Modeling Electrostatics and Low-field Electron Mobility in GaN FinFETs

by

Viswanathan Naveen Kumar

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2022 by the
Graduate Supervisory Committee:

Dragica Vasileska, Chair
Stephen Goodnick
Robert Nemanich
Michael Povolotskyi
Ivan Sanchez Esqueda

ARIZONA STATE UNIVERSITY

August 2022

ABSTRACT

Gallium Nitride (GaN) is uniquely suited for Radio Frequency (RF) and power electronic applications due to its intrinsically high saturation velocity and high mobility compared to Silicon and Silicon Carbide (SiC). High Electron Mobility Transistors (HEMTs) have remained the primary topology for GaN transistors in RF applications. However, GaN HEMTs suffer from a variety of issues such as current crowding, lack of enhancement mode (E-Mode) operation and non-linearity. These drawbacks slow the widespread adoption of GaN devices for ultra-low voltage (ULV) applications such as voltage regulators, automotive and computing applications. E-mode operation is especially desired in low-voltage high frequency switching applications. In this context, Fin Field Effect Transistors (FinFETs) offer an alternative topology for ULV applications as opposed to conventional HEMTs.

Recent advances in material processing, high aspect ratio epitaxial growth and etching methods has led to an increased interest in 3D nanostructures such as Nano-FinFETs and Nanowire FETs. A typical 3D nano-FinFET is the AlGaN/GaN Metal Insulator Semiconductor (MIS) FET wherein a layer of Al_2O_3 surrounds the AlGaN/GaN fin. The presence of the side gates leads to additional lateral confinement of the 2D Electron Gas (2DEG). Theoretical calculations of transport properties in confined systems such as AlGaN/GaN Finfets are scarce compared to those of their planar HEMT counterparts.

A novel simulator is presented in this dissertation, which employs self-consistent solution of the coupled 1D Boltzmann – 2D Schrödinger – 3D Poisson problem, to yield the channel electrostatics and the low electric field transport characteristics of

AlGaN/GaN MIS FinFETs. The low field electron mobility is determined by solving the Boltzmann transport equation in the Quasi-1D region using 1D Ensemble Monte Carlo method. Three electron-phonon scattering mechanisms (acoustic, piezoelectric and polar optical phonon scattering) and interface roughness scattering at the AlGaN/GaN interface are considered in this theoretical model. Simulated low-field electron mobility and its temperature dependence are in agreement with experimental data reported in the literature.

A quasi-1D version of alloy clustering model is derived and implemented and the limiting effect of alloy clustering on the low-field electron mobility is investigated for the first time for MIS FinFET device structures.

DEDICATION

Dedicated to

My loving parents

Viswanathan,

Nalini

and my brother Anuj

for staying with me on this journey

ACKNOWLEDGMENTS

I begin foremost by thanking my advisor Dr. Dragica Vasileska. She has guided and mentored me right from the time I started at Arizona State University. She is extremely patient and was always there to constantly motivate and push me in the right direction throughout my PhD, especially when I inevitably became stuck in my research. I am extremely grateful for her support, both academically and mentally.

I would also like to thank Dr. Michael Povolotskyi, who is a great researcher with a treasure trove of knowledge. His key insights during our weekly meetings were very necessary for my progress. I am also grateful to my other committee members, Dr. Robert Nemanich, Dr. Steven Goodnick and Dr. Ivan Sanchez for spending crucial time to review my research.

I am also thankful to Dr. James Aberle. Working as a teaching assistant under him gave me crucial experience as well as provided me with critical financial support throughout my PhD.

I would also like to all the current and past members of my research group, Abdul, Chi, Izak, Robin, Harshad…. who have been great colleagues and were always present to exchange ideas and provided valuable feedback.

I want to acknowledge the love and support of my parents and my brother. I am only able to pursue and achieve my dreams only because of their hard work and sacrifices.

I also thank my friends, Kiran, Shiladitya, and Gaurav for being there whenever I needed support.

TABLE OF CONTENTS

Page

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

1.1 Gallium Nitride HEMTs

Gallium Nitride is one of the strongest candidates for high power, high frequency device applications. GaN has wide applications in optoelectronic devices (such as Blue LEDs)[1], [2] as well as RF [3]–[5] and power electronics [6]–[8]. GaN also holds good potential for developing radiation-hard devices due to its high bandgap and high critical electric field. Power devices are widely sought after for fast chargers, wireless chargers and EVs.

Heterostructure interfaces of III-V materials allow formation of quantum wells, leading to formation of 2D Electron Gas (2DEG) at the interface. The 2DEG has excellent electrical conductivity because of the presence of Nitrides with large spontaneous polarization that leads to high electron density in the triangular potential well at the heterointerface. The strain present at the interfaces due to lattice mismatch also induces piezoelectric polarization, which further enhances the 2DEG density.

Table 1.1 Material properties of GaN compared with conventional semiconductors[3][9]

Property	**GaN**	**Si**	**GaAs**	**4H-SiC**
Bandgap (eV)	3.39	1.12	1.42	3.25
Critical Electric Field (MV/cm)	4.0	0.25	0.40	3.0
Saturation velocity (10^7 cm/s)	3.0	1.0	2.0	2.0
Electron Mobility (cm^2/Vs)	1300	1350	6000	800
Thermal conductivity (W/cm K)	1.3	1.4	0.5	4.9
Dielectric Constant	9.0	11.8	12.8	9.7

The excellent material properties of III-nitrides, such as high electron mobility, high saturation velocity and high critical electric field (see Table 1.1 for more details) lead to GaN being a material of choice for RF and power devices working in the high frequency, high voltage regimes. Power device applications can be categorized into low-voltage, medium-voltage and high-voltage classes. The scope for GaN in medium and high voltage class devices is well researched. Currently, GaN HEMTs (Figure 1.2) have been commercialized for operation in the 15-650V range[10]–[13]. The high switching speeds of HEMTs compared to Si devices have enabled application of GaN HEMTs in the charging infrastructure for electrified transportation[14]. Because of their widespread usage, GaN HEMTs have been extensively studied experimentally and theoretically. Indeed, this is evidenced by the availability of a vast amount of literature on modeling mobility in GaN HEMTs [15]–[18].

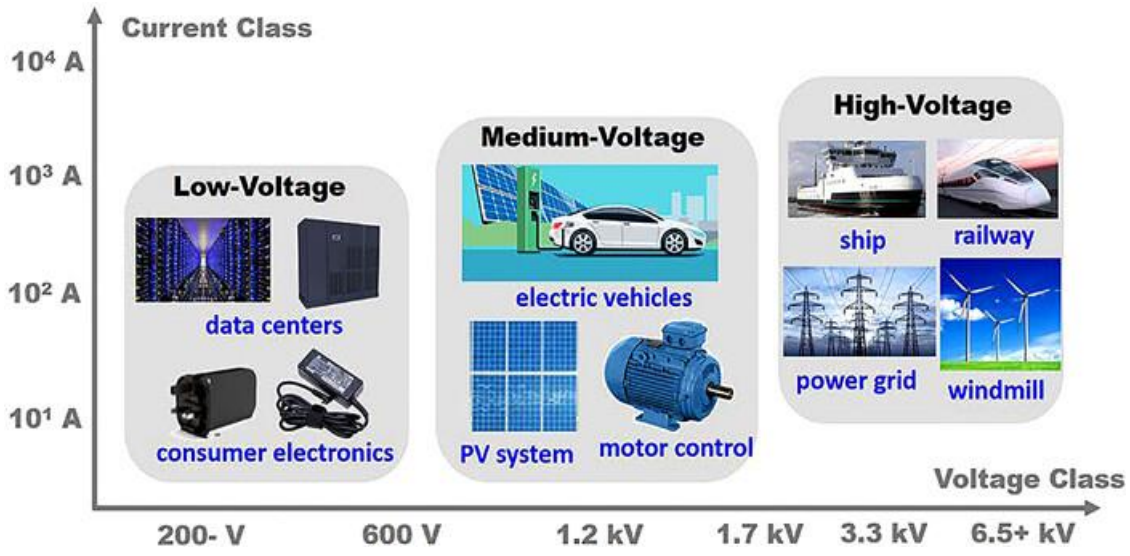

Figure 1.1 Voltage classes of power device applications. Reprinted from [14] , with the permission of AIP Publishing

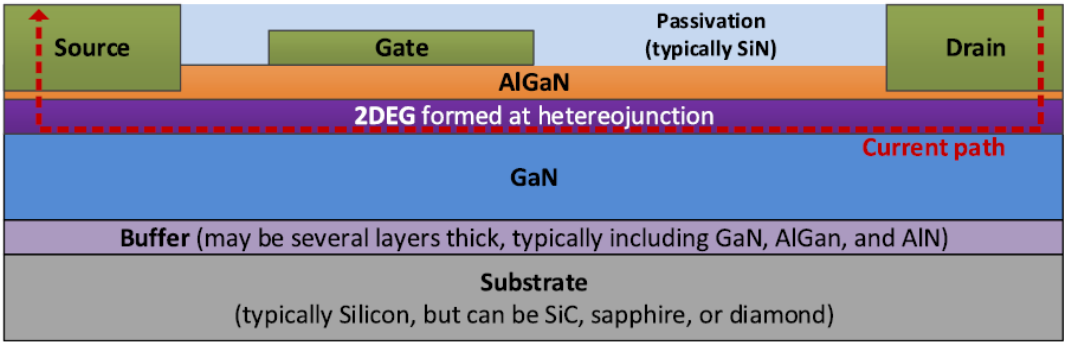

Figure 1.2 Typical Gallium Nitride HEMT structure. An AlN layer may or may not be present (depending on the design).

The high polarization field at the AlGaN (AlN)/GaN hetero-interface induces 2D electron gas (2DEG) which is responsible for the high conductivity of the channel in HEMTs. However, the ever-present electron gas, limits HEMTs to only D-mode operation. (HEMTs are depletion mode devices because the 2DEG channel exists at zero gate bias and requires application of negative gate bias to be turned off.) Widespread application of HEMTs in RF applications is also limited by the non-linear transconductance exhibited by these devices[10].

In addition to medium and high voltage devices, several opportunities exist for extending the application of GaN in the low and ultra-low voltage (ULV) device applications (<50 V) traditionally dominated by Silicon lateral MOS devices. Some key low voltage power applications include voltage regulators (VR) commonly used in data centers, mobile devices and wearable electronics [19]. The VRs which are mounted directly next to the load (CPUs and GPUs) consume valuable real estate. One of the approaches to increase efficiency and reduce space is to increase the switching frequency of the VRs. GaN FETs have been explored for design of high frequency DC-DC

converters [20], [21] for use in automotive applications such as hybrid and electric vehicles [22]. While Si LDMOS is most widely used in ultra-low voltage high frequency applications, GaN provides a favorable alternative here too due to the small reverse recovery times, thus allowing operations at higher switching frequencies. Current GaN devices, which are being considered in this voltage regime, are based on the Schottky p-gate HEMT (SP-HEMT) [14]. However, the difficulty of realizing a high-quality p-type GaN prevents the full realization of the material potential of GaN. Additionally, in order to realize E-mode devices (necessary in power converters to have safe switching), either thick p-GaN layers or MOS-HEMT approaches are necessary. Note that both technologies increase the on-state resistance and exhibit a low transconductance.

For ULV devices, the critical targets are to reduce on resistance and increase channel mobility, gate capacitance and transconductance. In this context, Finfets and other non-planar GaN FETs are better candidates for ULV operation, as they can achieve E-mode operation due to lateral gates, rather than using a p-GaN layer. Heterojunction GaN Finfets retain the 2DEG, thus reducing the on resistance, increasing channel mobility, transconductance and gate capacitance.

There is a significant knowledge gap which exists regarding the designs of such GaN Finfets, particularly for the above mentioned ULV applications. E-mode operation in Finfets can be achieved by varying the critical design parameters such as the fin width. However, there is a scarcity of models which predict the transport behavior of these GaN Finfets. A unified low-field transport model for GaN Finfets can thus allow one to predict channel mobility, transconductance, on resistance and gate capacitance, all of which are key design features of ULV GaN device technology. With the increasing market of

wearables and mobile devices, it is expected that there will be sufficiently large need for ULV high frequency GaN devices, such as those examined in this work.

1.2 Need for GaN Finfets

Although GaN has superior material properties over Silicon, the performance of commercial devices have not yet reached the material limit. Recent advances in high aspect ratio etching have led to the development of 3D nanostructures[23]. These 3D GaN nanostructures can be classified into NanoFin (NF) FETs and Nanowire (NW) FETs. GaN Finfets offer an alternative topology to realizing GaN transistors. Like Si Finfets, GaN Finfets offer significant advantages over planar devices such as improved electrostatic gate control and alleviated short channel effects. Unlike Si Finfets, GaN Finfet bodies can be made of doped GaN alone, AlGaN/GaN heterojunction or multiple heterojunctions. GaN Finfets also improve upon the HEMT disadvantages such as non-linearity[24] and poor on/off characteristics[25][26].

FinFETs have significant advantages over their HEMT counterparts. Figure 1.3 shows the superiority of FinFETs over HEMTs in power transistor applications. One can see that vertical GaN FinFETs have significantly lower specific on resistance and possess BVs of over 1000 V. For RF applications, devices are used as amplifiers in the saturation regime, which requires high linearity. Lateral nanowire FETs display a flatter transconductance curves compared to the bell curves for planar HEMTs[10]. This is primarily due to the presence of sidewall channels in lateral Finfets. The broad transconductance plot preserves linearity and, thus, will have high f_T across a wide range of gate biases[27]. Nanowire and Nanofin FET devices are preferable for large signal

operation of GaN transistors over complex circuit and material level approaches. The current distribution in non-planar FETs also aids in better heat management.

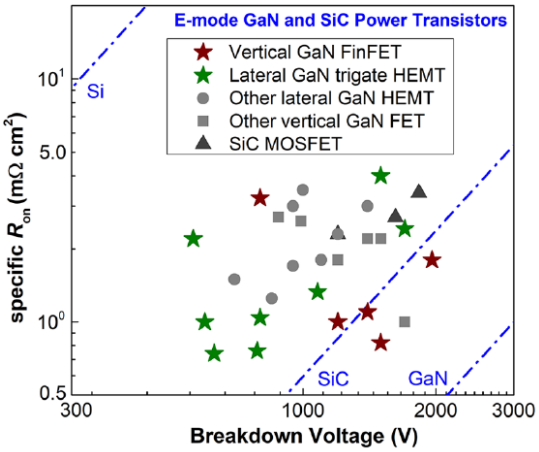

Figure 1.3 R_{ON} vs BV trade-off for lateral and vertical GaNFETs compared to HEMTs [10]. © IOP Publishing. Reproduced with permission. All rights reserved

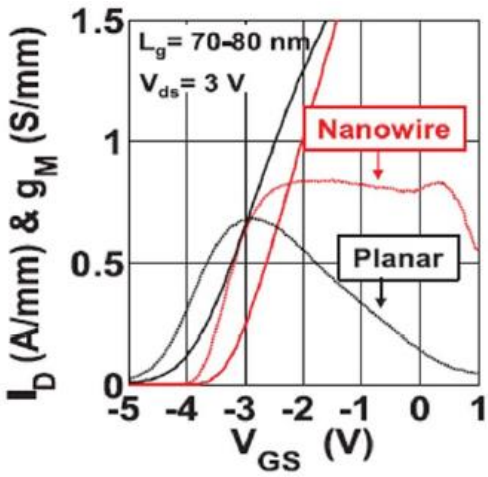

Figure 1.4 Transfer characteristics and transconductance plots of Nanowire FETs (red curves) and planar HEMT devices (black curves). © IOP Publishing. Reproduced with permission from [10]. All rights reserved

NanoFin FETs can be subdivided into Lateral and Vertical NF FETs. Lateral devices are used in Low voltage (LV) and Medium Voltage (MV) applications, such as RF devices, whereas vertical devices are favored for power devices due to the presence of a large die area.

1.2.1 Lateral and Vertical NanoFin FETs

Lateral Finfets can be heterojunction free or consist of an AlGaN/ GaN heterojunction. The conduction mechanism in the heterojunction free FET is through the bulk GaN. FinFETs allow E-mode operation by varying the fin width. This is different from other methods of achieving E-mode in GaN FETs, such as recess gate and p-GaN gate, in that the 2DEG is still present in the FinFETs, thus preserving the excellent transport properties. The heterojunction free GaNFET has better off-state characteristics, however, the AlGaN/GaN NF MISFET has a lower on resistance due to the presence of the 2DEG layer.

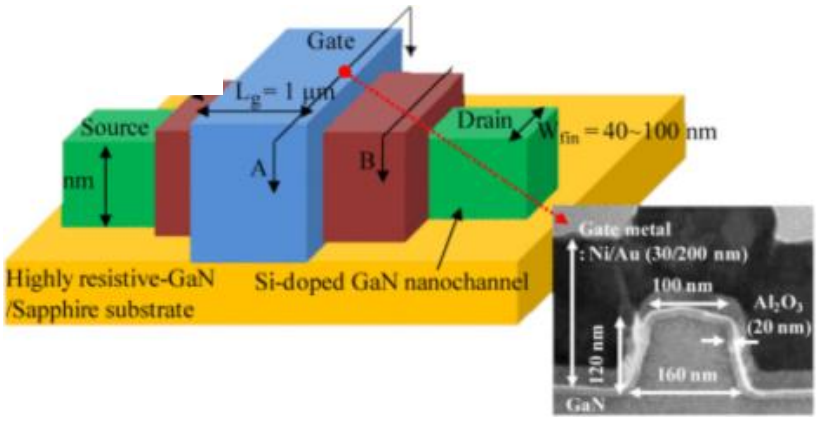

Figure 1.5 A GaN nanoFin FET without heterojunction. The conduction path is through the bulk and sidewalls. © 2013 IEEE. Reprinted with permission, from[28].

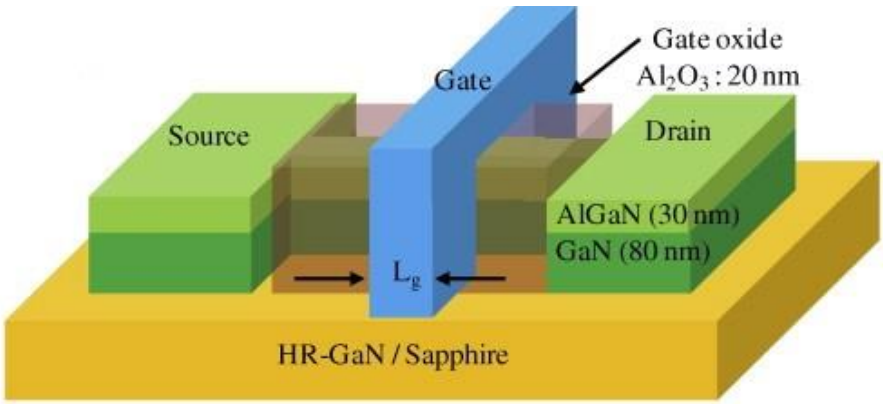

Figure 1.6 The AlGaN/GaN nanoFin MISFET. Reprinted from [29] with permission from Elsevier.

The device structure under study in this work is the AlGaN/GaN MISFET structure shown in figure 1.6. The fin channel is made up of an 80nm thick *n*-doped GaN layer grown over a Sapphire or a high resistive GaN substrate [29]. A 30 nm *n*-doped AlGaN layer with Aluminum composition (20-30%) is grown on top of the GaN layer. An Al_2O_3 gate oxide layer then wraps the fin on all three sides. The width of the fin varies from 20-50 nm (narrow fin device) to up to 100s of nm (wide-Fin device). Typically, many of these fin channels may exist as parallel nanoribbons between the same source and drain ohmic contacts [30].

Vertical Finfets are characterized by the presence of a thick drift region and a vertical conduction channel. The small area of the fin-channels enables higher current density in the channel as compared to conventional power MOSFETs. Heterojunctions are not typically present in vertical FinFETs. There are two approaches to designing a vertical FinFET depending on the gate stack: the Junction FET and the MOS FET approach (see figure 1.7). While P-N junctions in Fin-JFETs allow for stronger depletion of the vertical channel leading to a positive threshold voltage, the sidewall channels in the

MOS-Fin devices decrease the on resistance. Vertical devices are, thus, suitable for MV and HV applications. Vertical GaN power MOS-Finfets have been shown to achieve 1.2 kV breakdown voltages [31]–[33].

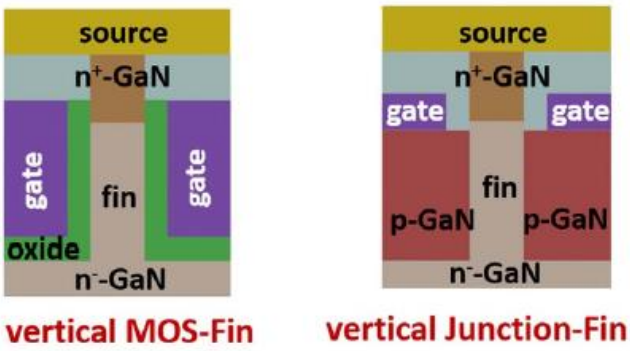

Figure 1.7 Two different implementations of the vertical Finfet (cross-section shown). MOS-FinFET (left) and the JFET (right). © IOP Publishing. Reproduced from [10] with permission.

1.2.2 Temperature Dependent Characteristics of Lateral GaN Finfets

Ki-Sik Im *et al.* [34] in their work describe the variation of threshold voltage and the electron mobility of an AlGaN/GaN MIS FinFET with temperature. The results are peculiar in that the temperature dependence reverses as one shifts from a negative to a positive gate bias. The authors compare Id-Vg characteristics at different temperatures from 100-350K for the MIS FinFET with fin width of 400 nm (wide fin) and for a MIS FinFET with a fin width of 50 nm (narrow fin). In narrow fin devices, one notes the presence of two different regimes: $Vg < 0$ and $Vg > 0$. The transconductance has a negative temperature coefficient for the former and a positive temperature coefficient for the latter. The focus of this work is to develop and implement theoretical model that will allow us to understand this temperature dependence in the negative gate bias regime.

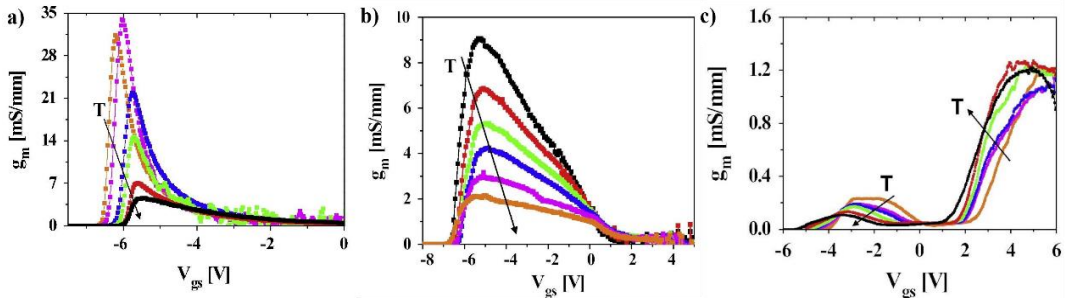

Figure 1.8 Experimental measurements of transconductance values for Finfets and planar MISFETs [31]. a) Transconductance vs. temperature variation for a planar MISFET. b) Transconductance vs. temperature variation for 400nm wide Finfet. c) Transconductance vs. temperature variation for 50nm wide Finfet. Reprinted from [34] © 2016 with permission from Elsevier.

1.3 Previous Work in Mobility modeling of nanoscale devices

The low field electron mobility is a critical transport property which is responsible for the drift component of the current. In nanoscale devices, the reduction in dimensionality from 3D to 2D and 1D changes the mobility of the system. Rode's method [35] and the Kubo-Greenwood formalism are some of the common methods used to determine the electron mobility[36], [37]. Monte Carlo approaches for modeling electron mobility have been employed for GaN nanowires[38], [39] and Silicon nanowires[40]–[43].

1.3.1 Determining electron mobility using the relaxation time approximation (Kubo-Greenwood approach)

To understand how electron mobility can be calculated from Kubo-Greenwood Formula, one needs to begin from the relaxation-time approximation of the Boltzmann

transport equation. The Boltzmann transport equation, in the effective mass approximation and for parabolic bands, is written as

$$\frac{df}{dt} + \mathbf{v}\cdot\nabla_{\mathrm{r}} f + \mathbf{F}\cdot\nabla_p f = \left.\frac{\partial f}{\partial t}\right|_{scatt} \tag{1.1}$$

If one assumes steady state conditions and a homogeneous material, the first two terms on the left-hand side vanish. For small electric fields, the distribution function is expanded as

$$f(k) = f_o(k) + g(k) \tag{1.2}$$

Here $g(k)$ is the perturbation of the distribution function, i.e. the deviation of the distribution function from its equilibrium value.

At **low electric fields** and if the scattering process is either **elastic or isotropic,** the right hand side of the above equation can be written as

$$\left.\frac{\partial f}{\partial t}\right|_{scatt} = -\frac{f_A}{\tau(k)} = -\frac{f - f_o}{\tau(k)} \tag{1.3}$$

Thus,

$$\frac{-eE}{\hbar}.\nabla_k f = -\frac{f - f_o}{\tau(k)} \tag{1.4}$$

$$f(k) = f_o + \frac{\tau(k)eE}{\hbar}.\nabla_k f_o \tag{1.5}$$

The current density can be extracted from the first moment of the distribution function as

$$\mathbf{J}(\boldsymbol{r}, t) = -\frac{e}{V}\sum_{k} \mathbf{v}(\mathbf{k}) f(\mathbf{r}, \mathbf{k}, \mathrm{t}) \tag{1.6}$$

Using the integral form of equation 1.6, one has (for nanowires)

$$\mathbf{J} = -e\frac{2}{2\pi}\int f(\mathbf{k})\mathbf{v}(\mathbf{k})dk \tag{1.7}$$

Substituting $f(k)$ from equation 1.5 leads to

$$\mathbf{J} = -e\frac{2}{2\pi}\int\left(f_o + \frac{\tau(k)eE}{\hbar}\frac{\partial f_o}{\partial \epsilon}\frac{\partial \epsilon}{\partial k}\right)\mathbf{v}(\mathbf{k})dk \tag{1.8}$$

The derivative of the equilibrium Fermi-Dirac distribution function with respect to energy evaluates to $\frac{\partial f_o}{\partial \epsilon} = \frac{-1}{K_B T} f_o(\epsilon)(1 - f_o(\epsilon)$

Thus, equation 1.7 reduces to

$$\mathbf{J} = \frac{e}{K_B T}\frac{2}{2\pi}\int f_o(\epsilon)(1 - f_o(\epsilon))\left(\frac{\tau(\epsilon)eE}{\hbar}\right)\mathbf{v}(\boldsymbol{\epsilon})d\epsilon \tag{1.9}$$

For quasi-1D system, expressing the velocity in terms of the kinetic energy $(\epsilon - \epsilon_n)$, where ϵ_n is the nth subband energy, gives for the current density of the nth channel the following result:

$$\mathbf{J_n} = \frac{e}{K_B T}\frac{2}{2\pi}\sqrt{\frac{2}{m^*}}\int f_o(\epsilon)(1 - f_o(\epsilon))\left(\frac{\tau(\epsilon)eE}{\hbar}\right)\sqrt{\epsilon - \epsilon_n}d\epsilon \tag{1.10}$$

Since there is no spatial gradient in the distribution function, the current consists only of the drift component, thus

$$\mathbf{J_n} = N_l^n e\mu_i E = \frac{e}{K_B T}\frac{2}{2\pi}\sqrt{\frac{2}{m^*}}\int f_o(\epsilon)(1 - f_o(\epsilon))\left(\frac{\tau(\epsilon)eE}{\hbar}\right)\sqrt{\epsilon - \epsilon_n}d\epsilon \tag{1.11}$$

Where, N_l^i and μ_i are the line density and the electron mobility of the ith subband, respectively. Hence

$$\mu_n = \frac{e}{K_B T N_l^n}\frac{2}{2\pi}\sqrt{\frac{2}{m^*}}\int f_o(\epsilon)(1 - f_o(\epsilon))\left(\frac{\tau(\epsilon)}{\hbar}\right)\frac{\epsilon - \epsilon_n}{\sqrt{\epsilon - \epsilon_n}}d\epsilon \tag{1.12}$$

The 1-D density of states expression is of the form:

$$g_{1D}(\epsilon) = \sqrt{\frac{m^*}{2}}\frac{1}{\pi\hbar}\frac{1}{\sqrt{\epsilon - \epsilon_n}} \tag{1.13}$$

Using equation 1.13 in 1.12, gives

$$\mu_n = \frac{2e}{K_B T N_l^n m^*}\int g_{1D}(\epsilon) f_o(\epsilon)(1 - f_o(\epsilon))\tau(\epsilon)(\epsilon - \epsilon_n)d\epsilon \tag{1.14}$$

Equation 1.14 can be used to determine the electron mobility assuming that the relaxation-time approximation holds. This is also referred to as Kubo-Greenwood formalism in the literature [44], [45]. Kotliyar et al. [46] and Fischetti et al. [47] use this method to determine phonon-limited electron mobility in silicon nanowires. Kumar et al. [36] use Kubo-Greenwood formalism to determine the electron mobility for an AlGaN/GaN split gate FET. The mobility is computed by calculating the momentum relaxation time for each scattering mechanism and the 1D electron density. 1D scattering rate expressions are used to determine the momentum relaxation rate.

1.3.2 Determining Electron Mobility using the Rode's iterative method

If the scattering process is inelastic and anisotropic, as in the case of polar materials, the RTA cannot be used since it is impossible to define a simple relaxation time $\tau_f(\epsilon)$ that does not depend upon the distribution function [48]. Returning to equation 1.2,

$$f_n(k) = f_{n_o} + g_n(k)$$

The BTE now becomes,

$$\frac{-eE}{\hbar}.\nabla_k f_n = \left.\frac{\partial f}{\partial t}\right|_{scatt} = \sum_{m,k'}[S_{k'k} f_m(1 - f_n) - S_{kk'} f_n(1 - f_m)] \tag{1.15}$$

Substituting $f_n(k)$ in equation 1.15 leads to

$$\frac{-eE}{\hbar}\frac{\partial}{\partial k}\left(f_{n_o}+g_n(k)\right)=$$
$$\sum_{m,k'}\left[S_{k'k}(f_{m_o}+g_m)\left(1-f_{n_o}-g_n\right)-S_{kk'}(f_{n_o}+g_n)\left(1-f_{m_o}-g_m\right)\right] \qquad (1.16)$$

Please note that we use the short-hand notation $f_m = f_m(k')\ and\ f_n = f_n(k)$.

The principle of detailed balance states that equilibrium scattering rate from state $k\ to\ k'$, from subband $m\ to\ n$ equals the scattering rate for the reverse process. In other words, under equilibrium conditions the change of the distribution function due to scattering processes has to be zero. The principle of detailed balance reads:

$$\frac{S_{k'k}}{S_{kk'}}=\frac{f_{n_o}\left(1-f_{m_o}\right)}{f_{m_o}\left(1-f_{n_o}\right)}$$

Using the principle of detailed balance, and assuming that higher order terms of $g_n(k)$ are negligible, equation 1.16 reduces to,

$$\frac{-eE}{\hbar}\frac{\partial}{\partial k}\left(f_{n_o}\right)=\sum_{m,k'}S_{k'k}\left[\left(1-f_{n_o}\right)g_m-f_{m_o}g_n\right]-$$
$$\sum_{m,k'}S_{kk'}\left[\left(1-f_{m_o}\right)g_n-f_{n_o}g_m\right] \qquad (1.17)$$

Expressing in terms of g_n gives

$$g_n=\frac{\frac{-eE}{\hbar}\frac{\partial}{\partial k}\left(f_{n_o}\right)+\sum_{m,k'}S_{k'k}\left[1-f_{n_o}g_m\right]+g_m f_{n_o}S_{kk'}}{\sum_{m,k'}\left[f_{m_o}S_{k'k}+S_{kk'}\left(1-f_{m_o}\right)\right]} \qquad (1.18)$$

Equation 1.18 is the Rode's iterative method. The equation is solved iteratively and converges in a few iterations, yielding the value of perturbation $g_n(k)$.

Returning to equation 1.7, the current density is expressed as

$$\mathbf{J}=-e\frac{2}{2\pi}\int f(k)\mathbf{v}(\mathbf{k})dk$$

$$\mathbf{J} = -e\frac{2}{2\pi}\int (f_o + g_n(k))\mathbf{v}(\mathbf{k})dk \tag{1.19}$$

or

$$\mathbf{J} = -\frac{em^*}{\pi\hbar\sqrt{2m^*}}\int (g_n(k))\mathbf{v}(\mathbf{k}))/\sqrt{\epsilon - \epsilon_n}\ d\epsilon \tag{1.20}$$

Using the 1D Density of states expression (Equation 1.13), one can write

$$\mathbf{J} = -e\int (g_n(\epsilon))\mathbf{v}(\boldsymbol{\epsilon})g_{1D}(\epsilon))\ d\epsilon = \ N_l^n e\mu_n E \tag{1.21}$$

Expressing velocity in terms of energy, the expression for electron mobility becomes

$$\mu_n = \frac{1}{N_l^n E}\sqrt{\frac{2}{m^*}}\int g_n(\epsilon)\sqrt{\epsilon - \epsilon_n}\, g_{1D}(\epsilon))\ d\epsilon \tag{1.22}$$

The total mobility is then a weighted sum of the subband mobility, i.e.:

$$\mu = \frac{\sum_n \mu_n N_l^n}{\sum_n N_l^n} \tag{1.23}$$

Equation 1.22 gives the electron mobility determined using Rode's Iterative method. Rode's iterative method has been used to model the low field mobility in InAlAs solar cells [49] as well as nanowires.

Although Rode's method and the relaxation time approximation approach are effective ways to determine the low-field electron mobility, Monte Carlo methods provide a very generic framework to solve the Boltzmann Transport Equation. Knezevic et al. [38] determine the electron mobility of rectangular GaN nanowires using Ensemble Monte Carlo simulations. Monte Carlo method also enables one to build a device simulator and generate device performance characteristics.

1.4 Motivation for this work

There are several reasons which form the motivation for this work:

a) To understand the peculiar temperature dependence of the transconductance in lateral Finfets as a function of bias and width.

b) To calculate I-V characteristics, one would need to couple 3D Poisson with 2D Schrödinger solved in slices (from source to drain). These are next coupled with a quasi-1D Monte Carlo transport solver.

Since the focus is on the temperature dependence of transconductance in this work, which comes from the transfer characteristics of the device at low drain bias, one can assume low-field transport and, thus, low-field mobility is the desired quantity to be determined. For that purpose, one needs to solve coupled Schrödinger-Poisson-Monte Carlo problem illustrated below.

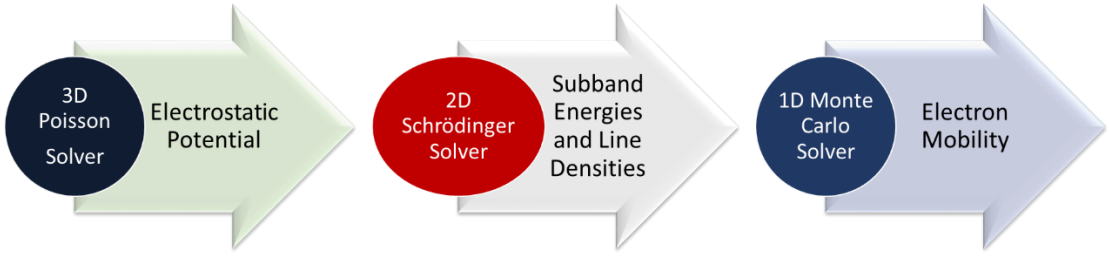

Figure 1.9 The theoretical modeling approach used in this work.

Solution of Schrödinger-Poisson Problem for heterostructures has been done by solving 3D Poisson Equation coupled with a 2D Schrödinger Solver. For the solution of the Poisson equation, BiCGSTAB method is used and implemented in MATLAB. For the solution of the Schrödinger equation, this work used canned Eigenvalue solvers in MATLAB. In the Boltzmann solver, this work considered quasi-1D system and included

the following scattering mechanisms; acoustic phonon scattering, piezoelectric scattering, polar optical phonon scattering and Interface roughness scattering. The impact of alloy clustering is also studied in details.

1.5 Outline of the dissertation

The technical details for the solution of the 3D Poisson Equation coupled with a 2D Schrödinger equation are given in chapter 2. Chapter 3 of the dissertation contains details on the choice of the band-structure and the implementation of the quasi-1D Monte Carlo transport solver. The simulation results for the MIS FinFET devices introduced in section 1.2 are presented in the latter half of chapter 3. At the end we present conclusive comments related to the work accomplished so far and discus prospective future work.

CHAPTER 2

MODELING ELECTROSTATICS

In a typical FinFET device, the charge densities and the material compositions vary along the length, along the fin width and along the device thickness. Therefore, to accurately model the electrostatics, it is necessary to couple a 3D Poisson equation solver and a 2D Schrödinger solver and solve them self-consistently. The following section introduces the Poisson equation and discusses efficient solution of the same for the case of 3D problems. The solution of the 2D Schrödinger equation is discussed next. Finally, self-consistent results of the solution of the 3D Poisson-2D Schrödinger problem are presented, and the overall convergence of the self-consistent scheme is discussed.

2.1 3D Poisson Equation Solver

The solution of the 3D Poisson Equation, described by equation 2.1, yields the potential profile, and, thus, the charge densities in a given region.

$$\nabla \cdot \left(\epsilon(\boldsymbol{r})\nabla\phi(\mathbf{r})\right) = -\rho(\boldsymbol{r}) \tag{2.1}$$

2.1.1 Linearization and discretization of the Poisson Equation

Under equilibrium conditions and non-degenerate statistics, the charge density term on the right-hand-side (RHS) of equation 2.1 is expressed as

$$\rho(\phi) = q(p_o - n_o + Dop) = q(n_i \exp\left(\frac{-\phi}{V_T}\right) - n_i \exp\left(\frac{\phi}{V_T}\right) + Dop)$$

i.e., the Poisson equation can be written as

$$\nabla(\epsilon\nabla\phi) = -\frac{\rho(\phi)}{\epsilon_o}. \tag{2.2}$$

The linearization of the RHS of equation 2.2, and the finite difference discretization of the resultant equation, leads to the following system of linear equations:

$$F_{ijk}\phi_{i,j,k-1} + N_{ijk}\phi_{i,j-1,k} + E_{ijk}\phi_{i-1,j,k} + C_{ijk}\phi_{ijk} + W_{ijk}\phi_{i+1,j,k} + S_{ijk}\phi_{i,j+1,k} + B_{ijk}\phi_{i,j,k+1} = \not{b}_{ijk} \quad (2.3)$$

The off-diagonal coefficients are given by:

$$N_{ijk} = \frac{\epsilon_{i-1,j-1,k} + \epsilon_{i-1,j-1,k-1} + \epsilon_{i,j-1,k-1} + \epsilon_{i-1,j-1,k}}{2Y_{j-1}(Y_j + Y_{j-1})}$$

$$S_{ijk} = \frac{\epsilon_{i-1,j,k} + \epsilon_{i-1,j,k-1} + \epsilon_{i,j,k-1} + \epsilon_{i-1,j,k}}{2Y_j(Y_j + Y_{j-1})}$$

$$E_{ijk} = \frac{\epsilon_{i-1,j-1,k} + \epsilon_{i-1,j-1,k-1} + \epsilon_{i-1,j,k-1} + \epsilon_{i-1,j,k}}{2X_{i-1}(X_i + X_{i-1})}$$

$$W_{ijk} = \frac{\epsilon_{i,j-1,k-1} + \epsilon_{i,j,k-1} + \epsilon_{i,j-1,k} + \epsilon_{i,j,k}}{2X_i(X_i + X_{i-1})}$$

$$F_{ijk} = \frac{\epsilon_{i-1,j-1,k-1} + \epsilon_{i,j-1,k-1} + \epsilon_{i-1,j,k-1} + \epsilon_{i,j,k-1}}{2Z_{k-1}(Z_k + Z_{k-1})}$$

$$B_{ijk} = \frac{\epsilon_{i-1,j-1,k} + \epsilon_{i,j-1,k} + \epsilon_{i-1,j,k} + \epsilon_{i,j,k}}{2Z_k(Z_k + Z_{k-1})}$$

The central coefficient is calculated using

$$C_{ijk} = -\left(B_{ijk} + F_{ijk} + E_{ijk} + W_{ijk} + N_{ijk} + S_{ijk}\right) - \left(n_{ir}\exp\left(\frac{\phi_{ijk}^{old}}{V_T}\right) + n_{ir}\exp\left(-\frac{\phi_{ijk}^{old}}{V_T}\right)\right)$$

and the forcing function is given by

$$f_{ijk} = V_T\left(n_{ir}\exp\left(\frac{\phi_{ijk}^{old}}{V_T}\right) - n_{ir}\exp\left(-\frac{\phi_{ijk}^{old}}{V_T}\right) - Dop_{ijk}\right)$$
$$- \phi_{ijk}^{old}\left(n_{ir}\exp\left(\frac{\phi_{ijk}^{old}}{V_T}\right) + n_{ir}\exp\left(-\frac{\phi_{ijk}^{old}}{V_T}\right)\right)$$

Here, $i, j, and\ k$ are the nodes along the x, y and z directions respectively. ϕ_{ijk}^{old} corresponds to the old (previous) solution.

2.1.2 Modifying the Poisson Equation for Heterostructures

For the case where the material parameters are inhomogeneous and position dependent, such as for the case of heterostructures, the expressions for the electron and hole concentrations, and for the electrostatic potential need to be modified. If the density of states function is position dependent, the electron concentration also becomes position dependent and is calculated using

$$n(x) = \int_{E_C}^{E_C^{Top}} g_c(E - E_c, x) f_{FD}\big(E, E_{fn}\big) dE \tag{2.4}$$

For a non-degenerate semiconductor, Boltzmann statistics can be used, for which

$$n(x) = n_i \exp\left(\frac{F_n - E_I}{\kappa T}\right) \tag{2.5}$$

If ϕ_n is the quasi-Fermi level in the semiconductor, then $F_n = E_F - \phi_n$. At equilibrium, $\phi_n = \phi_p = 0$ and $F_n = F_p = E_F$.

To accurately express the intrinsic Fermi level in heterostructures in devices with position dependent material parameters, one needs to adopt the local vacuum level

approach [50], [51]. The local vacuum level must follow electrostatic potential within the device. In this approach, all energy levels are defined with respect to the reference level (E_o), thus

$$E_C(x) = E_o - q\phi(x) - \chi(x) \tag{2.6}$$

$$E_V(x) = E_o - q\phi(x) - \chi(x) - E_G\,(x) \tag{2.7}$$

$$E_I(x) = E_o - \chi(x) - q\phi(x) - \frac{E_G(x)}{2} + \frac{\kappa T}{2}\log\left(\frac{N_V(x)}{N_C(x)}\right) \tag{2.8}$$

Defining a reference location in the heterostructure [52][53] and writing the electron (and hole) concentrations in equation 2.5 as

$$n(x) = n_{iref}\exp\left(\frac{E_F - E_I}{\kappa T}\right)\exp\left(\log\left(\frac{n_i}{n_{iref}}\right)\right) = n_{iref}\exp\left(\frac{E_F - E_I + \kappa T\log\left(\frac{n_i}{n_{iref}}\right)}{\kappa T}\right) \tag{2.9}$$

Substituting Equation 2.8 into 2.9, gives

$$n(x) = n_{iref}\exp\left(E_F - E_o + \chi(x) + \kappa T\log\left(\frac{N_c(x)}{n_{iref}}\right) + \phi(x)\right) \tag{2.10}$$

Let us define the electron concentration as

$$n(x) = n_{iref}\exp\left(\frac{q}{\kappa T}(\phi(x) + V_n(x)\right) \tag{2.11}$$

And the hole concentration as

$$p(x) = n_{iref}\,exp\left(-\frac{q}{\kappa T}\left(\phi(x) - V_p(x)\right)\right)$$

Then, the band parameters V_n and V_p are calculated using

$$V_n = \frac{\chi(x)}{q} - E_0 + \frac{\kappa T}{q}\log\left(\frac{N_c(x)}{n_{iref}}\right) \tag{2.12}$$

$$V_p = -\frac{\chi(x)}{q} - \frac{E_G}{q} + E_0 + \frac{\kappa T}{q}\log\left(\frac{N_v(x)}{n_{iref}}\right)$$

One can choose the reference energy level E_0 such that the electrostatic potential ϕ = 0 in the reference material [54] (the reference material being intrinsic). Thus,

$$E_o = \chi_{ref} + E_F + \kappa T \ln\left(\frac{N_{cref}}{n_{iref}}\right) \tag{2.13}$$

Using this result, the heterostructure band parameters become

$$V_n = \frac{\chi - \chi_{ref}}{q} + \left(\frac{\kappa T}{q}\log\frac{N_c(x)}{N_{cref}}\right) \tag{2.14}$$

$$V_p = -\frac{(\chi - \chi_{ref})}{q} - \frac{\left(E_G - E_{Gref}\right)}{q} + \left(\frac{\kappa T}{q}\log\frac{N_V(x)}{N_{Vref}}\right)$$

If Fermi-Dirac statistics are used, similar analysis [55] can lead to the following modification in the band parameters

$$V_n = \frac{\chi - \chi_{ref}}{q} + \left(\frac{\kappa T}{q}\log\frac{N_c(x)}{N_{cref}}\right) + \frac{\kappa T}{q}\ln\left(\frac{F_{\frac{1}{2}}(\eta_c)}{\exp(\eta_c)}\right) \tag{2.15}$$

$$V_p = -\frac{\left(\chi - \chi_{ref}\right)}{q} - \frac{\left(E_G - E_{Gref}\right)}{q} + \left(\frac{\kappa T}{q}\log\frac{N_V}{N_{Vref}}\right) + \frac{\kappa T}{q}\ln\left(\frac{F_{\frac{1}{2}}(\eta_v)}{\exp(\eta_v)}\right)$$

Where, $\eta_c = \frac{E_F - E_C}{\kappa T}$.

2.1.3 Boundary conditions for Schottky and Ohmic contacts

Schottky contacts

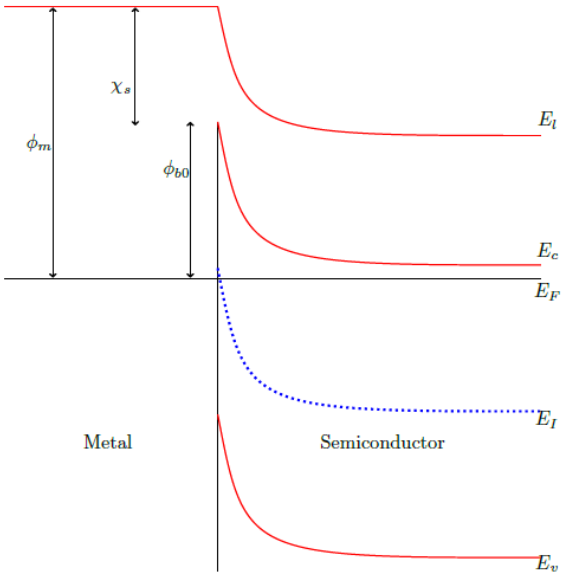

Figure 2.1 Energy band diagram at a Schottky interface.

Referring to fig 2.1, the Schottky barrier height ϕ_{b0} can be given as $\phi_{b0} = E_C(x) - E_F$.

Note also that

$$E_F - E_I(x) = E_F - E_C(x) + E_c(x) - E_I(x)$$

$$= -\phi_{b0} + \frac{E_G(x)}{2} - \frac{\kappa T}{2}\ln\left(\frac{N_V(x)}{N_C(x)}\right) \qquad (2.16)$$

Substituting $E_I(x)$ and E_0 from equations (2.8) and (2.13) into the LHS of equation (2.16), one arrives at the following boundary condition for the potential ϕ at the Schottky contact

$$q\phi(x) = -\phi_{b0} + \left(\chi_{ref} - \chi(x)\right) + \kappa T \log\left(\frac{N_{cref}}{n_{iref}}\right) \qquad (2.17)$$

$$q\phi(x) = -\left(\phi_m - \chi(x_{int})\right) + \left(\chi_{ref} - \chi(x_{int})\right) + \frac{\kappa T}{2}\log\left(\frac{N_{Cref}}{N_{Vref}}\right) + \frac{E_G(x_{int})}{2} \qquad (2.18)$$

Ohmic contacts

At Ohmic contacts, Dirichlet boundary conditions are used for the electrostatic potential (ϕ). Invoking charge neutrality conditions at the Ohmic contact, gives

$$\mathrm{p} - \mathrm{n} + \mathrm{Dop} = 0$$

with $\mathrm{Dop} = \mathrm{N_D} - \mathrm{N_A}$. Substituting the results given in equation (2.11) leads to

$$n_{iref} \exp\left(\frac{q}{\kappa T}(-\phi + V_p)\right) \; - \; n_{iref} \exp\left(\frac{q}{\kappa T}(\phi + V_n)\right) + Dop = 0 \qquad (2.19)$$

Solving for ϕ, gives

$$\phi(\mathrm{x}) = \frac{\mathrm{V_p} - \mathrm{V_n}}{2} + \mathrm{sgn(D)}\frac{\mathrm{KT}}{\mathrm{q}} \ln\left(\frac{\left|\frac{\mathrm{Dop}}{2\mathrm{n_{iref}}}\right| \exp\left(-\frac{\mathrm{q}\left(\mathrm{V_p(x)} + \mathrm{V_n(x)}\right)}{2\mathrm{KT}}\right) + \sqrt{\frac{1}{4}\left(\frac{\mathrm{Dop}}{\mathrm{n_{iref}}}\right)^2 \exp\left(-\frac{\mathrm{q}\left(\mathrm{V_p(x)} + \mathrm{V_n(x)}\right)}{2\mathrm{KT}}\right) + 1}}\right) \qquad (2.20)$$

2.1.4 Introducing polarization-induced charge density into the Poisson Equation

Origin of spontaneous and piezoelectric polarization in III-nitrides

Group III-nitride materials naturally occur as a Wurtzite crystal structure with lattice constants a_0 and c_0 along the base hexagon and along the height, respectively. Since the growth along the c plane {0001} is a highly mature standard, the discussion presented here will be restricted to polarization in the c plane.

The atoms in group III-nitrides are arranged in closely spaced hexagonal bilayers with alternating anion and cation layers creating a dipole moment. Due to the ionic nature of the Ga-N bond, the positive and negative charge centers do not coincide, leading to the formation of a dipole. This forms the basis of spontaneous polarization present in the c plane of Ga (Al) N. The spontaneous polarization is positive in the direction from Ga (Al) to N atom.

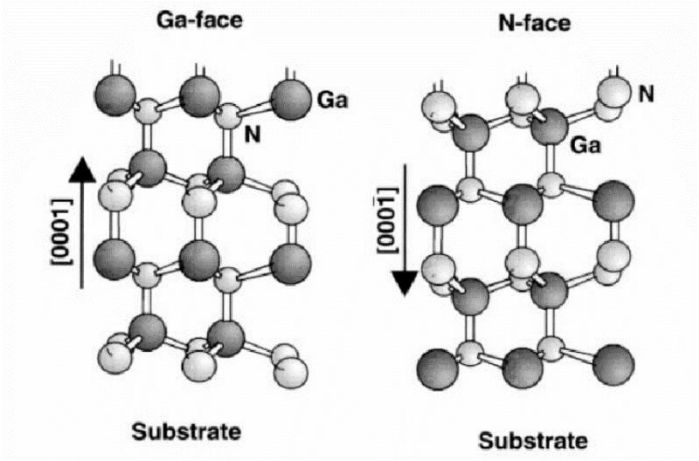

Figure 2.2 Crystal structure of wurtzite GaN indicating the Ga –face and N- face crystals along growth direction of [0001] and [000$\bar{1}$] respectively[56].

Piezoelectric polarization on the other hand arises from the dipole moment created in response to stress applied to the crystal. The stress arises typically due to lattice mismatch between different materials in the growth layers. Piezoelectric polarization along the c-axis ([0001]) is given by the contribution of piezoelectric tensors and the in-plane and c-plane strain components [15].

$$P_{pe} = e_{33}\epsilon_z + e_{31}(\epsilon_x + \epsilon_y) \tag{2.21}$$

Where, e_{33} and e_{31} are piezoelectric coefficients and ϵ_x, ϵ_y and ϵ_z are strain components along the basal (a) plane and along the c plane, respectively. The respective strain components are given by

$$\epsilon_x = \epsilon_y = \frac{a_s - a_0}{a_0} \tag{2.22}$$

$$\epsilon_z = \frac{c - c_0}{c_0}$$

The unstrained lattice constants are given by a_0 and c_0, and satisfy the following relationship

$$\frac{c - c_0}{c_0} = -2\frac{C13}{C33}\frac{a_s - a_0}{a_0} \tag{2.23}$$

where, C_{13} and C_{33} are the elastic constants. Using equations 2.22 and 2.23 into 2.21, the piezoelectric polarization expression becomes,

$$P_{pe} = 2\frac{a_s - a_0}{a_0}\left(e_{31} - \frac{C_{13}}{C_{33}}e_{33}\right) \tag{2.24}$$

The strained lattice constant a_s is determined by averaging the lattice constants of the layers directly above and below the strained layer. This approach is common in commercial device simulators such as Silvaco [57]. In case of a tensile strain, the piezoelectric polarization is parallel to the spontaneous polarization (for Ga- face crystals) since $e_{31} - \frac{C_{13}}{C_{33}}e_{33}$ is always negative. Piezoelectric polarization is antiparallel to spontaneous polarization if the layer in question is under a compressive strain.

The AlGaN/GaN MISFET presented here is grown on the polar c-plane ([0001] direction). Thus, the spontaneous polarization for both AlGaN and GaN layer and the piezoelectric polarization for the AlGaN layer need to be accounted for in the model.

Strain relaxation in nano-FinFETs In nanostructures, such as finfets and nanoribbons[58], it has been observed that the heterostructure layers grown epitaxially tend to relax, especially if the lateral dimensions (for example, the fin width) is less than 100nm. This directly results in a reduction of the piezoelectric polarization charge density. To account for this phenomenon, one needs to modify equation 2.24 by introducing a 'strain relaxation parameter' (the value of which is set to 1 if there is no strain relaxation).

Total polarization charge in a layer and coupling to the total charge density

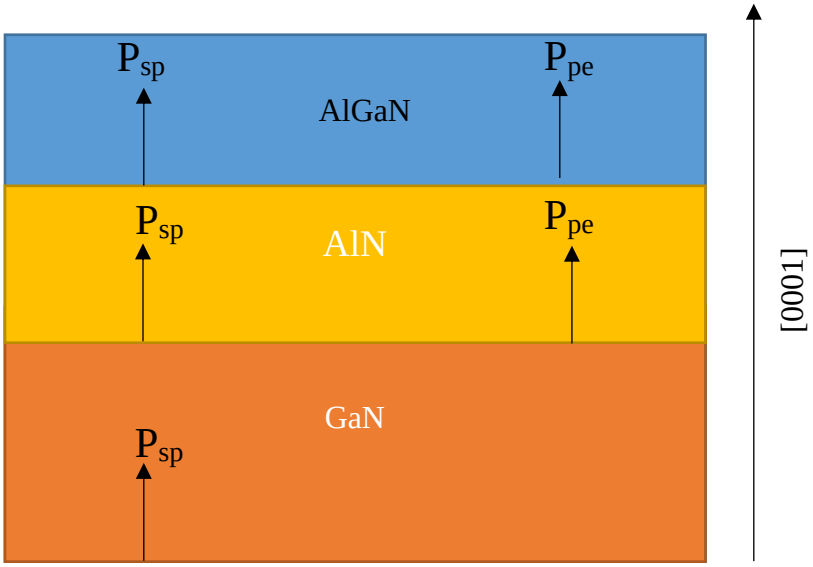

Figure 2.3 Polarization charge densities in an AlGaN/AlN/GaN heterostructure

Consider an AlGaN/AlN/GaN heterostructure system presented in figure 2.3. The total polarization charge density in a given layer is the sum of the spontaneous and the piezoelectric polarization charge densities in that layer, i.e.

$$\boldsymbol{P_{tot}} = P_{sp} + P_{pe} \tag{2.25}$$

The bottom GaN layer is relaxed and hence is absent of any piezoelectric polarization. The AlN and AlGaN layers are under tensile strain and the total polarization

in those layers is a sum of piezoelectric and spontaneous polarization charge densities. The net polarization charge density that appears at the interfaces between the two layers is

$$\sigma_{pol} = \boldsymbol{P}_{\boldsymbol{tot}}^{\boldsymbol{bot}} - \boldsymbol{P}_{\boldsymbol{tot}}^{\boldsymbol{top}} \tag{2.26}$$

The polarization charge density is accounted for in the Poisson equation by adding it as a sheet charge density at the heterojunction interfaces.

2.2 Numerical solution methods for the Poisson equation

The linearized (and discretized) Poisson equation (Equation 2.3) is of the form

$$A\boldsymbol{x} = f \tag{2.27}$$

where, A is the coefficient matrix, $\boldsymbol{x}$ is the unknown potential and f is the forcing function. A variety of numerical methods exist in the literature for solving equations of the form 2.27. Numerical solution methods for Poisson's equation can be divided into ***Direct*** and ***Iterative*** methods. Direct methods include LU decomposition and Gaussian elimination, whereas Iterative methods can be ***mesh relaxation methods***, such as Jacobi, Gauss-Seidel and Successive over Relaxation (SOR), Alternating direction Implicit (ADI) or ***Matrix methods***, which consist of Sparse Matrix methods, conjugate gradient methods and Multi-Grid methods.

2.2.1 Direct Methods

Direct methods such as LU decomposition and Gaussian elimination involve matrix factorization and inversions. As such, these methods are limited to use in solution of 1D or 2D Poisson equations where the coefficient matrix is smaller in size, and the inversion of the matrix does not pose a significant computational burden. For 3D systems, the

coefficient matrix becomes too large. At that size, matrix inversion becomes computationally expensive and, hence, time consuming.

2.2.2 Iterative Methods

Since the coefficient matrix is large, but often sparse, one can utilize the power of iterative solution methods. Iterative (or relaxation) methods start with an approximation which is successively improved by the repeated iteration of the same algorithm, until a sufficient accuracy is obtained. In this way, the original approximation is "relaxed" toward the exact solution which is numerically more stable. Iterative methods are used most often for large sparse system of equations, and always when a good approximation of the solution is known. Error analysis and convergence rate are two crucial aspects of the theory of iterative methods. All iterative methods follow a similar algorithm as outlined below:

- For the equation $= f$, a sequence of approximations $\boldsymbol{v}^0, \boldsymbol{v}^1 \ldots \boldsymbol{v}^n$ converging to **x** is constructed. If $\boldsymbol{v}^i$ is the i^{th} approximation to **x**, then the residual is defined as a measure of deviation of the i^{th} approximation from **x.** The residual for the i^{th} iteration is given by

$$r^i = f - Av^i \tag{2.28}$$

- The algebraic error of the i^{th} approximation is defined by

$$e^i = x - v^i \tag{2.29}$$

- The error e^i obeys the residual equation

$$Ae^i = r^i \tag{2.30}$$

Mesh relaxation methods such as SOR and ADI methods are popular iterative methods used when the coefficient matrix is small, such as for 2D Poisson equations. One can switch to matrix methods when dealing with large sparse matrices such as those arising from discretization of the 3D Poisson Equation.

Mesh relaxation methods are slower to converge. Although the high frequency errors are damped quickly in mesh relaxation methods, convergence of the low frequency errors is quite slow. This is not the case in matrix methods, which are popular for fast convergence. This work makes use of conjugate gradient methods which fall under the category of matrix methods.

Conjugate gradient Methods: general algorithm: For all descent algorithms, the following describes the fundamental structure[59]:

- Start with an initial point determined according to a fixed rule, a direction of movement.
- The direction is toward a relative minima of the objective function.
- At this now point, a new direction is determined and the process is repeated.

The major difference between different matrix method algorithms is the rule by which successive directions of movement are selected.

Method of Steepest Descent

In the method of steepest descent, one starts with an arbitrary point x_0 and takes a series of steps x_1 and x_2 and move toward the solution. The direction of movement is toward what minimizes the forcing function most quickly. That is

$$-f'(x_i) = b - Ax(i)$$

From the definition of residuals, it follows that,

$$r_i = -Ae_i = -f'(x_i)$$

From x_0 , the next vector will fall along the solid line given by

$$x_1 = x_0 + \alpha r_0$$

Since the magnitude of the step must be such that it minimizes f along the line, then

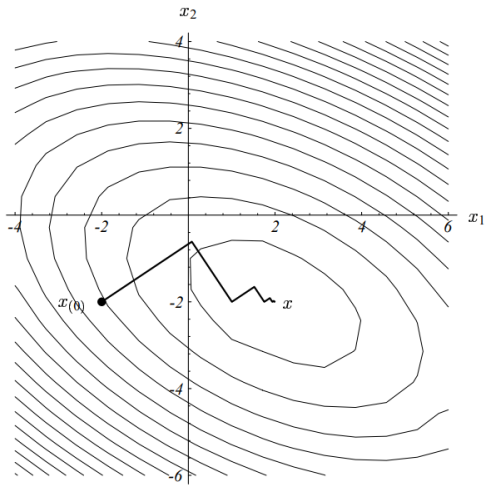

Figure 2.4 Path of the search vector x_i toward the solution x [59].

$$\frac{d}{d\alpha} f(x_1) = f'(x_1)^T \frac{dx_1}{d\alpha} = 0$$

$$-r_1^T r_0 = 0$$

$$\alpha = \frac{r_0^T r_0}{r_0^T A r_0}$$

Thus, the algorithm for steepest descent is

$$x_{i+1} = x_i + \alpha_i r_i$$

$$\alpha_i = \frac{r_i^T r_i}{r_i^T A r_i} \tag{2.31}$$

To reduce matrix multiplication, one can simplify the above as

$$r_{i+1} = r_i - \alpha_i A r_i \tag{2.32}$$

A major disadvantage of the steepest descent algorithm is that (as seen in figure 2.4) a large number of steps has to be taken before converging to the solution. This is in part because this algorithm often results in the descent steps repeating the same direction as earlier steps. This issue is resolved in the Conjugate directions and Conjugate gradient algorithms.

Method of conjugate directions

The core idea in the method of conjugate directions is that instead of steps in successive iterations repeating the direction of previous steps, the search directions ($\boldsymbol{d_0, d_1, d_{n-1}}$) will be orthogonal [59]. Taking exactly one step in each search direction to line up with the solution $\boldsymbol{x}$ ensures convergence in precisely n steps for an $n \times n$ coefficient matrix A.

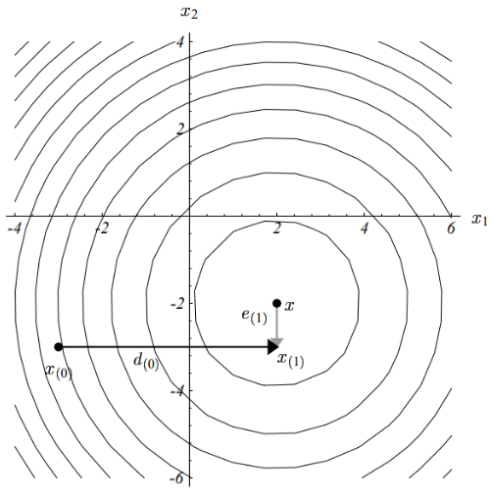

Figure 2.5 Illustration of the basic idea of conjugate directions. [59]

At each step, the next iteration of $\boldsymbol{x}$ is computed as

$$x_{i+1} = x_i + \alpha_i d_i \tag{2.33}$$

Since $\boldsymbol{e_{i+1}}$ needs to be orthogonal to $\boldsymbol{d_i}$

$$d_i^T e_{i+1} = 0$$

$$d_i^T (e_i + \alpha_i d_i) = 0$$

$$\alpha_i = -\frac{d_i^T e_i}{d_i^T d_i} \tag{2.34}$$

Since e_i is unknown, instead of demanding the search directions to be orthogonal, one can force the directions to be **A-orthogonal** to each other. In other words,

$$d_i^T A d_j = 0 \tag{2.35}$$

Equation 2.35 is a condition of A-orthogonality. Thus, the new requirement is that e_{i+1} needs to be **A-orthogonal** to d_i**.** Ensuring the minimization of $f(x_i)$, one has

$$\frac{d}{d\alpha} f(x_{i+1}) = f'(x_{i+1})^T \frac{dx_{i+1}}{d\alpha} = 0$$

Substituting $\boldsymbol{x_{i+1}}$ from equation 2.33 gives

$$-r_{i+1}^T d_i = 0$$

$$d_i^T A e_{i+1} = 0$$

$$d_i^T A(e_i + \alpha_i d_i) = 0$$

$$\alpha_i = -\frac{d_i^T A e_i}{d_i^T A d_i}$$

$$\alpha_i = -\frac{d_i^T r_i}{d_i^T A d_i} \tag{2.36}$$

Finally, one needs to compute the A-orthogonal search directions ($\boldsymbol{d_i}$). This is accomplished using the conjugate Gram-Schmidt process. The algorithm is outlined below.

- Take a set of linearly independent vectors $\boldsymbol{u_0, u_1, u_{n-1}}$, assuming that $\boldsymbol{d_o = u_0}$

- For $i > 0$, take an $\boldsymbol{u_i}$ and subtract all the components from it that are not A-orthogonal to the previous search directions.

$$\boldsymbol{d_i} = \boldsymbol{u_i} + \sum_{j=0}^{i-1} \beta_{ij} \boldsymbol{d_j}$$

$$\beta_{ij} = -\frac{\boldsymbol{u_i^T A d_j}}{\boldsymbol{d_j^T A d_j}} \tag{2.37}$$

A key point in the method of conjugate directions is that while the error vector is A-orthogonal to all previous search directions, the residual vector is orthogonal to previous search directions as well as all previous basis vectors.

Method of conjugate gradients

The method of Conjugate Gradients (CG) expands upon the conjugate directions method where the search directions are constructed by conjugation of the residuals. That is, $u_i = r_i$. This simplifies the calculation of the new search direction.

$$\beta_{ij} = \begin{cases} \dfrac{1}{\alpha_{i-1}} \dfrac{r_i^T r_i}{d_j^T A d_j} = \dfrac{r_i^T r_i}{r_{i-1}^T r_{i-1}} & i = j+1 \\ 0 & i > j+1 \end{cases}$$

The new search direction then is determined by the linear combination of the previous search direction and the new residual:

$$d_{i+1} = r_{i+1} + \beta_{i+1} d_i$$

Thus, the method of conjugate gradients is represented with the following pseudocode:

$$d_i = r_i = b - Ax_i$$

$$\alpha_i = -\frac{r_i^T r_i}{d_i^T A d_i}$$

$$x_{i+1} = x_i + \alpha_i d_i$$

$$r_{i+1} = r_i - \alpha_i A d_i$$

$$\beta_{i+1} = \frac{r_{i+1}^T r_{i+1}}{r_i^T\, r_i}$$

$$d_{i+1} = r_{i+1} + \beta_{i+1} d_i \tag{2.38}$$

For the CG method, there is one matrix-vector product, two vector dot products per iteration and a need to store four n-vectors. The CG method converges in at most n steps. However, rounding errors in floating point arithmetic can lead to loss of orthogonality among the residuals. The CG method will converge in m steps if the **A** matrix has only m distinct eigenvalues.

Methods for non-symmetric Matrices

The conjugate gradient methods discussed above are valid when the **A** matrix is symmetric. However, this is seldom the case when discretizing Poisson equations. For example, in the case of heterostructures, the coefficient matrix is asymmetric due to the differences in relative permittivity (ϵ_r) . Asymmetric coefficient matrix can also result from the discretization implemented, such as the presence of non-uniform meshing over the simulation domain. Since this work deals with heterostructure systems, the A-matrix is always asymmetric. For such cases, conjugate gradient type algorithms for non-symmetric matrices based on Krylov subspace methods have to be used.

The **Bi-Conjugate gradient (Bi-CG)** algorithm was proposed by Lanczos in 1954 and can solve the original system $\mathbf{Ax} = \mathbf{b}$ as well as the dual linear system $\mathbf{A^T x}^* = \mathbf{b}^*$**.** Each step of this algorithm involves a matrix-vector product with both **A** and $\mathbf{A^T}$. The update relations for residuals in the CG method (equation 2.38) are augmented in the Bi-

CG method by involving $\mathbf{A^T}$ as well as **A.** The search direction $\boldsymbol{p_j^*}$ does not directly contribute to the solution directly [60].

The **Conjugate gradient squared** (CGS) algorithm was developed by Sonneveld in 1984. The CGS method requires the same number of iterations as the BCG method but does not involve computations with $\mathbf{A^T}$**.**

The BiCGSTAB method was developed by van der Vorst in 1992. This method improves upon the Bi-CG method which can be numerically unstable. BiCGSTAB converges as fast (and sometimes faster) than CGS method. The residual vector is minimized locally which leads to a considerably smoother convergence behavior. The unpreconditioned BiCGSTAB method proceeds according to the following algorithm [61].

Let x_0 be an initial guess for the linear system $\mathbf{Ax = B}$**.** The residual is then $\mathbf{r_0 = b - Ax_0}$. The algorithm is initialized as follows

1. Choose an arbitrary vector $\widehat{\boldsymbol{r_0}}$ **,** such that $\widehat{\boldsymbol{r_0}}^{\boldsymbol{T}}\boldsymbol{r_0} \neq \mathbf{0}$.
2. Set $\rho_0 = \alpha = \omega_0 = 1$.
3. Set the search direction $\boldsymbol{v_0 = p_0 = 0}$.

For $i > 1$, the iterations proceed as

1. $\rho_i = \widehat{\boldsymbol{r_0}}^{\boldsymbol{T}}\boldsymbol{r_{i-1}}$
2. $\beta = (\rho_i/\rho_{i-1})(\alpha/\omega_{i-1})$
3. $\boldsymbol{p_i = r_{i-1} + \beta(p_{i-1} - \omega_{i-1}v_{i-1})}$
4. $\boldsymbol{v_i = Ap_i}$
5. $\alpha = \rho_i/(\widehat{\boldsymbol{r_0}}^{\boldsymbol{T}}\boldsymbol{v_i})$

6. $\boldsymbol{h} = \boldsymbol{x_{i-1}} + \alpha \boldsymbol{p_i}$
7. $First\ convergence\ test; test\ accuracy\ of\ h.\ If\ |h - x_{i-1}| < tol\ , set\ \boldsymbol{x_i} = \boldsymbol{h}\ and\ quit.$
8. $\boldsymbol{s} = \boldsymbol{r_{i-1}} - \alpha \boldsymbol{v_i}$
9. $\boldsymbol{t} = \boldsymbol{As}$
10. $\omega_i = (\boldsymbol{t^T s})/(\boldsymbol{t^T t})$
11. $\boldsymbol{x_i} = \boldsymbol{h} + \omega_i \boldsymbol{s}$
12. $Second\ convergence\ test;\ .\ If\ |x_i - x_{i-1}| < tol\ ,\ quit.$
13. $\boldsymbol{r_i} = \boldsymbol{s} - \omega_i \boldsymbol{t}$

BiCGSTAB method has two stopping tests: Stopping on the first test reduces the number of iterations. BiCGSTAB method requires two matrix-vector products and four vector dot products.

In this work, the algorithm described above is implemented in MATLAB to solve the discretized Poisson equation, $A\phi = f$ and determine the electrostatic potential. The coefficient matrix resulting from discretization of the 3D Poisson equation is stored as a sparse matrix. MATLAB allows for operations on matrices and vectors directly, which significantly reduces the computation time spent per iteration. Figure 2.6 shows the convergence behavior of the BICGSTAB algorithm when implemented to solve a 3D equilibrium Poisson equation in an AlGaN/GaN FinFET. The convergence plot is shown for two different applications: 3D Poisson equation for a narrow (40nm) Finfet and a wide (160nm) Finfet.

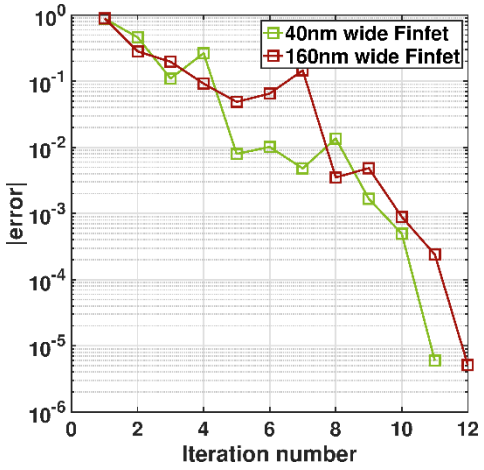

Figure 2.6 BiCGSTAB convergence plot for solving a 3D equilibrium Poisson equation for narrow and wide Finfet.

2.3 The 2D Schrödinger equation solver

The 3D Poisson equation solver discussed in section 2.1 provides the electrostatic potential, from which one can compute the carrier densities using semi-classical equilibrium Boltzmann statistics. The semi-classical charge description is not appropriate for the case when carriers are confined in a potential well in nanostructure devices like FinFETs, and/or quantum wells and Si/SiO_2 inversion layers. The confinement of electrons in the potential wells leads to a quasi-1D (Q1D) or a quasi-2D (Q2D) behavior. In the case of Q1D or Q2D carrier confinement, one needs to solve the Schrödinger equation in conjunction with the previously described Poisson equation to get accurate description of the charge densities. Within this self-consistent solution procedure, the Schrödinger equation is solved within a domain that is a subset of the entire simulation domain (hereafter referred to in this work as Schrödinger domain). The solution of the Schrödinger equation yields the electron wavefunction and the subband energies. These quantities are, in turn, used to determine the quantum electron density within the

Schrödinger domain. The self-consistent procedure is discussed in more details in the subsequent section. Here, the focus is on the solution of the Schrödinger equation only.

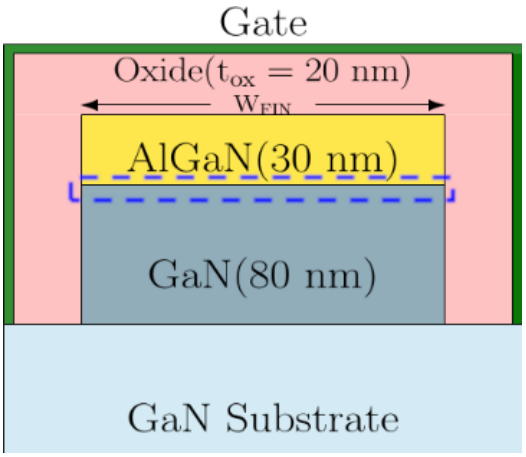

Figure 2.7 AlGaN/GaN FinFET cross section. The blue dotted box highlights the Schrödinger domain.

Figure 2.7 shows the cross section of the FinFET device under consideration in which the Schrödinger domain is highlighted (blue box). The presence of the triangular well region near the AlGaN/GaN interface and the reduced dimensions of the width of the device result in a quasi-1D behavior of the electrons in this system, thus necessitating the solution of a 2D Schrödinger equation.

The time independent Schrödinger wave equation is

$$-\frac{\hbar^2}{2m^*}\nabla\psi + \phi\psi = \mathcal{E}\psi \tag{2.39}$$

Where, m^* is the effective mass of the electron, ϕ is the confining electrostatic potential energy (related to the electrostatic potential ϕ' discussed in the previous section via the relationship $\phi = -q\phi'$, where q is the elementary charge). In heterostructures, where the electrons possess a varying effective mass depending on the bulk material of the region, the generalized effective mass Schrödinger equation is:

$$\frac{-\hbar^2}{2}\nabla\cdot\left(\frac{1}{m^*(x,y)}\nabla\psi(\mathrm{x},\mathrm{y})\right)+\phi(x,y)\psi(x,y)=\mathcal{E}\psi(x,y) \tag{2.40}$$

Since the effective mass is a property of a bulk, it is not well defined in the neighborhood of a sharp material transition. In the hypothesis of slow material composition variations in space, one can adopt the Schrödinger equation with a spatially varying effective mass, taken to be the mass of a bulk with the local material properties.

2.3.1 Discretization of the Schrödinger Equation using Finite Volume method

The Schrödinger equation given in equation 2.40 can be conveniently written in the form of the eigenvalue equation as shown below

$$\left(\frac{-\hbar^2}{2}\nabla.\left(\frac{1}{m^*(x,y)}\nabla\right)+\phi(x,y)\right)\psi(x,y)=\mathcal{E}\psi(x,y)$$

$$\mathcal{H}\psi=\mathcal{E}\psi \tag{2.41}$$

To arrive at the Hamiltonian $\mathcal{H}$ and employ eigenvalue solvers on equation 2.41 one needs to discretize the Schrödinger equation. This section describes briefly the Finite Volume (FV) discretization method [62].

The Schrödinger equation is a linear PDE consisting of a second order and a zeroth order differential operator. Consider a 2D grid depicted below in figure 2.8. Next, consider a "Control Volume", divided up into N octants around each mesh point. For a 2-dimensional system, the number of octants is $N=2^2=4$.

The second order differential operator in the Schrödinger equation is of the form

$$D^2(\vec{r})=-\frac{\hbar^2}{2}\sum_{i,j\in\{x,y\}}\partial_i\left(1/m^*_{ij}(r)\right)\partial_j\psi(\vec{r})$$

Writing the differential equation as an integral one by integrating over the control volume Ω, equation 2.41 can be rewritten as

$$\int_{\Omega} -\frac{\hbar^2}{2} \sum_{i,j \in \{x,y\}} \partial_i \left(1/m_{ij}^*(r)\right) \partial_j \psi(\vec{r}) \; dV + \int_{\Omega} \phi(\vec{r})\psi(\vec{r}) dV \; = \int_{\Omega} \mathcal{E}\psi(\vec{r}) dV \quad (2.42)$$

Let I_1 be the first integral term in equation 2.42

$$I_1 = \int_{\Omega} -\frac{\hbar^2}{2} \sum_{i} \partial_i \sum_{j} (1/m_{ij}^*(\vec{r})) \, \partial_j \psi(\vec{r}) \; dV \quad (2.43)$$

Using Gauss Divergence theorem, $\int \boldsymbol{\nabla}.\mathbf{F} \, \mathrm{dV} = \int \boldsymbol{F}.d\boldsymbol{S}$ and considering the terms within the second summation in equation 2.43 as flux **F,** one can write I_1 as

$$I_1 = \int_{S} -\frac{\hbar^2}{2} \sum_{i,j \in \{x,y\}} (1/m_{ij}^*(\vec{r})) \, \partial_j \psi(\vec{r}) \; d\boldsymbol{S} \, \hat{\boldsymbol{e}}_i \quad (2.44)$$

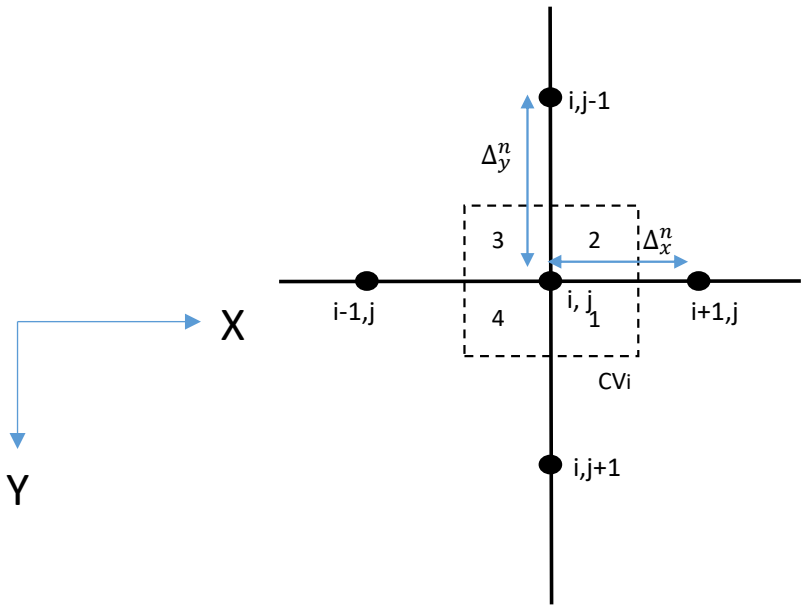

Figure 2.8 Five-Point stencil for finite volume discretization grid. Dotted lines enclose the control volume around the grid point (i,j) which is sub-divided into 4 octants. V_{ij} is the volume of the Control Volume at the grid point (i,j).

The surface integral in I_1 is next written as sum of surface integrals over each of the N octants.

$$I_1 = -\frac{\hbar^2}{2}\sum_{n=1}^{N=4}\int_{S_n}\sum_{i,j\in\{x,y\}}\left(\frac{1}{m_{ij}^*(\vec{r})}\right)^n \left[\partial_j\psi(\vec{r})\right]_n \, d\boldsymbol{S_n}\,\hat{\boldsymbol{e}}_i \qquad (2.45)$$

The surface integral over each octant ($\boldsymbol{S_n}$) can be further broken down into sum of surface integrals over each surface of the octants. For a 2D system, each octant has two such surfaces. The second summation in equation 2.45 is sum of surface integrals over p^{th}-surface of the n^{th}-octant.

$$I_1 = -\frac{\hbar^2}{2}\sum_{n=1}^{N=4}\sum_{p=1}^{2}\int_{S_n^p}\sum_{i,j\in[x,y\}}\left(\frac{1}{m_{ij}^*(\vec{r})}\right)^n \left[\partial_j\psi(\vec{r})\right]_n^p \, ds_n^p\hat{\boldsymbol{e}}_i\hat{\boldsymbol{e}}_{\boldsymbol{n}}^{\boldsymbol{p}} \qquad (2.46)$$

$\widehat{\boldsymbol{e_p}}$ is the unit normal vector of the p[th] surface of the n[th] octant. Since $\hat{\boldsymbol{e}}_i$ is also a unit vector, one can use basis orthonormality ($\hat{\boldsymbol{e}}_i . \hat{\boldsymbol{e}}_{\boldsymbol{n}}^{\boldsymbol{p}} = \delta_{ip} \hat{\boldsymbol{e}}_{\boldsymbol{n}}^{\boldsymbol{p}}(p)$) to eliminate the summation on p in equation 2.46, i.e.

$$I_1 = -\frac{\hbar^2}{2} \sum_{n=1}^{N=4} \sum_{i,j \in \{x,y\}} \int_{S_n^i} \left(\frac{1}{m_{ij}^*(\vec{r})} \right)^n \left[\partial_j \psi(\vec{r}) \right]_n^i \hat{\boldsymbol{e}}_{\boldsymbol{n}}^{\boldsymbol{i}}(i)\, ds_n^i \tag{2.47}$$

$$I_1 = -\frac{\hbar^2}{2} \sum_{n=1}^{N=4} \sum_{i,j \in \{x,y\}} \left(\frac{1}{m_{ij}^*(\vec{r})} \right)^n \left[\partial_j \psi(\vec{r}) \right]_n^i \hat{\boldsymbol{e}}_{\boldsymbol{n}}^{\boldsymbol{i}}(i) S_n^i \tag{2.48}$$

Here, S_n^i is the surface area of the surface in the i[th]-direction of the n[th]-octant. The term $\left[\partial_j \psi(\vec{r}) \right]_n^i$ describes the differential change in the wavefunction ψ in the j^{th} - direction on the surface along i^{th} - direction of the n^{th} octant of the control volume. For electron wavefunctions in the Schrödinger equation, this quantity exists only when $i = j$. Furthermore, if the effective mass can be assumed as being isotropic, then one can set $i = j$ in equation 2.48. For the case of $i = j$, the term $\left[\partial_j \psi(\vec{r}) \right]_n^i$ is discretized as

$$[\, \partial_i \psi(\vec{r})]_n^i = \frac{\psi\left(\Delta_n^i \hat{e}_n^i\right) - \psi(\vec{r})}{\Delta_n^i} \hat{e}_n^i \tag{2.49}$$

Using equation 2.49, the final discretized expression for the second order differential operator can be written as

$$I_1 = -\frac{\hbar^2}{2} \sum_{n=1}^{N=4} \sum_{i \in \{x,y\}} \left(\frac{1}{m_i^*(\vec{r})} \right)^n \frac{\psi\left(\Delta_n^i \hat{e}_n^i\right) - \psi(\vec{r})}{\Delta_n^i} S_n^i \tag{2.50}$$

Here Δ_n^i is the mesh spacing along the i[th] direction on the n[th] octant. The surface area S_n^i can be expressed as

$$S_n^i = \frac{\Delta_n^j}{2} \tag{2.51}$$

The zeroth order operator (the second integral in equation 2.42) discretization can be written as

$$I_2 = \int_{\Omega} \phi(\vec{r})\psi(\vec{r})dV$$

$$I_2 = \psi(\vec{r}) \sum_{n=1}^{4} \phi_n V_n \tag{2.52}$$

Here, ϕ is the confining potential and V is the control volume. Since the potential ϕ is defined around the grid point (i, j) and not in the octants, the summation can thus be eliminated, i.e.

$$I_2 = \psi(\vec{r})\phi(\vec{r})V_{i,j} \tag{2.53}$$

Here V_{ij} is the volume of the control volume at grid point (i, j) and is given as

$$V_{ij} = \left(\frac{X_i + X_{i-1}}{2}\right)\left(\frac{Y_j + Y_{j-1}}{2}\right) \tag{2.54}$$

where X_i and Y_j describes the mesh spacing along the x and y directions at grid point (i, j).

Thus, the final discretized Schrödinger equation (referring to the grid in figure 2.8) is given as

$$N_{i,j}\psi_{i,j-1} + E_{i,j}\psi_{i+1,j} + W_{i,j}\psi_{i-1,j} + S_{i,j}\psi_{i,j+1} + C_{i,j}\psi_{i,j} = \mathcal{E}V_{i,j}\psi_{i,j} \tag{2.55}$$

The off-diagonal coefficients are calculated using the following expressions

$$N_{i,j} = \left(-\frac{\hbar^2}{2}\right)\left[\sum_{n=1,4} \frac{\Delta_n^x}{\Delta_n^y} \frac{1}{m_n^*(i,j)}\right]$$

$$E_{i,j} = \left(-\frac{\hbar^2}{2}\right)\left[\sum_{n=1,2} \frac{\Delta_n^y}{\Delta_n^x}\frac{1}{m_n^*(i,j)}\right]$$

$$W_{i,j} = \left(-\frac{\hbar^2}{2}\right)\left[\sum_{n=3,4} \frac{\Delta_n^y}{\Delta_n^x}\frac{1}{m_n^*(i,j)}\right]$$

$$S_{i,j} = \left(-\frac{\hbar^2}{2}\right)\left[\sum_{n=2,3} \frac{\Delta_n^x}{\Delta_n^y}\frac{1}{m_n^*(i,j)}\right]$$

And the central coefficient is of the form,

$$C_{i,j} = \left(\frac{\hbar^2}{2}\right)\left[\sum_{n=1}^{4} \frac{\Delta_n^x}{\Delta_n^y}\frac{1}{m_n^*(i,j)} + \sum_{n=1}^{4} \frac{\Delta_n^y}{\Delta_n^x}\frac{1}{m_n^*(i,j)}\right] + \phi(i,j)V_{i,j}$$

Here, Δ_n^x and Δ_n^y define the mesh spacing in the n^{th} octant along the x and y directions respectively. $V_{i,j}$ is the control volume defined by equation 2.54.

The discretized equation leads to an eigenvalue equation of the form

$$\mathcal{H}\Psi = \mathcal{E}S\Psi \tag{2.56}$$

Where, $\mathcal{H}$ is the Hamiltonian matrix, Ψ is the collection of the electron wavefunctions, $\mathcal{E}$ is the diagonal eigenvalue matrix and S is the volume matrix containing the control volumes $V_{i,j}$ as diagonal elements. In some cases (such as for non-uniform meshing), the Hamiltonian may be non-Hermitian, thus leading to significantly higher computation times. The following transformation leads to a Hermitian Hamiltonian matrix.

$$\mathcal{H}\Psi = \mathcal{E}S\Psi$$

Left multiplying the above expression by $S^{-0.5}$ gives

$$S^{-0.5}\mathcal{H}\Psi = S^{-0.5}\mathcal{E}S\Psi$$

Right multiplying by $S^{-0.5}S^{0.5}$ leads to

$$S^{-0.5}\mathcal{H}S^{0.5}S^{-0.5}\Psi = S^{0.5}\mathcal{E}S^{-0.5}S^{0.5}\Psi$$

Since S matrix is a positive diagonal matrix, $S^{0.5}S^{0.5} = \mathrm{S}$ holds; thus

$$S^{-0.5}\mathcal{H}S^{-0.5}S^{0.5}\Psi = \mathcal{E}S^{0.5}\Psi$$

Writing

$$S^{0.5}\Psi = \Psi'$$

and

$$S^{-0.5}\mathcal{H}S^{-0.5} = \mathcal{H}'$$

The eigenvalue equation becomes

$$\mathcal{H}'\Psi' = \mathcal{E}\Psi' \tag{2.57}$$

This simple transformation leads to a Hermitian eigenvalue problem and the solution is less computationally expensive, compared to other discretization methods such as finite element method. Please note that the eigenvectors in Equation 2.56 and 2.57 are related by the expression

$$\Psi = S^{-0.5}\Psi' \tag{2.58}$$

2.3.2 Eigenvalue Solvers

For solving the eigenvalue equation 2.57, this work employs the canned Eigenvalue solvers provided in MATLAB for Hermitian matrices. Specifically, this work uses the `eigs` function in MATLAB. This function is a special version of the eigenvalue solver for Hermitian matrices. The MATLAB function uses an implementation of the QZ algorithm for determining the eigenvalues and eigenvectors if the matrix is Hermitian. The smallest *k* eigenvalues are requested from the solver. This corresponds to the lowest k subbands of the Q1D system. The value *k* depends on the dimensions of the

Schrödinger domain and hence the fin width and the depth of the well. The resulting eigenvectors are normalized such that their 2-norm is 1.

2.3.3 Quantum electron density

Once the subband energies and the corresponding electron wavefunctions are known, the subband occupation (or line electron density) is determined using general Fermi-Dirac statistics as

$$N_l^i = \int_{E_i}^{\infty} f(E)g(E)dE \tag{2.59}$$

Here, N_l^i is the line density of subband *i* and $f\ (E)$ and $g(E)$ are the Fermi-Dirac distribution function and the density of states function, respectively. Substituting the expression for 1D density of states function leads to

$$N_l^i = \frac{\sqrt{2m^*\kappa T}}{\pi\hbar} F_{-\frac{1}{2}}(\eta_i) \tag{2.60}$$

where

$$\eta_i = \frac{E_F - E_i}{\kappa T} \tag{2.61}$$

Here E_i is the electron energy in the i^{th} subband and $F_{-\frac{1}{2}}$ is the Fermi integral of order -1/2 which is evaluated using standard analytical approximations. The quantum electron density in the 2D slice is thus given as

$$n(x,y) = \sum_{i=1}^{k} N_l^i |\psi_i(x,y)|^2 \tag{2.62}$$

2.4 The Self Consistent Schrödinger Poisson Equation Solver

Once the quantum electron density for a given 2D slice is determined as described in the previous section, it is substituted back into the 3D Poisson equation. The quantum electron density calculation is repeated for each 2D slice along the z direction. As described previously, the quantum electron density is determined in the Schrödinger domain. The boundaries of the Schrödinger domain over the device need to be properly set as this ensures that electrostatics are accurately calculated wherever electron confinement is present.

2.4.1 Extent of Schrödinger domain

The electrostatic potential resulting from the 3D Poisson equation solver (with semi-classical charge description) determines the extent of the Schrödinger which depends on the device dimensions, i.e., fin width and the type of structure (MISFET or MESFET). The Schrödinger domain should encompass all regions where there is formation of a potential well. The potential in the transition region between Schrödinger and classical domains should not be discontinuous. Finally, one also needs to account for wavefunction decay onto surrounding regions.

These conditions can be further exemplified by considering the AlGaN/GaN MISFET structure shown in figure 2.7. One can note that the active regions in this structure are the AlGaN and GaN channel regions.

1. When a negative gate bias is applied, both the bulk of GaN channel and AlGaN channel regions are under depletion regime. Thus, the electron density is confined to the potential well created at the AlGaN/GaN interface owing to

polarization-induced charge density. Therefore, the Schrödinger domain will comprise of the potential well at the AlGaN/GaN interface.

2. When a positive gate bias is applied, the AlGaN and GaN channel regions shift to accumulation regime, thus the Schrödinger domain will include the entire AlGaN and GaN channel regions under the gate oxide.

2.4.2 Predictor-Corrector Method

The quantum electron density, at this step, needs to be inserted into the charge density term on the right hand side of the Poisson equation, as described in section 2.1. Referring back to equation 2.2,

$$\nabla(\epsilon\nabla\phi) = -\frac{\rho(\phi)}{\epsilon_0} \tag{2.63}$$

$$\text{Where } \rho(\phi) = -\frac{q}{\epsilon_o}(p(\phi) - n(\phi) + Dop)$$

In section 2.1, the carrier densities were calculated using semi-classical Boltzmann approximations or Fermi-Dirac statistics. Similar to section 2.1, one can use Taylor series expansion of the right-hand side of equation 2.63 around ϕ_k (where k denotes the current iteration)

$$f(\phi_{k+1}) = f(\phi_k) + \frac{\partial f(\phi_{k+1})}{\partial \phi_{k+1}}(\phi_{k+1} - \phi_k) + O(\phi_k) \tag{2.64}$$

$$f(\phi_k) = \frac{-q}{\epsilon_0}\left(p_k^c - n_k^q + Dop\right) \tag{2.65}$$

The superscripts c and q indicate if the terms are calculated using semi-classical and quantum mechanical expressions respectively. The expression for quantum electron density is given as

$$n_k^q = N_{1D} \sum_{i=1}^{N} |\Psi_i|^2 F_{-\frac{1}{2}}\left(\frac{\mathcal{E}^F - \mathcal{E}_k^i}{V_T}\right)$$

$$N_{1D} = \frac{(2m^* \kappa T)^{1/2}}{\pi \hbar}$$

Where N is the total number of subbands and $\mathcal{E}_k^i$ is the energy of the i[th] subband. The second term in equation 2.64 can be written as

$$\frac{\partial f(\phi_{k+1})}{\partial \phi_{k+1}}(\phi_{k+1} - \phi_k) = \frac{-q}{\epsilon_0}\left(\frac{-p_k^c}{V_T} - \frac{\partial n_{k+1}^q}{\partial \phi_{k+1}}\right)(\phi_{k+1} - \phi_k) \tag{2.66}$$

$$n_{k+1}^q = N_{1D} \sum_{i=1}^{N} |\Psi_i|^2 F_{-\frac{1}{2}}\left(\frac{\mathcal{E}^F - \mathcal{E}_{k+1}^i}{V_T}\right) \tag{2.67}$$

One can assume that the potential in the next iteration shifts each subband energy by a small perturbation, but does not change the electron wavefunction. The electron wavefunctions are determined exactly in the previous iteration of the Schrödinger equation solver. Using this approximation,

$$\mathcal{E}_{k+1}^i \approx \mathcal{E}_k^i - \delta\phi_{k+1} \tag{2.68}$$

Where, $\delta\phi_{k+1} = (\phi_{k+1} - \phi_k)$

Using the property of Fermi integrals,

$$\frac{dF_j(\eta)}{dx} = jF_{j-1}(\eta)$$

One can write,

$$\left.\frac{\partial n_{k+1}^q}{\partial \phi_{k+1}}\right|_{\phi_k} = \frac{1}{V_T}\left(-\frac{N_{1D}}{2} \sum_{i=1}^{N} |\Psi_i|^2 F_{-\frac{3}{2}}\left(\frac{\mathcal{E}^F - \mathcal{E}_k^i}{V_T}\right)\right) \tag{2.69}$$

Since analytical expressions for Fermi integral of order (-3/2) is not available, one can instead calculate the differential of $F_{-\frac{1}{2}}$ using the standard definition of derivative.

$$F_{-\frac{3}{2}}(\eta_F) = F'_{-\frac{1}{2}}(\eta_F) = \frac{d\,F_{-\frac{1}{2}}(\eta_F)}{d\eta} \approx \frac{F_{-\frac{1}{2}}(\eta_F + h) - F_{-\frac{1}{2}}(\eta_F)}{h} \tag{2.70}$$

Once equation 2.70 is numerically computed, one can rewrite equations 2.69 as

$$\left.\frac{\partial n_{k+1}^{q}}{\partial \phi_{k+1}}\right|_{\phi_k} = \frac{1}{V_T}\left(N_{1D}\sum_{i=1}^{N}|\Psi_i|^2 F'_{-\frac{1}{2}}\left(\frac{\mathcal{E}^F - \mathcal{E}_k^i}{V_T}\right)\right) = \frac{1}{V_T}n_k^Q \tag{2.71}$$

Where $n_k^Q = N_{1D}\sum_{i=1}^{N}|\Psi_i|^2 F'_{-\frac{1}{2}}\left(\frac{\mathcal{E}^F - \mathcal{E}_k^i}{V_T}\right)$.

Finally, substituting equation 2.71 in 2.65, the charge density term of the Poisson equation becomes

$$\rho(\phi_{k+1}) = -\frac{q}{\epsilon_0}\left(p_k^c - n_k^q + Dop\right) - \frac{q}{\epsilon_0 V_T}(p_k^c + n_k^Q)\phi_k + \frac{q}{\epsilon_0 V_T}(p_k^c + n_k^Q)\phi_{k+1}$$

Thus, for $(i,j,k) \in Schrödinger\ domain$, the central coefficient and the forcing function of the discretized Poisson equation (Equation 2.3) is updated to

$$C(i,j,k) = -\big(B(i,j,k) + F(i,j,k) + E(i,j,k) + W(i,j,k) + N(i,j,k) + S(i,j,k)\big) - \left(n_{ir}\exp\left(-\frac{\phi(i,j,k)_k}{V_T}\right) + n(i,j,k)_k^Q\right)$$

$$\not{f}(i,j,k) = V_T\left(n(i,j,k)_k^Q - n_{ir}\exp\left(-\frac{\phi(i,j,k)_k}{V_T}\right) - Dop(i,j,k)\right) - \phi(i,j,k)_k\left(n(i,j,k)_k^Q + n_{ir}\exp\left(-\frac{\phi(i,j,k)_k}{V_T}\right)\right)$$

All the other coefficients remain unchanged. The Poisson equation is solved using BICGSTAB method described previously.

States in the continuum

If one considers the triangular potential well at the AlGaN/GaN interface (along the depth) as shown in figure 2.9, one encounters bound states (with closed boundary conditions on both sides) as well as free states (with open boundary conditions).

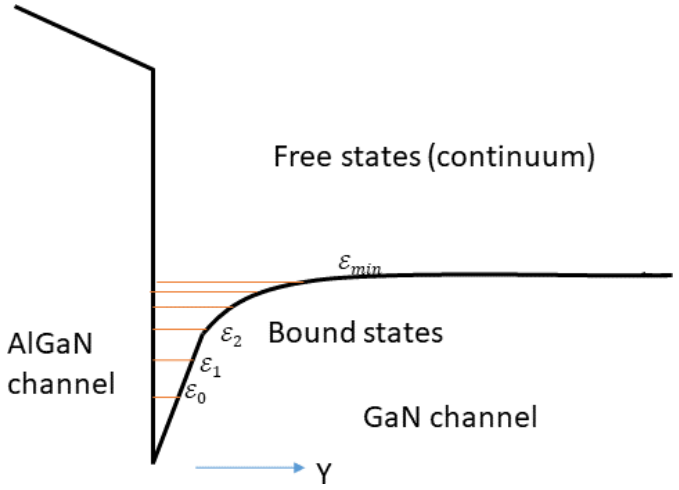

Figure 2.9 Bound and free states in the potential well.

One needs to account for electrons in the bound states as well as states in the continuum. If the electrons in the continuum are discounted, then the quantum electron density fluctuates from iteration to iteration. In the Schrödinger domain, the electrons in the bound states obey 1D density of states, while the electrons occupying the continuum states are not confined and thus obey the 3D density of states expression. It should be noted that the occupancy of the continuum states will be significant at positive gate biases.

Referring to figure 2.9, one can assume $\mathcal{E}_{min}$ to be the transition point between bound and free states such that the density of states for $\mathcal{E} > \mathcal{E}_{min}$, will obey 3D characteristics. Thus, the electron density in the Schrödinger domain (n_{sch}) is given as

$$n_{sch} = n_q + n_{cont} \tag{2.72}$$

n_q is the electron density of the bound states and determined using equation 2.67 and the continuum electron density is given as

$$n_{cont} = N_{3D} \frac{2}{\sqrt{\pi}} F_{\frac{1}{2}}(\mathcal{E}_F - \mathcal{E}_{min}) \tag{2.73}$$

$$N_{3D} = 2\left(\frac{2\pi m^* \kappa T}{(2\hbar\pi)^2}\right)^{3/2}$$

Applying the approach similar to section 2.4.1, one can modify equation 2.68 as

$$\left.\frac{\partial n_{k+1}^{sch}}{\partial \phi_{k+1}}\right|_{\phi_k} = \left.\frac{\partial n_{k+1}^{q}}{\partial \phi_{k+1}}\right|_{\phi_k} + \left.\frac{\partial n_{k+1}^{cont}}{\partial \phi_{k+1}}\right|_{\phi_k} \tag{2.74}$$

Using the similar approach as described previously and assuming that between k^{th} and $(k+1)^{th}$ iteration, $\mathcal{E}_{min}$ changes as

$$\mathcal{E}_{k+1}^{min} \approx \mathcal{E}_k^{min} - \delta\phi_{k+1}$$

The second term in equation 2.73 can be expanded to

$$\left.\frac{\partial n_{k+1}^{cont}}{\partial \phi_{k+1}}\right|_{\phi_k} = \frac{N_{3D}}{V_T}\frac{1}{\sqrt{\pi}} F_{-\frac{1}{2}}(\mathcal{E}_F - \mathcal{E}_{min}) = \frac{1}{V_T} n_k^C \tag{2.75}$$

One can combine electron density terms in equations 2.71 and 2.75 and write

$$n_k^S = n_k^C + n_k^Q \tag{2.76}$$

One can thus use equations 2.70, 2.74 and 2.75 to modify the charge density term of the Poisson equation to

$$\rho(\phi_{k+1}) = -\frac{q}{\epsilon_0}\left(p_k^c - n_k^{sch} + Dop\right) - \frac{q}{\epsilon_0 V_T}(p_k^c + n_k^S)\phi_k + \frac{q}{\epsilon_0 V_T}(p_k^c + n_k^S)\phi_{k+1}$$

$$n_k^{sch} = N_{1D}\sum_{i=1}^{N}|\Psi_i|^2 F_{-\frac{1}{2}}\left(\frac{\mathcal{E}^F - \mathcal{E}_k^i}{V_T}\right) + N_{3D}\frac{2}{\sqrt{\pi}}F_{\frac{1}{2}}\left(\mathcal{E}_F - \mathcal{E}_{min}^k\right)$$

$$n_k^S = N_{1D}\sum_{i=1}^{N}|\Psi_i|^2 F'_{-\frac{1}{2}}\left(\frac{\mathcal{E}^F - \mathcal{E}_k^i}{V_T}\right) + N_{3D}\frac{1}{\sqrt{\pi}}F_{-\frac{1}{2}}\left(\mathcal{E}_F - \mathcal{E}_{min}^k\right) \qquad (2.77)$$

The central coefficient and the forcing function is then modified accordingly, and the Poisson equation is solved using previously discussed numerical solution methods.

2.4.3 Initialization of the solver and the self-consistent solver loop

One can summarize the self-consistent Schrödinger-Poisson solver process described in the previous section with the flowchart shown in figure 2.10.

- Device data and Simulation parameters: The Finfet dimensions, structure and doping density of respective regions is taken as user input. Material parameter database is loaded and updated for given temperature. The mesh is refined near interfaces, if necessary.
- Initialization: Determine polarization induced charge density, heterostructure band parameters and initialize potential at all nodes using charge neutrality. Apply Dirichlet and Neumann boundary conditions.
- Define Schrödinger domain: Depending on the initial output of Poisson solver, define the boundaries of the Schrodinger domain. Set the requested number of eigenvalues depending on the size of the Schrödinger domain.

- Run Schrödinger Solver: Solve the Schrödinger equation over the Schrödinger domain and determine the electron wavefunctions and electron line densities.
- Update Central coefficient and forcing function: Modify the charge density term in Poisson equation after calculating n_k^S and n_k^{sch}. This leads to an update in the forcing function and central coefficient of the discretized Poisson equation. Run the Poisson Solver using BICGSTAB method.
- Check convergence and exit: If the convergence criteria is satisfied, exit loop. Save electron wavefunctions, line densities and subband energies. Transfer data to transport kernel.

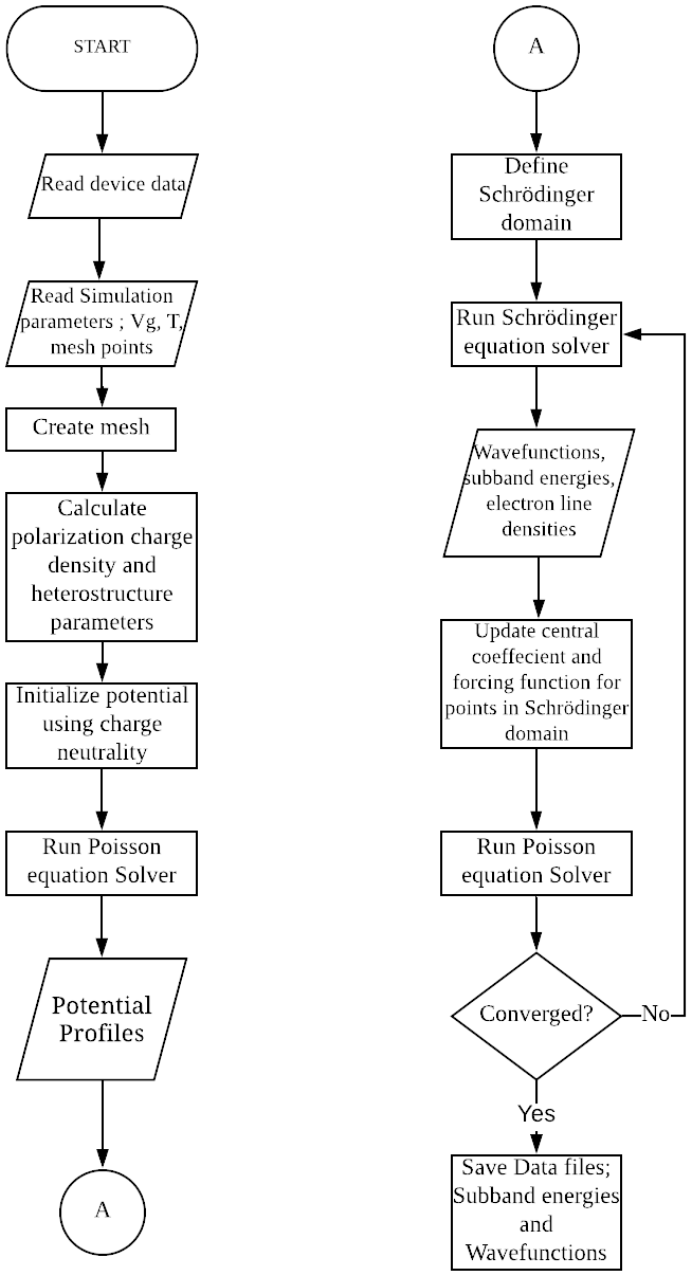

Figure 2.10 Implementation of Self-consistent Schrödinger-Poisson solver loop.

Convergence plot

Figure 2.11 shows the implementation of the Schrödinger-Poisson equation solver for Finfets of two different widths. The convergence plot shows the error steadily

decreasing in both cases. The wide Finfet (160nm) takes 15 iterations to converge as compared to 10 for the narrow Finfet (40nm).

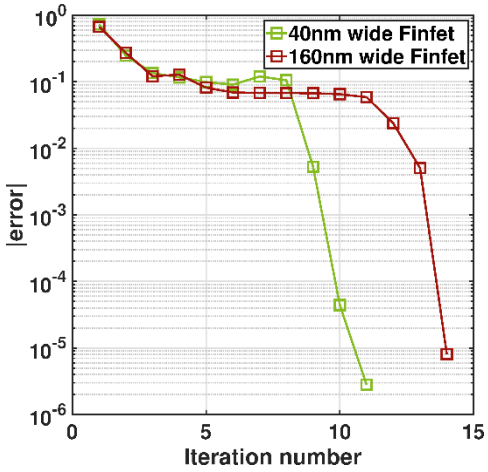

Figure 2.11 Convergence plot for the self-consistent Schrödinger-Poisson solver for narrow and wide Finfets.

The increased iterations for wide Finfet can be attributed to the higher number of subbands, which are also closer to each other, in the wider lateral potential well of the 160nm Finfet.

2.5 Results of the Schrödinger-Poisson Equation solver

2.5.1 Simulated device structure

The Schrödinger-Poisson solver described in the previous section is used to study AlGaN/GaN MIS FinFET discussed in Chapter 1. The Finfet device simulated here is similar to that fabricated by the authors in [29] and [34], which consists of an 80 nm-thick GaN channel layer and an AlGaN layer of thickness 30 nm grown on a highly resistive GaN substrate. The fin is wrapped around by 20m thick Aluminum oxide layer and a tri-gate. The width of the fin is varied from 20 nm to 200 nm. The fin width

variation is simulated to determine the change in behavior between narrow and wide Fin devices. The simulated cross section is shown in figure 2.12.

Table 2.1 Simulated device design parameters

Parameter	**Value**
W_{FIN}	20 nm -200 nm
T_{AlGaN}	25 nm, 30 nm
T_{ox}	20 nm
Al mole fraction x	20%, 30%

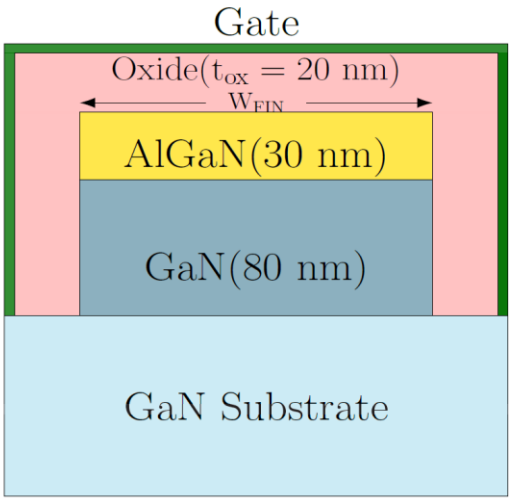

Figure 2.12 Cross section of simulated AlGaN/GaN Finfet. This is the cross section of the AlGaN/GaN MIS Finfet shown in figure 1.6 (channel region).

2.5.2 Subband Energies and Wavefunctions

At negative gate biases, the 2D Schrödinger domain is set near the AlGaN/GaN interface allowing 5 nm penetration depth into the AlGaN layer and into the surrounding oxide layers. For positive gate biases, the Schrödinger domain extends over the entire

GaN channel region. The shape of the confining potential along the depth of the fin, for V_G =0 V, is shown in fig. 2.13 which suggests a triangular confinement. The inset of fig 2.13 shows the wavefunctions of first 4 eigenstates.

Along the width of the fin, the confining potential is a truncated parabola. Therefore, the subband energies are a superposition of a triangular and parabolic confinement. This is clearly seen in the plot shown in fig. 2.14 which shows the first (lowest) 15 subband energy levels.

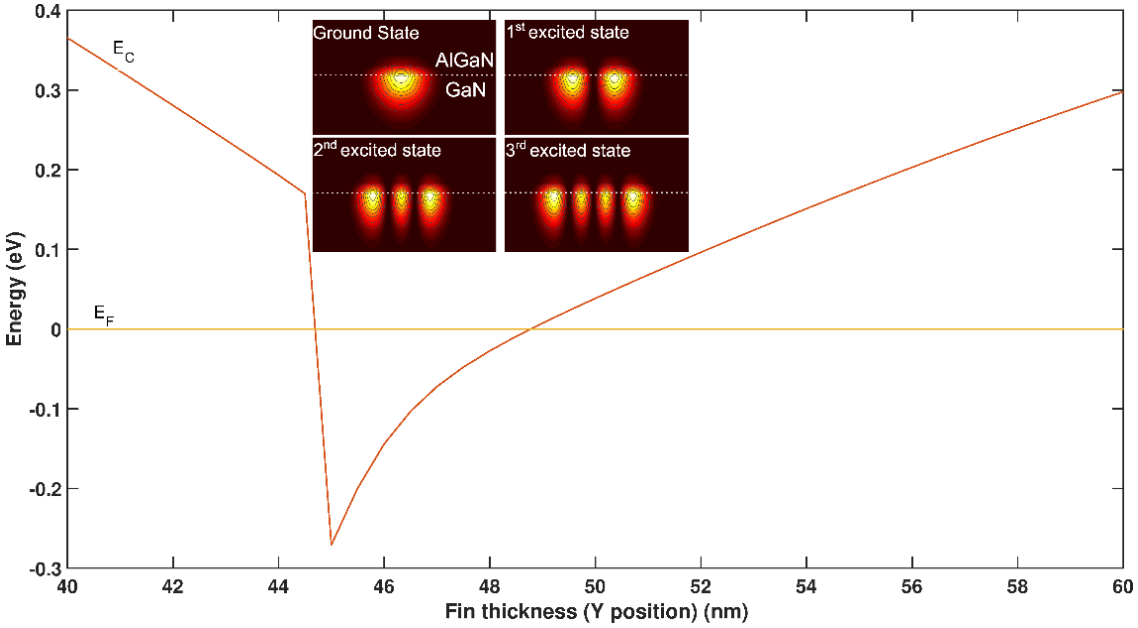

Figure 2.13 The triangular confining potential (Ec) along the fin depth near the AlGaN/GaN interface (only the Schrödinger domain is shown). The magnitude squared of the wavefunctions of the lowest four states is shown in the inset (V_G = 0V)

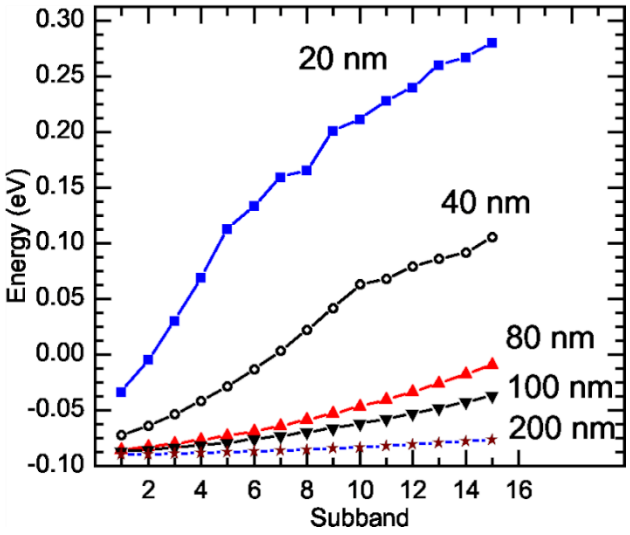

Figure 2.14 Subband energies vs. fin width at zero bias.

One can see that as the fin width reduces, the number of energy levels above the Fermi level (E = 0 eV) increases substantially. This results in an exponential reduction in occupancy of these states. Thus, one can conclude that the threshold voltage of a Finfet device reduces with fin width and the device shifts from a D-mode to an E-mode device.

2.5.3 Electron density contour plots

As one increases the gate bias above zero, one notes the formation of sidewall channels in the device. Figure 2.15 (a) shows the electron density in the cross-section when the gate is biased at 0 V. The electrons are confined in a quasi-1D channel peaking at the center of the fin.

As the gate bias shifts to positive, we see the formation of side-wall channels laterally in addition to channel present at the AlGaN/GaN interface. Although the electron density in the sidewall channels is an order of magnitude lower than the channel near the interface, the sidewall channel density increases sharply with increasing gate bias.

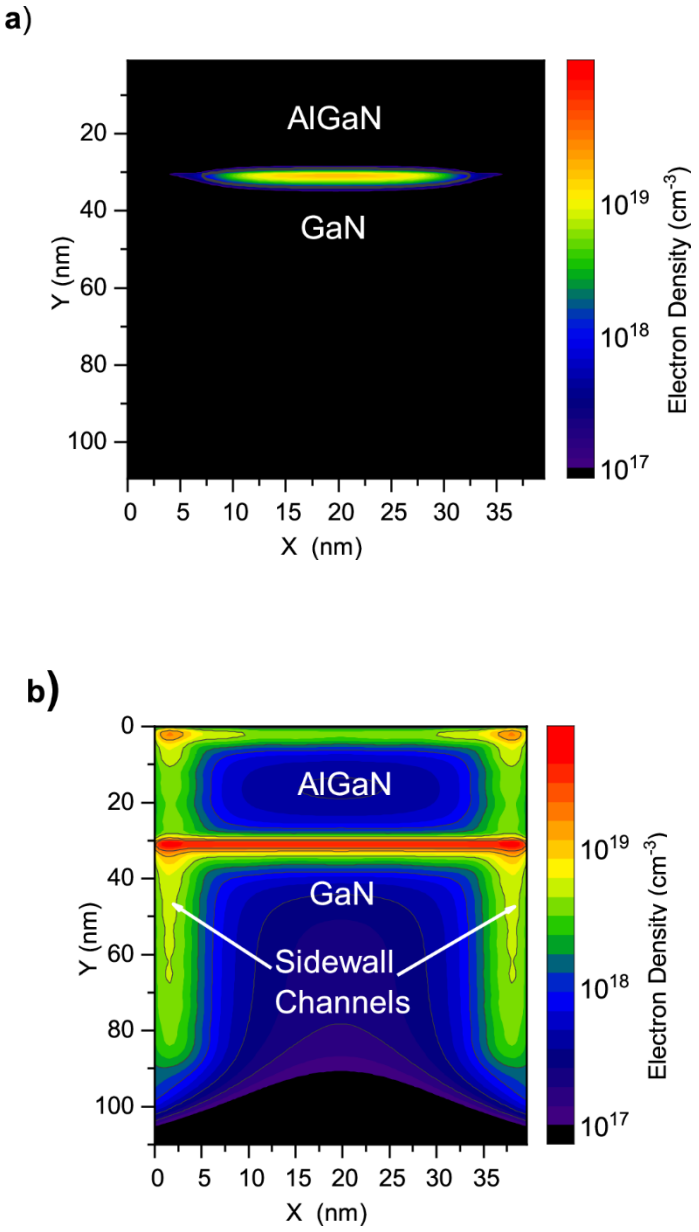

Figure 2.15 Electron density contour plots. a) Gate bias = 0 V. b) Gate bias = 1.0V. The GaN channel shifts to accumulation regime when gate bias is positive

.

2.5.4 Electron Line density and Threshold Voltage Extraction

The variation of line density with gate bias can be taken as a way of determining the threshold voltage. Since the drop in electron line density by an order of magnitude will lead to proportional decrease in drift current, one can estimate V_T from the line density – V_G plot.

The procedure starts with extracting the channel electron density along the 2DEG channel. It is then integrated over the Schrödinger domain to determine the corresponding line density. This line density is then plotted as a function of the gate bias and the procedure is repeated for a range of temperature from 100-350K. The threshold voltage is next taken as the gate bias at which the line density drops below $10^1\ cm^{-1}$. This value is somewhat arbitrary, and it is chosen as such because the line density increases exponentially beyond this point. The threshold voltage is plotted against temperature in figure 2.16 (b). The simulation is repeated for a lower fin width of 40 nm as well.
The experimental threshold voltage data is taken from the work of Ki Sik et al. [34]. The threshold voltage, as determined by the Schrödinger-Poisson Solver used in this work, matches closely the experimental values and follows the same decreasing trend with temperature for a given fin width. This can be attributed to the behavior of the Fermi-Dirac distribution function which increases with temperature. This leads to a higher electron density in the 2DEG channel at lower V_G, thus decreasing the threshold voltage (making it more negative).

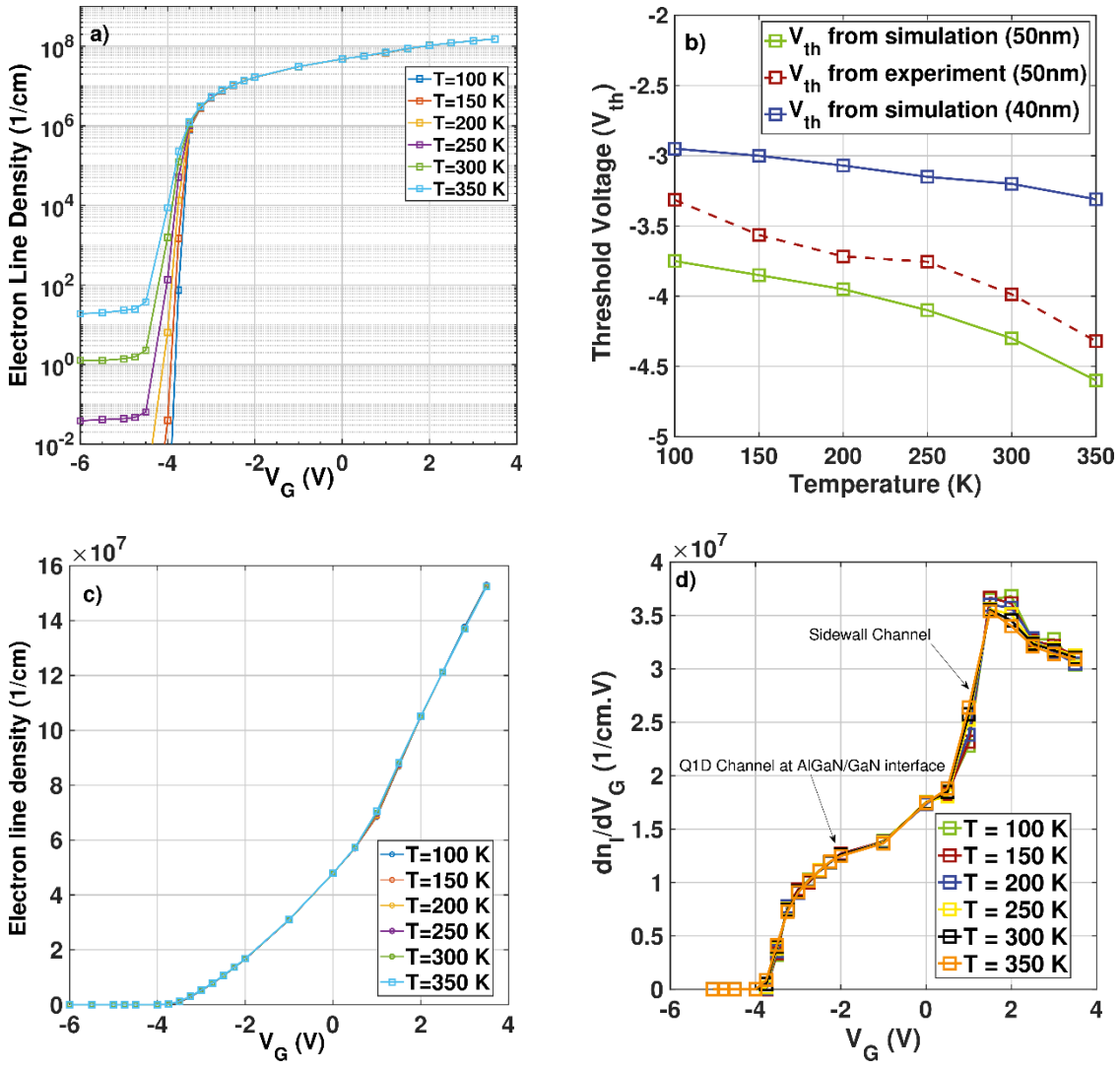

Figure 2.16 Electron line density and threshold voltage of an AlGaN/GaN Finfet. a) Variation of the line density vs. gate bias for different temperatures. b) Simulated threshold voltage data vs experimental results. Experimental data are taken from Ki-sik et al. [34] c) Line density on a linear scale illustrating the impact of the sidewall channels. d) Derivative of the line density with respect to the gate voltage to illustrate the conduction in the Q1D and the sidewall channels.

One can note that the V_{th} dependence on temperature reduces with decreasing fin width, as the 40 nm Finfet shows a lower variation in threshold voltage with temperature. This is due to the increasing impact of the side-gates in narrow fin devices. As the fin width reduces, one observes that the fringing fields from the side gates are the major factor in depleting the conducting channel, which do not show a dependence on temperature. For narrow fin width devices, the sidewall channels play a more critical role in conduction, and this leads to a 'flat' profile in the V_{th} -T curves. Thus, one can argue that with the proper design parameters, it is possible to design a Finfet with temperature-insensitive current.

CHAPTER 3

MODELING MOBILITY

For low electric fields, mobility is the most important transport parameter as it determines the drift component of the current. For confined systems, the electron mobility depends upon the electron wavefunctions and the corresponding subband energies. The overlap between wavefunctions of initial and final states along with the density of states determine the scattering rates. This work determines electron mobility by solving the Boltzmann transport equation using Monte Carlo method. Each scattering event is treated independently of each other, and Fermi's Golden rule is assumed to hold.

Conduction regimes in the AlGaN/GaN FinFet device

Referring to the electron density contour plots in figure 2.12 and the subsequent discussion in the previous chapter, one recalls the formation of sidewall channels along the lateral gates under the application of a positive gate bias. Only 2DEG channel at the AlGaN/GaN interface exists when $V_G \leq 0$. The sidewall channels may have different transport characteristics as compared to AlGaN/GaN interface channel, i.e. different scattering mechanisms will dominate in each regime. Thus, it is imperative to treat each conduction channel separately. It is also assumed that electrons in both channels are independent of each other. This enables one to treat both regions as parallel conduction channels and the total mobility can be calculated as a weighted sum of the individual mobility components.

Therefore, one would need to develop two different transport models; one for the interface channel, which is applied when $V_G \leq 0$ and one for the sidewall channels which are primarily present when $V_G > 0$. The work accomplished as part of this dissertation

focuses on the transport model for the conduction channel present at the AlGaN/GaN interface, which will be discussed in the current chapter.

3.1 Transport Model for the AlGaN/GaN 2DEG channel

As introduced previously, the Schrödinger domain, which includes the AlGaN/GaN channel, is quasi-1D in nature. It is important to reiterate here that, for $V_G \leq 0$ V, the Schrödinger domain only encompasses the AlGaN/GaN interface and does not extend to cover the sidewall channels. Thus, being a Q1D system, one needs to employ 1D scattering rate expressions for the various scattering mechanisms considered. The electrons are confined in the *x* and *y* directions, while the direction of propagation is along *z*. That is, if one considers the 3D structure of the lateral Finfet, the electrons will be free to propagate along the length of the fin channel. Accordingly, the applied electric field will be from source to drain along the *z* direction. The electron wavefunctions are plane waves in the unconfined *z* plane, and are solutions of the Schrödinger-Poisson problem in the confined x-y plane. Therefore, the electron wavefunctions for the initial and final states are given by

$$\Psi_n(k_z) = \frac{1}{\sqrt{L_z}}\psi_n(x,y)\exp(i\mathrm{k_z}.\mathrm{z})\Big|_{initial\ state} \tag{3.1}$$

$$\Psi_m(k_z') = \frac{1}{\sqrt{L_z}}\psi_m(x,y)\exp(ik_z'.\mathrm{z})\Big|_{final\ state} \tag{3.2}$$

3.1.1 The Bandstructure

For the low-field electron mobility calculation in wurtzite GaN, it is sufficient to consider only a single valley (Γ_1). In the theoretical model, the non-parabolicity parameter for this valley is $\alpha = 0.189\ eV^{-1}$.

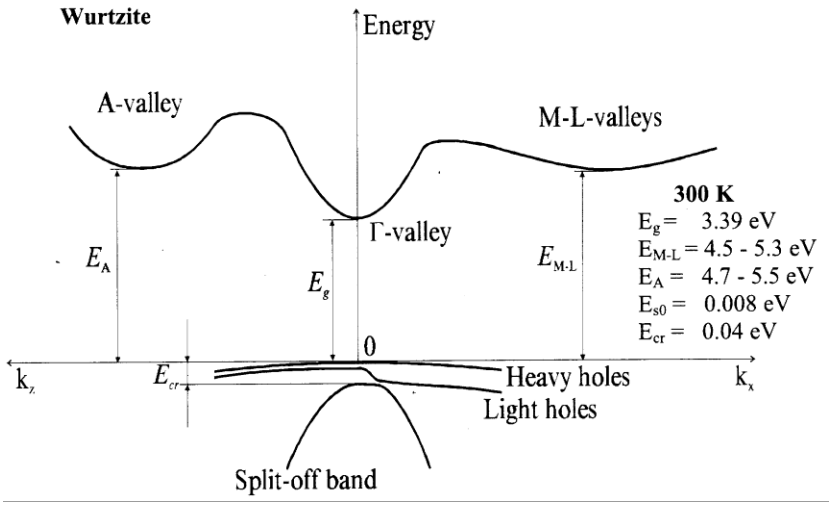

Figure 3.1 Bandstructure of Wurtzite GaN. For low-field electron mobility calculation, it is sufficient to only consider the Γ valley.

3.1.2 Relevant Scattering Mechanisms

Acoustic Phonon Scattering

Acoustic phonons originate from the lattice atoms oscillating in the same direction leading to a deformation (strain) in the unit cell. The perturbation potential is, thus, proportional to the differential displacement

$$\hat{H}_{e-p}(r) = \Xi_{ac}\, \nabla \cdot \boldsymbol{u}(\boldsymbol{r}) \tag{3.3}$$

Where, Ξ_{ac} is the acoustic phonons deformation potential and $\boldsymbol{u}(\boldsymbol{r})$ is the phonon unit vector given by

$$u(r) = \sum_q \sqrt{\frac{\hbar}{2\rho\omega_q\Omega}}\boldsymbol{e_q}\left(a_q\boldsymbol{e^{iq.r}} + {a_q}^+\boldsymbol{e^{-iq.r}}\right) \tag{3.4}$$

Here $a_q, {a_q}^+$ are the annihilation and creation operators, q is the phonon wave vector, ρ is the material density, $\boldsymbol{e_q}$ is the unit polarization vector and $\hbar\omega_q$ is the phonon energy. Using elastic and equipartition approximation, one can assume that the phonon energy is much less than the thermal energy ($\hbar\omega_q \ll \kappa T$). Thus, in addition to the interactions with electrons being elastic, one can also assume that

$$N_q \cong N_q + 1 \cong \frac{\kappa T}{\hbar\omega_q} \tag{3.5}$$

For low energies, the dispersion curve for acoustic phonons is linear (Debye model); therefore, one can write

$$\omega_q = v_s q \tag{3.6}$$

Where v_s is the longitudinal sound velocity in GaN. As described in [41], the matrix element for this scattering process is

$$|M(k_z, k_z')|^2 = \frac{\Xi^2\kappa T}{2\rho\Omega v_s^2}\left|{I_{nm}}^{ac}(q_x, q_y)\right|^2\delta(k_z - k_z' \pm q_z) \tag{3.7}$$

The overlap integral in equation 3.7 is of the form

$${I_{nm}}^{ac}(q_x, q_y) = \iint \psi_m(x,y)\exp\left(i\left(q_x x + q_y y\right)\right)\psi_n(x,y)dxdy$$

And the final expression for acoustic scattering rate from subband n to subband m is given as

$$\Gamma_{nm}^{ac} = 2\sum_{kz'q} S_{nm}(k_z, k_z') \tag{3.8}$$

Where the prefactor of 2 accounts for absorption and emission of phonons. Using Fermi's Golden rule and converting the summation to an integral, leads to the following result

$$\Gamma_{nm}^{ac} = 2\frac{2\pi}{\hbar}\frac{\Xi^2\kappa T}{2\rho\Omega v_s^2}\frac{1}{2\pi}$$

$$\iint \frac{1}{(2\pi)^2}\left|I_{nm}{}^{ac}(q_x, q_y)\right|^2 dq_x dq_y \int \delta(k_z - k_z' \pm q_z)\,\delta(E - E')dk_z' \tag{3.9}$$

Evaluating the integral of the overlap function $\left|I_{nm}{}^{ac}(q_x, q_y)\right|^2$ one arrives at

$$I_{overlap}^{acoustic} = \iint \frac{1}{(2\pi)^2}\left|I_{nm}{}^{ac}(q_x, q_y)\right|^2 dq_x dq_y = \iint |\psi_m(x,y)|^2|\psi_n(x,y)|^2 dxdy \tag{3.10}$$

The overlap integral is evaluated over the entire Schrödinger domain. The final scattering rate for acoustic phonons is given by

$$\Gamma_{nm}^{ac} = \frac{\Xi^2\kappa T\sqrt{2m^*}}{2\rho v_s^2\hbar^2}\, I_{overlap}^{acoustic}\frac{(1 + 2\alpha E_f)}{\sqrt{E_f(1 + \alpha E_f)}}\Theta(E_f) \tag{3.11}$$

Where the step function $\Theta(E_f)$ allows only those transitions when the final energy is positive. Acoustic phonon scattering rate decreases with increasing energy. Hence, acoustic phonons are primarily active at low electron energies. The cumulative scattering rate for acoustic phonon scattering out of subband 1 is plotted in figure 3.2.

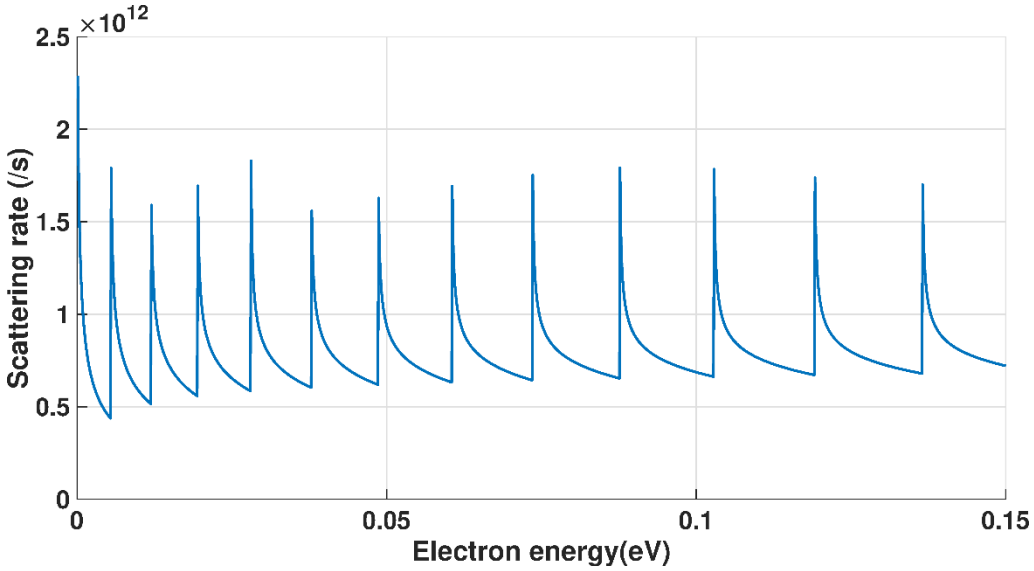

Figure 3.2 Cumulative Acoustic phonon scattering rate out of subband 1

Piezoelectric Scattering

Polar scattering occurs in materials having two different kinds of atoms in their basis. In such crystals, there is charge transfer between the two atoms, thus creating a dipole moment. Polar Scattering can be due to optical phonons (Polar Optical Phonon scattering) and acoustic phonons (Piezoelectric scattering). Second order interaction of electrons with these fields gives rise to piezoelectric scattering which is discussed here. Piezoelectric polarization is proportional to the acoustic strain which further depends on the differential displacement, i.e.

$$\boldsymbol{P} = e_{pz} \nabla . \mathrm{u}(\mathbf{r}) \tag{3.12}$$

In eq. 3.12, e_{pz} is the piezoelectric constant which depends on the piezoelectric and elastic properties of the material [63], [64]. The perturbing Hamiltonian is given by

$$\widehat{H}_{piezo} = -\frac{e e_{pz}}{\varepsilon_{\infty}} \sum_{q} \sqrt{\frac{\hbar}{2\mu N \omega_q}} \left(a_q \boldsymbol{e^{iq.r}} + {a_q}^{+} \boldsymbol{e^{-iq.r}} \right) \tag{3.13}$$

Since one is dealing with acoustic phonons, one can use the elastic and equipartition approximation as well as the linear dispersion of acoustic phonons at low energies. Integrating equation (3.13) over phonon coordinates, one can write the matrix element as

$$|M(k_z, k_z')|^2 = \left(\frac{e e_{pz}}{\varepsilon_\infty}\right)^2 \frac{\kappa T}{2\rho v_s^2} \frac{1}{q^2} \left|I_{nm}{}^{piezo}(q_x, q_y)\right|^2 \delta(k_z - k_z' \pm q_z) \tag{3.14}$$

The final scattering rate from state *n* to *m* is then given as

$$\Gamma_{nm}^{piezo} = (K_{av})^2 \frac{e^2 \kappa T}{4\pi^2 \hbar^2 \epsilon_\infty} \sqrt{\frac{m^*}{2}} I_{overlap}^{piezo}(q_z) \frac{(1 + 2\alpha E_f)}{\sqrt{E_f(1 + \alpha E_f)}} \Theta(E_f) \tag{3.15}$$

Where $E_f = E_n - E_m + E_i$ and

$$I_{overlap}^{piezo}(q_z) = \iint \frac{1}{q_x^2 + q_y^2 + q_z^2} \left|I_{nm}{}^{piezo}(q_x, q_y)\right|^2 dq_x\, dq_y \tag{3.16}$$

One has to evaluate the integral in equation 3.16 numerically.

The piezoelectric constant (e_{pz}) can be expressed in terms of electromechanical coupling coefficient (K_{av}). K_{av} is expressed in literature [65] in terms of the longitudinal and transverse elastic and piezoelectric constants as

$$K_{av}^2 = \frac{e_{pz}^2}{\rho v_s^2 \varepsilon_\infty} \Rightarrow \frac{e_L^2}{C_L \varepsilon_\infty} + \frac{e_T^2}{C_T \varepsilon_\infty} \tag{3.18}$$

The cumulative scattering rate for piezoelectric scattering out of subband 1 is plotted in figure 3.3.

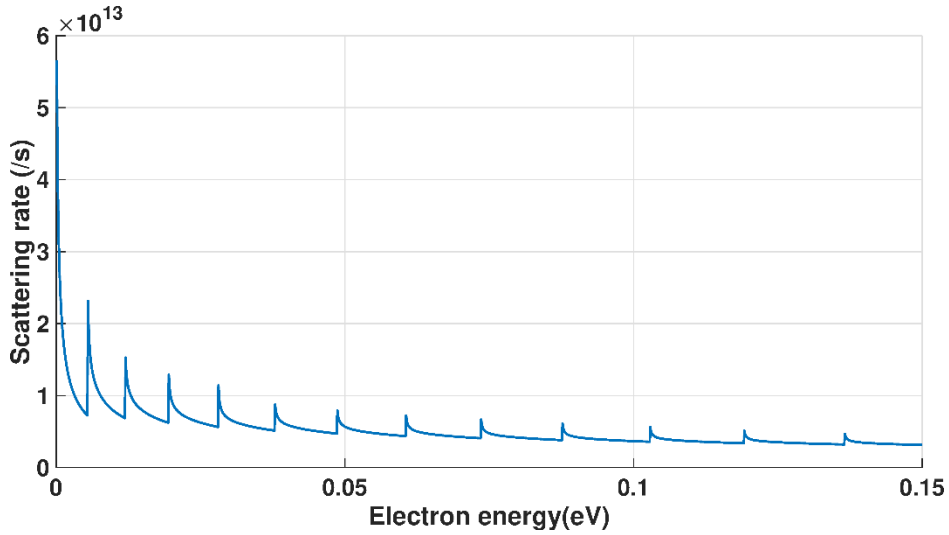

Figure 3.3 Cumulative Piezoelectric scattering rate out of subband 1

Polar optical phonon Scattering

Optical Phonons arise out of lattice vibrations wherein the constituent atoms oscillate in opposite directions creating a relative displacement. This does not induce lattice strain, rather alters the size of the unit cell [66]. Therefore, any perturbation potential arising out of optical phonon interaction is a product of the optical deformation potential and the lattice displacement

$$\hat{H} = \mathrm{D}_o \boldsymbol{u}(\boldsymbol{r}) \tag{3.19}$$

In polar materials, the optical phonons cause oscillatory behavior in the existing dipole. Electrons are scattered by this long-range dipole field which oscillates in time and space [67]. This forms the basis of the perturbing potential. Although there is no clear distinction between longitudinal mode (LO) and transverse mode (TO) optical phonons, in Wurtzite materials, it is seen that LO phonons have higher scattering rates than TO phonons[38]. This work considers only the longitudinal mode (LO) optical phonons as participating in polar optical scattering. The interaction between LO optical phonons and

electrons is called Fröhlich interaction and the Fröhlich Hamiltonian[68] is used in determination of the matrix element. The Fröhlich Hamiltonian can be expressed as

$$\widehat{H}_{pop} = \frac{iee^*}{\varepsilon_\infty \Omega} \sum_q \sqrt{\frac{\hbar}{2\mu N \omega_q}} \frac{1}{q} \left(a_q \boldsymbol{e^{iq.r}} - a_q{}^+ \boldsymbol{e^{-iq.r}} \right) \tag{3.20}$$

Where, e* is the effective charge and is given by

$$e^{*^2} = \mu \Omega \omega_{LO}^2 \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon(0)} \right) \tag{3.21}$$

Using the fact that $\omega_q = \omega_{LO}$ for optical phonons, the matrix element is given as

$$|M(k_z, k_z')|^2 = \frac{\hbar e^2 \omega_{LO}}{2V} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon(0)} \right) \frac{1}{q^2} \left(N_q + \frac{1}{2} \pm \frac{1}{2} \right) \left| I_{nm}{}^{pop}(q_x, q_y) \right|^2 \delta(k_z - k_z' \mp q_z) \tag{3.22}$$

Thus, the final scattering rate from subband n to m is given as

$$\Gamma_{nm}^{pop} = \frac{e^2 \omega_{LO}}{8\pi^2} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon(0)} \right) \sqrt{\left(\frac{2m}{\hbar^2} \right)} \left(N_q + \frac{1}{2} \pm \frac{1}{2} \right) I_{overlap}^{pop}(q_z) \frac{\left(1 + 2\alpha E_f\right)}{\sqrt{E_f\left(1 + \alpha E_f\right)}} \tag{3.23}$$

The overlap integral $I_{overlap}^{pop}(q_z)$ is a function of the final electron wavevector and is evaluated using

$$I_{overlap}^{pop}(q_z) = \iint \frac{1}{q_x^2 + q_y^2 + q_z^2} \left| I_{nm}{}^{pop}(q_x, q_y) \right|^2 dq_x \, dq_y \tag{3.24}$$

where

$$I_{nm}{}^{pop}(q_x, q_y) = \iint \psi_m(x, y) \exp\left(i\left(q_x x + q_y y \right) \right) \psi_n(x, y) dx dy \tag{3.25}$$

Since the final electron wavevector is limited to only a component along the *z* axis, this results in q_z taking on only two possible values

$$q_z = k_z' \mp k_z$$

These are termed as forward and backward scattering, respectively. The cumulative scattering rate for polar optical phonon scattering out of subband 1 is plotted in figure 3.4.

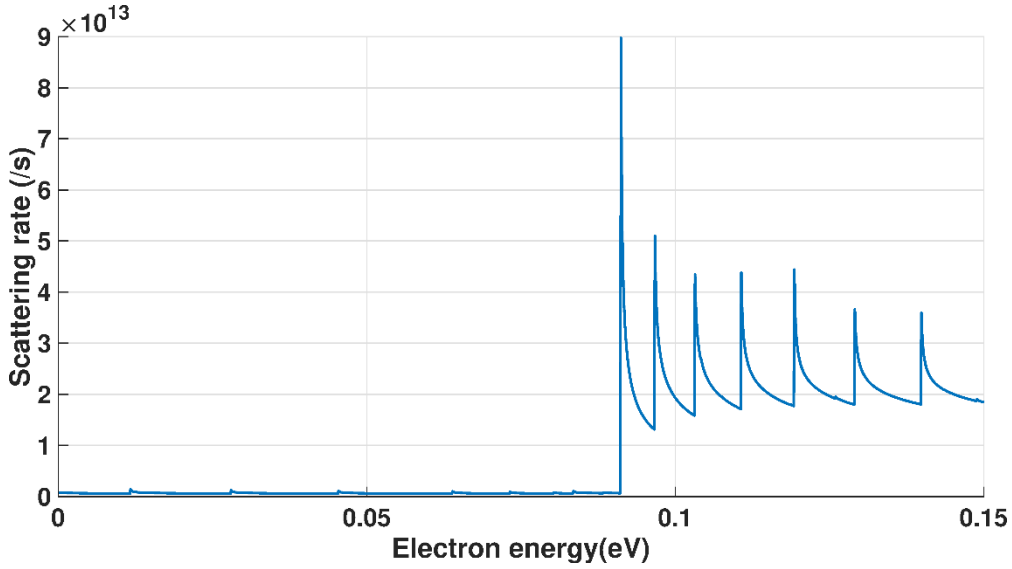

Figure 3.4 Cumulative Polar optical phonon scattering rate out of subband 1

Interface Roughness Scattering

Interface roughness scattering arises due to surface irregularities at the interface between AlGaN/GaN regions. The AlGaN/GaN interface exists along the growth direction, that is, the *y* direction. This invariably leads to variations in confining potential ϕ along the growth direction (y). These variations lead to scattering of electrons as they move closer to the interface. Since the electrons in the Q1D gas exist close to the interface, interface roughness scattering plays a key role in impacting the mobility of electrons in the channel. One can adopt the approach used by Goodnick et al. [69] to model Si-SiO_2 interfaces to this work as well. To that end, one can assume that the interface between AlGaN and GaN is an abrupt boundary with $\Delta(z)$ describing the random interface fluctuations. Secondly, the

electron wavefunction is assumed to be not perturbed by the random interface fluctuations. Under these assumptions, expanding the confining potential around the interface,

$$\phi\big(x, y + \Delta(z)\big) = \phi(x, y) + \frac{\partial \phi}{\partial y}\Delta(z) + \cdots \tag{3.26}$$

$$\phi\big(x, y + \Delta(z)\big) \cong \phi(x, y) + e\varepsilon_y \Delta(z) \tag{3.27}$$

Where ε_y is the electric field along y. Thus the perturbing potential is given as

$$H_{IR} = e\varepsilon_y(x, y)\Delta(z) \tag{3.28}$$

The matrix element of scattering from subband n to m is then expressed as

$$|M(k_z k_z')|_{nm} = e \iint \psi_n(x, y)\varepsilon_y(x, y)\psi_m(x, y)\mathrm{dxdy}$$
$$\frac{1}{L_z}\int \exp(i(k_z - k_z') \cdot z)\, \Delta(z)\mathrm{dz} \tag{3.29}$$

Grouping together the terms that make up the overlap integral leads to

$$I_{nm}^{IR} = \iint \psi_n(x, y)\varepsilon_y(x, y)\psi_m(x, y)\mathrm{dxdy}$$

One can then write the expectation value of the matrix element as

$$\langle |M(k_z, k_z')|^2 \rangle = e^2 |I_{nm}^{IR}|^2 \, L_z^{-2} \int dz_1 \int dz_2 \, e^{i(z_1 - z_2)(k_z - k_z')} \langle \Delta(z_1)\Delta(z_2) \rangle \tag{3.30}$$

As done in [69], one can assume that the autocorrelation function is exponential,

$$\langle \Delta(z_1)\Delta(z_2) \rangle = \Delta^2 exp\left(-\frac{\sqrt{2}|z|}{L}\right)$$

Where, Δ is the rms roughness and L is the correlation length for GaN interfaces

Thus, the square of the matrix element can be written as

$$|M(k_z, k_z')|^2 = e^2 |I_{nm}^{IR}|^2 \, L_z^{-2} \int dz_1 \int dz_2 \, e^{i(z_1 - z_2)(k_z - k_z')} \Delta^2 \exp\left(-\frac{\sqrt{2}|z|}{L}\right) \tag{3.31}$$

Writing $q_z = k_z \mp k_z', we\ have$

$$S(q_z) = \int e^{iq_z z}\, \Delta^2 \exp\left(-\frac{\sqrt{2}|z|}{L}\right) dz = \int e^{iq_z z}\, \Delta^2 \exp\left(-\frac{|z|}{L'}\right) dz$$

Where $L' = L/\sqrt{2}$. This leads to

$$|M(k_z, k_z')|^2 = e^2 |I_{nm}^{IR}|^2\, L_z^{-1} S(q_z) \tag{3.32}$$

Where, the power spectrum $S(q_z)$ of the autocorrelation function is

$$S(q_z) = \frac{\Delta^2}{iq_z + \frac{1}{L'}} - \frac{\Delta^2}{iq_z - \frac{1}{L'}}$$

$$S(q_z) = \frac{2\Delta^2 L'}{1 + q_z^2 L'^2} = \frac{4\Delta^2 L}{\sqrt{2}(2 + q_z^2 L^2)} \tag{3.33}$$

Thus, using Equations 3.32 and 3.33, the final scattering rate from subband m to n is written as

$$\Gamma_{nm}^{IR} = \frac{2\sqrt{m^*}e^2}{\hbar^2} \frac{\Delta^2 L}{2 + q_z^2 L^2} |I_{nm}^{IR}|^2 \frac{\left(1 + 2\alpha E_f\right)}{\sqrt{E_f\left(1 + \alpha E_f\right)}} \Theta(\mathrm{E_f}) \tag{3.34}$$

The cumulate interface roughness scattering rate out of subband 1 is given in figure 3.5

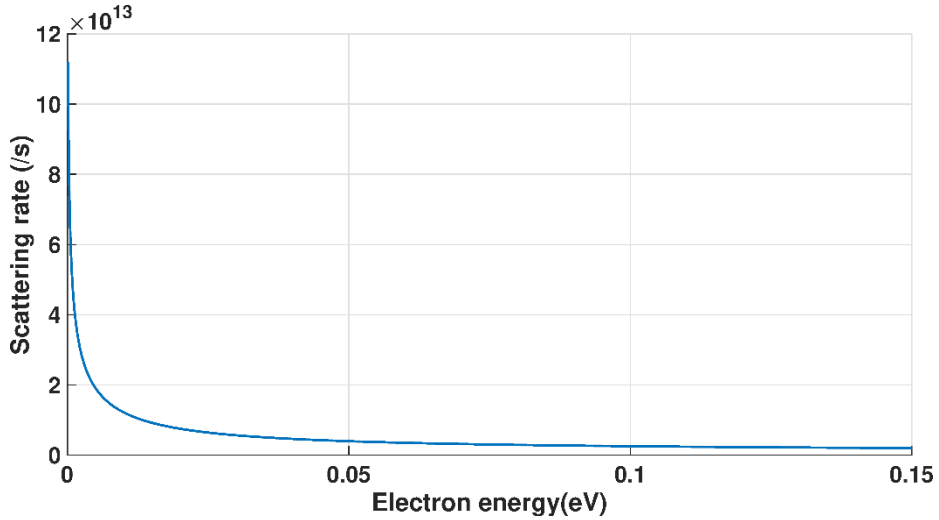

Figure 3.5 Cumulative Interface Roughness scattering rate out of subband 1.

Scattering due to Alloy Clustering

Substantial progress in AlGaN/GaN and InAlN/GaN research has been made over the past ten years, but crystal imperfections are still a major obstacle to realize high performance HEMTs and Finfets. Due to large differences in binding energies between Al-N (2.74 eV) and In-N (2.08 eV) bonds, InAlN/GaN heterostructures, especially those grown by molecular beam epitaxy, generally exhibit severe lateral compositional inhomogeneity. The effect is less pronounced for the AlGaN/GaN heterointerfaces because the binding energy of the Ga-N bond is 2.45 eV, but the effect is still there [70]. In compositionally inhomogeneous InAlN (AlGaN) layers, alternate AlN-rich and InN (GaN)-rich regions form clusters. Alloy clustering causes 2DEG subband energy fluctuations and InAlN (AlGaN) conduction band fluctuations (see figure 3.6), both affecting the electron transport properties of InAlN/GaN and AlGaN/GaN heterostructures.

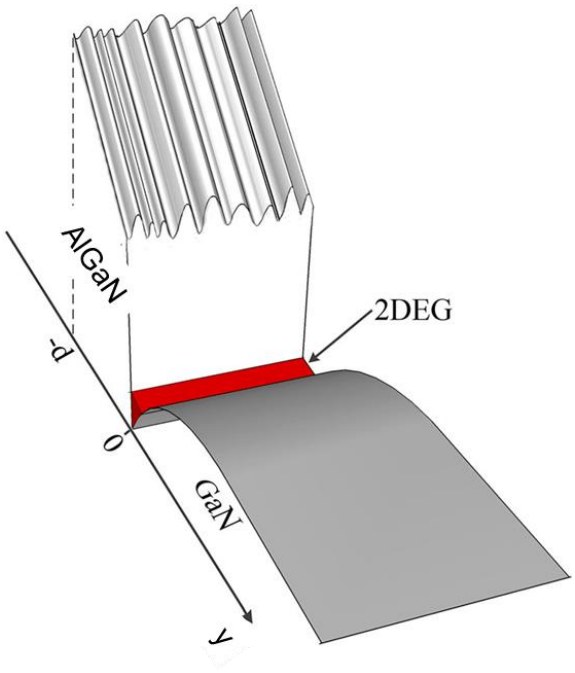

Figure 3.6 Schematic diagram of the conduction band in an AlGaN/GaN heterostructure with columnar clusters. [71]

Ahmadi et al. [72] and Liu et al. [73] studied the subband energy fluctuation scattering (SEFS) and suggested that the SEFS played an important role in the electron transport, especially for lower 2DEG sheet densities. The SEFS, although closely related to the InAlN and AlGaN conduction band fluctuations, is independent of electrons penetrating into the InAlN or AlGaN barrier layers. Hence, the interactions of InAlN and AlGaN conduction band fluctuations and electrons penetrating into the InAlN (AlGaN) barrier layers behaves as another scattering mechanism. In contrast with the SEFS, the conduction band fluctuation scattering (CBFS) taking place in InAlN (AlGaN) barrier layers is directly related to the degree of the overlap between the InAlN (AlGaN)

conduction band fluctuations and the electronic wave function. The CBFS has been studied in the literature in the works of Li et al. [71].

In this dissertation a theoretical model is presented for the SEFS for the AlGaN/GaN MISFET device structure that is in the focus of the present research, and the theoretical results and comparisons with experimental data are analyzed for evaluating the role of the SEFS in electron transport in the nano-Finfet. According to our knowledge, this is the first study of the effect of SEFS for nanoscale Finfets where we have confinement in two spatial directions and carriers are free to move only in the z-direction. Also, the approach of Ahmadi et al. [72] is expanded for a multi-subband case. The effect of CBFS, even though less important for AlGaN/GaN heterointerfaces, will be studied in future work.

Following the work of Ahmadi et al. [72] , one can assume that SEFS is similar to interface-roughness scattering (discussed in the work of Goodnick et al.[69]). In the case of interface roughness, changes in the width of the quantum well cause fluctuations in the energy levels of the 2DEG, whereas in the case of alloy clustering, variations in the depth of the quantum well change the energy levels. Therefore, a local change in the composition causes perturbing potential of the form:

$$\Delta E_n(x,z) = \frac{\partial E_n}{\partial X}\Delta X(x,z) \tag{3.35}$$

Where the vector $r = r(x,z)$ is a two-dimensional vector in the plane parallel to the heterointerface. It is difficult to obtain the exact form for $\Delta x(r)$ in a real heterostructure due to complicacy of the Al compositional distribution. For mathematical convenience,

this work has adopted the Gaussian autocorrelation function to describe the statistical properties of Al composition distribution, which can be expressed as:

$$\langle \Delta x(r)\Delta x(r')\rangle = \Delta^2 \exp\left(-\frac{(r-r')^2}{L^2}\right) \tag{3.36}$$

Using Gaussian model for the autocorrelation function allows one to describe the Al composition distribution by two parameters: the standard deviation Δ and the correlation length L of the Al composition distribution.

Following the work of Goodnick and co-workers [69], the matrix element for SEFS, for a nanowire system, is of the form:

$$|M(k_z,k_z')|_{nm} = \frac{1}{L_z}\iiint \psi_m(x,y)\Delta E_n(x,z)e^{-iq_z.z}\,\psi_n(x,y)dxdydz \tag{3.37}$$

And the magnitude of the matrix element squared for alloy clustering scattering is:

$$|M(k_z,k_z')|^2_{nm} = \frac{1}{L_z^2}\iiint dxdydz \iiint dx'dy'dz'\,\psi_m(x',y')\psi_n(x',y') \Delta E_n(x,z)\Delta E_n(x',z')\psi_m(x,y)\psi_n(x,y)e^{-iq_z.(z-z')} \tag{3.38}$$

Doing statistical averaging gives

$$\langle \Delta E_n(x',z')\Delta E_n(x,z)\rangle = \left|\frac{\partial E_n}{\partial X}\right|^2 \Delta^2 e^{-\frac{(x-x')^2}{L^2}-\frac{(z-z')^2}{L^2}} \tag{3.39}$$

Where Δ is the r.m.s amplitude of the fluctuation and *L* is the autocorrelation (AC) length. Using equation 3.39 into statistically averaged equation 3.38 gives

$$|M(k_z,k_z')|^2_{nm} = \left|\frac{\partial E_n}{\partial X}\right|^2 \Delta^2 \frac{1}{L_z^2}\iint dxdy \iint dx'\,dy'\psi_m(x,y)\,\psi_m(x',y') e^{-\frac{(x-x')^2}{L^2}}\psi_n(x,y)\psi_n(x',y') \int dz \int dz'\, e^{-iq_z.(z-z')-\frac{(z-z')^2}{L^2}} \tag{3.40}$$

Writing the 4-dimensional overlap integral as

$$I_{nm}^{alloy} = \iint dxdy \iint dx'\,dy'\psi_m(x,y)\,\psi_m(x',y')e^{-\frac{(x-x')^2}{L^2}}\psi_n(x,y)\psi_n(x',y') \quad (3.41)$$

The scattering matrix element becomes

$$|M(k_z,k_z')|_{nm}^2 = \left|\frac{\partial E_n}{\partial X}\right|^2 \Delta^2 \frac{1}{L_z}\sqrt{\pi}Le^{-\frac{L^2q_z^2}{4}}I_{nm}^{alloy} \quad (3.42)$$

Using Fermi's golden rule, and assuming that the scattering process is elastic, the total scattering rate out of subband *n* to subband *m*, due to this scattering mechanism is:

$$\Gamma_{\mathrm{nm}}^{\mathrm{alloy}}(k_z) = \sum_{\mathrm{k_z'}} \frac{2\pi}{\hbar}\,|\mathrm{M(k_z,k_z')}|_{\mathrm{nm}}^2\delta(\mathrm{E'} - E)$$

The final scattering rate expression, after some algebra, reduces to:

$$\Gamma_{\mathrm{n}}^{\mathrm{alloy}}(k_z) = \sum_m \sqrt{\frac{\mathrm{m^*\pi}}{2\hbar}}\Delta^2\mathrm{L}\left|\frac{\partial E_n}{\partial X}\right|^2 I_{nm}^{alloy}\frac{\left(1+2\alpha E_f\right)}{\sqrt{E_f\left(1+\alpha E_f\right)}}e^{-\frac{L^2q_z^2}{4}} \quad (3.43)$$

Where $q_z = \pm k_z' - k_z$ denoting forward and backward scattering respectively. The final energy $E_f = E_i + E_m - E_n$.

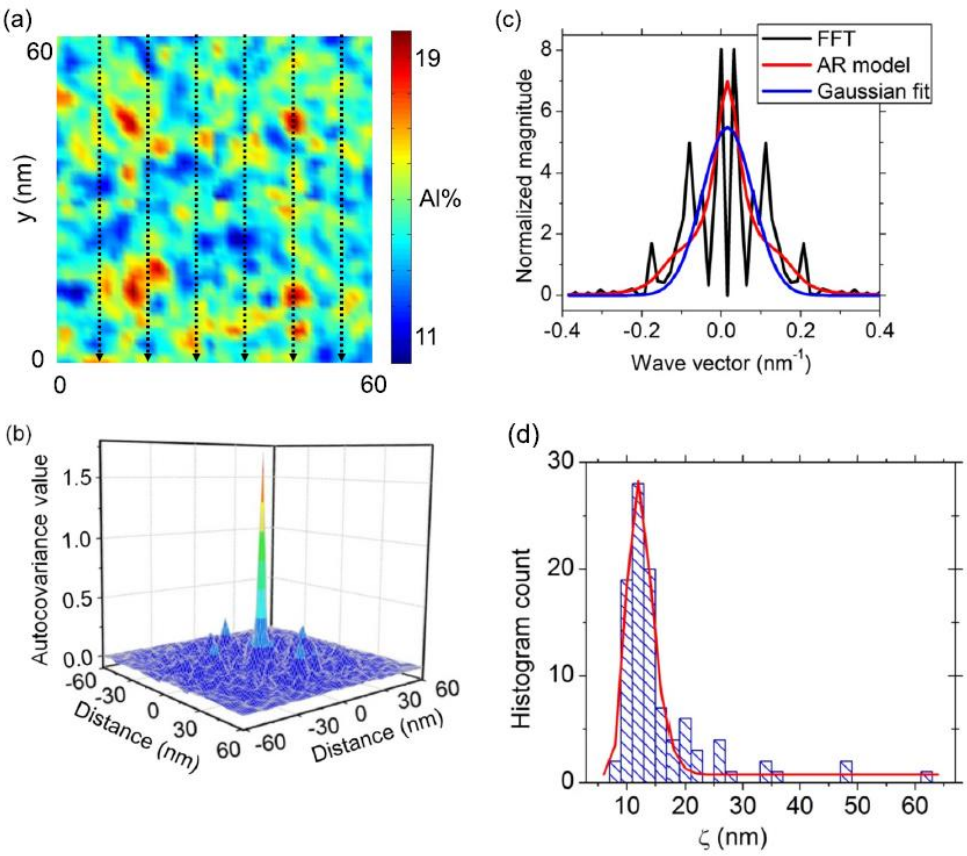

Figure 3.7 (a) Composition map of the in-plane Al distribution in an $Al_{0.15}Ga_{0.85}N$ layer (Black arrows illustrate the directions along which the AC lengths were calculated). (b) AC sequence of digitized data shown in part (a). (c) An example of the power spectrum calculated using both the FFT and AR methods and the fitted Gaussian function. (d) Histogram of the AC lengths obtained from different areas on the 2D III-site composition map. Reprinted with permission from [72].

To characterize alloy clustering, in ref. [72] , Atomic Probe Tomography (APT) was performed on the $Al_{0.15}Ga_{0.85}N$ electron blocking layer of a commercial c-plane (0001) GaN LED. The in-plane Al distribution in the $Al_{0.15}Ga_{0.85}N$ layer was reconstructed by averaging the Al mole fraction over 3 nm along the c-axis. To obtain a

significant number of sampling points, a 60×60 nm^2 in-plane composition map was generated by combining 30×30 nm^2 composition maps extracted from different regions in the AlGaN layer. The digitalized data were then used to calculate the 2D autocorrelation function (AC). The root mean square (rms) value of Al composition fluctuations (Δ) was obtained from the zeroth coefficient of the 2D AC sequence. To estimate the AC length, the composition profile was sampled along different directions, and the AC function was calculated for each profile. Since the Fourier transform of the AC function (the power spectrum) is the quantity that appears in the scattering rate rather than the AC function itself, the power spectrum was calculated next using Fast Fourier Transform (FFT). Autoregressive (AR) model was used to smooth out the power spectral density. The AC length was obtained by fitting a Gaussian function to the power spectrums of the 1D composition sequences. The distribution of f-values was next characterized by a log-normal function, and, for this particular sample, the AC length was estimated to be 12.3 nm from the expectation value of the log-normal distribution fit. The results from ref. [72] are summarized in figure 3.7.

In the results section below (Section 3.2), we vary both the rms value of the Al composition fluctuations and the AC length to get estimates of the limiting effect of alloy clustering (SEFS) on the low-field electron mobility in the AlGaN/GaN MISFET structure.

To calculate the variation of the subband energy E_n with the mole fraction X (the scattering potential), the variation of the energies of the lowest 10 subbands as a function of the mole fraction in the $Al_xGa_{1-x}N$ layer was calculated (figure 3.8). As expected, one can observe almost linear variation of the subband energy with the Al mole fraction. The

slope of this curve gives the strength of the scattering potential, and for the structure investigated, it equals approximately to -0.596 eV for each of the lowest 10 subbands at T=300 K. By entering these values in equation 3.43, the scattering rate out of subband *n* due to alloy clustering for different AC lengths and fluctuation amplitudes is calculated, as discussed later in the text.

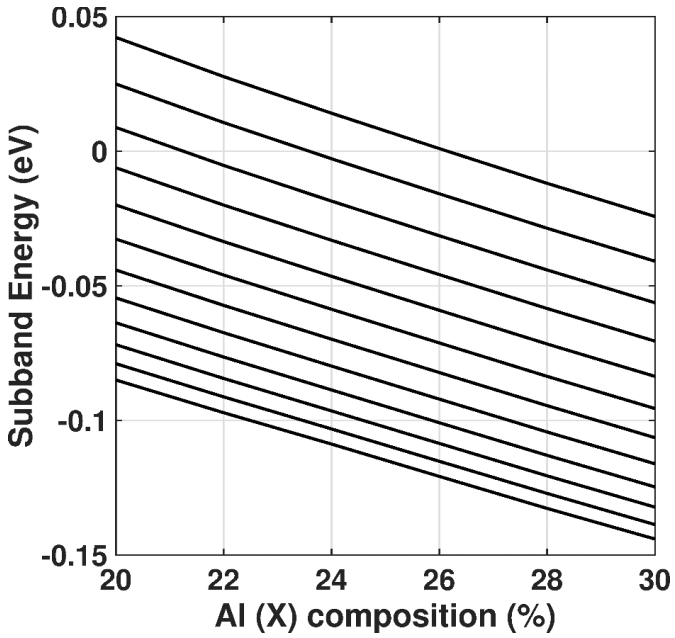

Figure 3.8 Variation of the 10 lowest subband energies as a function of the mole fraction at T=300 K. The slope of the curves gives the strength of the scattering potential.

The scattering rate for scattering out of subband 1, for the case when the overlap integral is calculated using the result given in equation 3.41 and when the overlap integral is assumed to be unity only when $n = m$, is shown in figure 3.9.

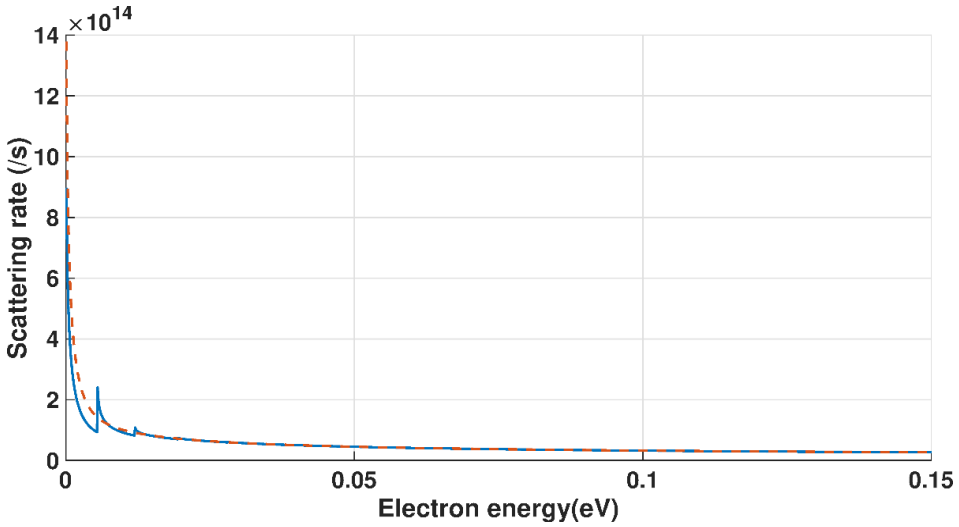

Figure 3.9 Scattering rate vs. energy for SEFS for the case when the overlap integral is calculated using the result given in equation 3.41 (most general result) (solid line) and, when the overlap integral is assumed to be unity only when n = m (intra-subband scattering only) (dashed line).

3.1.3 Brief Summary of the 1D Monte Carlo Algorithm

This section briefly describes the algorithm used to solve the Boltzmann transport equation for Q1D low-field transport in a nanowire Finfet. The 1D Monte Carlo code calculates mobility by simulating transport along the unconfined *z*-direction. Electrons are moved in k space in response to the electric field applied along the positive z-direction. Since this work is concerned with low field mobility calculation, the applied electric field is restricted to a few kV/cm (1-10) in the transport direction. Overlap integrals between the initial and final states are a crucial factor in the calculation of the scattering rates. The number of electrons taking part in the simulation is set to 5×10^6.

Since the motion of electrons is restricted to only one direction, the average energy of the carriers without any applied electric field is, according to Boltzmann distribution,

equal to $\frac{1}{2}\kappa T$. The momentum vector is determined according to non-parabolic E-k relationship and assigned either a forward (+z) or a backward (-z) direction according to a uniformly distributed random variable. The number of subbands being considered depends on the size of the Schrödinger domain and, hence, the fin width. The electrons are distributed among different *subbands* by using the percentage population of the *subbands* obtained from the *subband density* file output by the Schrodinger Poisson Solver. The rest of the algorithm follows generic ensemble Monte Carlo process of free-flight (drift) and scatter [74]. At steady state, ensemble averages of the drift velocity and the average carrier energy are calculated, which, in turn, allows calculation of the low-field electron mobility.

3.2 Mobility Modeling Results (without the effect of alloy clustering)

From the 1D Monte Carlo Simulation, the subband mobility and the total mobility are extracted in the following manner. Upon reaching steady state, the electron mobility is calculated by a weighted average of the mobility of each subband, which is, in turn, calculated by dividing the time-averaged subband velocity with the applied electric field ε. Thus, mobility of the n^{th} subband is calculated as

$$\mu_n = \frac{v_n^{avg}}{\varepsilon} \tag{3.44}$$

For M subbands, the weighted average of the mobility is

$$\mu = \frac{\sum_{i=1}^{M} \mu_i N_i}{\sum_{i=1}^{M} N_i} \tag{3.45}$$

Where N_i is the population of the i^{th} subband. The electron mobility calculation is repeated for Finfet structures with fin widths ranging from 40 nm (narrow fin) to 200 nm (wide fin)

devices. The variation of the electron mobility with the gate bias and for different temperatures is also simulated.

3.2.1 Validation with experimental results

The mobility results presented in this work are compared to experimental data presented in the work of Ki Sik et al. [34]. The cited work presents the variation of the transconductance with temperature in addition to the electron mobility results. These trends are presented for a wide as well as narrow Finfets. The transconductance variation plots from the paper are shown in figure 3.10. As previously mentioned, the transport model is split into two different regimes based on the conduction channel. The results presented in this work deal with the conduction along the interface channel when $V_G \leq 0$ V. Thus, it is sensible to validate the results with corresponding experimental data for $V_G \leq 0$V.

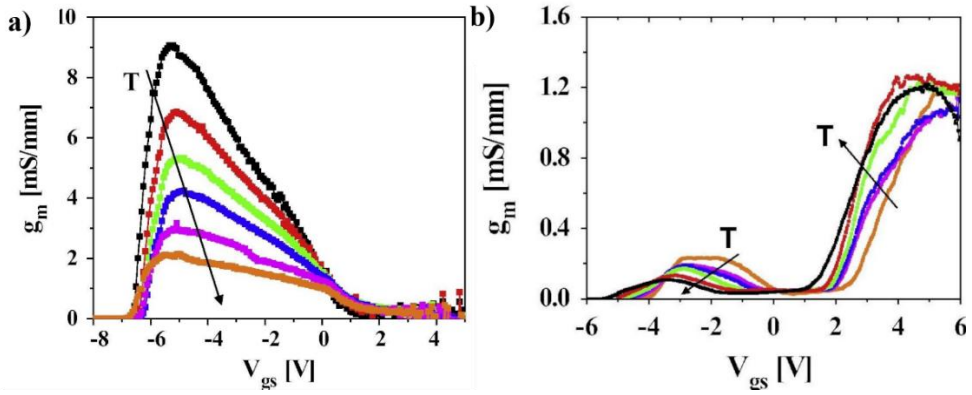

Figure 3.10 Transconductance vs. gate bias plots at different temperatures for a) Wide Fin device (400nm). b) narrow fin device (50nm) Reprinted from [34]© 2016 with permission from Elsevier.

The relation between transconductance and electron mobility is.

$$g_m = \frac{\delta I_D}{\delta V_G}\Big|_{V_{DS}} = A\mu V_{DS} \tag{3.46}$$

Where the various constant terms have been grouped together under the constant 'A'.

Thus,

$$\frac{\delta \log g_m}{\delta T} \sim \frac{\delta \log(\mu)}{\delta T} \tag{3.47}$$

Therefore, the simulation results can be deemed validated if the transconductance plots vs. temperature T and the electron mobility variation vs. T have the same slopes on a semi-log plot. Note that the transconductance values are extracted at the gate bias where g_m peaks. Therefore, the transconductance and the mobility are compared at $V_G = -3V$.

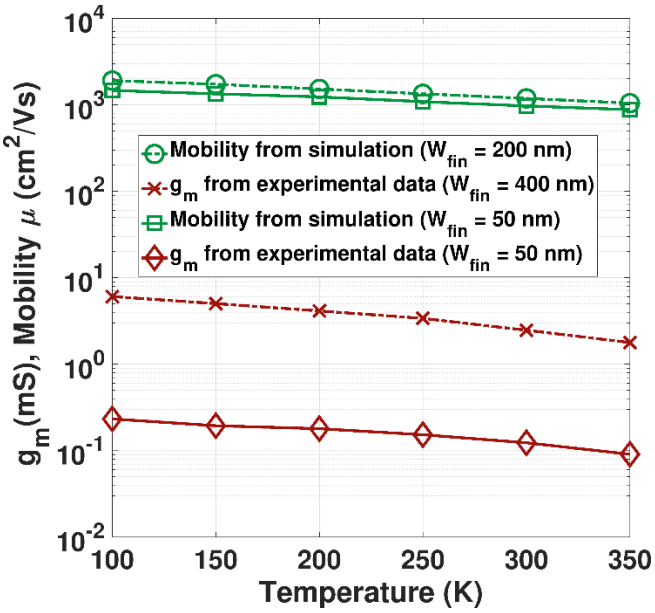

Figure 3.11 Simulated mobility compared with experimental values of transconductance for wide and narrow Finfets

3.2.2 Contribution to the electron mobility at Room Temperature

The electron mobility in the Finfet is influenced by the three phonon scattering processes considered and by interface roughness. The contribution of each mechanism can be determined by selectively 'turning on' only that particular scattering mechanism in the Monte Carlo module. One can then employ Mathiessen's rule to determine the contribution of the said mechanism to the overall electron mobility. At a gate bias of 0V and at room temperature (300K), the contribution of each scattering mechanism is shown in figure 3.12

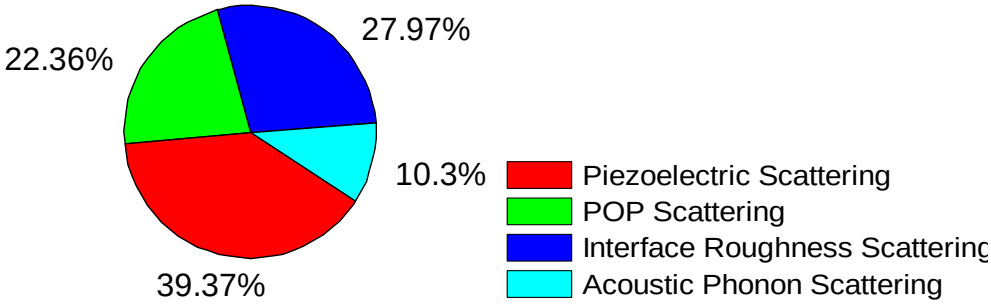

Figure 3.12 Contribution of each scattering mechanism at Room Temperature, $V_G = 0V$.

One can note that piezoelectric scattering rate is the dominant mobility scattering mechanism at 300K. Interface roughness scattering significantly affects the electron mobility as well. This is expected as the electrons are situated close to the AlGaN/GaN interface in the conduction channel. The polar optical phonon scattering does not dominate at room temperature due to the high energy of the polar optical phonon ($\sim$90 meV).

3.2.3 Electron Mobility as a function of the fin width.

The fin width is a critical design parameter in AlGaN/GaN Finfets. One can note the change in the transport properties for different fin width devices by observing the trend

in the electron mobility as a function of the fin width. Figure 3.13 shows the electron mobility vs. fin width plot for an AlGaN/GaN MIS Finfet device with an Al composition of 20%. The fin width range is varied from narrow fin (< 50 nm) to a wide (> 100nm) fin devices.

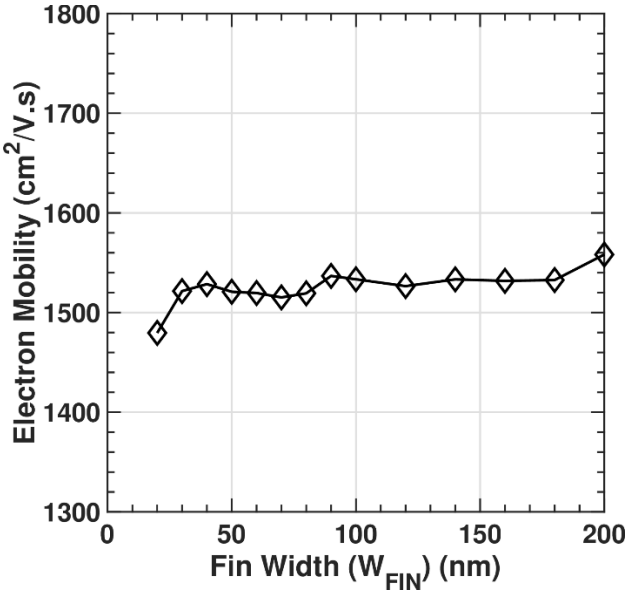

Figure 3.13 Electron Mobility plotted as a function of fin width for AlGaN/GaN Finfet. Al composition 20%.

As the fin width increases, the size of the lateral well also increases, increasing the number of subbands present below the Fermi level. This must be accounted for in the Schrödinger equation solver by increasing the number of eigenvalues requested. One must also include a higher number of subbands in the input to the Monte Carlo kernel. One finds that the mobility is almost invariant upon the variation of the fin width, however, one can note a definite transition point when the fin width increases beyond 100nm. Beyond 100 nm, it can be said that the lateral confinement reduces, and one starts

observing Quasi-2D behavior instead of a Quasi-1D behavior (which is the case for narrow fin devices). As the fin width increases, the impact of interface roughness scattering reduces as the average distance of an electron from the interface increases slightly [75]. It should be also noted that, as the fin width increases, the electron mobility approaches the experimental Hall mobility value for planar AlGaN/GaN HEMT devices[29], [76].

3.2.4 Electron Mobility as a function of the gate bias

The gate bias directly affects the induced electron density in the AlGaN/GaN interface channel. Furthermore, as seen in previous chapters, it also affects the lateral confinement, and hence the shape of the 2D well region. Thus, it is imperative to know the variation of the electron mobility in the interface channel as a function of the gate bias (V_G). It is to be noted that the gate bias is varied only in the range ($V_G \leq 0$ V) since this transport model is valid for conduction in the interface channel only. The electron mobility is plotted as a function of the gate bias in figure 3.14

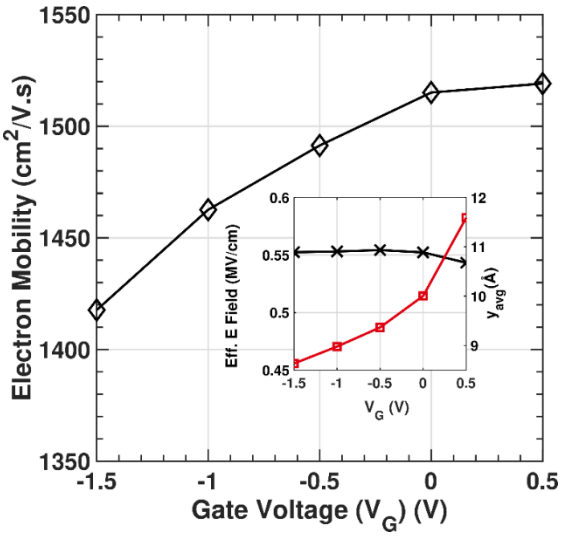

Figure 3.14 Variation of Electron mobility vs. Gate voltage V_G for AlGaN/GaN Finfet (Fin width = 40 nm. Al composition =20%). Inset: Average electron distance and Effective electric field as a function of V_G

The simulated device is an AlGaN/GaN MIS Finfet device similar to the one shown in figure 1.6. This is a narrow fin width device with $W_{Fin} = 40\,nm$ and with 20% Al composition. The gate bias is varied from 0.5V down to -1.5 V. The gate bias is not varied beyond 0.5 V as sidewall channels start to form when V_G > 0.5 V. The electron mobility drops as the gate bias reduces. As explained previously, the effective width (x) of the 2D well in the Schrödinger domain reduces with decreasing V_G. As the well width reduces, the electron wavefunctions are confined in narrower widths, increasing the overlap integral. The increase in the overlap integral increases the scattering rates, thus reducing the electron mobility as V_G reduces.

Average electron distance and effective electric field

Similar to the approach used in the analysis of the electron mobility in Si inversion layers in MOS devices[77], one can determine the effective electric field and the average

electron distance from the interface. This can help one analyze the impact of interface roughness. For the i[th]-subband, we let $\langle y \rangle$ represent the average position of electron along the well depth that is calculated using

$$\langle y \rangle_i = \iint \psi_i^*(x,y)\, y\, \psi_i(x,y) dxdy = \iint y \mid \psi_i(x,y)|^2 dxdy \tag{3.48}$$

$$y_{avg} = \frac{\sum_{i=1}^{N} \langle y \rangle_i n_i^l}{\sum n_i^l} \tag{3.49}$$

n_i^l represents electron line density of the i[th] subband and y_{avg} is the average carrier position. Let the interface be located at y_{int}. Thus, average carrier distance from the interface $\Delta y = y_{avg} - y_{int}$. This quantity is plotted in the inset of figure 3.14. Similarly, E_y is the vertical electric field seen by the electrons in the 2D well

$$\langle E_y \rangle_i = \iint E_y(x,y) \mid \psi_i(x,y)|^2 dxdy \tag{3.50}$$

$$Effective\ Electric\ Field = \frac{\sum_{i=1}^{N} \langle E_y \rangle_i n_i^l}{\sum n_i^l} \tag{3.51}$$

Unlike the case for Si, the effective electric field is almost invariant with decreasing gate bias. However, the average carrier distance decreases with decreasing gate bias. This implies that the electrons move closer to the interface with reducing gate bias, thus increasing the impact of interface roughness scattering and thus reducing electron mobility.

3.2.5 Impact of strain relaxation on electron mobility

The lateral strain experienced by the epitaxially grown AlGaN layer due to lattice mismatch is key to the induced electron density at the AlGaN/GaN interface. As

discussed previously, piezoelectric polarization is directly dependent on the lateral strain in the given layer.

In thin films and nanostructures, lateral strain relaxation has been observed[58], [78], specifically as the dimensions of the nanopillars reduce. Palacios et al.[30] investigated strain relaxation in AlGaN/GaN nano Finfets similar to the ones discussed in this work. Figure 3.15 is one of the key observations from this work, where they plot the bi-axial strain and the sheet resistance as a function of the nanoribbon width. One can relate the NR width to the fin width of the Finfets discussed in this work (refer to the inset in figure 3.16 showing the structure of their nanoribbons). As the nanoribbon width drops below 300 nm, the biaxial strain reduces. There is a corresponding increase seen in the sheet resistance.

One can investigate this effect using the simulation developed in this work. (It should be noted that the results discussed so far do not take any strain relaxation into account.) A strain relaxation parameter is introduced in the Schrödinger-Poisson equation solver which modifies calculation of the piezoelectric charge density. A strain relaxation parameter of 1 corresponds to 'no strain relaxation'; a parameter of 0.4 implies a strain relaxation of 60 %, and so on. Using this method, the electron line density and the electron mobility is calculated for three different values of strain relaxation: no strain relaxation, 30 % strain relaxation in the fin and 60 % strain relaxation, respectively. This simulation is run for an AlGaN/GaN Finfet device with a fin width of with $W_{Fin} = 40\ nm$ and a 20% Al composition. The electron mobility variation vs. gate bias is plotted in figure 3.16.

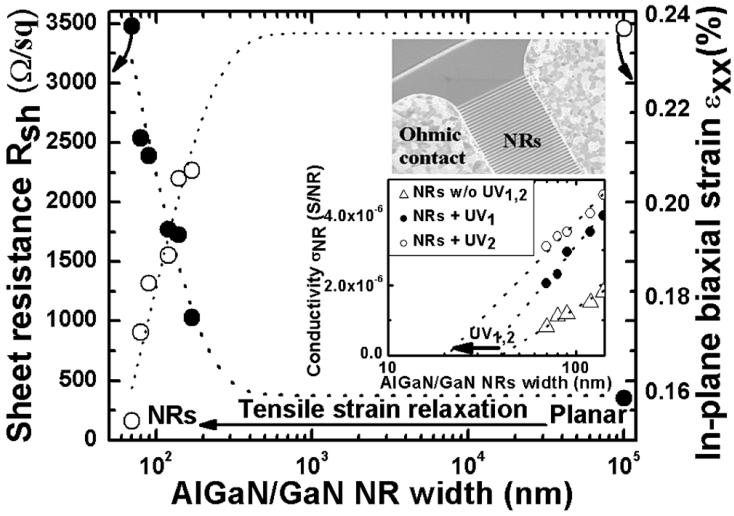

Figure 3.15 Tensile strain relaxation and corresponding sheet resistance change in AlGaN GaN nanoribbons. Reprinted from [30] with the permission of AIP Publishing.

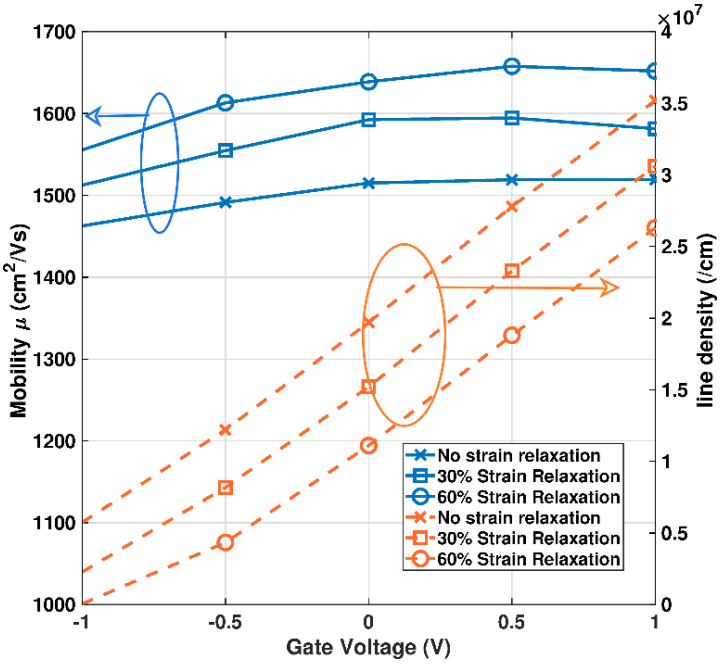

Figure 3.16 Impact of Strain relaxation. Electron Mobility (left) and electron line density (right) vs gate bias for different values of strain relaxation (W_{Fin} = 40 nm and a 20% Al composition).

As the strain relaxation parameter is increased, less piezoelectric polarization density is induced at the AlGaN/ GaN interface (this is seen in the decrease in electron line densities in figure 3.16). A small increase in electron mobility is observed as the lateral strain is reduced. This effect can be explained by the reduction in interface roughness scattering which is primarily influenced by the vertical electric field.

To test this hypothesis, the potential well along the depth is plotted in figure 3.17 for three different values of the strain relaxation parameter. The well becomes shallower as the strain relaxation increases. This results in the increase of average electron separation from the interface, as seen in the inset of figure 3.18. The average electron separation is calculated using equations 3.48 and 3.49. As discussed previously, the increase in average electron separation from the interface lowers the impact of interface roughness scattering and leads to a corresponding increase in electron mobility.

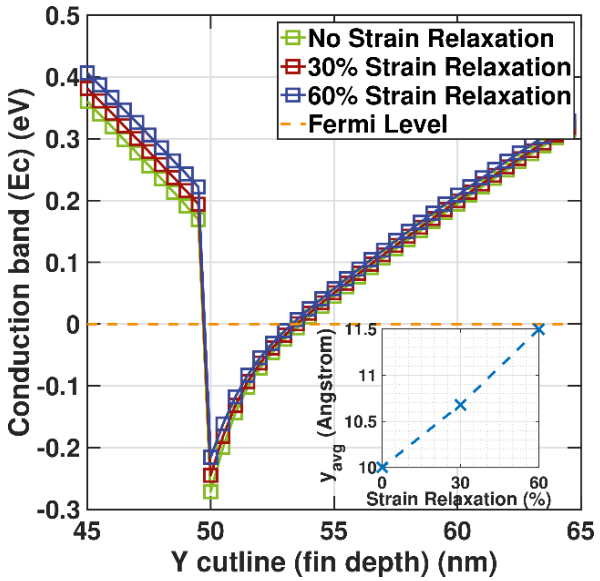

Figure 3.17 Change in the shape of the potential well with increasing strain relaxation. Inset: Average electron separation from the AlGaN/GaN interface.

Calculation of the Sheet resistance

One can calculate the sheet resistance of the conduction channel in the Quasi 1D region in the following manner. If y_{avg} is the average electron separation from the AlGaN/GaN interface, then one can estimate the thickness of the channel below the interface as $\sim 2\, y_{avg}$. If W_{Fin} is the width of the fin, n_l is the electron density and μ is the channel mobility, then the conductivity of the channel can be given as

$$\sigma = \frac{1}{\rho} = \frac{e\mu n_l}{2 y_{avg} W_{fin}} \tag{3.52}$$

Thus, assuming a thickness t of 1 nm, the sheet resistance is

$$R_{sh} = \frac{\rho}{t} = \frac{2\, y_{avg}\, W_{FIN}}{e\, n_l \mu} \tag{3.53}$$

The sheet resistance for different values of the strain relaxation parameter is given in Table 3.1. Comparing with the experimental results shown in figure 3.16, one can note that at 60% strain relaxation, the sheet resistance determined from the simulation of 4000 Ω/square matches well with the measured value of about 3500 Ω/square for a NR width of ~50 nm.

Table 3.1 Sheet resistance for different strain relaxation parameters.

Strain (%)	**Line Density, (10^7/cm)**	**Mobility μ (cm^2/V-s)**	**y_{avg} (Å)**	**R (Ohms/sq)**
0	**2**	**1500**	**10.0**	**2080**
30	**1.6**	**1600**	**10.7**	**2608**
60	**1.1**	**1630**	**11.5**	**4002**

3.3 The Role of Alloy Clustering on the low-field electron mobility

Figure 3.18 shows the mobility limited by alloy clustering (SEFS) as a function of the fluctuation amplitude for different AC lengths. As expected, the mobility limited by alloy clustering decreases as the fluctuation amplitude increases. The decrease of the autocorrelation length of the alloy clustering also leads to decrease in the mobility as lower values of L mean more pronounced clustering effect on a shorter length scale.

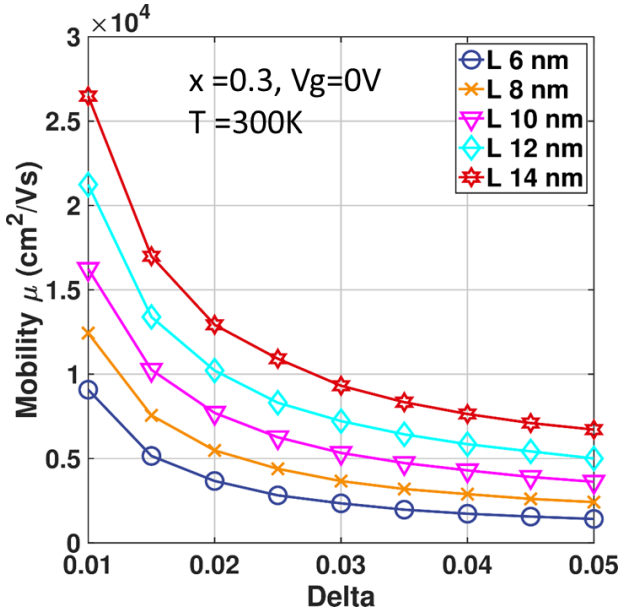

Figure 3.18 . Low field electron mobility limited by alloy clustering as a function of the rms (Delta) of the Al mole fraction variation and the autocorrelation (AC) length L

The variation of the low-field electron mobility as a function of the mole fraction x (Al composition) in the AlGaN layer is shown in Figure 3.19. The overall mobility decreases with increasing x due to the increased importance of piezoelectric scattering. As a result of this trend, the mobility degradation due to the additional alloy clustering is

about 8% for x=20% and about 6% for x=30%. We use L=10 nm for the autocorrelation length (AC). The rms of the alloy clustering scattering is Delta=0.03.

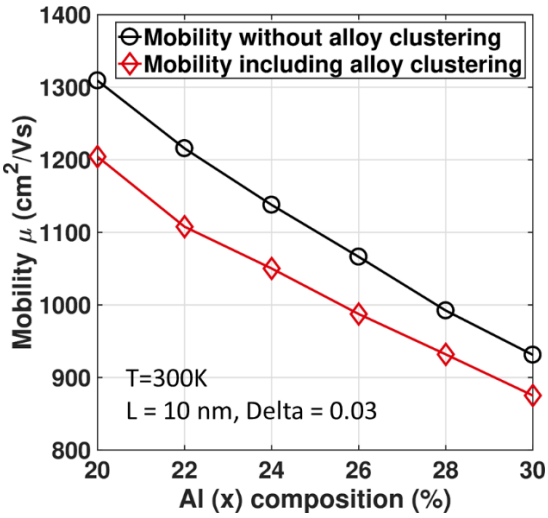

Figure 3.19 Low-field electron mobility as a function of Al mole fraction for the case when alloy clustering is excluded (open circles) and included (open diamonds) in the theoretical model.

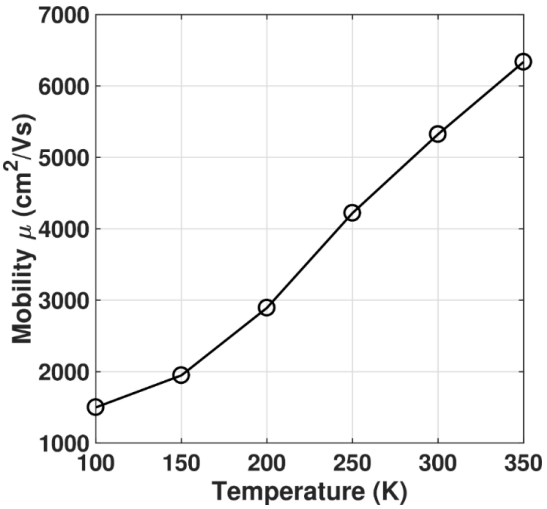

Figure 3.20 Alloy clustering (SEFS) limited mobility as a function of temperature. (L = 10 nm, $\Delta = 0.03$)

The variation of the alloy clustering limited mobility vs. temperature is shown in Figure 3.20. One can note that alloy clustering scattering has significant effect on the overall mobility values at low temperatures, and has very small effect at T=300K, as already shown in Figure 3.19. The larger impact of alloy clustering scattering at low temperatures is because, at low T, carriers occupy few of the lowest subbands, which makes the average separation of the carriers from the heterointerface to be smaller at T=100K as opposed to the case of T=350K, when the alloy clustering limited mobility is over 6,000 cm^2/V-s. At T=350K, the low-field electron mobility is dominated, i.e. limited by interface roughness, piezoelectric and polar optical phonon scattering (not shown on this figure).

3.4 Impact of Interface Roughness on the Low-Field Electron Mobility for the 50 nm width Finfets

The overall variation of the electron low-field mobility for a MIS Finfet with 50nm fin width, is shown in figure 3.21. All relevant scattering mechanisms, except for Coulomb scattering in the side-wall channels, are included in the theoretical model. Also shown in this figure are the experimental data from ref [34]. Parameter in these curves is the r.m.s. height (Δ) of the interface roughness (0.11, 0.3 and 0.35 nm). From the results shown, it is evident that, at high temperatures, interface-roughness is dominant scattering mechanism in this particular structure. Since at low temperatures Coulomb scattering (not included into our theoretical model), in addition to roughness scattering, limits the low-field mobility, there is a discrepancy between the simulated and the experimental data. Incorporation of Coulomb scattering will be done in future work.

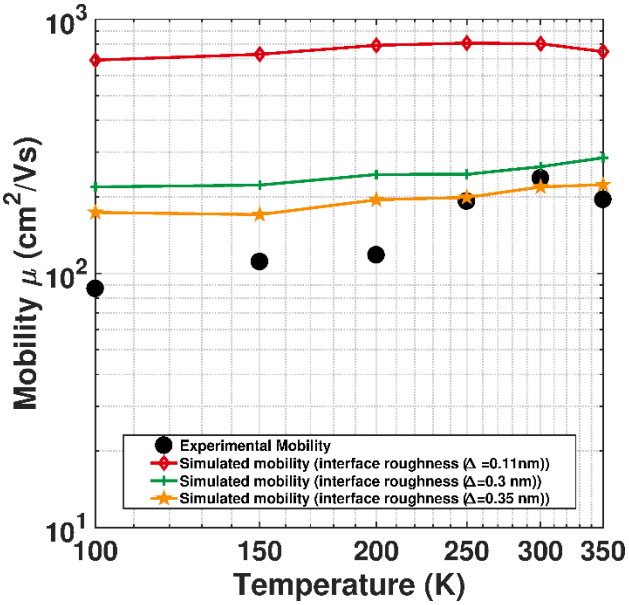

Figure 3.21 Comparison of the expermental low-field electron mobility data for MIS Finfet with 50nm fin width. The theoretical model includes interface roughness, alloy clustering (SEFS), and all modes of phonon scattering (acoustic, piezoelectric and polar). Parameter in the simulated data is the r.m.s height of interface roughness. For SEFS, L = 6 nm and Δ = 0.03 is used. The Al compostion (x) in the Finfet is 30%.

CHAPTER 4

CONCLUSIONS AND FUTURE WORK

The properties of GaN make it an excellent wide bandgap (WBG) material of choice for applications ranging from power and RF devices to optoelectronics. GaN has also been proposed as an alternative to Si in the high-speed ultra-low voltage (<50V) class of switching devices such as those needed for EVs and computing applications. Extending GaN to ULV necessitates a shift from HEMTs to Finfet devices. Compared to HEMTs, modeling studies on Finfets have been scarce. Previous studies used Kubo-Greenwood approach for determining electron mobility.

A novel simulator was developed in this work to solve the coupled Schrödinger-Poisson-Boltzmann solver for modeling mobility in lateral AlGaN/GaN MIS Finfets. A self-consistent 2D Schrödinger – Poisson solver along with a 1D Monte Carlo transport kernel is used for determining the electron mobility in the confined Q1D channel. The linearized and discretized 3D Poisson equation is solved using the Biconjugate Gradient stabilized (BiCGSTAB) method as the discretization results in a non-symmetric coefficient matrix. The Poisson solver results determine the extent of the Schrödinger domain, over which the 2D Schrödinger equation is solved. The self-consistent solution of Schrödinger-Poisson equation results in determination of subband energies in the 2D well, the quantum electron density and the electron wavefunctions. The electron line densities are determined using Fermi-Dirac statistics. Using line densities as a measure for the threshold voltage of the Finfet device, the simulator results are validated with experimental measurements from literature. The results from the solver show that narrow fin devices can function as E-mode devices with a positive threshold voltage due to

additional depletion of the channel from the side gates. Temperature variance of V_{TH} also reduces with fin width. The solver results also show distinct regimes of conduction; AlGaN/GaN interface channel (predominantly for $V_G \leq 0V$) and a combination of sidewall channels and interface channel (for $V_G > 0V$).

Presently, the transport kernel is developed only for the interface channel conduction. Thus, gate bias is restricted to $V_G \leq 0V$. The Schrödinger-Poisson solver results are transferred to the 1D Monte Carlo kernel. Three phonon scattering mechanisms (acoustic, polar optical phonon and piezoelectric scattering) and interface roughness scattering are considered. Total electron mobility is determined as a weighted average of mobility of each subband and is derived from the steady state drift velocities. The mobility results are validated by using transconductance values of a comparable Finfet structure. The equivalence between derivatives of transconductance and mobility with temperature is used to compare and validate the mobility modeling. At 300K and zero gate bias, piezoelectric and interface roughness scattering emerge as dominant scattering mechanisms. Electron mobility in the channel is found to reduce with gate bias due to increasing impact of interface roughness scattering.

The impact of lateral strain relaxation in narrow fin devices on the electron mobility in Finfets is discussed. The simulation results predict that, although increase in electron mobility due to decreased impact of interface roughness scattering is observed, strain relaxation reduces the induced electron density in the channel significantly which, in turn, leads to increase of the sheet resistance.

The impact of alloy clustering (Subband Energy Fluctuation Scattering) and interface roughness on the 1D mobility is studied for the case of narrow Finfets. Alloy clustering

largely affects the low temperature electron mobility. At room temperature, the effect of alloy clustering ranges from 8% drop in electron mobility for Finfet with Al mole fraction of 0.2 to 6% drop for Finfet with Al mole fraction of 0.3. Throughout the whole temperature range, interface roughness is the most dominant scattering mechanism. An r.m.s height of 0.35nm for the interface roughness leads to a good agreement with experimental measurements of low-field electron mobility of similar Finfets at higher temperatures.

The primary future development in this work is the incorporation of sidewall channels into the transport kernel. A separate and independent transport model will need to be developed for the sidewall conduction and the positive gate bias regime. Since the sidewall conduction would be affected primarily by electrons in the bulk GaN region, Coulomb scattering is predicted to play a major role on the magnitude of the electron mobility, especially at low T. In addition, one can also extend this work to include a full particle-based device simulator to determine the channel current and predict *IV*-characteristics.

REFERENCES

[1] B. Zhang and Y. Liu, “A review of GaN-based optoelectronic devices on silicon substrate,” *Chinese Sci. Bull.*, vol. 59, no. 12, pp. 1251–1275, 2014, doi: 10.1007/s11434-014-0169-x.

[2] K. H. Teo, N. Chowdhury, Y. Zhang, T. Palacios, K. Yamanaka, and Y. Yamaguchi, “Recent Development in 2D and 3D GaN devices for RF and Power Electronics Applications,” *2020 IEEE Int. Symp. Radio-Frequency Integr. Technol. RFIT 2020*, pp. 22–24, 2020, doi: 10.1109/RFIT49453.2020.9226187.

[3] Microsemi, “Gallium Nitride (GaN) versus Silicon Carbide (SiC) In The High Frequency (RF) and Power Switching Applications,” *Microsemi PPG*, p. 8, 2014, [Online]. Available: http://www.digikey.co.uk/Web Export/Supplier Content/Microsemi_278/PDF/Microsemi_GalliumNitride_VS_SiliconCarbide.pdf

[4] K. S. Im *et al.*, “Normally off single-nanoribbon Al2O3GaN MISFET,” *IEEE Electron Device Lett.*, vol. 34, no. 1, pp. 27–29, 2013, doi: 10.1109/LED.2012.2222861.

[5] S. Nakajima, “GaN HEMTs for 5G Base Station Applications,” *Tech. Dig. - Int. Electron Devices Meet. IEDM*, vol. 2018-Decem, pp. 14.2.1-14.2.4, 2019, doi: 10.1109/IEDM.2018.8614588.

[6] T. J. Flack, B. N. Pushpakaran, and S. B. Bayne, “GaN Technology for Power Electronic Applications: A Review,” *J. Electron. Mater.*, vol. 45, no. 6, pp. 2673–2682, 2016, doi: 10.1007/s11664-016-4435-3.

[7] H. Amano *et al.*, “The 2018 GaN power electronics roadmap - IOPscience,” *J. Phys. D Appl. Phys.*, vol. 51, 2018, [Online]. Available: https://iopscience.iop.org/article/10.1088/1361-6463/aaaf9d%0Ahttps://iopscience.iop.org/article/10.1088/1361-6463/aaaf9d/meta

[8] T. Ueda, “Gan power devices: Current status and future challenges,” *Jpn. J. Appl. Phys.*, vol. 58, no. SC, pp. 0–9, 2019, doi: 10.7567/1347-4065/ab12c9.

[9] Y. Zhang, “Comparison between competing requirements of gan and sic family of power switching devices,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 738, no. 1, 2020, doi: 10.1088/1757-899X/738/1/012004.

[10] Y. Zhang *et al.*, “GaN FinFETs and trigate devices for power and RF applications: Review and perspective,” *Semicond. Sci. Technol.*, vol. 36, no. 5, 2021, doi: 10.1088/1361-6641/abde17.

[11] U. K. Mishra, P. Parikh, and Y. F. Wu, “AlGaN/GaN HEMTs - An overview of device operation and applications,” *Proc. IEEE*, vol. 90, no. 6, pp. 1022–1031, 2002, doi: 10.1109/JPROC.2002.1021567.

[12] F. Zeng *et al.*, “A comprehensive review of recent progress on GaN high electron mobility transistors: Devices, fabrication and reliability,” *Electron.*, vol. 7, no. 12, 2018, doi: 10.3390/electronics7120377.

[13] E. A. Jones, F. F. Wang, and D. Costinett, “Review of Commercial GaN Power Devices and GaN-Based Converter Design Challenges,” *IEEE J. Emerg. Sel. Top. Power Electron.*, vol. 4, no. 3, pp. 707–719, 2016, doi: 10.1109/JESTPE.2016.2582685.

[14] K. Hoo Teo *et al.*, “Emerging GaN technologies for power, RF, digital, and quantum computing applications: Recent advances and prospects,” *J. Appl. Phys.*, vol. 130, no. 16, 2021, doi: 10.1063/5.0061555.

[15] O. Ambacher *et al.*, “Two dimensional electron gases induced by spontaneous and piezoelectric polarization in undoped and doped AlGaN/GaN heterostructures,” *J. Appl. Phys.*, vol. 87, no. 1, pp. 334–344, 2000, doi: 10.1063/1.371866.

[16] M. A. Alsharef, R. Granzner, and F. Schwierz, “Theoretical investigation of Trigate AlGaN/GaN HEMTs,” *IEEE Trans. Electron Devices*, vol. 60, no. 10, pp. 3335–3341, 2013, doi: 10.1109/TED.2013.2279264.

[17] S. Gökden, R. Baran, N. Balkan, and S. Mazzucato, “The effect of interface roughness scattering on low field mobility of 2D electron gas in GaN/AlGaN heterostructure,” *Phys. E Low-Dimensional Syst. Nanostructures*, vol. 24, no. 3–4, pp. 249–256, 2004, doi: 10.1016/j.physe.2004.04.042.

[18] J. Fang, M. V. Fischetti, R. D. Schrimpf, R. A. Reed, E. Bellotti, and S. T. Pantelides, “Electron Transport Properties of AlxGa1-x N/Ga N Transistors Based on First-Principles Calculations and Boltzmann-Equation Monte Carlo Simulations,” *Phys. Rev. Appl.*, vol. 11, no. 4, p. 1, 2019, doi: 10.1103/PhysRevApplied.11.044045.

[19] M. H. Ahmed, C. Fei, F. C. Lee, and Q. Li, “48-V Voltage Regulator Module with PCB Winding Matrix Transformer for Future Data Centers,” *IEEE Trans. Ind. Electron.*, vol. 64, no. 12, pp. 9302–9310, Dec. 2017, doi: 10.1109/TIE.2017.2711519.

[20] F. C. Lee, L. Fellow, and Q. Li, “High-Frequency Integrated Point-of-Load Converters : Overview,” vol. 28, no. 9, pp. 4127–4136, 2013.

[21] D. Reusch, J. Strydom, and J. Glaser, “Improving high frequency DC-DC converter performance with monolithic half bridge GaN ICs,” *2015 IEEE Energy Convers. Congr. Expo. ECCE 2015*, pp. 381–387, 2015, doi: 10.1109/ECCE.2015.7309713.

[22] “See How GaN is Leading the 48 V Revolution Across Multiple Industries at APEC 2022,” 2022. https://epc-co.com/epc/GaNTalk/Post/21455

[23] M. F. Fatahilah, K. Strempel, F. Yu, S. Vodapally, A. Waag, and H. S. Wasisto, “3D GaN nanoarchitecture for field-effect transistors,” *Micro Nano Eng.*, vol. 3, no. December 2018, pp. 59–81, 2019, doi: 10.1016/j.mne.2019.04.001.

[24] Y.-W. Jo *et al.*, “AlGaN/GaN FinFET With Extremely Broad Transconductance by Side-Wall Wet Etch,” *IEEE Electron Device Lett.*, vol. 36, no. 10, pp. 1008–1010, Oct. 2015, doi: 10.1109/LED.2015.2466096.

[25] K. S. Im *et al.*, “High-performance GaN-based nanochannel FinFETs With/Without AlGaN/GaN heterostructure,” *IEEE Trans. Electron Devices*, vol. 60, no. 10, pp. 3012–3018, 2013, doi: 10.1109/TED.2013.2274660.

[26] J. W. Yu *et al.*, “Short channel effects on gallium nitride/gallium oxide nanowire transistors,” *Appl. Phys. Lett.*, vol. 101, no. 18, 2012, doi: 10.1063/1.4764554.

[27] D. H. Son *et al.*, “Effects of sidewall MOS channel on performance of AlGaN/GaN FinFET,” *Microelectron. Eng.*, vol. 147, pp. 155–158, Nov. 2015, doi: 10.1016/j.mee.2015.04.101.

[28] K. Im *et al.*, “Heterojunction-Free GaN Nanochannel,” vol. 34, no. 3, pp. 381–383, 2013.

[29] J. H. Lee *et al.*, “Characteristics of GaN and AlGaN/GaN FinFETs,” *Solid. State. Electron.*, vol. 97, pp. 66–75, 2014, doi: 10.1016/j.sse.2014.04.033.

[30] M. Azize and T. Palacios, “Top-down fabrication of AlGaN/GaN nanoribbons,” *Appl. Phys. Lett.*, vol. 98, no. 4, 2011, doi: 10.1063/1.3544048.

[31] Y. Zhang *et al.*, “Large-Area 1.2-kV GaN vertical power FinFETs with a record switching figure of merit,” *IEEE Electron Device Lett.*, vol. 40, no. 1, pp. 75–78, 2019, doi: 10.1109/LED.2018.2880306.

[32] S. Chowdhury, B. L. Swenson, M. H. Wong, and U. K. Mishra, “Current status and scope of gallium nitride-based vertical transistors for high-power electronics application,” *Semicond. Sci. Technol.*, vol. 28, no. 7, 2013, doi: 10.1088/0268-1242/28/7/074014.

[33] M. Xiao, T. Palacios, and Y. Zhang, “ON-Resistance in Vertical Power FinFETs,” *IEEE Trans. Electron Devices*, vol. 66, no. 9, pp. 3903–3909, 2019, doi: 10.1109/TED.2019.2928825.

[34] J. H. Lee *et al.*, “Temperature-dependent characteristics of AlGaN/GaN FinFETs with sidewall MOS channel,” *Solid. State. Electron.*, vol. 120, pp. 47–51, Jun. 2016, doi: 10.1016/j.sse.2016.03.007.

[35] R. Hathwar, “Full Band Monte Carlo Simulation of Nanowires and Nanowire Field Effect Transistors.” doi: 10.1017/CBO9781107415324.004.

[36] A. S. Kumar, N. S. Garigapati, and D. Saha, “Low field mobility in electrostatically evolved AlGaN/GaN one-dimensional channel from a two-dimensional electron gas system,” *Appl. Phys. Lett.*, vol. 115, no. 3, 2019, doi: 10.1063/1.5098864.

[37] S. Jin, M. V. Fischetti, and T. W. Tang, “Modeling of electron mobility in gated silicon nanowires at room temperature: Surface roughness scattering, dielectric screening, and band nonparabolicity,” *J. Appl. Phys.*, vol. 102, no. 8, 2007, doi: 10.1063/1.2802586.

[38] A. H. Davoody, E. B. Ramayya, L. N. Maurer, and I. Knezevic, “Ultrathin GaN Nanowires: Electronic, Thermal, and Thermoelectric Properties,” 2014.

[39] V. N. Kumar and D. Vasileska, “Phonon-limited mobility modeling of gallium nitride nanowires,” *J. Appl. Phys.*, vol. 125, no. 11, p. 114301, Mar. 2019, doi: 10.1063/1.5072759.

[40] M. Mohamed, A. Godoy, and U. Ravaioli, “3D Monte Carlo simulation of current trends and performance in scaled trigate MOSFET,” *J. Comput. Electron.*, vol. 7, no. 3, pp. 217–221, 2008, doi: 10.1007/s10825-008-0187-5.

[41] E. B. Ramayya, D. Vasileska, S. M. Goodnick, and I. Knezevic, “Electron transport in silicon nanowires: The role of acoustic phonon confinement and surface roughness scattering,” *J. Appl. Phys.*, vol. 104, no. 6, 2008, doi: 10.1063/1.2977758.

[42] E. B. R. I. Knezevic, “Self-consistent Poisson-Schrödinger-Monte Carlo solver : electron mobility in silicon nanowires,” *J. Comput. Electron.*, pp. 206–210, 2010, doi: 10.1007/s10825-010-0341-8.

[43] T. Sadi *et al.*, “One-dimensional multi-subband Monte Carlo simulation of charge transport in Si nanowire transistors,” *Int. Conf. Simul. Semicond. Process. Devices, SISPAD*, pp. 23–26, 2016, doi: 10.1109/SISPAD.2016.7605139.

[44] R. Kubo, “Statistical mechanical theory of irreversible processes. 1. General theory and simple applications in magnetic and conduction problems,” *J. Phys. Soc. Jap.*, vol. 12, pp. 570–586, 1957, doi: 10.1143/JPSJ.12.570.

[45] D. A. Greenwood, “The Boltzmann equation in the theory of electrical conduction in metals,” *Proc. Phys. Soc.*, vol. 71, no. 4, pp. 585–596, 1958, doi: 10.1088/0370-1328/71/4/306.

[46] R. Kotlyar, B. Obradovic, P. Matagne, M. Stettler, and M. D. Giles, “Assessment of room-temperature phonon-limited mobility in gated silicon nanowires,” *Appl. Phys. Lett.*, vol. 84, no. 25, pp. 5270–5272, 2004, doi: 10.1063/1.1762695.

[47] S. Jin, M. V. Fischetti, and T. W. Tang, “Modeling of electron mobility in gated silicon nanowires at room temperature: Surface roughness scattering, dielectric screening, and band nonparabolicity,” *J. Appl. Phys.*, 2007, doi: 10.1063/1.2802586.

[48] D. Vasileska, “Computational Electronics Relaxation Time Approximation,” *Nanohub.Org*, [Online]. Available: http://nanohub.org/resources/1518/download/relaxationtimeapproximation_word.pdf

[49] I. Welland and D. K. Ferry, “Electron transport in the solar-relevant InAlAs,” *Semicond. Sci. Technol.*, vol. 34, no. 6, 2019, doi: 10.1088/1361-6641/ab16fd.

[50] A. H. Marshak, “On the Inappropriate Use of the Intrinsic Level as a Measure of the Electrostatic Potential in Semiconductor Devices,” *IEEE Electron Device Lett.*, vol. 6, no. 3, pp. 128–129, 1985, doi: 10.1109/EDL.1985.26069.

[51] A. H. Marshak, “Modeling Semiconductor Devices with Position- Dependent Material Parameters,” vol. 36, no. 9, pp. 1764–1772, 1989.

[52] M. Lundstrom, “Heterostructure Fundamentals,” vol. 1995, pp. 1–43, 2006.

[53] M. S. Lundstrom and R. J. Schuelke, “Modeling semiconductor heterojunctions in equilibrium,” *Solid State Electron.*, vol. 25, no. 8, pp. 683–691, 1982, doi: 10.1016/0038-1101(82)90195-2.

[54] J. E. Sutherland and J. R. Hauser, “A Computer Analysis of Heterojunction and Graded Composition Solar Cells,” *IEEE Trans. Electron Devices*, vol. ED-24, no. 4, pp. 363–372, 1977, doi: 10.1109/T-ED.1977.18742.

[55] Pranay Kumar Reddy Baikad, “Efficient Schr¨odinger-Poisson Solvers for Quasi 1D Systems That Utilize PETSc and SLEPc,” *MS thesis*, no. December, 2020.

[56] B. Padmanabhan, “Modeling Reliability of Gallium Nitride High Electron Mobility Transistors,” Arizona State University, 2013.

[57] “Atlas User’s Manual,” in *Atlas User’s Manual*, 2016, pp. 1598-1600,1606-1609. [Online]. Available: www.silvaco.com

[58] W. J. Tseng *et al.*, “Strain relaxation in GaN nanopillars,” *Appl. Phys. Lett.*, vol. 101, no. 25, 2012, doi: 10.1063/1.4772481.

[59] J. R. Shewchuk, “An Introduction to the Conjugate Gradient Method Without the Agonizing Pain,” *Sch. Comput. Sci. Carnegie Mellon Univ.*, vol. 1.5, pp. 6–30, 1994.

[60] N. Black, S. Moore, and Eric W. Weisstein., “Biconjugate Gradient Method,” *MathWorld--A Wolfram Web Resource,*.

[61] N. Black, S. Moore, and E. W. Weisstein, “Biconjugate Gradient Stabilized Method,” *MathWorld--A Wolfram Web Resource*.

[62] M. Povolotskyi, “Theoretical Study of Electronic and Optical Properties of Low-Dimensional Semiconductor Nanostructures,” 2004.

[63] D. Vasileska, “Piezoelectric Scattering Lecture Notes.” https://nanohub.org/resources/11524

[64] D. Vasileska, “Scattering Rates for Q2D systems.” https://nanohub.org/resources/11532

[65] S. Vitanov, M. Nedjalkov, and V. Palankovski, “A Monte Carlo model of piezoelectric scattering in GaN,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4310 LNCS, pp. 197–204, 2007, doi: 10.1007/978-3-540-70942-8_23.

[66] D. Vasileska and S. M. Goodnick, *Computational Electronics : Semiclassical and quantum device modeling and simulation*, vol. 1, no. 1. 2006. doi: 10.2200/S00026ED1V01Y200605CEM006.

[67] M. V Fischetti, “Part IV : Scattering and Interactions,” pp. 203–271, 2010, [Online]. Available: http://www.ecs.umass.edu/ece/ece618/

[68] D. Vasileska, “Polar Optical Phonon Scattering.”

https://nanohub.org/resources/11522

[69] S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, "Surface roughness at the Si(100)-SiO2 interface," *Phys. Rev. B*, vol. 32, no. 12, pp. 8171–8186, 1985, doi: 10.1103/PhysRevB.32.8171.

[70] A. K. Kandalam, R. Pandey, M. A. Blanco, A. Costales, J. M. Recio, and J. M. Newsam, "First Principles Study of Polyatomic Clusters of AlN, GaN, and InN. 1. Structure, Stability, Vibrations, and Ionization," *J. Phys. Chem. B*, vol. 104, no. 18, pp. 4361–4367, May 2000, doi: 10.1021/JP994308S.

[71] Q. Li, Q. Chen, and J. Chong, "Conduction band fluctuation scattering due to alloy clustering in barrier layers in InAlN/ GaN heterostructures," *AIP Adv.*, vol. 7, p. 125103, 2017, doi: 10.1063/1.5003195.

[72] E. Ahmadi, H. Chalabi, and S. W. Kaun, "Contribution of alloy clustering to limiting the two-dimensional electron gas mobility in AlGaN/GaN and InAlN/GaN heterostructures: Theory and experiment ARTICLES YOU MAY BE INTERESTED IN," *J. Appl. Phys*, vol. 116, p. 133702, 2014, doi: 10.1063/1.4896967.

[73] G. Liu, J. Zhang, K. Lu, W. Chen, Y. Tian, and J. Yang, "Two-dimensional electron gas (2DEG) mobility affected by the in mole fraction fluctuation in InxAl1−xN/GaN heterostructures," *Phys. E Low-dimensional Syst. Nanostructures*, vol. 83, pp. 207–210, Sep. 2016, doi: 10.1016/J.PHYSE.2016.05.016.

[74] R. Hathwar and D. Vasileska, "Generalized Monte Carlo Tool for Investigating Low-Field and High Field Properties of Materials Using Non-parabolic Band Structure Model," no. June, 2011.

[75] D. Vasileska, D. K. Schroder, and D. K. Ferry, "Scaled silicon MOSFET's: Degradation of the total gate capacitance," *IEEE Trans. Electron Devices*, vol. 44, no. 4, pp. 584–587, 1997, doi: 10.1109/16.563362.

[76] S. Elhamri *et al.*, "An electrical characterization of a two-dimensional electron gas in GaN/AIGaN on silicon substrates," *J. Appl. Phys.*, vol. 95, no. 12, pp. 7982–7989, 2004, doi: 10.1063/1.1736327.

[77] G. Kannan and D. Vasileska, "The impact of surface-roughness scattering on the low-field electron mobility in nano-scale Si MOSFETs," *J. Appl. Phys.*, vol. 122, no. 11, Sep. 2017, doi: 10.1063/1.5003253.

[78] S. Sankaranarayanan *et al.*, "Determination of strain relaxation in InGaN/GaN nanowalls from quantum confinement and exciton binding energy dependent photoluminescence peak," *Sci. Rep.*, vol. 8, no. 1, pp. 1–8, 2018, doi: 10.1038/s41598-018-26725-6.