_2 , μ_3 , and series resistances) is validated with the SOI MOSFET measurement data [45] as shown in Figs. 4.19–4.24. The λ , v_{sat} , and v_T used are 15nm, 0.7×10^7 cm/s, and 0.7×10^7 cm/s, respectively. The model is first fitted to the $I_d - V_g$ characteristics of the long-channel (980 nm) SOI device. The short-channel parameters are extracted from the $I_d - V_d$ and $I_d - V_g$ characteristics of the 30nm device. The parameters required for mobility are extracted from the low-field (at low V_g and V_d) $I_d - V_g$ characteristics. Once the parameters are extracted (key parameters are shown in Table. 4.1), they are invariant with channel length. The model predicts the current characteristics for 980nm, 180nm, 100nm, 50nm, 40nm, and 30nm devices accurately without any refitting. This can be a good validation for the model.

Figure 4.20: Model comparison (lines) with SOI measurement data [45] (symbols) for $V_{ds} = 1V$ and various channel lengths, as indicated.

Table 4.1: Key extracted parameters

Symbol	Unit	Extracted Value
λ	cm	15×10^{-7}
v_T	cm/s	10^7
v_{sat}	cm/s	10^7
η_x	-	0.5
η_y	-	0.03
μ_2	$\text{cm}^{5/3}\text{MV}^{-2/3}\text{s}^{-1}$	1.6×10^9
μ_3	$\text{cm}^{2-\nu}\text{MV}^{\nu-1}\text{s}^{-1}$	10^9
ν	-	1
α_{dibl}	-	0.3
ξ_c	-	0.03
δ_L	-	1

Figure 4.21: Model comparison (lines) with SOI measurement data [45] (symbols) for $L = 30\text{nm}$ and various V_{ds} , as indicated.

Figure 4.22: Model comparison (lines) with SOI measurement data [45] (symbols) for $L = 980\text{nm}$ and various V_{ds} , as indicated.

Figure 4.23: $I_d - V_{ds}$ model comparison (lines) with SOI measurement data [45] (symbols) for $L = 30\text{nm}$ and various V_{gs} , as indicated.

Figure 4.24: $I_d - V_{ds}$ model comparison (lines) with SOI measurement data [45] (symbols) for $L = 980\text{nm}$ and various V_{gs} , as indicated.

4.4 Conclusion

This chapter extends the bulk velocity–field characteristics to ballistic regime. The impact of the device parameter t_{Si} and t_{ox} on the transition region encountered during the channel-length scaling is systematically studied using the δ -based mobility model. The model behavior is studied for various combinations of v_T and v_{sat} . The model is in accordance with the physics of ballistic and long-channel devices whose ON current is determined by v_T and v_{sat} , respectively. The impact of λ on the linear-to-quadratic I_{ON} current is studied and found to tally with the expected physical behavior. The proposed model is also tested by varying temperatures since the v_T and v_{sat} are temperature dependent. The model is validated with SOI MOSFET measurement data. The model predicts the current characteristics of the SOI devices without any refitting from long to short channel length.

Chapter 5

Low Frequency Noise in Deep Submicron Devices

5.1 Low Frequency Noise, $1/f$ Noise

Noise is inherent in nature and can be observed in the power spectrum of a signal because of the involvement of physical phenomenon, some of which can be associated with different time constants. At low frequencies, noise in semiconductors show an inverse frequency dependence [61], [62] [63]. The low frequency noise ($1/f$) has been well studied experimentally. LFN is used as a method to quantify the quality of a particular material causing the noise [61], [64], [65]. For field effect transistors, there are many LFN models in literature [62], [66], [67], [68] [69]. The following section gives a brief overview of the well accepted LFN models.

5.1.1 Hooge model

Typical noise characteristics of a MOSFET are shown in Figs. 5.1–5.3. For a MOSFET, the normalized I_d noise power is given and S_{Id}/I_d^2 [69]. It is seen that S_{Id}/I_d^2 is inversely proportional to the carrier concentration. Also, the S_{Id} shows a $1/f$ dependence. Hooge noise model is heuristically based empirical relationship given as [70], [71]

$$\frac{S_{Id}}{I_d^2} = \frac{\alpha_H}{Nf} \quad (5.1)$$

where N and α_H are the carrier number and Hooge parameter, respectively.

Figure 5.1: Typical $1/f$ noise behavior in an n channel MOSFET. Adapted from [69].

Some of the past studies tried to associate α_H to bulk mobility fluctuation, thus giving a semi-empirical description to the Hooge model [72], [73] and is

Figure 5.2: Typical gate bias dependence of the $1/f$ noise. Adapted from [69].

commonly known as Hooge mobility fluctuation (HMF) noise. In literature, α_H is used as an estimate of material quality [74], [75]. For high quality materials α_H values are much smaller compared to materials with many defects. This is a direct observation from (5.1) since increased defect density is associated to larger noise power.

5.1.2 Charge number fluctuation (CNF) model

The charge number fluctuation (CNF) or McWorther model is a physically-based model unlike the HMF model. The CNF is based on carrier number fluctuation according to which, the noise is caused by fluctuation of carrier

due to their capture and subsequent emission by a trap center [66], [75]. In MOSFETs, traps are distributed within oxide region [76], as shown in Fig. 5.4. Let us consider the oxide traps at a fixed position relative to the energy and oxide/channel interface. Such traps could be associated with fixed carrier capture and emission time constants, τ_c and τ_e , respectively. Schematic of the trapping and detrapping process is given in Fig. 5.5 [77]. The carrier capture and emission process would directly reflect in the I_d of the device, as shown in the Fig. 5.6. The power spectrum of such a signal is given as [62]

$$S(\omega) = \frac{1}{\pi} \frac{\tau_c \tau_e}{(\tau_c + \tau_e)^2} \frac{1/T}{\omega^2 + (1/T)^2} \quad (5.2)$$

$$\frac{1}{T} = \frac{1}{\tau_c} + \frac{1}{\tau_e} \quad (5.3)$$

Figure 5.3: Typical drain bias dependence of the $1/f$ noise. Adapted from [69].

where $\omega = 2\pi f$ where f is the frequency. Such a fluctuation is commonly known as a Random Telegraphic Noise (RTN) and its power spectrum, given by (5.2), is a Lorentzian function and is shown in Fig. 5.6. Note that, in this case the $S(\omega)$ is inversely dependent on $1/f^2$ and not $1/f$. In a more practical situation, we could have traps distributed in position and energy within the oxide [76]. These traps differ and would have different capture and emission time constants and would result in a fluctuating signal. The power spectrum of such a signal could be associated to many Lorentzians'. The overall noise, which is a sum of Lorentzian, results in the $1/f$ dependency of the noise power spectrum, as shown in Fig. 5.7 [64].

Figure 5.4: Traps distributed in space and in energy within the gate oxide. Adapted from [76].

Figure 5.5: Schematic of the carrier capture and emission process. Adapted from [77].

Figure 5.6: RTN associated with fixed capture and emission time constants. The power spectrum of such a signal is a Lorentzian, as shown. Adapted from [79].

Figure 5.7: Multiple traps with varying time constants of capture and emission results in the LFN as illustrated. Adapted from [79].

5.1.3 Number fluctuation and correlated mobility fluctuation in MOSFET

Figure 5.8: The possible trapping detrapping mechanisms involved in the LFN are illustrated. Tunneling and capture mechanism is shown by (i) and (ii). Capture and tunneling mechanism is shown by (a–f). Electron transition is indicated by the arrows. Adapted from [78].

In an n channel MOSFET, at inversion, the electrons in the low energy notch region are in constant interaction with the traps distributed in the oxide region. The interface traps could act as mediators for the electrons to tunnel into the oxide traps [63]. The two possible trapping and detrapping processes are illustrated in Fig. 5.8 [78]. In the first mechanism known as elastic tunneling, the electrons tunnel to an intermediate state, and then

Figure 5.9: Tunneling involved in the capture and emission process is illustrated. The tunneling probability decreases exponentially with distance, into the oxide. Adapted from [75].

energetically relax to the final energy location governed by the Fermi level. Earlier studies were against this tunneling mechanism as the signature of the extra energy dissipation in oxide could not be supported experimentally [63]. In the second mechanism, commonly known as inelastic tunneling or capture and tunneling mechanism, electrons undergo a Shockley-Read-Hall (SRH) process and get trapped in the interface traps [63]. These trapped charges tunnel into the oxide trap, thus resulting in current fluctuation. There is no energy dissipation involved in the oxide. This theory was widely accepted until the detailed RTN measurement results on highly scaled nMOSFETs which proved the inelastic tunneling mechanism to be the process governing trapping/detrapping of charges in the oxide [79].

The time constant τ_c associated with the tunneling, illustrated in Fig. 5.8, is given as [66], [75]

$$\tau_c = \tau_0 \exp^{2\alpha_t x} \quad (5.4)$$

Figure 5.10: Typical volume element that is considered in the LFN model derivation. The noise in this volume element is integrated over volume and energy to calculate the full LFN spectrum. Adapted from [69].

where

$$\alpha_t = \sqrt{\frac{2m^*\phi_b}{\hbar^2}} \quad (5.5)$$

and y , m^* , and ϕ_b are the locations of the traps within the oxide relative to the oxide/channel, effective mass of electron, and the potential barrier seen by the electron during tunneling, respectively.

Let us consider a small volume element ΔV with an energy volume ΔE as shown in Fig. 5.10. The power spectrum of fluctuating trap density within this volume element is given as [69], [67]

$$S_{N_T \Delta V \Delta E} = \frac{\tau(x)}{1 + \omega^2 \tau^2(x)} N_T(x, E) f_t (1 - f_t) \Delta V \Delta E \quad (5.6)$$

where $N_T(x, E)$ and $f_t (1 - f_t)$ are the trap density and occupancy factor within a few kT of the Fermi level, respectively. $N_T(y, E) f_t (1 - f_t) \Delta V \Delta E$

gives the net traps occupied within the volume element.

The number fluctuation part of current noise power is given as [67]

$$S_{I_d \Delta V \Delta E} = \frac{q^2 I_d^2}{W^2 L^2 Q_n^2} S_{N_T \Delta V \Delta E} \quad (5.7)$$

Here it is assumed that fluctuation in channel charge density produces equal fluctuation in the trapped charge density. This may not be true for this oxides and thick trap distribution within gate oxide as shown in Fig. 5.4. The gate voltage fluctuation noise is given as [67]

$$S_{V_g \Delta V \Delta E} = \frac{S_{I_d \Delta V \Delta E}}{g_m^2} \quad (5.8)$$

where g_m is the transconductance. Assuming a uniform trap distribution within the oxide and integrating over the energy and volume gives the total noise power as [67]

$$S_{V_g} = \frac{q^2}{W L C_{ox}^2} \int_0^{d_m} \int_{E_v}^{E_c} \frac{\tau(x)}{1 + \omega^2 \tau^2(x)} N_T(y, E) f_t (1 - f_t) dE dx \quad (5.9)$$

where d_m , E_c , and E_v are the depth of trap distribution in oxide, conduction band, and valance band energy, respectively. Assuming a uniform distribution of traps in space and energy, (5.9) can be evaluated as [67]

$$S_{V_g} = \frac{kT q^2}{8W L C_{ox}^2 \alpha_t} \frac{N_t}{f} \quad (5.10)$$

Here N_t is the equivalent trap density evaluated around the Fermi level where the traps are actively involved in the LFN process, as illustrated by the $f_t (1 - f_t)$ in the Fig. 5.8. It should be noted that the non-uniform trap distribution results in a $1/f^\gamma$ dependence that is well studied in literature. The equation (5.10) is the full CNF model based on the McWorther carrier

number fluctuation as discussed in the previous subsection.

From the simple drift current density equation

$$I_d = WqN_i\mu\vec{E} \quad (5.11)$$

where N_i is the inversion charge density. The fluctuation in I_d is given as [67], [69]

$$\frac{\delta I_d}{I_d} = - \left[\frac{1}{\Delta N_i} \frac{\delta \Delta N_i}{\delta \Delta N_t} \pm \frac{1}{\Delta \mu} \frac{\delta \Delta \mu}{\delta \Delta N_t} \right] \quad (5.12)$$

This means that the total current noise power could depend on both the carrier number and mobility fluctuation. Different studies have dealt with mobility fluctuation as such, for example the HMF discussed in section 5.0.1.1 which was associated to bulk mobility fluctuation. Mobility fluctuation could actually be correlated to the CNF. From the vertical-field dependent mobility in (2.83), one could see that the trapping dynamics modulates actually the vertical field, thus causing the mobility to fluctuate adding to the LFN. This is termed as correlated mobility fluctuation (CMF) and, together with CNF, it is known as CNF/CMF model. Based on the Mathissens rule, mobility fluctuation in (5.12) is given as [67], [69]

$$\frac{\delta \mu}{\delta N_t} = \frac{-\alpha_s \mu^2}{W \Delta y} \quad (5.13)$$

where α_s is the Coulomb scattering parameter. The power spectrum in the volume element is given as [67], [69]

$$S_{\Delta I_d \Delta V \Delta E} = \left[\frac{I_d}{\Delta N_i} (1 + \pm \alpha_s \mu N_i) \right]^2 S_{N_T \Delta V \Delta E} \quad (5.14)$$

which, upon integration over the space and energy, results in [67], [69]

$$S_{Id} = \frac{kT I_d^2 \alpha_t}{W L f} \left(\frac{1}{N_i} + \alpha_s \mu \right)^2 N_t \quad (5.15)$$

Both HMF (5.1) and CNF/CMF (5.15) models are able to predict the LFN spectrum. Later, it was shown that HMF fails to model the weak-inversion region accurately. Reimbold *et al.* [80] and Jindal *et al.* [81] derived a charge fluctuation formalism valid from subthreshold to strong inversion and predicted a constant S_{Id}/I_d^2 spectrum with I_d . The plateau in the S_{Id}/I_d^2 spectrum was explained by neglecting the inversion charge density, which is not always true.

An easier way of interpreting the noise mechanism was developed through the transconductance (g_m) based formulations of the already existing full spectrum models and is given as [82]

$$\frac{S_{Id}}{I_d^2} = \left(\frac{g_m}{I_d} \right)^2 \left(1 + \frac{\alpha \mu_{eff} C_{ox} I_d}{g_m} \right)^2 S_{V_{fb}} \quad (5.16)$$

where $S_{V_{fb}}$ is the power spectrum of flat band voltage fluctuation. If there is no CMF, (5.16) become

$$\frac{S_{Id}}{I_d^2} = \left(\frac{g_m}{I_d} \right)^2 S_{V_{fb}} \quad (5.17)$$

The change in electrostatics due to trapping-detrapping could be incorporated as flat band voltage fluctuation since V_{fb} includes the effect of oxide trapped charges. This is the full CNF/CMF model used in literature to interpret and study the LFN measurement data for conventional and advanced MOSFETs. An important application of the LFN study is in determining the nature of carrier transport in devices based on new materials like MoS₂ and WSe₂. It should be noted that in subthreshold regime, effect of mobility fluc-

Figure 5.11: S_{I_d}/I_d^2 for a MOSFET at various gate bias. It can be noted that in subthreshold regime, effect of mobility fluctuation is negligible. Adapted from [83].

tuation is negligible, as shown in Fig. 5.11 [83]. It should also be noted that the noise associated with series resistances, given by R_{acc} in Fig. 5.11, is also negligible in subthreshold regime. In the subthreshold regime, if the S_{I_d}/I_d^2 is proportional to $(g_m/I_d)^2$ as given in (5.16) and (5.17), the noise is due to CNF, or CNF/CMF. This interprets the current flowing near the gate and channel interface to have higher chances for carrier trapping and de-trapping in the traps within the gate dielectric [66], [75]. If S_{I_d}/I_d^2 is proportional to $1/I_d$, HMF is used for noise analysis, and the transport is assumed to be bulk dominated since the origin of HMF is considered to be due to bulk mobility fluctuation [84]. Based on CNF and HMF, or a combination of both [67], [69], conclusions have been made on the nature of carrier transport in devices based on new materials operating in subthreshold and strong

inversion [85]–[87]. Similar studies have been performed on III-V nanowire devices to evaluate the surface and core contributions in the transport mechanism [88].

As discussed before, the initial physics-based formulations on LFN are derived for the linear operating regime of MOSFETs which means a linear fluctuation between the carrier number and the resulting δV_{fb} . This is not true in the subthreshold regime and will be discussed in detail in the next section where a modified $(g_m/I_d)^2$ model is proposed to accurately describe the subthreshold-LFN [89].

5.2 Noise in Subthreshold Regime

From the g_m , given as $\partial I_d / \partial V_{gs}$, the normalized noise power is calculated as [66]

$$\frac{S_{I_d}}{I_d^2} = \left(\frac{g_m}{I_d} \right)^2 S_{V_{gs}} \quad (5.18)$$

$$V_{gs} - V_{fb} - \phi_s = -\frac{Q_{sc}}{C_{ox}} \quad (5.19)$$

where ϕ_s and Q_{sc} are the surface potential and space charge (per unit area) which include both depletion and inversion charges, respectively. The flat-band voltage is defined as

$$V_{fb} = \Phi_{MS} - \frac{Q_{ox}}{C_{ox}} \quad (5.20)$$

Q_{ox} is the trapped oxide charge. The fluctuation in the trapped oxide charge is given as

$$\delta Q_{ox} = t_{ox} / (t_{ox} - d_t) \delta Q_n \quad (5.21)$$

where Q_n , d_t , and t_{ox} are the average inversion charge per unit area, trapping distance, and oxide thickness, respectively [75]. For simplicity of calculation and to focus on the effect of nature of inversion charge population on the noise, let us assume $\delta Q_{ox} \sim \delta Q_n$.

Figure 5.12: S_{I_d}/I_d^2 vs. I_d model behavior described by (5.30) for varying Λ . The inset shows the schematic of source side of the channel with the $f(E) \times \text{DOS}$ and its dependency on V_{gs} . For low V_{gs} , the carrier injection into the channel is governed by the exponential tail of the Fermi–Dirac distribution.

The Fig. 5.12 shows the schematic of a MOSFET source region with charge following the Fermi–Dirac distribution, $f(E)$, in energy with the density of states (DOS) $\sim \sqrt{E}$. This distribution is obtained as

$$N(E) \sim f(E) \times \sqrt{E} \quad (5.22)$$

A similar distribution occurs at the drain region if both source/drain (S/D) contacts are identical. At strong inversion, the average inversion charge is

linearly dependent on the gate bias, given as $Q_n \sim C_{ox} (V_{gs} - V_t)$, and results in the linearly-related fluctuation dependency of δQ_n , δQ_{ox} , and δV_{fb} on δV_{gs} .

In the subthreshold regime, as shown in the inset of Fig. 5.12, the carriers are populated in the channel by the gate bias following the exponential tail of $f(E)$ distribution along the channel y , given (for nanowire (NW) of radius R) as $Q_i(y) = -q\pi R^2 n_i e^{q(V_{gs} - V_{fb} - V(y))/kT}$ [90] with $V(y) = 0/V_{ds}$ for $y = 0/L$ at S/D ends. The average inversion charge per unit area Q_n is defined as

$$\begin{aligned} Q_n &\equiv -\frac{Q_i(L) - Q_i(0)}{L} \\ &= -Q_{n0} e^{q(V_{gs} - V_{fb})/kT} (1 - e^{-qV_{ds}/kT}) \end{aligned} \quad (5.23)$$

where Q_{n0} is $(\pi R^2/L)qn_i$. For doped-body bulk MOSFETs, a similar expression can be derived with $Q_{n0} = (W/L)C_d(kT/q)$ and an identity factor added to the exponential term, where C_d is the depletion capacitance. From (5.23),

$$\frac{\partial Q_n}{\partial V_{fb}} = -Q_{n0} \frac{-q}{kT} e^{q(V_{gs} - V_{fb})/kT} (1 - e^{-qV_{ds}/kT}) \quad (5.24)$$

which results in the fluctuation of subthreshold charge as

$$\delta Q_n = -\frac{q}{kT} Q_n \delta V_{fb}. \quad (5.25)$$

This shows that the linear fluctuation relation between the δV_{fb} and δV_{gs} is not valid in the subthreshold regime. From (5.19) and (5.25), we have

$$\frac{\partial V_{gs}}{\partial V_{fb}} = 1 + \frac{q}{kT} \frac{Q_n}{C_{ox}}. \quad (5.26)$$

From the diffusion-dominated drain current in the subthreshold regime [90], [91],

$Q_n = -I_d/(\mu_{eff} \times kT/q)$ gives

$$\frac{\partial V_{gs}}{\partial V_{fb}} = \left(1 - \frac{I_d}{\left(\frac{kT}{q}\right)^2 C_{ox} \mu_{eff}} \right). \quad (5.27)$$

After a first-order Taylor series expansion

$$\partial V_{gs} = \frac{1}{\left(1 + \frac{I_d}{\left(\frac{kT}{q}\right)^2 C_{ox} \mu_{eff}} \right)} \partial V_{fb}. \quad (5.28)$$

The power spectrum is calculated as

$$S_{V_{gs}} = \frac{1}{\left(1 + \frac{I_d}{\left(\frac{kT}{q}\right)^2 C_{ox} \mu_{eff}} \right)^2} S_{V_{fb}}. \quad (5.29)$$

From (5.18) and (5.29), we have

$$\frac{S_{I_d}}{I_d^2} = \left(\frac{g_m}{I_d} \right)^2 \times \frac{1}{(1 + \Lambda I_d)^2} S_{V_{fb}} \quad (5.30)$$

in which Λ is also dependent on temperature through the μ_{eff} and the kT/q factor and can be denoted as

$$\Lambda(T) = \frac{1}{\left[C_{ox} \mu_{eff}(T) \left(\frac{kT}{q} \right)^2 \right]}. \quad (5.31)$$

The μ_{eff} denotes the transverse-field-dependent mobility, given as [92]

$$\mu_{eff} = \frac{\mu_{1,T}}{1 + (\mu_{1,T}/\mu_{2,T}) E_{eff}^{1/3} + (\mu_{1,T}/\mu_{3,T}) E_{eff}^\nu} \quad (5.32)$$

$$\mu_{r,T} = \mu_r \left(\frac{300}{T} \right)^{\alpha_{r,T}} \quad (r = 1, 2, 3). \quad (5.33)$$

It can also be noted that the $S_{V_{fb}}$ term includes the temperature dependent $kT \times N_t$, which denotes the active traps in the oxide. The active traps are spread $\sim kT$ around the Fermi level at the oxide/substrate interface in the MOSFET. The $S_{V_{fb}}$ term is constant for a fixed temperature.

The slope m is defined as

$$m = \frac{d [\log(S_{Id}/I_d^2)]}{d [\log(I_d)]}. \quad (5.34)$$

For MOSFETs operating in nanoscale regime, the quantization effects in the channel region could also have some effect on the discussed slope change in the high gate bias regime. In subthreshold regime, the quantized subbands are closer and quantum effects on the carrier capture and emission could thus be neglected.

Depending on the Λ value, different slopes can be identified in the normalized drain-current noise for a conceptual MOS device at 300K, as shown in Fig. 5.12. If $\Lambda \times I_d \ll 1$, the S_{Id}/I_d^2 saturates as shown. In the subthreshold regime, for MOSFETs it is safe to neglect the higher-order I_d terms. From (5.29) with a first-order Taylor approximation, we have

$$\frac{S_{Id}}{I_d^2} = (g_m/I_d)^2 \times \frac{1}{(1 + 2\Lambda(T)I_d)} \times S_{V_{fb}}. \quad (5.35)$$

In the literature, this “correction factor” $[1/(1 + \Lambda(T)I_d)^2]$ in (5.30) was not considered and (5.17) was used for CNF-based noise analyses in the subthreshold regime, validity of which is questionable. Fig. 5.13 and its inset show the normalized noise power and m variation for varying Λ . It can be noted that the slope increases while moving from deep-subthreshold towards weak-inversion region. The rate of slope change varies with Λ , depending on the relative magnitude of $1/(1 + \Lambda I_d)^2$. The carrier population dependency

Figure 5.13: S_{Id}/I_d^2 vs. I_d and the slope m variation for varying Λ . m varies from $m \sim 0$ to -2 depending on the C_{ox} and μ_{eff} that determine the Λ . The temperature is fixed to 300K resulting in fixed Λ .

on V_{gs} shifts from exponential in subthreshold to linear in strong inversion. Hence, S_{Id}/I_d^2 follows the typical $(g_m/I_d)^2$ dependency and was used in the literature to study noise even in the subthreshold region.

Since, in the subthreshold regime, $I_d/g_m \sim \text{constant}$, the effect of CMF is negligible (shown in Fig.5.11) in determining the slope of the normalized current noise spectrum and can be given as

$$\frac{S_{Id}}{I_d^2} \sim \text{const} \times \frac{1}{(1 + 2\Lambda(T)I_d)} S_{V_{fb}}. \quad (5.36)$$

5.3 Results and Discussion

Fig. 5.14 shows the S_{Id}/I_d^2 measurement data for MOSFETs with Hafnium dioxide (HfO_2) and Zirconium oxide (ZrO_2) gate stack with a channel length of 40nm [93]. C_{ox} was extracted to be $2.7 \times 10^{-6} \text{ F/cm}^2$ and $2.8 \times 10^{-6} \text{ F/cm}^2$ for HfO_2 and ZrO_2 , respectively [93]. The Λ 's are estimated to be $0.68 \times 10^7 \text{ A}^{-1}$ and $0.66 \times 10^7 \text{ A}^{-1}$ for HfO_2 and ZrO_2 , respectively. Since $\Lambda \times I_d \ll 1$ for the I_d towards 10^{-8} A in the subthreshold regime, the slope is roughly constant for both the gate stacks, as illustrated in Fig. 5.12.

Figure 5.14: S_{Id}/I_d^2 vs. I_d for HfO_2 and ZrO_2 based devices [93]. The extracted Λ 's of 0.68×10^7 and 0.66×10^7 for HfO_2 and ZrO_2 , respectively, result in a constant S_{Id}/I_d^2 plateau.

Figure 5.15: S_{I_d}/I_d^2 vs. I_d for InGaAs based devices *B* and *C* [94].

Figure 5.16: The blown-up of weak-inversion region for InGaAs device *B*, showing increased slope due to the increasing contribution from parabolic nature of the $(1 + \Lambda I_d)^2$ term.

Figure 5.17: The S_{Id}/I_d^2 vs. I_d for InAs NW in the subthreshold regime [88]. The dashed line with symbols fitted with the old $(g_m/I_d)^2$ model does not follow the measurement data [16] (dash-star line). The new model (solid line with symbols) with the additional factor described in (5.30) follows the measurement data.

Fig. 5.15 shows S_{Id}/I_d^2 for the InGaAs-based devices B and C of gate stacks with effective oxide thicknesses of 1.2nm and 1.7nm, respectively [94]. It can be noted that $\Lambda_C > \Lambda_B$. This difference results in the S_{Id}/I_d^2 slope change in the subthreshold regime where device B saturates whereas device C maintains a slope $m \sim -1$ as predicted by (5.31). The blown-up region of device B in the weak-inversion region [94] is shown in Fig. 5.16. Here $m \sim -1.3$ and matches the slope variation predicted by the new model as illustrated in the inset of Fig. 5.13. This can be a good validation for the proposed model.

Figure 5.18: The S_{Id}/I_d^2 vs. I_d for a silicon NW biochemical FET at 100K and 300K [95].

Fig. 5.17 shows the S_{Id}/I_d^2 vs. I_d for InAs NW in the subthreshold regime [88]. The dashed line with symbols is fitted with the old $(g_m/I_d)^2$ model based on (5.17). The measurement data does not follow the model and is fitted with the Hooge model in the literature [88]. The solid line with symbols is the new model with the additional factor described in (5.30). It is

Figure 5.19: The $\Lambda(T)$ predicts the m to change by ~ 0.33 , as indicated.

seen that the modified $(g_m/I_d)^2$ model follows the S_{Id}/I_d^2 measurement data. The Λ extracted is $9.2 \times 10^5 \text{ A}^{-1}$ with the C_{ox} and μ_{eff} in the ranges shown in literature [88]. Hence, the $(g_m/I_d)^2$ method cannot predict conclusively the nature of carrier transport.

Fig. 5.18 shows the S_{Id}/I_d^2 against I_d for a silicon nanowire biochemical field-effect transistor at 100K and 300K [95]. It can be noted that the subthreshold S_{Id}/I_d^2 is roughly constant for 300K, while at 100K, the spectrum varies with I_d . Fig. 5.19 shows the slope variation in the subthreshold regime. The temperature dependence of Λ , given by (5.31), predicts the variation in $m \sim \mu_{eff}(T)T^2$. From the temperature dependence of μ_{eff} and literature [92], it is seen that mobility drops by a factor of ~ 3 as temperature rises from 100K to 300K, which predicts the m to change by ~ 0.33 , and is in accordance with the slope extracted from the literature [95].

Fig. 5.20 shows measured S_{Id}/I_d^2 for the PMOS and NMOS silicon

Figure 5.20: S_{Id}/I_d^2 vs. I_d for PMOS and NMOS SNWTs and their comparison with models given by (5.17) and (5.35). The values of N_t (in $\text{eV}^{-1}\text{cm}^{-3}$) for PMOS and NMOS are 5.6×10^{17} and 3.7×10^{17} , respectively.

nanowire transistor (SNWT) of core radius R of 40nm discussed in [90], [96] and the comparison with measured g_m and I_d following equations (5.17) and (5.35). Significant bulk inversion is seen only if the radius of NW devices is reduced, as discussed in [97]. It is seen that the normalized noise power follows the modified model (5.35) accurately, clearly showing the significance of the exponential V_{gs} -dependent carrier population in determining the noise power in the subthreshold regime. The $(g_m/I_d)^2$ -dependent model (5.17) in literature underestimates the noise, and under such situations, the HMF model was used to interpret the noise to be due to bulk mobility fluctuation [90]. In the new model, a fixed low-field mobility [μ_r in (5.33)] is assumed like in [90]. This is justified since the operation of the device is in the subthreshold regime where the vertical-field effects are negligible due to the low gate bias. Based on our revised CNF model result in Fig. 5.20, the measurement data for PMOS and NMOS SNWTs [90], [96] were fitted with CNF

Figure 5.21: Measured and modeled $\langle S_{Id}/I_d \rangle$ vs. V_{ds} at $f = 10\text{Hz}$ for PMOS and NMOS SNWTs of $L = 90\text{nm}$ and $R = 40\text{nm}$ operating in the subthreshold region with $V_{gs} = -0.22\text{V}$ and $V_{gs} = 0.1\text{V}$, respectively. The respective values of N_t (in $\text{eV}^{-1}\text{cm}^{-3}$) and C_{ox} (in F/cm^2) are indicated.

model as shown in Fig. 5.21 (whose $I_d - V_g$ are shown in the inset of Fig. 5.22) shows the LFN spectrum for PMOS SNWT fitted with the new model. These data were previously interpreted using HMF as discussed in [90].

Figure 5.22: The measured (symbols for different devices on the wafer) and modeled [dashed line, by (5.35)] drain current noise power spectral density S_{Id} for a PMOS SNWT with $L = 90nm$ and biased in the subthreshold regime at $I_d = 3.1nA$. The inset shows $I_d - V_{gs}$ characteristics of the p-type and n-type SNWTs biased at $V_{ds} = 50mV$.

5.4 Conclusion

This chapter presents an overview of the exiting LFN models. The noise formulation in subthreshold regime is re-derived. The presence of an extra term in the subthreshold region as a correction factor for S_{Id}/I_d^2 that significantly affects the CNF noise is shown. The present literature uses $(g_m/I_d)^2$ -based interpretations to study the nature of transport in devices based on new materials (WSe₂, MoS₂). The new formulation is validated using data extracted from the literature for InGaAs, InAs NW, NW sensor, and high-K based (HfO₂ and ZrO₂) devices. The reinterpreted data for PMOS and NMOS SNWT-based devices fitted with the new model formulation is also discussed, which supports the revised CNF model in subthreshold rather than interpreting with the HMF model [90].

The LFN models discussed in this chapter are based on the DD transport mechanism. In the next chapter, the impact of QB transport mechanism arising from the channel length scaling (discussed in detail in chapters 3 and 4) on the LFN is discussed.

Declaration

Content of this chapter has been partly published in IEEE Transactions on Electron Devices DOI: 10.1109/TED.2017.2670615.

Chapter 6

Low Frequency Noise in Drift-Diffusion to Ballistic Devices

6.1 Introduction

In Chapters 3 and 4, impact of channel length scaling in terms of the DD to QB transition were discussed in detail. The LFN modeling in terms of the physics of fluctuating charge or mobility were discussed in detail in Chapter 5. The inherent assumption involved in all the LFN models is DD dominant transport. The literature also lacked studies of LFN exclusively for ballistic devices. Even though devices operating in QB regime were studied for LFN, the models used were still based on DD formalism well established in literature. The DD based models could still be used for analyzing the LFN for individual devices. To study the collective effect of scaling and the DD to QB transition, the DD based LFN models may need to be rederived or reinterpreted. In this chapter, a new model is proposed to study the noise behavior under DD to QB transition.

In the linear regime, the CNF/CMF model could be written in terms of the inversion charge as [67], [69]

$$\frac{S_{Id}}{I_d^2} = (1/N_i)^2 (1 + \alpha_s \mu_0 N_i)^2 S_{V_{fb}} \quad (6.1)$$

and HMF model is given as $S_{Id}/I_d^2 = \alpha_H/fN_{tot}$. Both models predict $S_{Id}/I_d^2 \sim 1/L$ with L scaling.

Recent measurement results on III-V [98] and Ge [99] based devices, (shown in Fig. 6.1) show a very interesting scaling effect that cannot be explained by the conventional LFN theories. It is seen that the LFN decreases with L contrary to the $1/L$ dependence predicted by the CNF/CMF and HMF models. Some researchers tried to give a heuristic explanation to this behavior by citing the reduced scattering in QB regime as the main reason for this anomalous behavior [98], [99]. Nevertheless, this effect has not been modeled yet, and the literature lacks a proper explanation to this anomalous behavior. It should be noted that there are a few models that study noise in very small dimensions. A model based on the DOS fluctuation was proposed by Walczak for RTN [100]. The main limitation of this model is that it fails to clearly study the impact of transport mechanism on the LFN. Another model was proposed by Tresoff [101] for carbon nanotube based devices. This model does not clearly explain the scaling effects of LFN nor does it discuss the difference in scaling trends in DD and QB regimes. Recent model presented by Dasgupta *et al.* [102] derived an LFN model for QB devices from their ballistic drain current model. There are many discrepancies in the model behavior even though it is able to fit the noise measurement data. So there clearly exists a gap in literature for physics-based models to study the LFN in QB devices.

Figure 6.1: Normalized noise power for Ge nanowire MOSFETs operating in ballistic regime. Adapted from [99].

In this chapter, a new apparent MFP-based model is proposed to capture the LFN in QB regime. The proposed model tallies with the QB regime device physics and predicts the anomalous scaling trend accurately. The model is also validated for III-V based QB device LFN data [98].

6.2 Model Formulation

In QB devices, as discussed in Chapters 3 and 4, $\mu_{app} = \mu_B \mu_D / (\mu_B + \mu_D)$ extends a DD-based I_{ds} model to QB regime. Keeping the same form, the CNF/CMF models can be extended into the ballistic regime by replacing μ_0

with μ_{app} and is given as

$$\frac{S_{Id}}{I_d^2} = \left(\frac{1}{N_i} \right)^2 (1 + \alpha \mu_{app} N_i)^2 S_{V_{fb}}. \quad (6.2)$$

The scattering-induced mobility fluctuation noise for DD devices is given as [103]

$$\frac{S_{Id}}{I_d^2} = \frac{4N_{t\mu}}{V} \frac{\tau w (1-w)}{1 + (\omega\tau)^2} \lambda^2 (\sigma_2 - \sigma_1)^2 \quad (6.3)$$

where $N_{t\mu}$ is the concentration of scattering centers, σ_1 to σ_2 is the scattering cross section change that results in the mobility fluctuation whose respective probabilities are w and $1-w$, τ is the characteristic time constant and V is the volume. The MFP in (6.3) is $\lambda = \tau_m v$, where τ_m is the momentum relaxation time and v is the velocity of electron chaotic motion.

As discussed in Chapters 3 and 4, scattering becomes a probability-based event in ballistic devices where $L \ll \lambda$. Since the odds of scattering depends on L itself, $L = \tau_m v$ is used for λ in (6.3) at ballistic limits. To extend the scattering-induced mobility fluctuation noise to QB regime, the apparent MFP (λ_{app}) given as [104]

$$\frac{1}{\lambda_{app}} = \frac{1}{\lambda} + \frac{1}{L} \quad (6.4)$$

is used to replace λ in (6.3). As discussed in Chapter 5, the integration of (6.3) over τ [66], [105], the new scattering-based LFN model is given by

$$\frac{S_{Id}}{I_d^2} = \frac{4N_{t\mu}}{V} \frac{w(1-w)}{f} \lambda_{app}^2 (\sigma_2 - \sigma_1)^2 \sim \varsigma \frac{\lambda_{app}^2}{V} \quad (6.5)$$

where ς is a constant. It can be noted that the proposed λ_{app}^2 model varies as λ^2 and L^2 in DD and QB regimes, respectively.

From (6.2), the mobility-fluctuation-dominant noise in QB regime is

given as

$$\frac{S_{Id}}{I_d^2} \sim \alpha^2 \mu_B^2 S_{V_{fb}} \sim L^2 \times 1/L \sim L \quad (6.6)$$

From (6.5), the scattering-limited noise in QB regime is given as

$$\frac{S_{Id}}{I_d^2} \sim \lambda_{app}^2 / V \sim L^2 / L \sim L. \quad (6.7)$$

It can be noted that both the models predict the noise power to scale with L . This confirms that the anomalous noise behavior in QB devices is actually due to the reduced scattering events and not due to the CNF/CMF. The probability of scattering induced noise decreases with channel length. This is in consonance with the physics of QB devices.

From (6.2), (6.4), and (6.5), we can calculate a rough estimate of $N_{t\mu}$ as

$$N_{t\mu} = \frac{\alpha^2 v_T^2 S_{V_{fb}} f V}{16 (kT/q)^2 w (1-w) (\sigma_2 - \sigma_1)^2}. \quad (6.8)$$

The total number of scatterers in a volume V is given as

$$N_{tr} = N_{t\mu} V \sim V^2 S_{V_{fb}} \sim L^2 / L \sim L \quad (6.9)$$

which is what one would expect by just scaling L , resulting in a direct relationship between the LFN and N_{tr} in ballistic devices.

6.3 Results and Discussion

The proposed CNF/CMF models behavior and comparison with the new λ_{app} noise model for a conceptual MOS device with different $\lambda = 90nm$ and $180nm$, at varying channel lengths, is shown in Fig. 6.2. The model pa-

Figure 6.2: Model behavior and comparison of the proposed CNF/CMF model with the new λ_{app} model for two conceptual MOS devices with $\lambda = 90nm$ and $180nm$, at varying channel lengths.

parameters are fitted to match the device at $L = 1\mu m$. CNF/CMF and λ_{app} model behaviors for varying channel lengths is shown with the dashed and the solid lines respectively. The nearly overlapping model behavior in the QB regimes show that the modified CNF/CMF model actually predicts a reduction in scattering rates, and hence the LFN, in accordance with the λ_{app} model. This behavior also tallies with the ballistic transport theory where the scattering becomes a probability-based event. The two models match in the DD regime since they are fitted to the long-channel behavior to study the scaling effect. The actual mobility fluctuation mechanism in this regime could be heavily affected by the CMF. It is also interesting to note from the figure that the onset of the transition from the conventional CMF/CNF to the new scattering-limited LFN during L scaling depends on the λ . For devices with larger λ , an earlier onset is seen.

Fig. 6.3 shows the S_{Id}/I_d^2 , at different gate biases, for a concep-

Figure 6.3: S_{Id}/I_d^2 vs. channel length. The model parameters used are indicated.

Figure 6.4: $\sqrt{S_{Id}}$ for 20nm QB device. The model parameters used are indicated.

tual MOS device with $\lambda = 180\text{nm}$ and varying channel lengths. Since, with increasing V_{gs} , the fluctuations are better screened due to increased carrier density, the LFN strength goes down for large devices. A similar trend, but

Figure 6.5: $\sqrt{S_{Id}}$ for $1\mu m$ DD device. The model parameters used are indicated.

with decreased impact, is seen at QB regimes and is attributed to the ballistic nature of carrier transport. Figs. 6.4 and 6.5 show the square root of LFN noise power for the 20nm QB and $1\mu m$ DD devices. It is worthwhile to note that the $\sqrt{S_{Id}}$ show striking similarity to the $I_d - V_{ds}$ characteristics of QB and DD devices (discussed in Chapter 4), with a linear and a quadratic V_{gs} dependence for QB and DD, respectively.

Fig. 6.6 shows the measurement data and model comparison for InGaAs nanowire devices [98]. The λ for III-V based devices is typically high, with 160–180 nm for InGaAs material [28]. The DIBL for the 20 nm device is low (~ 35 mV/V) [98] and this results in an improved short-channel behavior for even larger L . Hence, the contribution of short-channel effects to noise could be safely ignored. The extended CNF/CMF model (6.2) can be tested by any DC I_{ds} model. The $I_d - V_{gs}$ and $I_d - V_{ds}$ measurement data for the 20nm device (shown in Figs. 6.7 and 6.8) is fitted to our DC model discussed

Figure 6.6: Measurement data and model comparison for InGaAs nanowire devices. The model parameters used are indicated.

in [4]. An average charge is extracted from the DC model and is used for calculating the noise power based on (6.2). Also, the extracted λ_{app} model parameter $N_{t\mu} = 7.8 \times 10^{15} \text{cm}^{-3}$ (based on (6.8)). It should be noted that the estimated $N_{t\mu}$ value is in accordance with the unintentional doping in InGaAs. The unintentional dopants could act as potential scatterers [106]. The model accurately predicts the LFN for varying L , without any refitting for different channel lengths and could be a good validation for the model behavior. It could also be noted that both the models nearly overlap. Only difference is in the λ_{app} model that does not include any number-fluctuation effects. The modified CNF/CMF also caters to number-fluctuation effects, even though the magnitude of CNF is less due to ballistic nature of carrier transport.

By setting $d(6.2)/dL = 0$, an analytical equation for the channel

Figure 6.7: $I_d - V_{gs}$ fitted to the DC model described in [4].

Figure 6.8: $I_d - V_{ds}$ fitted to the DC model described in [4].

length at which S_{Id}/I_d^2 is peaked can be derived as

$$L_{pk} = k\lambda \quad (6.10a)$$

Figure 6.9: L_{pk} vs. λ with varying V_g at $T = 300K$. For large λ devices, $L_{pk} \approx \lambda$ with weak V_g dependence. For small λ devices, L_{pk} is dependent on both V_g and λ , as shown in the inset.

$$k = \frac{1}{2(1 + \Omega)} \left[\Omega - 2 + \sqrt{(\Omega - 2)^2 - 4(1 + \Omega)} \right] \quad (6.10b)$$

with $\Omega \equiv \alpha \varpi \lambda N$, $\varpi \equiv v_T / (k_B T / q) = \sqrt{2q / \pi m^* k_B T}$.

When $\Omega \gg 1$ [which means one of the following: large λ , large N (or V_g); low T], we have $k \rightarrow (\Omega + \sqrt{\Omega^2 - 4\Omega}) / 2\Omega \approx 1$, i.e., $L_{pk} \approx \lambda$. That means that under those conditions (for $\Omega \gg 1$), peak noise PSD always occurs around $L \approx \lambda$, which is the transition length between DD and QB regimes based on the apparent-mobility/MFP models. For devices with small λ (which implies more scatterers), peak noise PSD also depends on carrier numbers (or applied gate voltage). At low V_g , the number of carriers is less and, hence, the onset of scattering-limited QB noise behavior occurs at a smaller channel length ($L_{pk} < \lambda$). With increasing V_g , the “corner” peak-noise length increases to approach λ . The V_g and λ dependence of L_{pk} at $T = 300K$ is shown in Fig. 6.9, whose inset shows the V_g effect at low λ

more clearly. These predictions by the simple analytical model should be further verified by controlled LFN experiments with different geometries on the same wafer, and at varying temperatures.

6.4 Conclusion

In this chapter, impact of DD to QB transition with channel length on the LFN is studied. Recent measurements on Ge and InGaAs-based QB devices showed an anomalous LFN scaling trend with $S_{Id}/I_d^2 \sim L$. This contradicts the conventional CNF/CMF and HMF models that predict a $1/L$ dependence. The conventional CNF/CMF model is extended to the QB regime based on the apparent mobility. The model tallies with the QB device physics and is found to match the new scattering-limited noise model thus confirming the role of scatterers in determining the noise in QB regime. Both the models provide a new interpretation for accurately predicting LFN measurement results in QB devices.

Declaration

Content of this chapter has been partly published in IEEE Electron Device Letters DOI: 10.1109/LED.2017.2716411.

Chapter 7

Conclusion and future work

7.1 Conclusion

The main aim of this thesis was to study the quasi-ballistic transport and its impact in semiconductor devices. In Chapter 2, a unified regional Fermi-potential model for double heterostructure HEMT devices with rectangular quantum well has been developed. The model is scalable and matches the coupled Poisson-Schrödinger solution in the subthreshold, weak and strong inversion regimes with millivolt accuracy. It was also shown that the triangular-well model fails to capture the device physics and hence, cannot be used for compact modeling of rectangular-well devices. Since the Fermi-potential is analogous to the surface potential in conventional MOSFETs, the well studied DD-based current equation derived from the surface potential for MOSFETs were used to simulate the DC characteristics of HEMTs.

HEMTs encounter QB transition at a larger length scales than MOSFETs. Chapter 3 deals with ballistic phenomenon encountered at very small

length scales. In this chapter, the different models used to study the mobility in QB region were studied. Since these models were all L based, they do not have the device physical parameter dependence. A new critical length model based on quasi-2D Poisson solution was discussed. The model was validated with Monte Carlo simulation results from literature and could be used to study the onset of ballistic transport with L scaling.

In Chapter 4, impact of DD and QB transport regimes on the DC characteristic of the device was discussed in details. An overview and limitations of the existing models were given. A new model, which extended the well known bulk velocity-field characteristics to the QB regime, was used for an in-depth study of the impact of the injection and saturation velocities, MFP, and temperature on the transition region between DD to QB regimes. The effect of critical length based mobility model on the transition regime was also studied in detail. The full DC model was validated with SOI measurement data for devices with channel length spanning the DD and QB regimes.

It is also interesting to note the impact of transport mechanisms on the LFN in MOSFETs. In Chapter 5, the existing LFN models were discussed in detail. It was shown that since $(g_m/I_d)^2$ based LFN formalism used to classify the origin of noise needed reformulation in the subthreshold regime since the initially derived model was based on strong-inversion conditions. The proposed model was extensively studied and validated using measurement results and data extracted from the literature.

Chapter 6 further studies the impact of DD to QB transition on LFN in MOSFETs. The recent LFN data for devices operating in QB regime showed an anomalous scaling trend with an L -dependent LFN instead of the well known $1/L$ -dependence predicted by the conventional DD theory. The existing LFN model was extended to the QB regime and was used to

successfully explain the anomalous noise behavior. The scattering based LFN model was also extended to QB regime to confirm that the LFN in very small length scales was indeed due to the scattering effects.

7.2 Future Work

- (i) The QB model proposed does not include the degeneracy effects. Such effects has been well studied in literature. The new v_{avg} model along with the degeneracy effects could be used to do more accurate physics based studies on III-V based devices. It would also be interesting to re-examine the effect of temperature with a degeneracy model.
- (ii) For HEMTs, kink and impact ionization effects are important. These have not been included since the aim of this chapter is to study the transport characteristics from the scaling point of view. To get the full model, these effects need to be considered.
- (iii) The proposed critical-length-based model could be used for studying the reliability characteristics during MOSFET operation. Since the critical length model can be used to extract out the low field region, it can be used as an efficient tool to probe both trapping-detrapping characteristics in the low field and impact ionization happening in the high field regions of the channel.
- (iv) Impact of charge trapping and detrapping within the oxide could be studied in terms of the related surface potential fluctuations. Improved threshold voltage models for DD regime and LFN models for nano MOSFETs where coulombic effects are significant could be derived.

Publications

Journals:

- A. Ajaykumar, X. Zhou, S. B. Chiah and B. Syamal, "Quasi-2D Surface Potential-Based Critical Length for Drift-Diffusion," in IEEE Electron Device Letters, vol. 37, no. 8, pp. 1051-1054, Aug. 2016.
- A. Ajaykumar, X. Zhou, S. B. Chiah and B. Syamal, "Impact of Sub-threshold Carrier Statistics on the Low-Frequency Noise in MOSFETs," in IEEE Transactions on Electron Devices, vol. 64, no. 4, pp. 1702-1707, April 2017.
- A. Ajaykumar; X. Zhou; S. B. Chiah, "A New Interpretation for the Anomalous Channel-Length Dependence of Low-Frequency Noise in Quasi-Ballistic Transistors," in IEEE Electron Device Letters , vol. 38, no. 8, pp. 1113-1116, Aug. 2017.

Conference:

- A. Ajaykumar, Z. Xing, B. Syamal and S. Ben Chiah, "Compact Fermi potential model for heterostructure HEMTs with rectangular quantum

well,” 2014 44th European Solid State Device Research Conference (ESSDERC), Venice, 2014, pp. 266-269.

- A. Ajaykumar, X. Zhou, S. Ben Chiah and B. Syamal, "Unified regional fermipotential-based compact model for double heterostructure HEMTs," 2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Nuremberg, 2016, pp. 225-228.
- A. Ajaykumar, X. Zhou, B. Syamal, S. Ben Chiah and L. Zhang, "Compact electrical/optical model for InGaN/GaN quantum-well based LEDs," 2015 IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC), Singapore, 2015, pp. 766-769
- A. Ajaykumar, X. Zhou and S. B. Chiah, "Effect of metal work function on the DC characteristics of an asymmetric MOSFET with Schottky-based source," 2016 International Symposium on Integrated Circuits (ISIC), Singapore, 2016, pp. 1-4.

Bibliography

- [1] Feilong Liu, Jian Zhang, Frank He, Feng Liu, Lining Zhang, and Mansun Chan. A charge-based compact model for predicting the current-voltage and capacitance-voltage characteristics of heavily doped cylindrical surrounding-gate {MOSFETs}. *Solid-State Electronics*, 53(1):49 – 53, 2009.
- [2] Xing Zhou, Khee Yong Lim, and D. Lim. A general approach to compact threshold voltage formulation based on 2d numerical simulation and experimental correlation for deep-submicron ulsi technology development [cmos]. *IEEE Transactions on Electron Devices*, 47(1):214–221, Jan 2000.
- [3] H.C. Pao and C.T. Sah. Effects of diffusion current on characteristics of metal-oxide (insulator)-semiconductor transistors. *Solid-State Electronics*, 9(10):927 – 937, 1966.
- [4] Xing Zhou, Guojun Zhu, Guan Huei See, Karthik Chandrasekaran, Siau Ben Chiah, and Khee Yong Lim. Unification of MOS compact models with the unified regional modeling approach. *J. Comput. Electron.*, 10(1):121–135, 2011. DOI: 10.1007/s10825-011-0354-y.
- [5] T. J. Anderson, M. J. Tadjer, J. K. Hite, J. D. Greenlee, A. D. Koehler, K. D. Hobart, and F. J. Kub. Effect of reduced extended defect density in mocvd grown algan/gan hemts on native gan substrates. *IEEE*

- Electron Device Letters*, 37(1):28–30, Jan 2016.
- [6] Feng Zhihong, Xie Shengyin, Zhou Rui, Yin Jiayun, Zhou Wei, and Cai Shujun. A high-performance enhancement-mode algan/gan hemt. *Journal of Semiconductors*, 31(8):084001, 2010.
- [7] Dieter Fritzsche. Heterostructures in modfets. *Solid-State Electronics*, 30(11):1183 – 1195, 1987.
- [8] J. Zhang, B. Syamal, X. Zhou, S. Arulkumaran, and G. I. Ng. A compact model for generic mis-hemts based on the unified 2deg density expression. *IEEE Transactions on Electron Devices*, 61(2):314–323, Feb 2014.
- [9] Sourabh Khandelwal and T.A. Fjeldly. A physics based compact model of iv and cv characteristics in algan/gan HEMT devices. *Solid-State Electronics*, 76:60 – 66, 2012.
- [10] D. Delagebeaudeuf and Nuyen T. Linh. Metal-(n) algaas-gaas two-dimensional electron gas fet. *Electron Devices, IEEE Transactions on*, 29(6):955–960, Jun 1982.
- [11] F. F. Fang and W. E. Howard. Negative field-effect mobility on (100) si surfaces. *Phys. Rev. Lett.*, 16:797–799, May 1966.
- [12] J. A. del Alamo, D. A. Antoniadis, J. Lin, W. Lu, A. Vardi, and X. Zhao. Iii-v mosfets for future cmos. In *2015 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS)*, pages 1–4, Oct 2015.
- [13] J. Lin, D. A. Antoniadis, and J. A. del Alamo. Impact of intrinsic channel scaling on ingaas quantum-well mosfets. *IEEE Transactions on Electron Devices*, 62(11):3470–3476, Nov 2015.
- [14] A. Ajaykumar, Z. Xing, B. Syamal, and S. Ben Chiah. Compact fermi potential model for heterostructure hemts with rectangular quantum well. In *2014 44th European Solid State Device Research Conference (ESSDERC)*, pages 266–269, Sept 2014.

- [15] A. Ajaykumar, X. Zhou, S. Ben Chiah, and B. Syamal. Unified regional fermi-potential-based compact model for double heterostructure hemts. In *2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pages 225–228, Sept 2016.
- [16] Tatsushi Akazaki, K. Arai, T. Enoki, and Y. Ishii. Improved inalas/ingaas hemt characteristics by inserting an inas layer into the ingaas channel. *Electron Device Letters, IEEE*, 13(6):325–327, June 1992.
- [17] G. Bastard, E. E. Mendez, L. L. Chang, and L. Esaki. Variational calculations on a quantum well in an electric field. *Phys. Rev. B*, 28:3241–3245, Sep 1983.
- [18] Physical properties of indium arsenide InAs. <http://www.ioffe.ru/SVA/NSM/Semicond/InAs/>.
- [19] S. C. Lin and J. B. Kuo. Temperature-dependent kink effect model for partially-depleted soi nmos devices. *IEEE Transactions on Electron Devices*, 46(1):254–258, Jan 1999.
- [20] M. H. Somerville, A. Ernst, and J. A. del Alamo. A physical model for the kink effect in inalas/ingaas hemts. *IEEE Transactions on Electron Devices*, 47(5):922–930, May 2000.
- [21] R. Reuter, M. Agethen, U. Auer, S. van Waasen, D. Peters, W. Brockerhoff, and F. J. Tegude. Investigation and modeling of impact ionization with regard to the rf and noise behavior of hfet. *IEEE Transactions on Microwave Theory and Techniques*, 45(6):977–983, Jun 1997.
- [22] Naoki Yasuda, Shuichi Ueno, Kenji Taniguchi, Chihiro Hamaguchi, Yasuo Yamaguchi, and Tadashi Nishimura. Analytical device model of soi mosfets including self-heating effect. *Japanese Journal of Applied Physics*, 30(12S):3677, 1991.
- [23] Mark Lundstrom. Fundamentals of carrier transport, 2nd edn. *Measurement Science and Technology*, 13(2):230, 2002.

- [24] Kenji Natori. Ballistic metal-oxide-semiconductor field effect transistor. *J. Appl. Phys.*, 76(8):4879–4890, 1994. DOI: <http://dx.doi.org/10.1063/1.357263>.
- [25] Maziar Nekovee, Bernard J. Geurts, Henk M. J. Boots, and Martin F. H. Schuurmans. Failure of extended-moment-equation approaches to describe ballistic transport in submicrometer structures. *Phys. Rev. B*, 45:6643–6651, Mar 1992.
- [26] Jung-Hoon Rhew and Mark S. Lundstrom. Drift-diffusion equation for ballistic transport in nanoscale metal-oxide-semiconductor field effect transistors. *Journal of Applied Physics*, 92(9):5196–5202, 2002.
- [27] Jing Wang and M. Lundstrom. Ballistic Transport in High Electron Mobility Transistors. *IEEE Trans. Electron Devices*, 50(7):1604–1609, Jul. 2003. DOI: 10.1109/TED.2003.814980.
- [28] J. Lin, Y. Wu, J. A. del Alamo, and D. A. Antoniadis. Analysis of Resistance and Mobility in InGaAs Quantum-Well MOSFETs from Ballistic to Diffusive Regimes. *IEEE Trans. Electron Devices*, 63(4):1464–1470, Apr. 2016. DOI: 10.1109/TED.2016.2532604.
- [29] M. Lundstrom. Elementary scattering theory of the Si MOSFET. *IEEE Electron Device Lett.*, 18(7):361–363, Jul. 1997. DOI: 10.1109/55.596937.
- [30] M.S. Lundstrom and D.A. Antoniadis. Compact Models and the Physics of Nanoscale FETs. *IEEE Trans. Electron Devices*, 61(2):225–233, Feb. 2014. DOI: 10.1109/TED.2013.2283253.
- [31] Vijay K. Arora. Theory of scattering-limited and ballistic mobility and saturation velocity in low- dimensional nanostructures. *Current Nanoscience*, 5(2), 2009.
- [32] A. Ajaykumar, X. Zhou, S. B. Chiah, and B. Syamal. Quasi-2D surface-potential-based critical length for drift-diffusion. *IEEE Electron Device*

- Lett.*, 37(8):1051–1054, Aug. 2016. DOI: 10.1109/LED.2016.2578718.
- [33] A.A. Kastalsky and M.S. Shur. Conductance of small semiconductor devices. *Solid State Communications*, 39(6):715 – 718, 1981.
- [34] M. S. Shur. Low ballistic mobility in submicron HEMTs. *IEEE Electron Device Lett.*, 23(9):511–513, Sep. 2002. DOI: 10.1109/LED.2002.802679.
- [35] V. Arora. Quantum engineering of nanoelectronic devices: the role of quantum emission in limiting drift velocity and diffusion coefficient. *Microelectronics Journal*, 31(1112):853 – 859, 2000. DOI: 10.1016/S0026-2692(00)00085-9.
- [36] Vijay K. Arora. Drift diffusion and Einstein relation for electrons in silicon subjected to a high electric field. *Appl. Phys. Lett.*, 80(20), 2002. DOI: 10.1063/1.1480119.
- [37] Ismail Saad, Michael L.P. Tan, Ing Hui Hii, Razali Ismail, and Vijay K. Arora. Ballistic mobility and saturation velocity in low-dimensional nanostructures. *Microelectronics Journal*, 40(3):540 – 542, 2009. Workshop of Recent Advances on Low Dimensional Structures and Devices (WRA-LDSD).
- [38] V. K. Arora. Ballistic transport in nanoscale devices. In *Proceedings of the 19th International Conference Mixed Design of Integrated Circuits and Systems - MIXDES 2012*, pages 17–24, May 2012.
- [39] F. Berz. The bethe condition for thermionic emission near an absorbing boundary. *Solid-State Electronics*, 28(10):1007 – 1013, 1985.
- [40] K. Huet, J. Saint-Martin, A. Bournel, S. Galdin-Retailleau, P. Dollfus, G. Ghibaudo, and M. Mouis. Monte carlo study of apparent mobility reduction in nano-mosfets. In *ESSDERC 2007 - 37th European Solid State Device Research Conference*, pages 382–385, Sept 2007.
- [41] G. Ghibaudo, M. Mouis, L. Pham-Nguyen, K. Bennamane, I. Pappas,

- A. Cros, G. Bidal, D. Fleury, A. Claverie, G. Benassayag, P. F. Fazzini, C. Fenouillet-Beranger, S. Monfray, F. Boeuf, S. Cristoloveanu, T. Skotnicki, and N. Collaert. Electrical transport characterization of nano CMOS devices with ultra-thin silicon film. In *Proc. IWJT*, pages 58–63. DOI: 10.1109/IWJT.2009.5166220, Jun. 2009.
- [42] Kenji Natori, Hiroshi Iwai, and Kuniyuki Kakushima. Anomalous degradation of low-field mobility in short-channel metal-oxide-semiconductor field-effect transistors. *J. Appl. Phys.*, 118(23):234502–1–234502–7, 2015. DOI: <http://dx.doi.org/10.1063/1.4937548>.
- [43] Yeonghun Lee, Kuniyuki Kakushima, Kenji Natori, and Hiroshi Iwai. Modeling of quasi-ballistic transport in nanowire metal-oxide-semiconductor field-effect transistors. *J. Appl. Phys.*, 118(15):155105–1–155105–9, 2015. DOI: <http://dx.doi.org/10.1063/1.4933287>.
- [44] Z.-H. Liu, Chenming Hu, J.-H. Huang, T.-Y. Chan, Min-Chie Jeng, P.K. Ko, and Y.C. Cheng. Threshold voltage model for deep-submicrometer MOSFETs. *IEEE Trans. Electron Devices*, 40(1):86–95, Jan. 1993. DOI: 10.1109/16.249429.
- [45] A. Majumdar and D. A. Antoniadis. Analysis of carrier transport in short-channel MOSFETs. *IEEE Trans. Electron Devices*, 61(2):351–358, Feb. 2014. DOI: 10.1109/TED.2013.2294380.
- [46] P. Palestri, D. Esseni, S. Eminent, C. Fiegna, E. Sangiorgi, and L. Selmi. Understanding quasi-ballistic transport in nano-MOSFETs: Part I—scattering in the channel and in the drain. *IEEE Trans. Electron Devices*, 52(12):2727–2735, Dec. 2005. DOI: 10.1109/TED.2005.859593.
- [47] A. Rahman and M. S. Lundstrom. A compact scattering model for the nanoscale double-gate MOSFET. *IEEE Trans. Electron Devices*, 49(3):481–489, Mar. 2002. DOI: 10.1109/16.987120.

- [48] S. Rakheja, M. Lundstrom, and D. Antoniadis. A physics-based compact model for fets from diffusive to ballistic carrier transport regimes. In *2014 IEEE International Electron Devices Meeting*, pages 35.1.1–35.1.4, Dec 2014.
- [49] G. Mugnaini and G. Iannaccone. Physics-based compact model of nanoscale MOSFETs-part I: Transition from drift-diffusion to ballistic transport. *IEEE Trans. Electron Devices*, 52(8):1795–1801, Aug. 2005. DOI: 10.1109/TED.2005.851827.
- [50] P. Michetti, G. Mugnaini, and G. Iannaccone. Analytical model of nanowire FETs in a partially ballistic or dissipative transport regime. *IEEE Trans. Electron Devices*, 56(7):1402–1410, Jul. 2009. DOI: 10.1109/TED.2009.2021720.
- [51] B. R. Nag, S. R. Ahmed, and M. D. Roy. Electron velocity in short samples of $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ at 300 K. *IEEE Trans. Electron Devices*, 33(6):788–791, Jun. 1986. DOI: 10.1109/T-ED.1986.22569.
- [52] T. H. Windhorn, L. W. Cook, and G. E. Stillman. The electron velocity-field characteristic for n- $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ at 300 K. *IEEE Electron Device Lett.*, 3(1):18–20, Jan. 1982. DOI: 10.1109/EDL.1982.25459.
- [53] A. Majumdar, Y. Sun, C. W. Cheng, Y. H. Kim, U. Rana, R. M. Martin, R. L. Bruce, K. T. Shiu, Y. Zhu, D. B. Farmer, M. Hopstaken, E. A. Joseph, J. P. de Souza, M. M. Frank, S. L. Cheng, M. Kobayashi, E. A. Duch, D. K. Sadana, D. G. Park, and E. Leobandung. CMOS-compatible self-aligned $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ MOSFETs with gate lengths down to 30 nm. *IEEE Trans. Electron Devices*, 61(10):3399–3404, Oct. 2014. DOI: 10.1109/TED.2014.2335747.
- [54] S. A. Schwarz and S. E. Russek. Semi-empirical equations for electron velocity in silicon: Part I—bulk. *IEEE Trans. Electron Devices*, 30(12):1629–1633, Dec. 1983. DOI: 10.1109/T-ED.1983.21423.

- [55] D. Fleury, G. Bidal, A. Cros, F. Boeuf, T. Skotnicki, and G. Ghibaudo. New experimental insight into ballisticity of transport in strained bulk MOSFETs. In *2009 Symp. VLSI Technology*, pages 16–17, Jun. 2009.
- [56] M. Shin, M. Shi, M. Mouis, A. Cros, E. Josse, S. Mukhopadhyay, B. Piot, G. T. Kim, and G. Ghibaudo. Experimental and theoretical investigation of magnetoresistance from linear regime to saturation in 14-nm FD-SOI MOS devices. *IEEE Trans. Electron Devices*, 62(1):3–8, Jan. 2015. DOI: 10.1109/TED.2014.2366170.
- [57] M. Casse, L. Hutin, C. Le Royer, D. Cooper, J. M. Hartmann, and G. Reimbold. Experimental investigation of hole transport in strained pMOSFETs: Part II: Mobility and high-field transport in nanoscaled PMOS. *IEEE Trans. Electron Devices*, 59(3):557–564, Mar. 2012. DOI: 10.1109/TED.2011.2177985.
- [58] S. Khandelwal, H. Agarwal, P. Kushwaha, J. P. Duarte, A. Medury, Y. S. Chauhan, S. Salahuddin, and C. Hu. Unified compact model covering drift-diffusion to ballistic carrier transport. *IEEE Electron Device Lett.*, 37(2):134–137, Feb. 2016. DOI: 10.1109/LED.2015.2507519.
- [59] A. Dasgupta, A. Agarwal, S. Khandelwal, and Y. S. Chauhan. Compact modeling of surface potential, charge, and current in nanoscale transistors under quasi-ballistic regime. *IEEE Trans. Electron Devices*, 63(11):4151–4159, Nov. 2016. DOI: 10.1109/TED.2016.2603540.
- [60] K.Y. Lim and X. Zhou. A physically-based semi-empirical effective mobility model for MOSFET compact I-V modeling. *Solid-State Electron.*, 45(1):193–197, 2001. DOI: 10.1016/S0038-1101(00)00190-8.
- [61] P. Dutta and P. M. Horn. Low-frequency fluctuations in solids: $\frac{1}{f}$ noise. *Rev. Mod. Phys.*, 53:497–516, Jul 1981.
- [62] Stefan Machlup. Noise in semiconductors: Spectrum of a twoparameter random signal. *Journal of Applied Physics*, 25(3):341–343, 1954.

- [63] Horng-Sen Fu and Chih-Tang Sah. Theory and experiments on surface $1/f$ noise. *IEEE Transactions on Electron Devices*, 19(2):273–285, Feb 1972.
- [64] K. Clment, N. and Nishiguchi, A. Fujiwara, and D. Vuillaume. One-by-one trap activation in silicon nanowire transistors. *Nat Commun*, 1(92).
- [65] S. Christensson and I. Lundstrm. Low frequency noise in mos transistorsii experiments. *Solid-State Electronics*, 11(9):813 – 820, 1968.
- [66] S. Christensson, I. Lundstrm, and C. Svensson. Low frequency noise in MOS transistors-I theory. *Solid-State Electron.*, 11(9):797–812, 1968. DOI: 10.1016/0038-1101(68)90100-7.
- [67] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng. A unified model for the flicker noise in metal-oxide-semiconductor field-effect transistors. *IEEE Transactions on Electron Devices*, 37(3):654–665, Mar 1990.
- [68] L. K. J. Vandamme, Xiaosong Li, and D. Rigaud. $1/f$ noise in mos devices, mobility or number fluctuations? *IEEE Transactions on Electron Devices*, 41(11):1936–1945, Nov 1994.
- [69] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng. A physics-based mosfet noise model for circuit simulators. *IEEE Transactions on Electron Devices*, 37(5):1323–1333, May 1990.
- [70] F.N. Hooge. $1/f$ noise is no surface effect. *Physics Letters A*, 29(3):139 – 140, 1969.
- [71] F.N. Hooge. $1/f$ noise. *Physica B+C*, 83(1):14 – 23, 1976.
- [72] F N Hooge, T G M Kleinpenning, and L K J Vandamme. Experimental studies on $1/f$ noise. *Reports on Progress in Physics*, 44(5):479, 1981.
- [73] L.B. Kiss and T.G.M. Kleinpenning. Interpretation of hooge’s $1/f$ noise formula. *Physica B+C*, 145(2):185 – 189, 1987.
- [74] Vilius Palenskis and Kstutis Maknys. Nature of low-frequency noise in homogeneous semiconductors. *Scientific Reports*, 5:18305, Dec

2015. DOI: 10.1038/srep18305. Article.
- [75] R. Jayaraman and C. G. Sodini. A $1/f$ noise technique to extract the oxide trap density near the conduction band edge of silicon. *IEEE Trans. Electron Devices*, 36(9):1773–1782, Sep. 1989. DOI: 10.1109/16.34242.
- [76] Z. Celik-Butler and T.Y. Hsiang. Spectral dependence of 1 noise on gate bias in n-mosfets. *Solid-State Electronics*, 30(4):419 – 423, 1987.
- [77] Martin von Haartman and Mikael Östling. *1/F Noise in Mosfets*, pages 53–102. Springer Netherlands, Dordrecht, 2007.
- [78] Fan-Chi Hou, G. Bosman, and M. E. Law. Simulation of oxide trapping noise in submicron n-channel mosfets. *IEEE Transactions on Electron Devices*, 50(3):846–852, March 2003.
- [79] J. P. Campbell, J. Qin, K. P. Cheung, L. C. Yu, J. S. Suehle, A. Oates, and K. Sheng. Random telegraph noise in highly scaled nmosfets. In *2009 IEEE International Reliability Physics Symposium*, pages 382–388, April 2009.
- [80] G. Reimbold. Modified $1/f$ trapping noise theory and experiments in mos transistors biased from weak to strong inversion;influence of interface states. *IEEE Trans. Electron Devices*, 31(9):1190–1198, Sep. 1984. DOI: 10.1109/T-ED.1984.21687.
- [81] R.P. Jindal and A. van der Ziel. Carrier fluctuation noise in a mosfet channel due to traps in the oxide. *Solid-State Electron.*, 21(6):901–903, 1978. DOI: 10.1016/0038-1101(78)90317-9.
- [82] G. Ghibaudo, O. Roux, Ch. Nguyen-Duc, F. Balestra, and J. Brini. Improved analysis of low frequency noise in field-effect mos transistors. *physica status solidi (a)*, 124(2):571–581, 1991.
- [83] Philippe Gaubert, Akinobu Teramoto, Rihito Kuroda, Yukihiisa Nakao, Hiroaki Tanaka, Shigetoshi Sugawa, and Tadahiro Ohmi. Analysis

- of the low-frequency noise reduction in si(100) metaloxidesemiconductor field-effect transistors. *Japanese Journal of Applied Physics*, 50(4S):04DC01, 2011.
- [84] L. K. J. Vandamme and F. N. Hooge. What do we certainly know about 1/f noise in MOSTs? *IEEE Trans. Electron Devices*, 55(11):3070–3085, Nov. 2008. DOI: 10.1109/TED.2008.2005167.
- [85] Deepak Sharma, Abhishek Motayed, Pankaj B. Shah, Matin Amani, Mariela Georgieva, A. Glen Birdwell, Madan Dubey, Qiliang Li, and Albert V. Davydov. Transfer characteristics and low-frequency noise in single- and multi-layer MoS₂ field-effect transistors. *Appl. Phys. Lett.*, 107(16), 2015. DOI: 10.1063/1.4932945.
- [86] Junhong Na, Min-Kyu Joo, Minju Shin, Junghwan Huh, Jae-Sung Kim, Mingxing Piao, Jun-Eon Jin, Ho-Kyun Jang, Hyung Jong Choi, Joon Hyung Shim, and Gyu-Tae Kim. Low-frequency noise in multilayer MoS₂ field-effect transistors: the effect of high-k passivation. *Nanoscale*, 6:433–441, 2014. DOI: 10.1039/C3NR04218A.
- [87] In-Tak Cho, Jong In Kim, Yoonki Hong, Jeongkyun Roh, Hyeonwoo Shin, Geun Woo Baek, Changhee Lee, Byung Hee Hong, Sung Hun Jin, and Jong-Ho Lee. Low frequency noise characteristics in multilayer WSe₂ field effect transistor. *Appl. Phys. Lett.*, 106(2), 2015. DOI: 10.1063/1.4906141.
- [88] K.-M. Persson, B. G. Malm, and L.-E. Wernersson. Surface and core contribution to 1/f-noise in InAs nanowire metal-oxide-semiconductor field-effect transistors. *Appl. Phys. Lett.*, 103(3), 2013. DOI: 10.1063/1.4813850.
- [89] A. Ajaykumar, X. Zhou, S. B. Chiah, and B. Syamal. Impact of sub-threshold carrier statistics on the low-frequency noise in mosfets. *IEEE Transactions on Electron Devices*, 64(4):1702–1707, April 2017.

- [90] C. Wei, Y. Z. Xiong, X. Zhou, N. Singh, S. C. Rustagi, G. Q. Lo, and D. L. Kwong. Investigation of low-frequency noise in silicon nanowire MOSFETs in the subthreshold region. *IEEE Electron Device Lett.*, 30(6):668–671, Jun. 2009. DOI: 10.1109/LED.2009.2019975.
- [91] D. Jimenez, B. Iniguez, J. Sune, L. F. Marsal, J. Pallares, J. Roig, and D. Flores. Continuous analytic I-V model for surrounding-gate MOSFETs. *IEEE Electron Device Lett.*, 25(8):571–573, Aug. 2004. DOI: 10.1109/LED.2004.831902.
- [92] S. B. Chiah, X. Zhou, and L. Yuan. Compact zero-temperature coefficient modeling approach for mosfets based on unified regional modeling of surface potential. *IEEE Trans. on Electron Devices*, 60(7):2164–2170, Jul. 2013. DOI: 10.1109/TED.2013.2262713.
- [93] B. C. Wang, S. L. Wu, Y. Y. Lu, S. J. Chang, J. F. Chen, S. C. Tsai, C. H. Hsu, C. W. Yang, C. G. Chen, O. Cheng, and P. C. Huang. Comparison of the trap behavior between ZrO_2 and HfO_2 gate stack nMOSFETs by $1/f$ noise and random telegraph noise. *IEEE Electron Device Lett.*, 34(2):151–153, Feb. 2013. DOI: 10.1109/LED.2012.2226698.
- [94] M. Si, N. J. Conrad, S. Shin, J. Gu, J. Zhang, M. A. Alam, and P. D. Ye. Low-frequency noise and random telegraph noise on near-ballistic III-V MOSFETs. *IEEE Trans. Electron Devices*, 62(11):3508–3515, Nov. 2015. DOI: 10.1109/TED.2015.2433921.
- [95] Nitin K. Rajan, David A. Routenberg, Jin Chen, and Mark A. Reed. Temperature dependence of $1/f$ noise mechanisms in silicon nanowire biochemical field effect transistors. *Appl. Phys. Lett.*, 97(24), 2010. DOI: 10.1063/1.3526382.
- [96] C. Wei, Y. Jiang, Y. Z. Xiong, X. Zhou, N. Singh, S. C. Rustagi, G. Q. Lo, and D. L. Kwong. Impact of gate electrodes on $1/f$ noise of gate-all-around silicon nanowire transistors. *IEEE Electron Device Lett.*,

- 30(10):1081–1083, Oct. 2009. DOI: 10.1109/LED.2009.2027614.
- [97] S. H. Lee, C. K. Baek, S. Park, D. W. Kim, D. K. Sohn, J. S. Lee, D. M. Kim, and Y. H. Jeong. Characterization of channel-diameter-dependent low-frequency noise in silicon nanowire field-effect transistors. *IEEE Electron Device Lett.*, 33(10):1348–1350, Oct. 2012. DOI: 10.1109/LED.2012.2209625.
- [98] N. Conrad, M. Si, S. H. Shin, J. J. Gu, J. Zhang, M. A. Alam, and P. D. Ye. Low-frequency noise and RTN on near-ballistic III-V GAA nanowire MOSFETs. In *2014 IEDM*, pages 20.1.1–20.1.4. DOI: 10.1109/IEDM.2014.7047086., Dec. 2014.
- [99] Wangran Wu, H. Wu, M. Si, N. Conrad, Yi Zhao, and P. D. Ye. RTN and low frequency noise on ultra-scaled near-ballistic Ge nanowire nMOSFETs. In *2016 IEEE Symposium on VLSI Technology*, pages 1–2. DOI: 10.1109/VLSIT.2016.7573421., Jun. 2016.
- [100] K. Walczak and T. L. Reinecke. Scaling rules for telegraph noise. *IEEE Trans. on Nanotechnology*, 10(6):1224–1230, Nov. 2011.
- [101] J. Tersoff. Low-frequency noise in nanoscale ballistic transistors. *Nano Letters*, 7(1):194–198, 2007. PMID: 17212463.
- [102] A. Dasgupta, H. Agarwal, A. Agarwal, and Y. S. Chauhan. Modeling of flicker noise in quasi-ballistic devices. In *Int. Conf. on Emerging Electronics*, Dec. 2016.
- [103] A. P. Dmitriev, M. E. Levinshtein, and S. L. Rumyantsev. On the Hooge relation in semiconductors and metals. *J. Appl. Phys.*, 106(2):024514, 2009.
- [104] Mark Lundstrom and Changwook Jeong. *Near-equilibrium transport: fundamentals and applications*. World Scientific, 2013.
- [105] Yu. M. Galperin, V. L. Gurevich, and V. I. Kozub. Disorder-induced low-frequency noise in small systems: Point and tunnel contacts in

BIBLIOGRAPHY

- the normal and superconducting state. *EPL (Europhysics Letters)*, 10(8):753, 1989.
- [106] J. F. Vaitneia, J. A. Canido, and J. I. Izpura. 1/f noise in InGaAs/GaAs linear graded buffer layers. *IEEE Trans. Electron Devices*, 45(6):1201–1206, Jun. 1998. DOI: 10.1109/16.678512.