Quantum Capacitance in Scaled Down III-V FETs

by

Donghyun Jin

B.S. Electrical Engineering, Seoul National University, 2008

Submitted to the Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Electrical Engineering
at the
Massachusetts Institute of Technology
February 2010

© 2010 Massachusetts Institute of Technology
All Rights Reserved.

ARCHIVES
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
FEB 23 2010
LIBRARIES

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis and to grant others the right to so.

Author ______________________________ ______________________________

Department of Electrical Engineering and Computer Science

January 29, 2010

Certified by ____

Jesús A. del Alamo

Professor of Electrical Engineering

Thesis Supervisor

Accepted by ____________________________

Terry P. Orlando

Chairman, Department Committee on Graduate Students

Quantum Capacitance in Scaled Down III-V FETs

by

Donghyun Jin

Submitted to the Department of Electrical Engineering and Computer Science
January 29, 2010
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Electrical Engineering

ABSTRACT

As Si CMOS approaches the end of the roadmap, finding a new transistor technology that allows the extension of Moore's law has become a technical problem of great significance. Among the various candidates, III-V-based MOSFETs represent a very promising technology. In particular, low-effective mass materials with high electron velocities, such as InGaAs and InAs are of great interest.

A concern with this approach is the relatively small inversion-layer capacitance that is associated with a low-effective mass channel and the limits that this imposes on the gate capacitance that can be attained from barrier thickness scaling. This can seriously limit the current driving ability of scaled down devices.

In order to understand the scaling potential of III-V MOSFETs, we have built a physical gate capacitance model for III-V FETs that incorporates quantum capacitance and centroid capacitance in the channel. We verified its validity with simulations (Nextnano) and experimental measurements on High Electron Mobility Transistors (HEMTs) with InAs and InGaAs channels down to 30 nm in gate length. Our model confirms that in the operational range of these devices, the quantum capacitance significantly lowers the overall gate capacitance. In addition, our experiments suggest a large increase of the in-plane effective mass in very thin channel designs as a result of non-parabolicity, quantum confinement and biaxial compressive strain. This should help to achieve a relatively high electron concentration in future 10 nm high-k dielectric III-V MOSFETs. Our study provides a number of suggestions for capacitance scaling in future III-V MOSFETs.

Thesis supervisor: Jesús A. del Alamo

Title: Professor of Electrical Engineering

Acknowledgements

First of all, I would like to thank my research advisor Jesús A. del Alamo for giving me the great opportunity to work on this interesting project. He continuously provided me with invaluable knowledge and great intuition throughout this research. I am very grateful for his excellent supervision and persistent help. He has taught me to aim for a high standard of research and I have greatly benefited from his guidance and technical discussions. I sincerely appreciate his dedication and encouragement to my growth as a professional researcher.

I gratefully acknowledge Dr. Dae-Hyun Kim and Dr. Tae-Woo Kim (the Kim brothers) for their crucial contributions. Dae-Hyun Kim kindly gave me lots of advice when they were most required. As my second research advisor, I am much indebted to Tae-Woo Kim for his patient assistance with different kinds of fabrication processing problems in the TRL. He always helped me in obtaining valuable knowledge from the experimental failures. Their endless support and immense knowledge on InGaAs HEMTs were essential to accomplish my research.

I also would like to show my gratitude toward all the members of del Alamo's Xtreme semiconductor group. I gratefully thank Jungwoo Joh for his moral support. Whenever I met an obstacle in my MIT life, he gave me advice to build the right attitude for solving it and relieving stress. I am also thankful to Usha Gogineni for the enjoyable daily chats and the great trip to Walden Pond. I should not forget to thank Ling Xia for a fruitful discussion on this research. I owe a debt of gratitude to Sefa Demirtas for being the greatest officemate. We shared the most of joy, concern and happiness together. Our ping-pong match was one of the most competitive moments in my MIT life. I also thank Jianqiang Lin and hope to share many precious times together in the future.

I would like to thank all my 6th floor friends: Jinwook, Han, Bin, Omair, Dongseup and Kyungbum. Their strong enthusiasm towards research boosts me up everyday. I also would like to acknowledge my best friends: Jae-Byum, Joong Bum and Keun Sup. You guys are like my family. Our deep friendship has always energized me to move forward.

Finally, I appreciate my parents for their endless love and support throughout my entire life. I also thank my elder brother, Hyung Suk, for his moral care and persistent confidence in me. *This work was sponsored by Intel Corporation, FCRP-MSD Center and Samsung Scholarship.*

Contents

List of Figures 10

List of Tables 14

Chapter 1. Introduction 15

1.1. Introduction to III-V CMOS 15

1.2. Motivation - Gate capacitance in III-V MOSFET 18

1.3. Gate capacitance in III-V HEMTs 19

1.4. Thesis Outline 21

Chapter 2. Theoretical Gate Capacitance Model 23

2.1. Introduction 23

2.2. Gate capacitance model for III-V FETs 23

2.2.1. Inversion-layer capacitance model 23

2.2.2. Quantum capacitance model 26

2.2.3. Centroid capacitance model 29

2.3. Verification of physical model 29

2.3.1. 1-D Poisson-Schrodinger Solver (Nextnano) 29

2.3.2. Model and numerical simulation 31

2.3.2.a. Sheet carrier concentration (N_S) vs gate voltage (V_G) 32

2.3.2.b. Gate capacitance (C_G) vs gate voltage (V_G) 33

2.3.2.c. Low-pass filter for centroid capacitance (C_{cent}) 35

2.3.2.d. Inversion-layer capacitance (C_{inv}) vs sheet carrier concentration (N_S) 36

2.4. Summary 39

Chapter 3. Experimental Gate Capacitance Measurements in HEMTs 41

3.1. Introduction 41

3.2. Device technology and heterostructures 42

3.3. Gate capacitance extraction method 44

3.3.1. S-parameter analysis 44

3.3.2. Intrinsic and parasitic components of gate capacitance 46

3.4. S-parameter measurements results – Type B 47

3.4.1. Capacitance vs. Frequency 48

3.4.2. Capacitance vs. V_{GS} 49

3.5. Extraction of intrinsic gate capacitance 50

3.6. Summary 51

Chapter 4. Comparison of Measurements with Physical Model 53

4.1. Introduction 53

4.2. Schottky barrier height and effective mass in our model 53

4.3. C_G in Type A (InAs channel, t_{ch}= 10 nm, t_{ins}= 10 nm) 54

4.4. C_G in Type B (InAs channel, t_{ch}= 10 nm, t_{ins}= 4 nm) 55

4.5. C_G in Type C ($In_{0.7}Ga_{0.3}As$ channel, t_{ch}= 13 nm, t_{ins}= 4 nm) 58

4.6. Key findings from the comparison of Type A, B and C 58

4.7. Discrepancy in Type B 60

4.8. Summary 61

Chapter 5. Discussion and Projections 63

5.1. Introduction 63

5.2. Possible sources of discrepancy in type B structure 63

5.2.1. Experimental uncertainty 64
5.2.2. Channel quantization 64
5.2.3. Nonparabolicity of the conduction band 66
5.2.4. Biaxial channel strain 68
5.2.5. Role of channel cladding 69
5.3. Modified model C_G results 69
5.3.1. Channel cladding layer contribution to C_G 69
5.3.2. Reduced discrepancy in Type B 71
5.4. Projection for 10 nm III-V MOSFET 73
5.5. Summary 76

Chapter 6. Conclusions and Suggestions 77
6.1. Conclusions 77
6.2. Suggestions for future work 79

Bibliography 81

List of Figures

Figure 1-1. Electron mobility versus sheet electron density in different n-channel material FETs. [Fig. taken from R. Chau [2]]. 16

Figure 1-2. Sketch of a future III-V MOSFET structure. It shows key features of high-k gate dielectric layer and III-V quantum well channel layer on a Si substrate. 17

Figure 1-3. (a) Equivalent gate capacitance circuit diagram of III-V MOS gate structure and (b) sketch of conduction band diagram in strong inversion. 18

Figure 1-4. A sketch of an InGaAs High Electron Mobility Transistor. 20

Figure 2-1. Equivalent circuit diagram of gate capacitance in a III-V FET 24

Figure 2-2. Conduction band diagram of a III-V FET in strong inversion. Physical parameters (E_C, E_i, E_F, ψ_S, Q_S and $m_{\|}^*$) defined in eq. (1), (2) and (3) are shown 25

Figure 2-3. A sketch of the density of states for electrons in a 2D quantum well. 28

Figure 2-4. C_Q model characteristics with respect to E_F. C_{Q1} is expressed in a green line, and C_{Q2} in a pink line 28

Figure 2-5. Conduction band (E_C) and electron density ($n_S(z)$) profile of 1D HEMT structure. The Schottky barrier height is 0.63 eV and the applied voltage at the gate metal layer is 0.3 V in this simulation. 31

Figure 2-6. Sheet carrier concentration (N_S) as a function of gate voltage (V_G) for three HEMT structures in a linear scale. Solid lines are Nextnano simulation results and symbols are model results. 32

Figure 2-7. Sheet carrier concentration (N_S) as a function of gate voltage (V_G) for three HEMT structures in logarithm scale. Solid lines are numerical simulation results and symbols are model results. Subthreshold slopes (SS) are close to the ideal one (60 mV/dec) for three HEMTs. 33

Figure 2-8. Gate capacitance (C_G) as a function of gate voltage (V_G) for three HEMT structures. Solid lines are numerical simulation results and symbols are model results. 34

Figure 2-9. Gate capacitance (C_G) and capacitance component model results (insulator capacitance (C_{ins}), 1^{st} subband centroid capacitance (C_{cent1}), 1^{st} subband quantum capacitance (C_{Q1}), 1^{st} subband inversion-layer capacitance (C_{inv1}), and 2^{nd} subband inversion-layer capacitance (C_{inv2}). 35

Figure 2-10. 1st subband centroid capacitance from original calculation results (blue line) and the ones after low pass filter (red line). 36

Figure 2-11. Inversion-layer capacitance (C_{inv}) versus sheet carrier concentration (N_S) of all of three HEMTs in a logarithm scale. Solid lines are simulation results and symbols are our model results. All of slopes of the lines are equal to unity. 38

Figure 3-1. Experimental HEMT cross section and layer details of the three heterostructures explored in this work. 43

Figure 3-2. A TEM image of 30 nm gate length Type B device. [Figs. taken from Kim, [17]] .. 43

Figure 3-3. Small signal equivalent circuit model for a FET. This model is applied to our HEMT structure to extract C_{gs} and C_{gd} (R_i is assumed to be zero). [Fig. taken from Dambrine, [23]]. ... 45

Figure 3-4. Intrinsic and parasitic components of gate capacitance in a HEMT. Two different kinds of parasitic gate capacitance are defined in this model. C_{gext_outer} is associated with the top of the T-gate. C_{gext_inner} is associated with the sidewall of the gate stem. 47

Figure 3-5. Extracted small signal model components (C_{gs}, C_{gd} and C_g) results with respect to frequency in Type B device with L_G = 200 nm for different values of V_{GS}. 48

Figure 3-6. C_G (= C_{gs} + C_{gd}) vs. V_{GS} for different L_G in Type B devices. C_G (V_{GS} = -0.3 V) corresponds to C_{gext_outer} in Figure 3-4. 49

Figure 3-7. C_G –C_G (V_{GS} = -0.3 V) vs. V_{GS} for different L_G in Type B devices. 49

Figure 3-8. $C_G - C_G$ (V_{GS} = -0.3V) as a function of L_G at different V_{GS} for Type B devices. The slope of this line at each bias point corresponds to C_{gi}. The Y intercept is $2C_{gext\_inner}$ of Figure 3-4. 50

Figure 3-9. Intrinsic gate capacitance (C_{gi}) for Type A, B, and C heterostructures. 51

Figure 4-1. Experimental C_G and modeled C_G components of Type A (InAs channel, t_{ch} = 13 nm, t_{ins} = 4 nm) vs V_G. 56

Figure 4-2. Evolution of the 1st and 2nd subband energy levels with respect to the Fermi level in the Type A structure.. 56

Figure 4-3. Experimental C_G and modeled C_G components of Type B (InAs channel, t_{ch} = 10 nm, t_{ins} = 4 nm) vs V_G. 57

Figure 4-4. The evolution of the 1st and 2nd subband energy levels with respect to the Fermi level for the Type B structure. The 2nd subband contribution is negligible.. 57

Figure 4-5. Experimental C_G and modeled C_G components of Type C ($In_{0.7}Ga_{0.3}As$ channel, t_{ch} = 13 nm, t_{ins} = 4 nm) vs V_G. 59

Figure 4-6. The evolution of the 1st and 2nd subband energy levels with respect to the Fermi level for the Type C structure. The 2nd subband is populated in a minor way. 59

Figure 5-1. The in-plane effective mass ($m_{||}$*) as a function of well width (a) and electron quantization energy (b) in $In_{0.53}Ga_{0.47}As$/InP quantum well structures. Symbols represents reported experimental results and lines corresponds to calculated results from the model suggested by Wetzel [28]. [Figs. taken from Wetzel [28]].. 66

Figure 5-2. A sketch of nonparabolicity in E-k diagram inside a AlxGa1-xAs/GaAs quantum well. A indicates quantization energy shifts. B points out in-plane effective mass ($m_{||}$*) at the bottom of a subband energy level which is calculated from the deviated curvature (dashed line) compared with parabolic curvature (dotted line) at C. [Fig. taken from Ekenberg, [31]]............ 67

Figure 5-3. $In_xGa_{1-x}As$ in-plane (noted as par.) and perpendicular (noted as perp.) effective mass as a function of in-plane (biaxial) strain for different In mole fractions (x = 0, 0.47 and 1). [Fig. taken from Kopf [27]]. 68

Figure 5-4. Conduction band (E_C) and electron density ($n_S(z)$) profile of Type B at $V_G - V_T = 0.22$ V. Green shadow area under the n_S curve indicates electrons in the cladding layers. 70

Figure 5-5. Electron concentration proportion [%] in each channel layer of Type B for InAs layer $m_{||}^* = 0.026\ m_o, 0.050\ m_o$. 71

Figure 5-6. Gate capacitance as a function of gate voltage for the InAs-channel heterostructures (Type A and B). The dotted lines show calculations using heavier values of effective mass in the InAs layer. The error bars indicate the uncertainty in the calculated C_G that comes from an uncertainty of ±0.5 nm in t_{ins} 72

Figure 5-7. High-k dielectric 10 nm gate length III-V MOSFET prototype with t_{ch}= 3 nm and t_{ins} = 2.6 nm ($\varepsilon = 25\varepsilon_o$) 73

Figure 5-8. Modeled C_G and capacitance components in 10 nm III-V MOSFET prototype. 0.06 m_o is used for $m_{||}^*$ in the channel layer. 74

Figure 5-9. Gate capacitance of 10 nm III-V MOSFET prototype for different $m_{||}^*$ values in the channel layer. 75

Figure 5-10. Sheet carrier concentration (N_S) of 10 nm III-V MOSFET prototype for different $m_{||}^*$ values in the channel layer. V_T is defined as V_G at $N_S = 5 \times 10^{10}\ cm^{-2}$. 76

List of Tables

Table 3-1. Key parameters of the three different heterostructures studied in this work.............. 42

Chapter 1. Introduction

1.1. Introduction to III-V CMOS

CMOS has been the mainstream logic technology for several decades using silicon as main channel material and silicon dioxide as gate dielectric material. Scaling Si CMOS following Moore's Law has continuously improved performance, density, and power consumption. However, as Si CMOS scaling approaches fundamental physical limits especially beyond the 22 nm technology node, finding a new material technology that allows the extension of Moore's law has become a significant technical problem [1].

Among the various candidates, III-V based MOSFETs represent a very promising technology for future high-speed and low-power digital logic applications [2]. This bright hope comes from the much better electron transport properties of III-V materials with respect to Si which results from higher electron velocity. This directly contributes to high on-state current and transconductance [3]. Therefore, to obtain higher electron velocity in future scaled down devices, III-V FETs have been getting more emphasis nowadays. Mobility is a good exponent of the excellent transport characteristics of III-Vs since electron velocity is mostly correlated to the mobility property. Figure 1-1 shows different electron mobility in various III-V materials FETs as a function of electron density. It is apparent that most of III-V FETs have about a hundred times larger electron mobility than conventional Si MOSFETs.

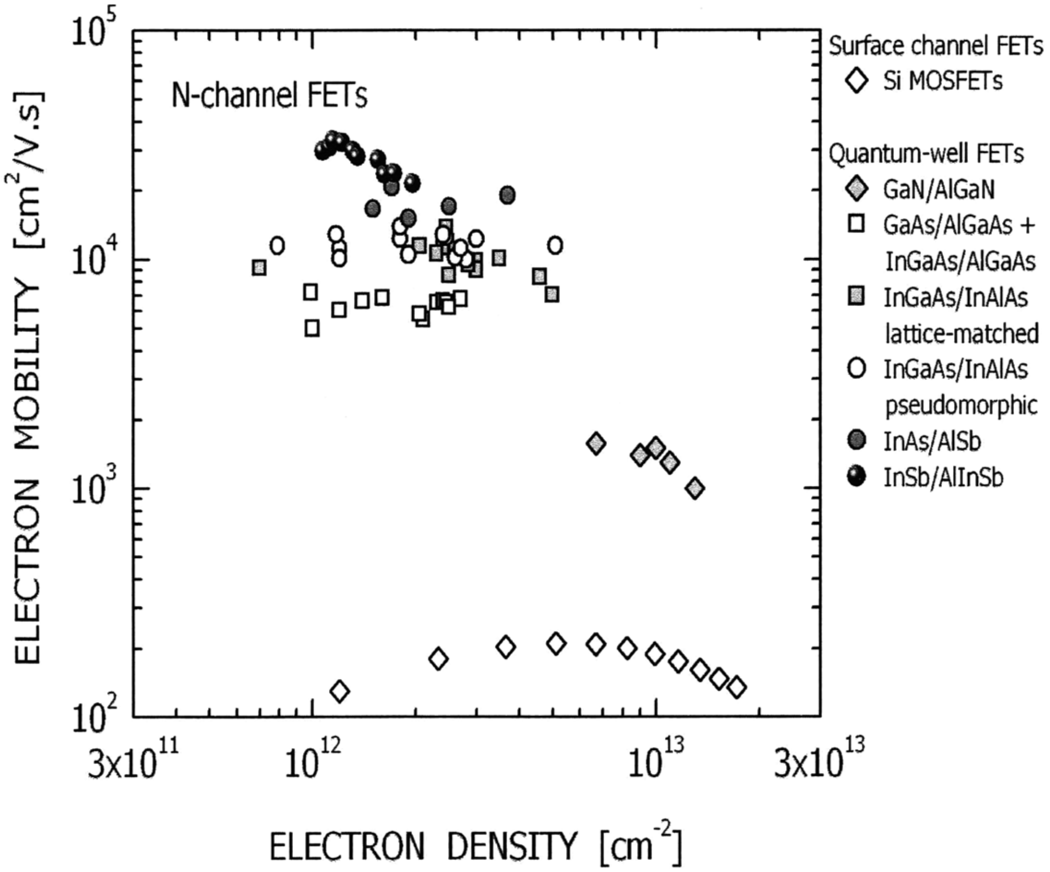

Figure 1-1. Electron mobility versus sheet electron density in different n-channel material FETs. [Fig. taken from R. Chau [2]].

Even though III-V's have this attractive advantage, there are still lots of technical difficulties to overcome before III-V MOSFETs can replace scaled Si MOSFETs as a mainstream logic technology. One big challenge is to identify a suitable high-k gate dielectric material that makes low interface-state density (D_{it}) near the conduction band edge [4] at the interface in semiconductor channel layer. It has another critical meaning because it can guarantee better gate control with reduced vertical gate leakage which is essential for enhancement mode devices [5]. Integration of the III-V materials onto the Si substrate should be also achieved [6]. Since III-V substrates are difficult to grow in large area and are easily broken, integration technology into the Si substrate are necessary for being processed in current matured Si manufacturing infrastructure.

In the long run, regardless of these problems or even advantages of the III-V substrate, we do really want to use Si as substrate so that a future III-V CMOS technology is very much silicon compatible.

On the assumption that these technical issues are solved, we can predict a schematic for a future III-V MOSFET device structure as shown in Figure 1-2. It features a III-V quantum well (QW) channel layer and a thin high-k gate dielectric layer directly on top of the QW channel. This QW channel extends below the source and drain regions. Source and drain regions are self-aligned and highly doped n+ so that an ohmic contact with very low contact resistance is realized. A δ doping layer is located under the QW channel and inside the wide band gap buffer layer. It is designed to provide sufficient carriers into the QW channel region. Buffer layers are grown on the Si substrate. Small gap between gate and source/drain is spaced for reducing parasitic capacitance. The undoped III-V QW channel layer enables carriers to have a higher velocity due to the reduced scattering and high mobility of the III-V material. High-k dielectric layer can be relatively thicker because it provides same gate capacitance as a thinner low-k dielectric layer like SiO_2 [7]. This advantage helps scaling to be continued and the extension of Moore's law.

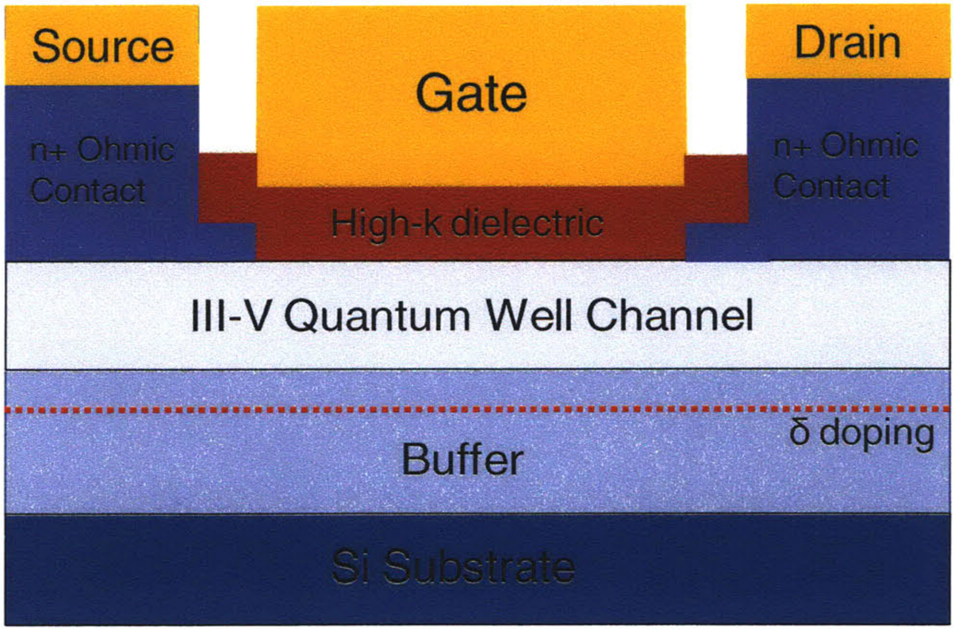

Figure 1-2. Sketch of a future III-V MOSFET structure. It shows key features of high-k gate dielectric layer and III-V quantum well channel layer on a Si substrate.

However, there is a key concern about this approach. High carrier velocity comes from the relatively low effective mass of the channel material but this also makes for a small density of states (DOS) in the channel layer [8]. This small DOS will directly result in a low sheet carrier concentration at a certain gate overdrive. This will require a larger gate voltage to generate enough sheet charge and drain current for appropriate device operation. It is a matter of some controversy whether, as a result of this issue, III-V MOSFETs at the 10 nm gate node will be able to drive enough current at low supply voltage. This thesis is about carrying out experimental and theoretical research involving state-of-the-art high-performance III-V FETs to sort out this issue.

1.2. Motivation - Gate capacitance in III-V MOSFET

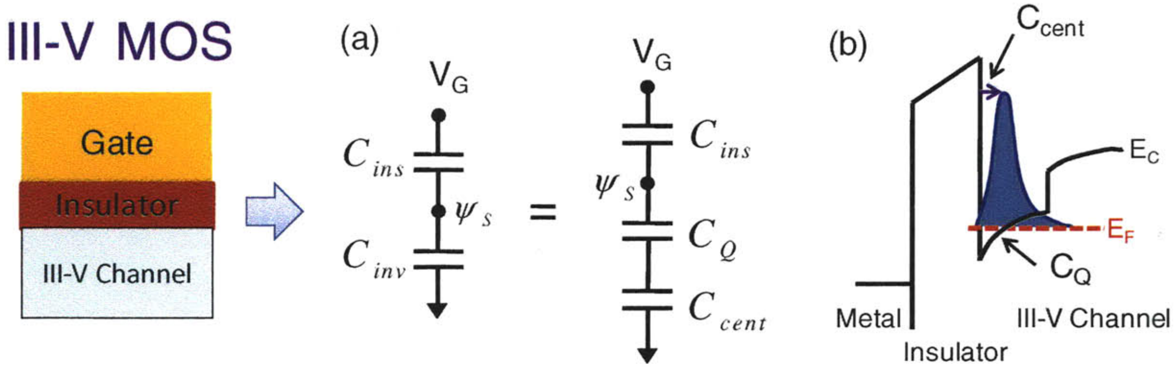

Figure 1-3. (a) Equivalent gate capacitance circuit diagram of III-V MOS gate structure and (b) sketch of conduction band diagram in strong inversion.

The impact of the low DOS on the sheet carrier concentration that can be attained in III-V MOS structures can be easily comprehended by a gate capacitance analysis. Figure 1-3 shows a sketch of III-V MOS gate structure and an equivalent gate capacitance circuit diagram. A sketch of the conduction band diagram for a bias point in strong inversion is also shown. The total gate capacitance of the III-V MOS can be expressed as a series of the insulator capacitance (C_{ins}) and the inversion-layer capacitance (C_{inv}) [9]. On the assumption that only the first electron subband in the channel is occupied, the inversion-layer capacitance can be represented as a series of the quantum capacitance (C_Q) and the centroid capacitance (C_{cent}) [10, 11]. This circuit diagram suggests that the total gate capacitance (C_G) is determined by the smallest gate capacitance component among the three of C_{ins}, C_Q, and C_{cent}.

Conceptually, the quantum capacitance physically originates in the Fermi-level (E_F) penetration into the conduction band as shown in Figure 1-3 (b). C_Q is proportional to the DOS of the channel material [12]. The centroid capacitance is related to the finite average distance of the electron channel from the insulator/channel interface [13]. The insulator capacitance is inversely proportional to the insulator thickness.

Ideally, the inversion-layer capacitance is much larger than the insulator capacitance in the strong inversion condition and the total gate capacitance approaches the insulator capacitance. In this situation, insulator thickness scaling increases the gate capacitance as needed for harmonious MOSFET scaling [9]. However, this ideal gate capacitance scaling does not hold for deeply scaled devices. As insulator thickness approaches the few nanometer range, the insulator capacitance becomes comparable to the inversion-layer capacitance, which means that the quantum capacitance and the centroid capacitance start to impact the gate capacitance [10].

This problem is especially severe in III-V MOSFET as a results of their relatively small effective mass which reduces the quantum capacitance significantly [8]. For example, the density of states electron effective mass of Si is 1.08 m_0 (where m_0 is the electron rest mass). On the other hand, for GaAs, the value is 0.067 m_0 and for InAs, 0.026 m_0. The low DOS that these small effective masses imply can potentially become a big bottleneck to attaining enough sheet carrier concentration in a future 10 nm gate node III-V CMOS technology.

1.3. Gate capacitance in III-V HEMTs

Today, it is not possible to carry out a fundamental study of gate capacitance in III-V MOSFEFs since these devices still suffer from a number of non-idealities [2] and are under development. Instead, we study the issues outlined in this chapter by taking advantage of the great progress that has taken place recently in scaling InGaAs High Electron Mobility Transistors (HEMTs) to very small dimensions. InGaAs HEMTs have been under development for some time for low-noise and high gain ICs for operation in communication and radar systems at more than hundred GHz. Recently, scaled down InGaAs HEMTs have also shown great logic performance. These devices constitute an excellent test bed to explore fundamental issues in a future III-V CMOS technology [14, 15].

A simple sketch of an InGaAs HEMT structure is shown in Figure 1-4. This device features a very thin InAlAs barrier layer ($\varepsilon_r \approx 12$) and an InGaAs quantum well channel layer which can have a relatively high InAs mole fraction. The InAlAs barrier typically includes a δ doping layer that sets the required sheet electron concentration in the channel. The gate has commonly a T shape and the gate length (L_G) is defined by the length of the gate stem. The InAlAs barrier thickness (t_{ins}) is controlled by a three-step recess process [15] and the width of the InGaAs QW channel defines channel thickness (t_{ch}). Our group at MIT has been making these devices with various heterostructures down to 30 nm in L_G [14]. These devices have shown a world record f_T [16] and superior logic application performance in terms of subthreshold slope, DIBL, g_m, etc [14, 15, 17].

In this thesis, we aim to develop a detailed understanding of the gate capacitance components of InAlAs/InGaAs HEMTs by experimentally and theoretically studying the gate capacitance of various device structures with different values of barrier thickness (t_{ins}) and channel thickness (t_{ch}). A detailed understanding of the quantum capacitance and centroid capacitance in HEMTs will allow us to project their role in future 10 nm gate length III-V MOSFETs.

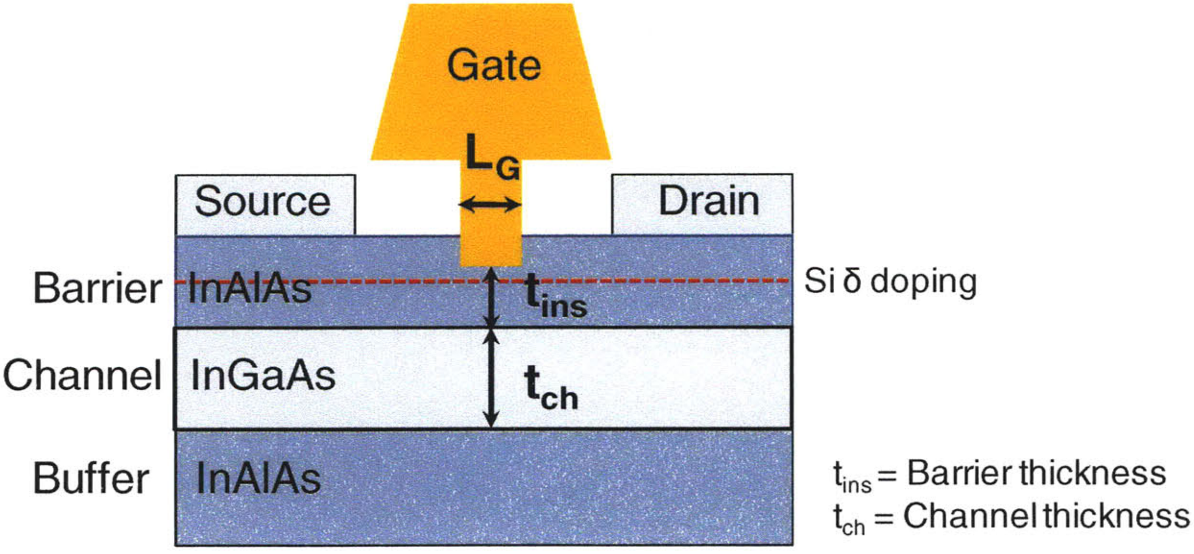

Figure 1-4. A sketch of an InGaAs High Electron Mobility Transistor.

1.4. Thesis Outline

In this research, we have built a physical gate capacitance model for III-V FETs that incorporates quantum capacitance and centroid capacitance in the channel. We verified its validity with simulations (Nextnano) and experimental measurements on HEMTs with InAs and InGaAs channels down to 30 nm in gate length. Our model confirms that in the operational range of these devices, the quantum capacitance significantly lowers the overall gate capacitance. In addition, the channel centroid capacitance is also found to have a significant impact on gate capacitance. From this work, we attempt to provide a number of suggestions for capacitance scaling in future III-V FETs and estimate sheet carrier concentration of future 10 nm node III-V FETs.

This thesis will be organized in the following way. Chapter 2 will start with introduction of a gate capacitance model for III-V FETs. The inversion-layer capacitance will be defined and modeled as a parallel combination of the inversion-layer capacitance associated with each occupied electron subband. These in turn are modeled as a series of quantum capacitance and centroid capacitance. Next, the physical origin of both capacitances will be described. We then verify our model using a 1-D Poisson-Schrodinger Solver (Nextnano) simulation tool. For this, we use three model heterostructures with different barrier and channel layer thicknesses.

In Chapter 3, the experimental InGaAs HEMT devices studied in this work will be described. Three types of device structures are considered. For each type, devices with different gate lengths were fabricated. This chapter describes the gate capacitance extraction method from S-parameter measurements and the determination of the intrinsic gate capacitance.

In Chapter 4, the experimental measurements of intrinsic gate capacitance will be compared with predictions from the physical model described in Ch. 2. From the comparison, a clear understanding emerges of the relative contributions of each gate capacitance component to the overall gate capacitance.

Chapter 5 discusses possible origins for the discrepancy between the measured and model gate capacitance observed in one of our devices. This chapter suggests that the in-plane effective mass in a thin channel layer can be increased by the combination of non-parabolicity, quantization and strain effects. It also shows that this increase of in-plane effective mass can play an important role in increasing the quantum capacitance and boosting the overall gate capacitance and sheet carrier concentration in the channel. We show that this effect will become very significant in future scaled down devices by projecting the gate capacitance and sheet carrier concentration of future 10 nm III-V MOSFETs.

Finally, in Chapter 6, we summarized all of conclusions from this gate capacitance analysis work and provide some suggestions for future studies.

Chapter 2. Theoretical Gate Capacitance Model

2.1. Introduction

In this chapter, we build a gate capacitance model for III-V FETs which includes DOS effects in the quantum capacitance. We verify our model by using a one-dimensional (1D) Poisson-Schrodinger Solver (Nextnano) simulator.

2.2. Gate capacitance model for III-V FETs

2.2.1. Inversion-layer capacitance model

The gate capacitance of a III-V FET in strong inversion can be modeled as the series combination of the insulator capacitance (C_{ins}) and the inversion-layer capacitance (C_{inv}) (Figure 2-1) [9]. This model assumes that there is no doping level underneath the channel such as Fully Depleted Silicon-On-Insulator (FDSOI) [18]. This one consists, in turn, of a parallel combination of the contributions of each occupied electron subband in the channel. In Figure 2-1, C_{inv1} and C_{inv2} indicate the 1st subband and 2nd subband inversion-layer capacitances, respectively.

For each subband *i*, the inversion-layer capacitance (C_{inv_i}) consists of the quantum capacitance (C_{Q_i}) and the centroid capacitance (C_{cent_i}) which are connected in series (Figure 2-1). This can be obtained from the definition of inversion-layer capacitance:

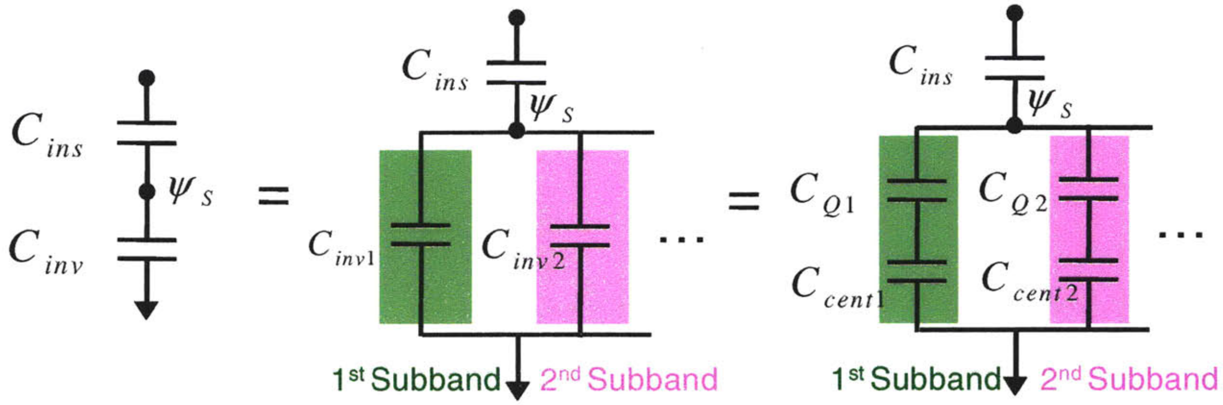
C_{ins}
ψ_S
C_{inv}
=
C_{ins}
ψ_S
C_{inv1}
C_{inv2}
...
=
C_{ins}
ψ_S
C_{Q1}
C_{Q2}
C_{cent1}
C_{cent2}
...
1st Subband
2nd Subband
1st Subband
2nd Subband

Figure 2-1. Equivalent circuit diagram of gate capacitance in a III-V FET

$$C_{inv}=\frac{\partial(-Q_S)}{\partial\psi_S}=\frac{q\partial(-Q_S)}{\partial(E_F-E_C)} \tag{1}$$

where ψ_S is the surface potential, and E_C is the conduction band edge at the barrier-channel interface on the channel side.

Q_S is the total electron charge in the channel which can be expressed as the sum of all the charge in each of the subbands. That can be easily computed through:

$$Q_S=\sum_i Q_i=\sum_i \int_{E_i}^{\infty}\frac{\frac{m_{\parallel}^*q}{\pi\hbar^2}}{1+\exp(\frac{E-E_F}{kT})}dE \tag{2}$$

where Q_i is the electron charge of subband *i* in the channel, E_i is the energy level of subband *i*, and $m_{\parallel}^*$ is the in-plane effective mass of the channel material. In this expression, conduction band degeneracy is set to two to account for spin degeneracy.

For a given subband, using the chain rule, we can in general write:

$$(\frac{q\partial(-Q_i)}{\partial(E_F-E_C)})^{-1}=(\frac{q\partial(-Q_i)}{\partial(E_F-E_i)})^{-1}+(\frac{q\partial(-Q_i)}{\partial(E_i-E_C)})^{-1} \tag{3}$$

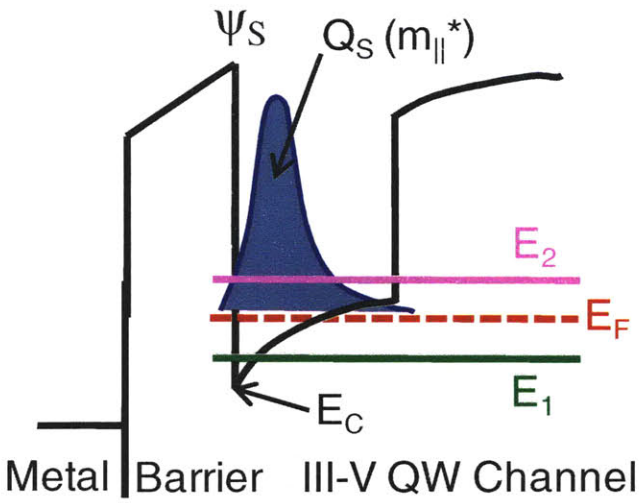

Figure 2-2. Conduction band diagram of a III-V FET in strong inversion. Physical parameters (E_C, E_i, E_F, ψ_S, Q_S and $m_{\parallel}{}^*$) defined in eq. (1), (2) and (3) are shown.

Physical parameters defined in eq. (1), (2), and (3) are described in Figure 2-2.

We can then define C_{Q_i} as:

$$C_{Q\_i} = \frac{q\partial(-Q_i)}{\partial(E_F - E_i)} \tag{4}$$

C_{Q_i} is the quantum capacitance of subband *i* and corresponds to the derivative of electron charge in subband *i* with respect to the energy difference between E_F and E_i.

We also define C_{cent_i} as:

$$C_{cent\_i} = \frac{q\partial(-Q_i)}{\partial(E_i - E_C)} \tag{5}$$

C_{cent_i} is the centroid capacitance of subband *i* and corresponds to the derivative of electron charge in subband *i* with respect to the energy difference between E_i and E_C.

Then, C_{inv} can be expressed as:

$$C_{inv} = \sum_i \left(\frac{1}{C_{Q\_i}} + \frac{1}{C_{cent\_i}} \right)^{-1} \tag{6}$$

We can derive additional formulas for C_{Q_i} and C_{cent_i} :

$$C_{Q\_i} = \frac{q\partial(-Q_i)}{\partial(E_F - E_i)} = \frac{q\partial\left(-\int_{E_i}^{\infty} \frac{\frac{m_{\parallel}^* q}{\pi\hbar^2}}{1+\exp(\frac{E-E_F}{kT})} dE\right)}{\partial(E_F - E_i)} = \frac{\frac{m_{\parallel}^* q^2}{\pi\hbar^2}}{1+\exp(\frac{E_i - E_F}{kT})} \tag{7}$$

$$C_{cent\_i} = \frac{q\partial(-Q_i)}{\partial(E_i - E_C)} = C_{Q\_i} \cdot \frac{\partial(E_F - E_i)}{\partial(E_i - E_C)} \tag{8}$$

The insulator capacitance is simply the permittivity (ε) over the barrier thickness (t_{ins}) :

$$C_{ins} = \frac{\varepsilon}{t_{ins}} \tag{9}$$

If the location of each subband energy level (E_i) and the Fermi level are known with respect to the conduction band edge, then all capacitance components can be evaluated. Rather than attempt to analytically solve the quantization problem for realistic FET structures, in this work, we used a 1D Poisson-Schrodinger solver (Nextnano) to obtain all of subband energy levels as a function of V_G.

2.2.2. Quantum capacitance model

The quantum capacitance concept was firstly introduced by S. Luryi [19]. A two-dimensional electron gas (2DEG) requires energy to be created in a semiconductor quantum well region due to the finite density of states. Since the density of states is finite in a semiconductor quantum well, the Fermi-level needs to move up above the conduction band edge as the charge in the quantum well increases. This movement of Fermi-level requires energy and this conceptually corresponds to quantum capacitance.

For example, when we induce channel charge in an MOS structure, we actually need to deliver an amount of energy equal to $\frac{Q_S^2}{2C_{ins}}+\frac{Q_S^2}{2C_Q}$ to the MOS (the centroid capacitance is not considered here; all charges are assumed to be located at the same position inside semiconductor layer). The first term is related to the required energy for the electric field in the oxide layer and the second term corresponds to the required energy to create 2DEG in the semiconductor layer which occurs due to the finite DOS of the semiconductor. Normally, C_{ins} is much smaller than C_Q, so the second term is considered negligible but as device scaling approaches a few nanometer scale, C_{ins} becomes very comparable to or even bigger than C_Q, and C_Q should be considered carefully in these scaled down devices.

Quantum capacitance is associated to the properties of the channel material itself and not to its geometrical structure. III-V materials have smaller DOS than Si and this implies a smaller quantum capacitance and higher required energy to produce a given amount of inversion charge.

Our analytical quantum capacitance model takes into account the effect from Fermi-level penetration into each subband energy level as well as DOS effect [20]. $\frac{m_{\parallel}^*}{\pi\hbar^2}$ in the numerator term of our quantum capacitance model is precisely the 2-dimensional (2D) DOS for one subband [21]. The important property of the 2D DOS is that it is independent of energy, as shown in Figure 2-3. Only the in-plane effective mass determines the 2D DOS of each subband. In a 3D bulk structure, electrons have three degrees of freedom in their movement and the effective mass can be considered as m_x*, m_y* and m_z* which corresponds to the each direction of movement. However, in 2D quantum well structure, the movement into the quantum well depth direction is confined and only parallel movement in the quantum well plane is allowed. Therefore, we can define two different kinds of effective mass: in-plane effective mass ($m_{\parallel}$*) and perpendicular effective mass ($m\perp$\*). $m_{\parallel}$* corresponds to the electron transport property in the quantum well plane and determines 2D DOS. On the other hand, $m\perp$\* determines the subband energy levels (E_i). If the quantum well plane is in the direction of the x-y plane, $m_{\|}^*$ equals to $\sqrt{m_x^* \cdot m_y^*}$ and m_z^* is equal to $m_{\perp}^*$ [22].

The denominator in eq. (7) accounts for the movement of Fermi level (E_F). In weak inversion condition ($E_F < E_i$), the denominator part becomes nearly infinite and C_{Qi} becomes negligible. However, in strong inversion condition ($E_F > E_i$), the denominator becomes 1 and C_{Qi} becomes equal to $\frac{m_{\|}^* q^2}{\pi \hbar^2}$. As E_F passes over each subband energy level one by one, each subband quantum capacitance jumps by an amount equal to $\frac{m_{\|}^* q^2}{\pi \hbar^2}$, as shown in Figure 2-4. This is why the quantum capacitance is sometimes referred to as the average DOS at the Fermi-level.

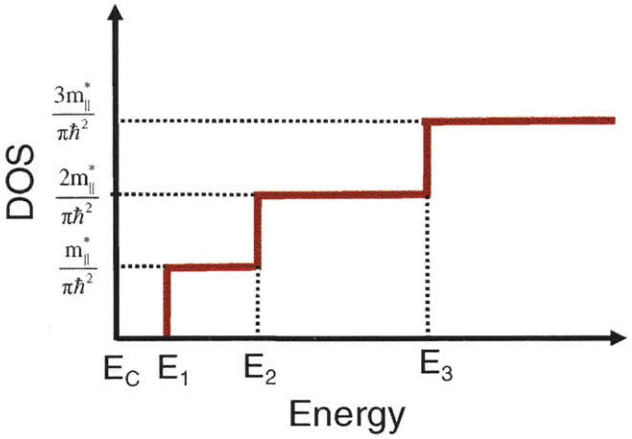

Figure 2-3. A sketch of the density of states for electrons in a 2D quantum well.

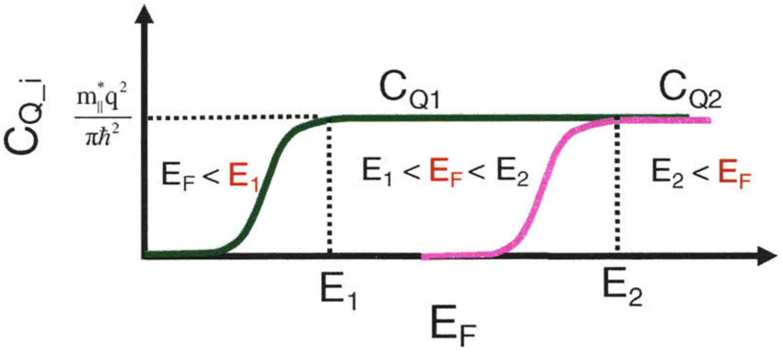

Figure 2-4. C_Q model characteristics with respect to E_F. C_{Q1} is expressed in a green line, and C_{Q2} in a pink line.

2.2.3. Centroid capacitance model

Charges in a quantum well do not distribute themselves in the form of a sheet with zero thickness but they typically make a bell-shape distribution. Therefore the physical distance from the metal gate to each charge is different. Furthermore, due to the confinement in the quantum well, the center of the charge distribution can be far from the barrier interface in the scale of the barrier thickness. These effects are relevant in the modeling of the total inversion-layer capacitance.

The common method to include this charge distribution effect into the gate capacitance model is to determine an average charge distance from the interface between the barrier and the channel. However, this requires that we model the charge distribution correctly. There are several papers that attempt to do this [6-9] but their methods are quite complicated. Instead of using this approach, we directly derived the capacitance term from the inversion-layer capacitance definition as described in section 2.2.1. This expression does not contain any assumption and is mathematically correct. Thus if we can solve 1-D Poisson-Schrodinger equation precisely and calculate correct E_i and E_C values with respect to E_F, we can obtain the centroid capacitance components that correspond to the charge distribution effect.

2.3. Verification of physical model

We have validated our gate capacitance model by simulating several structures using a 1D Poisson and Schrodinger solver and comparing its predictions with those of the model presented earlier in this chapter. The details are described in this section.

2.3.1. 1-D Poisson-Schrodinger Solver (Nextnano)

Our physical model requires correct values of E_i and E_C with respect to E_F. These physical values can be calculated from a self-consistent solution of the 1-D Poisson and Schrodinger equations, which respectively, are:

$$\nabla \varepsilon \nabla V = -\rho \tag{10}$$

$$\left(-\frac{\hbar^2}{2} \nabla \frac{1}{m^*} \nabla + \Delta E_C - qV \right) \cdot \varphi_i = E_i \varphi_i \tag{11}$$

In this equation, φ_i is the wave-function for electrons, ΔE_C is conduction band edge discontinuity, and E_i is eigenvalue of φ_i . Since it is hard to solve these two equations analytically, we used a 1-D Poisson-Schrodinger solver called Nextnano. Nextnano can solve this equation self-consistently and provide the conduction band profile, subband energy levels, and charge density together for a variety of heterostructures created using a wide range of III-V materials.

Figure 2-5 shows the conduction band profile and electron density profile of a 1D vertical cut through the gate of a typical HEMT structure studied in this work. The 1D simulation structure consists of gate metal layer, $In_{0.52}Al_{0.48}As$ barrier, δ doping layer, ($In_{0.53}Ga_{0.47}As$ + $In_{0.7}Ga_{0.3}As$ + $In_{0.53}Ga_{0.47}As$) channel layer and $In_{0.52}Al_{0.48}As$ buffer layer as described on the top of Figure 2-5. This heterostructure is a typical HEMT design that our group has fabricated [14, 15]. The Schottky barrier height is selected at 0.63 eV. We can also apply any voltage to gate with respect to the channel and in Figure 2-5, 0.3 V is applied. Therefore, we can sweep the voltage of the gate in this structure and calculate the charge and capacitance at each bias point. At each gate voltage point, subband energy levels and conduction band edge profiles are collected to generate our physical model.

Since Nextnano performs a self-consistent 1-D solution of the Poisson-Schrodinger equations, we can use it also to predict the overall capacitance and sheet charge in the channel for different gate voltages. The predictions of Nextnano can therefore be used to validate the physical model of capacitance presented above.

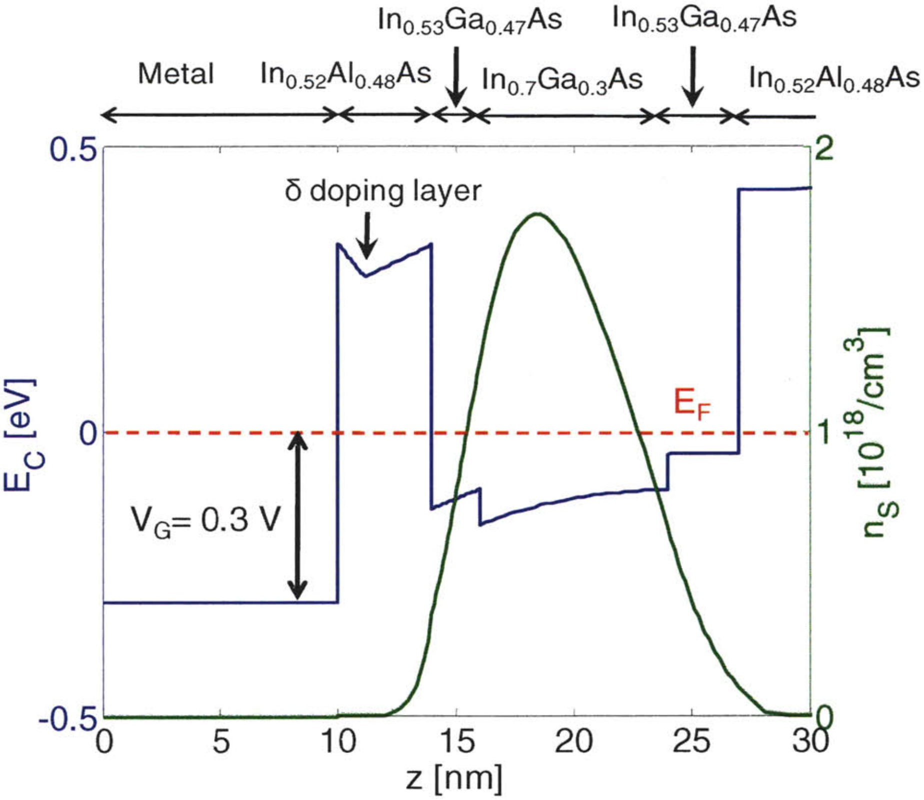

Figure 2-5. Conduction band (E_C) and electron density ($n_S(z)$) profile of 1D HEMT structure. The Schottky barrier height is 0.63 eV and the applied voltage at the gate metal layer is 0.3 V in this simulation.

2.3.2. Model and numerical simulation

We verified our model with Nextnano simulation results in three 1D HEMT structures with different channel and barrier thickness design. Channel material is either 8 nm $In_{0.7}Ga_{0.3}As$ or 5 nm pure InAs. There is a 2+3 nm $In_{0.53}Ga_{0.47}As$ channel cladding layer above and below, respectively, of the core channel layer. The total channel thickness (t_{ch}) is either 13 nm or 10 nm and the barrier thickness (t_{ins}) is either 4 nm or 10 nm. More details of these heterostructures will be given in section 3.2. Here, we first compare the predictions of our model against those of Nextnano in terms of sheet carrier concentration and gate capacitance.

2.3.2.a. Sheet carrier concentration (N_S) vs gate voltage (V_G)

Figure 2-6 shows the sheet carrier concentration of three HEMTs structures in a linear scale and Figure 2-7 shows in a logarithm scale. The solid lines are numerical simulation results which directly come from Nextnano and the symbols come from our analytical expression for N_S:

$$N_S = \sum_i \int_{E_i}^{\infty} \frac{\frac{m_{\|}^*}{\pi\hbar^2}}{1+\exp(\frac{E-E_F}{kT})} dE \tag{12}$$

In all three structures, the channel consists of a core layer and two cladding layers. Since our model uses a single $m_{\|}^*$ to calculate N_S in the entire channel layer, an average $m_{\|}^*$ is calculated by weighting different $m_{\|}^*$ of the subchannel layers accounting for the proportion of charge present in each channel layer. The details of the charge proportion profile are introduced in section 5.3.

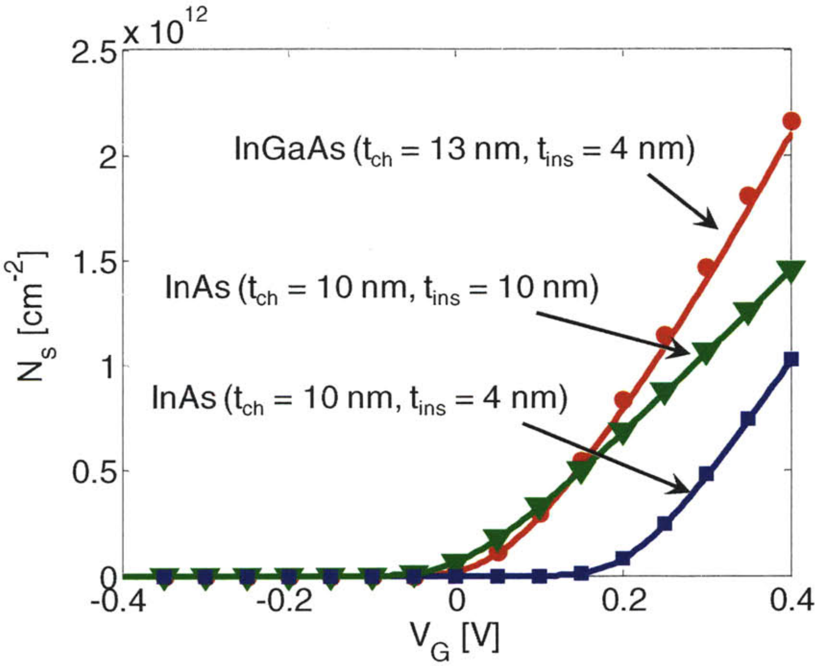

Figure 2-6. Sheet carrier concentration (N_S) as a function of gate voltage (V_G) for three HEMT structures in a linear scale. Solid lines are Nextnano simulation results and symbols are model results.

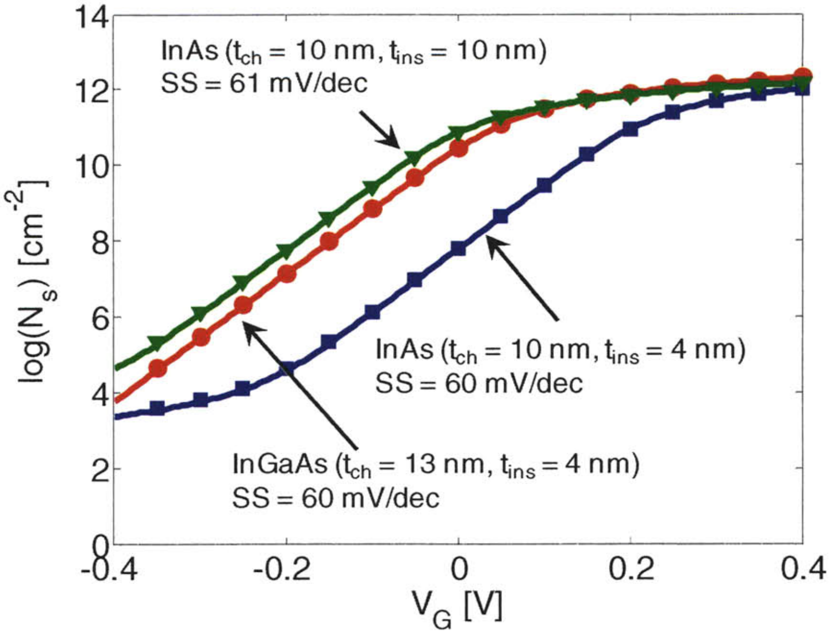

Figure 2-7. Sheet carrier concentration (N_S) as a function of gate voltage (V_G) for three HEMT structures in logarithm scale. Solid lines are numerical simulation results and symbols are model results. Subthreshold slopes (SS) are close to the ideal one (60 mV/dec) for three HEMTs.

In all of three cases, our analytic expression for N_S shows very good agreement with numerical simulation results. The Schottky barrier height in the Nextnano simulations has been adjusted to the experimental threshold voltages shown in Ch. 3. More details are given in Ch. 3. Subthreshold slopes (SS) of three HEMTs in Figure 2-7 are close to the ideal one (60 mV/dec). This is because the inversion-layer capacitance (C_{inv}) is much smaller than the insulator capacitance (C_{ins}) in weak inversion for all three HEMTs. The detailed capacitance analysis will be shown in Ch. 4.

2.3.2.b. Gate capacitance (C_G) vs gate voltage (V_G)

A comparison of the modeled and simulated gate capacitance for the three HEMTs is shown in Figure 2-8. Solid lines are Nextnano simulations results which come from the gate capacitance definition $C_G = \frac{q\partial N_S}{\partial V_G}$ (Nextnano calculates N_S directly as shown in Figure 2-8). The symbols are our model results which come from the combination of quantum capacitance, centroid capacitance, and insulator capacitance model described in section 2.2.1. Our model results agree very well with numerical calculation results in all three cases.

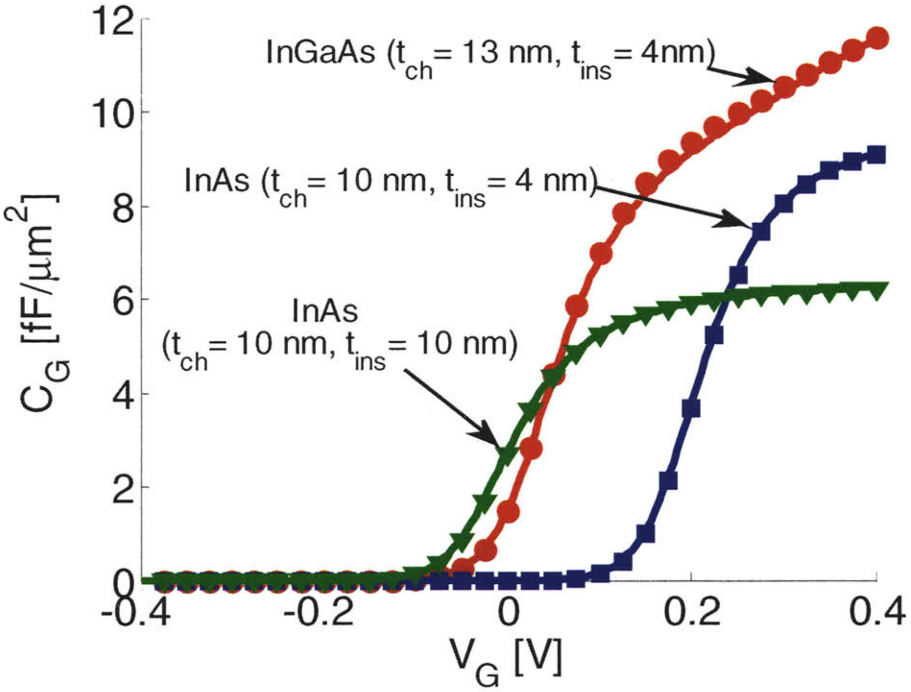

Figure 2-8. Gate capacitance (C_G) as a function of gate voltage (V_G) for three HEMT structures. Solid lines are numerical simulation results and symbols are model results.

Given the fact that our model uses input from Nextnano to evaluate the gate capacitance and that Nextnano can do this directly anyway, it might not seem obvious the need for our detailed model. The main reason to build our physical model is that we wish to decompose the total gate capacitance into each gate capacitance component. This allows us to obtain physical insight into the key dependencies of each component and also to compare them with each other. Figure 2-9 shows the total gate capacitance and its capacitance components for the heterostructure with an $In_{0.7}Ga_{0.3}As$ core channel with t_{ch}= 13 nm and t_{ins}= 4 nm. From this figure, we can understand the relative roles of the quantum capacitance and the centroid capacitance in this device. A more detailed gate capacitance analysis will be described together with experimental measurements in Chapter 4.

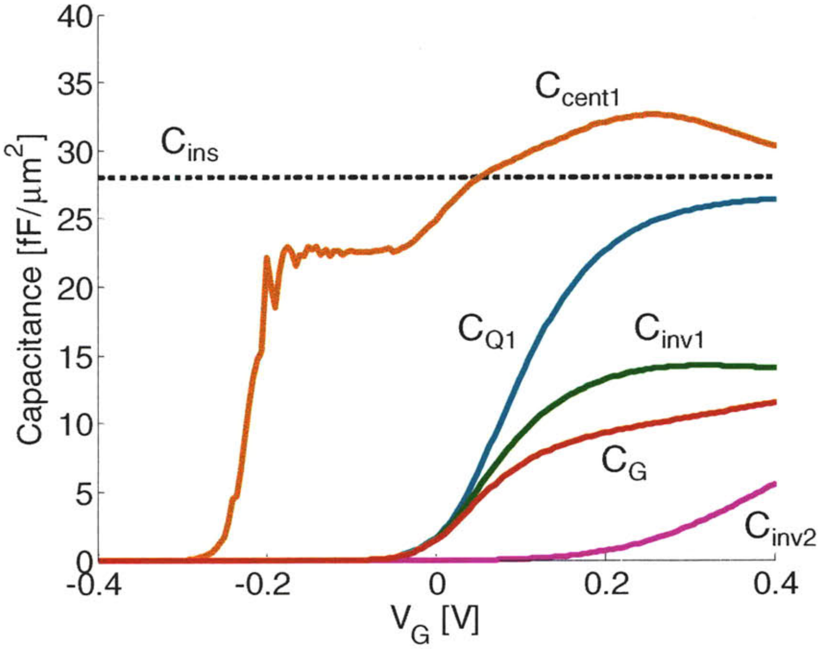

Figure 2-9. Gate capacitance (C_G) and capacitance component model results (insulator capacitance (C_{ins}), 1st subband centroid capacitance (C_{cent1}), 1st subband quantum capacitance (C_{Q1}), 1st subband inversion-layer capacitance (C_{inv1}), and 2nd subband inversion-layer capacitance (C_{inv2}).

2.3.2.c. Low-pass filter for centroid capacitance (C_{cent})

There was a technical issue in obtaining the centroid capacitance in our model results. Our centroid capacitance model includes a derivative part ($\frac{\partial(E_F\text{-}E_i)}{\partial(E_i\text{-}E_C)}$). When we calculate the increment of ($E_i - E_C$) with respect to V_G, in weak inversion condition, it sometimes goes very close to zero. Small calculation noise around zero in the denominator part of the derivative expression turn out huge spikes in the centroid capacitance which are not physical. These spikes are observed in all the heterostructure that we have studied. To eliminate this numerical noise, we have applied a low pass filter to the data, as shown in Figure 2-10. The final value of the total gate capacitance does not change in a significant way after the process for all structures that we have studied.

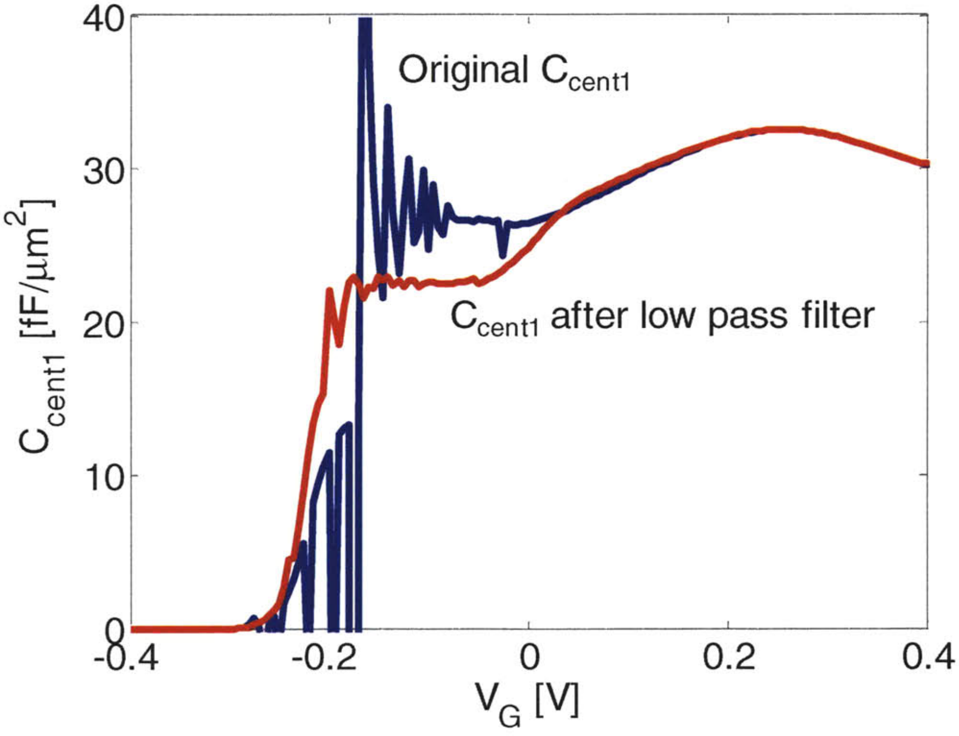

Figure 2-10. 1st subband centroid capacitance from original calculation results (blue line) and the ones after low pass filter (red line).

2.3.2.d. Inversion-layer capacitance (C_{inv}) vs sheet carrier concentration (N_S)

In the weak inversion condition, in most circumstances, only the first subband is partially occupied by electrons. If we assume that C_{cent1} is much bigger than C_{Q1}, we can find an interesting relation between C_{inv} and N_S from our physical model. The derivation of this relation is shown in detailed below.

In the first subband model, N_S is expressed as

$$N_S = \int_{E_1}^{\infty} \frac{\frac{m_{\parallel}^*}{\pi\hbar^2}}{1+\exp(\frac{E-E_F}{kT})} dE \tag{13}$$

Carrying out the integral:

$$\int_{E_1}^{\infty} \frac{1}{1+\exp(\frac{E-E_F}{kT})} dE = kT \cdot \ln(1+\exp(\frac{E_F-E_1}{kT})) \tag{14}$$

For $E_F < E_1$, we can do a Taylor Series expansion

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} \ldots \text{ for } |x| \leq 1 \tag{15}$$

This allows us to express N_S as

$$N_S = \frac{m_{\parallel}^* kT}{\pi\hbar^2} \exp(\frac{E_F-E_1}{kT}) \tag{16}$$

Under the same conditions ($E_F<E_1$):

$$C_{Q1} = \frac{\frac{m_{\parallel}^* q^2}{\pi\hbar^2}}{1+\exp(\frac{E_1-E_F}{kT})} \cong \frac{\frac{m_{\parallel}^* q^2}{\pi\hbar^2}}{\exp(\frac{E_1-E_F}{kT})} \tag{17}$$

From the combination of eqs.(16) and (17), we find a direct relation between N_S and C_{Q1}.

$$N_S = C_{Q1} \cdot \frac{kT}{q^2} = C_{inv} \cdot \frac{kT}{q^2} \tag{18}$$

$$\log(N_S) = \log(C_{inv}) + \log(\frac{kT}{q^2}) \tag{19}$$

Therefore, according to equation (19), C_{inv} becomes proportional to power one of N_S in weak inversion condition.

We tested this relation in our model and simulation results. Figure 2-11 shows the relation between C_{inv} and N_S in a logarithm scale for all three HEMT structures. The slopes in weak inversion are exactly unity for the three HEMTs and this fact means that C_{Q1} is the dominant component of C_{inv} in the weak inversion regime.

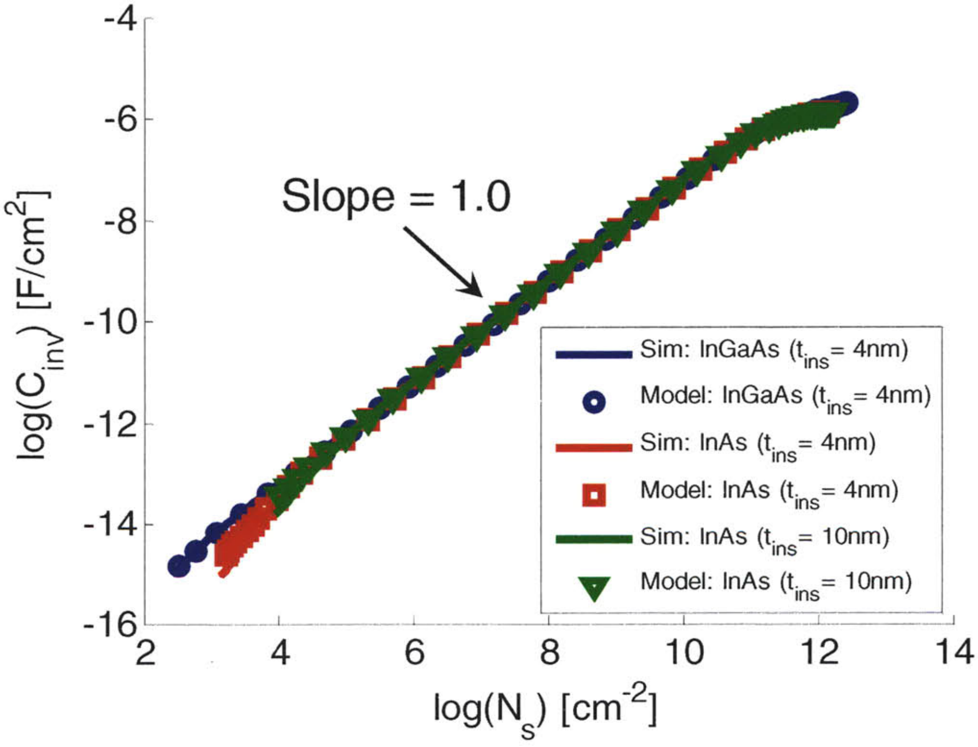

Figure 2-11. Inversion-layer capacitance (C_{inv}) versus sheet carrier concentration (N_S) of all of three HEMTs in a logarithm scale. Solid lines are simulation results and symbols are our model results. All of slopes of the lines are equal to unity.

2.4. Summary

In summary, we have built a physical model for the gate capacitance of III-V FETs and verified this with Nextnano simulations for three different HEMTs. In the next chapter, we will obtain experimental measurements of gate capacitance in three HEMT structures. In Chapter 4, the measured values will be compared with the calculated values using the model presented in this chapter.

Chapter 3. Experimental Gate Capacitance Measurements in HEMTs

3.1. Introduction

The impact of quantum capacitance and centroid capacitance on the overall gate capacitance is only observed in deeply scaled down devices. InGaAs HEMTs provide nearly ideal device conditions to explore this issue because of the recent great scaling progress that has been made in these devices [15]. Especially InGaAs HEMTs fabricated at MIT are some of the most aggressively scaled-down device structures down to 30 nm gate length with various heterostructure designs as well as world record f_T. In this work, we have extracted the gate capacitance of three different InGaAs HEMTs fabricated in our group. These have different barrier thickness and channel layer designs. We used S-parameter measurements to extract the gate capacitance in these devices. We have built a capacitance model for the HEMTs, and for this, we have obtained the intrinsic gate capacitance by taking measurements in devices with different gate lengths.

3.2. Device technology and heterostructures

We have investigated the gate capacitance in three HEMT structures with different channel and barrier designs (Table. 3-1 and Figure. 3-1). These devices are fabricated using InAlAs and InGaAs layers grown on InP substrates. In essence, we have two channels, one with $In_{0.7}Ga_{0.3}As$ at the center of a 13 nm thick channel [15] and another one with pure InAs at the center of a 10 nm channel [14]. In both cases, the channel cladding is $In_{0.53}Ga_{0.47}As$ (2 and 3 nm above and below the channel core, respectively). For the InAs-channel design, we have two $In_{0.52}Al_{0.48}As$ barrier thicknesses, t_{ins}, of 4 and 10 nm. For the InGaAs sample, t_{ins} is 4 nm. The final InAlAs barrier thickness is controlled by a three-step recess process [17]. A δ-doped layer is inserted in the barrier layer 3 nm away from the InAlAs/InGaAs interface. The δ doping concentration is 5 x 10^{12} cm^{-2} in all three cases. The device gate length is in the range of 200 to 30 nm. Two different kinds of gate stacks (Ti/Pt/Au and Pt/Ti/Mo/Au) are used. We have named these three different heterostructures as Type A, B, and C, as listed in Table 3-1.

Notation	t_{ins}(nm)	t_{ch}(nm)	Channel Core	Reference	L_G (nm)	Gate stack
Type A	10	10	InAs (5 nm)	Kim, unpublished	40 ~ 100	Ti/Pt/Au
Type B	4	10	InAs (5 nm)	Kim, [14]	30 ~ 200	Pt/Ti/Mo/Au
Type C	4	13	$In_{0.7}Ga_{0.3}As$ (8 nm)	Kim, [15]	40 ~ 100	Ti/Pt/Au

Table 3-1. Key parameters of the three different heterostructures studied in this work.

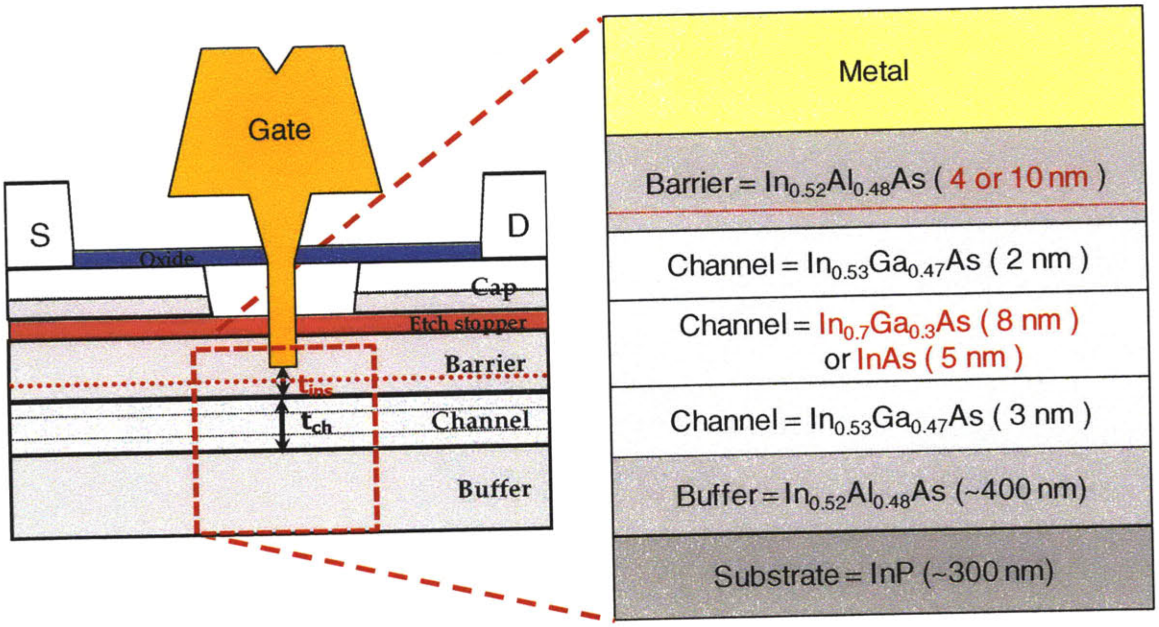
Gate
S
D
Oxide
Cap
Etch stopper
Barrier
t_{ins}
t_{ch}
Channel
Buffer
Metal
Barrier = $In_{0.52}Al_{0.48}As$ (4 or 10 nm)
Channel = $In_{0.53}Ga_{0.47}As$ (2 nm)
Channel = $In_{0.7}Ga_{0.3}As$ (8 nm) or InAs (5 nm)
Channel = $In_{0.53}Ga_{0.47}As$ (3 nm)
Buffer = $In_{0.52}Al_{0.48}As$ (~400 nm)
Substrate = InP (~300 nm)

Figure 3-1. Experimental HEMT cross section and layer details of the three heterostructures explored in this work.

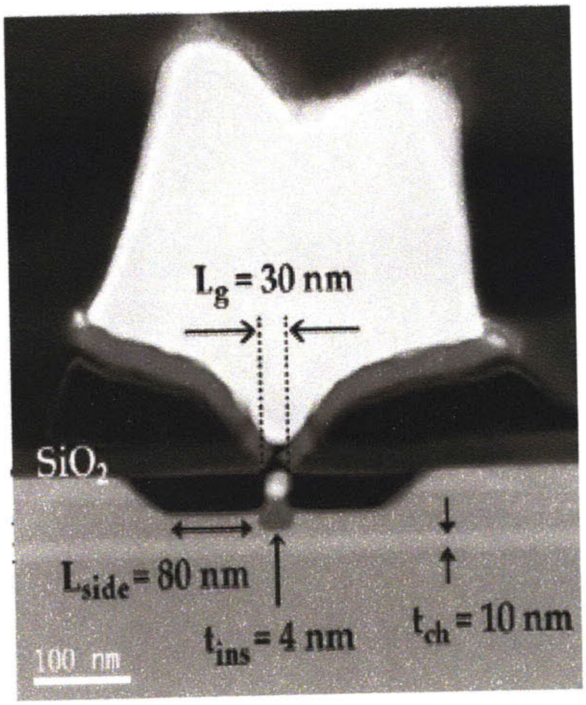
L_g = 30 nm
SiO_2
L_{side} = 80 nm
t_{ins} = 4 nm
t_{ch} = 10 nm
100 nm

Figure 3-2. A TEM image of 30 nm gate length Type B device. [Figs. taken from Kim, [17]]

Figure 3-2 shows a TEM image of a 30 nm gate length Type B device. This device has a very thin barrier layer of around 4 nm which is close to the vertical scaling limit of III-V HEMTs, as well as a rather thin 10 nm quantum well channel design. In our work, the use of a different channel material design such as $In_{0.7}Ga_{0.3}As$ and InAs will also illustrate the impact of the density of states on the gate capacitance. Therefore, a comparison of the gate capacitance among these three different scaled-down HEMTs will help us to build models that will allow us to estimate the gate capacitance characteristics and sheet carrier concentration of future ultra scaled-down III-V FETs.

3.3. Gate capacitance extraction method

This section presents small signal equivalent circuit model for a III-V FET and the detailed procedure of extracting C_{gs} and C_{gd} in this model from S-parameter measurements. We have built a capacitance model for the HEMTs which captures the intrinsic and parasitic gate capacitance. The intrinsic gate capacitance can be separated from the measurements of different gate length devices.

3.3.1. S-parameter analysis

We have used S-parameter characterization to extract the gate capacitance of the transistors. Figure 3-3 shows a small signal equivalent circuit model for a FET [23]. This model is also applicable to our HEMTs. In this model, $(C_{gs} + C_{gd})$ corresponds to the total gate capacitance that we want to extract. All of the S-parameter measurements have been performed using an Agilent 8517B S-parameter Test Set and an HP 8510C Network Analyzer.

The gate capacitance of these devices is obtained from a multi-step process. First, we eliminate the extrinsic circuit elements from the S-parameter measurements through well established de-embedding procedures. We then transform the de-embedded S-parameters to Z-parameters. Then, the intrinsic Z-parameters (Z') are calculated by subtracting R_g, R_s, and R_d following [24]:

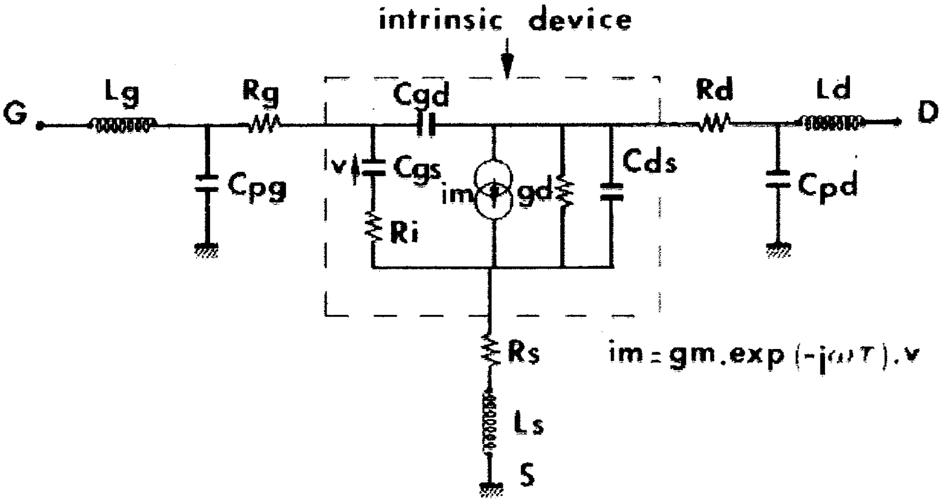
intrinsic device
G
Lg
Rg
Cgd
Rd
Ld
D
Cpg
v
Cgs
im
gd
Cds
Cpd
Ri
Rs
im = gm.exp (-jωτ).v
Ls
S

Figure 3-3. Small signal equivalent circuit model for a FET. This model is applied to our HEMT structure to extract C_{gs} and C_{gd} (R_i is assumed to be zero). [Fig. taken from Dambrine, [23]].

$$Z'_{11} = Z_{11} - (R_g + R_s) \tag{1}$$

$$Z'_{22} = Z_{22} - (R_d + R_s) \tag{2}$$

$$Z'_{12} = Z_{12} - R_s \tag{3}$$

$$Z'_{21} = Z_{21} - R_s \tag{4}$$

R_g, R_s, and R_d are obtained by zero-bias S-parameter measurements [24] and confirmed from TLM (Transmission-Line Modeling Method). Next, the intrinsic Y-parameters are obtained from the intrinsic Z parameters, Z', as:

$$Y = \frac{1}{Z'} \tag{5}$$

For the equivalent circuit of Fig. 3-3 (with $R_i = 0$), the intrinsic Y parameters have the following expressions:

$$Y_{11} = jw(C_{gs} + C_{gd}) \tag{6}$$

$$Y_{22} = jw(C_{ds} + C_{gd}) + \frac{1}{g_d} \tag{7}$$

$$Y_{12} = -jwC_{gd} \tag{8}$$

$$Y_{21} = g_m - jwC_{gd} \tag{9}$$

According to these equations,

$$C_g = C_{gs} + C_{gd} = \frac{\text{Im}(Y_{11}) + \text{Im}(Y_{22}) + 2\text{Im}(Y_{12})}{2\pi f} \tag{10}$$

By following this procedure, the total gate capacitance (C_g) of each HEMT is extracted from S-parameter measurements.

3.3.2. Intrinsic and parasitic components of gate capacitance

The gate capacitance extracted from S parameters as described in the previous subsection includes parasitic components. This needs to be separated in order to eventually isolate the intrinsic gate capacitance per unit area (C_{gi}) which is what should be compared with our physical model. In this work, we define two different kinds of parasitic gate capacitance as shown in Figure 3-4. There is an outer component, C_{gext_outer}, associated with the top of the T-gate. There is also an inner parasitic capacitance, C_{gext_inner}, that is associated with the sidewall of the gate stem. Unlike C_{gext_outer}, this parasitic component depends on V_G.

According to this model, the total gate capacitance can be expressed as:

$$C_G\,(\text{in fF/mm}) = C_{gi}\,(\text{in fF/}\mu\text{m}^2) \times L_G + 2 \times C_{gext\_inner}(V_G) + 2 \times C_{gext\_outer} \tag{11}$$

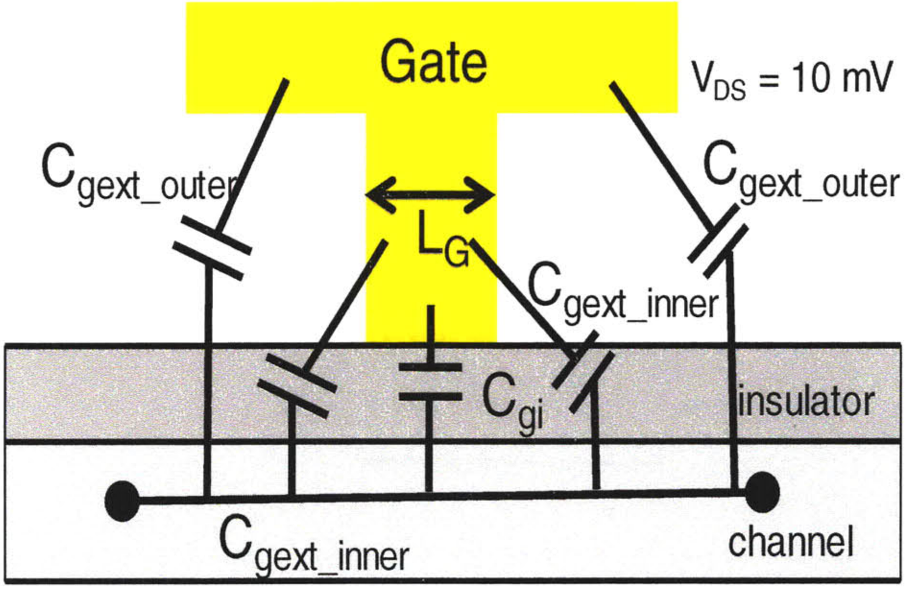

Figure 3-4. Intrinsic and parasitic components of gate capacitance in a HEMT. Two different kinds of parasitic gate capacitance are defined in this model. C_{gext_outer} is associated with the top of the T-gate. C_{gext_inner} is associated with the sidewall of the gate stem.

In order to eliminate these two parasitic terms, we measured C_G in devices with different gate lengths (L_G) as a function of V_G. For very low voltage such as V_{GS} = -0.3 V, C_{gi} and C_{gext_inner} becomes nearly zero since both terms depend on V_{GS}. Therefore, $2C_{gext\_outer}$ can be taken as C_G (V_{GS} = -0.3V). For a given value of V_{GS}, the resulting values of C_G - C_G (V_{GS} = -0.3 V) are linear in L_G. C_{gi} is then obtained from the slope of C_G - C_G (V_{GS}= -0.3 V) with L_G at each value of V_{GS}. The intersection of these lines corresponds to $2C_{gext\_inner}$. This experimental methodology separates all these components.

3.4. S-parameter measurements results – Type B

We carried out S-parameter measurements for the three HEMTs (Type A, B, and C) with different L_G in the linear regime (V_{DS}= 10 mV). The frequency used for small-signal model extraction ranges from 1 to 40 GHz. In this section, we illustrate the procedure for intrinsic gate capacitance extraction for Type B devices.

3.4.1. Capacitance vs. Frequency

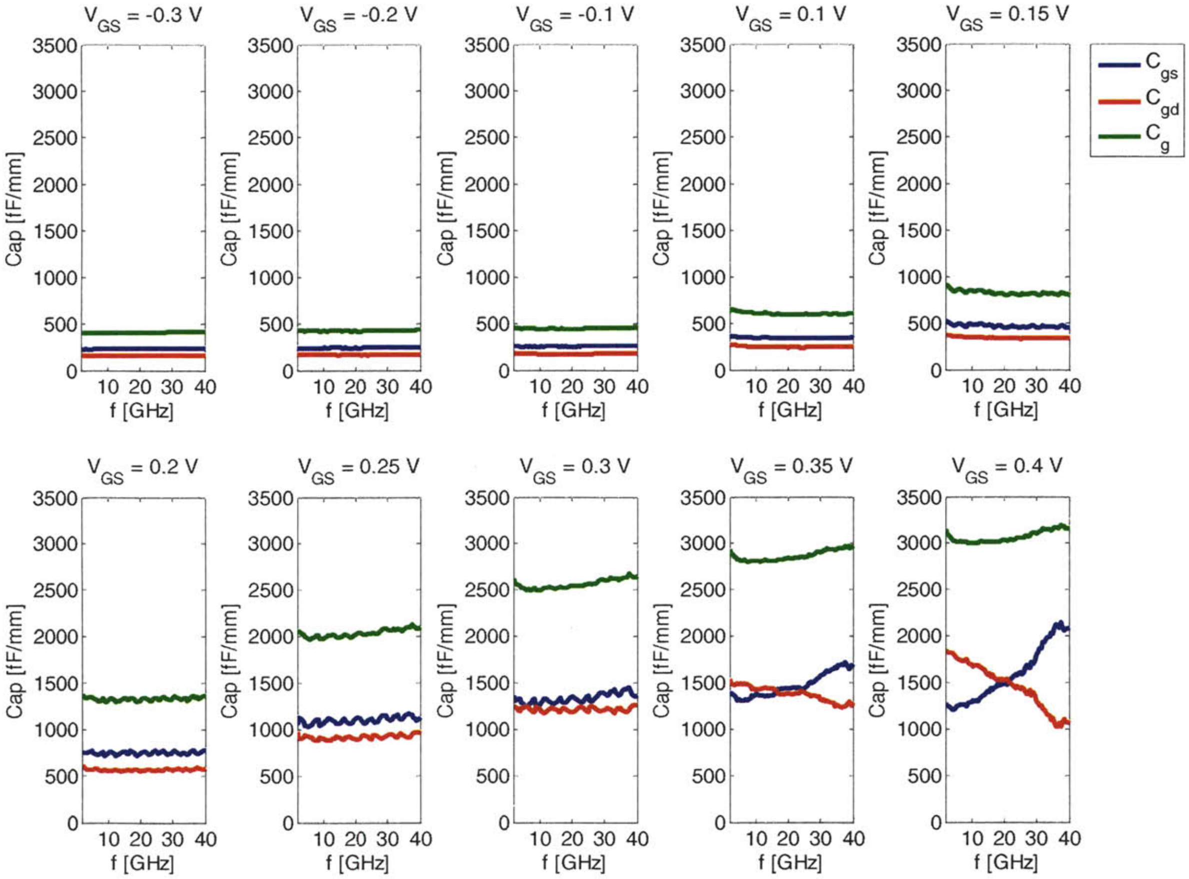

Figure 3-5. Extracted small signal model components (C_{gs}, C_{gd} and C_g) results with respect to frequency in Type B device with L_G = 200 nm for different values of V_{GS}.

There are four Type B devices with L_G = 30 nm, 60 nm, 100 nm and 200 nm. Figure 3-5 shows C_{gs}, C_{gd} and C_g as a function of frequency for the 200 nm gate length device at different values of V_{GS}. C_g (green line) is a summation of C_{gs} (blue line) and C_{gd} (red line), and remains nearly constant as a function of frequency. As V_{GS} increases, C_{gs} and C_{gd} exhibit some frequency dependence, but their sum, C_g, is relatively flat. C_g is averaged in this frequency range and used for the next gate capacitance extraction process.

3.4.2. Capacitance vs. V_{GS}

Figure 3-6 shows the extracted C_G (= $C_{gs} + C_{gd}$) as a function of V_{GS} for different L_G in Type B devices. In Figure 3-6, C_G remains nearly flat when V_{GS} is below -0.1V. This is because C_{gi} and C_{gext_inner} become nearly zero in this low voltage range. Therefore, C_G at V_{GS} = -0.3 V for each L_G is taken as $2C_{gext\_outer}$. Then, we subtracted $2C_{gext\_outer}$ from C_G for each L_G and illustrated results in Figure 3-7. Now, we need one more step to remove C_{gext_inner} from these results.

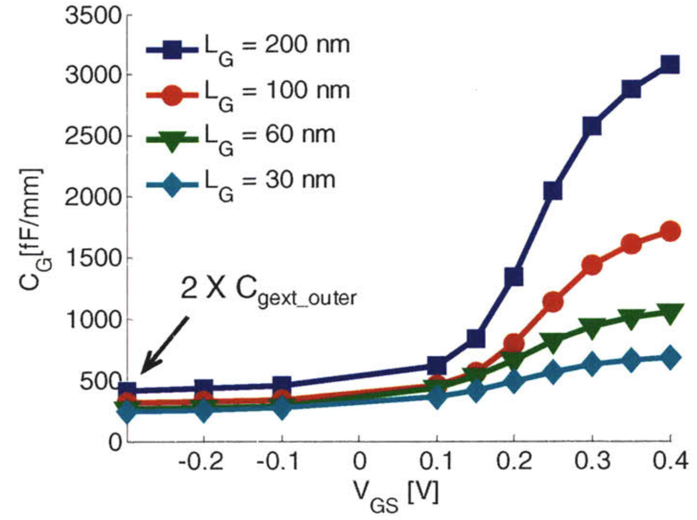

Figure 3-6. C_G (= $C_{gs} + C_{gd}$) vs. V_{GS} for different L_G in Type B devices. C_G (V_{GS} = -0.3 V) corresponds to C_{gext_outer} in Figure 3-4.

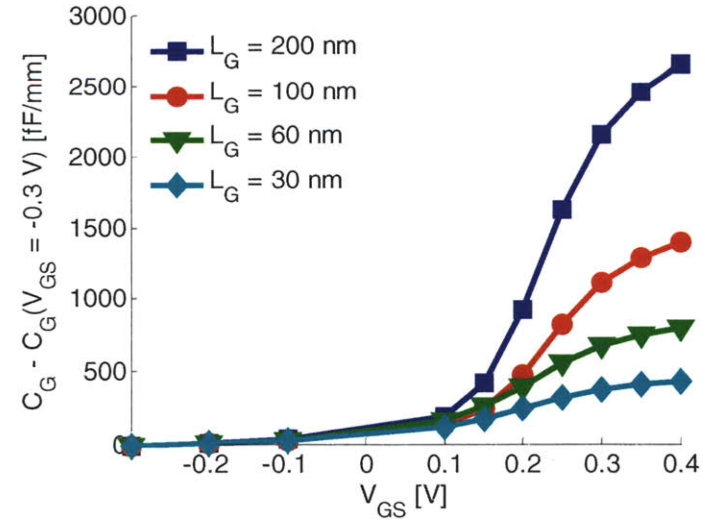

Figure 3-7. C_G –C_G (V_{GS} = -0.3 V) vs. V_{GS} for different L_G in Type B devices.

3.5. Extraction of intrinsic gate capacitance

Figure 3-8 is a redrawn picture of Figure 3-7 by expressing $C_G - C_G$ ($V_{GS} = -0.3$ V) as a function of L_G for different values of V_{GS}. According to the equation (11), the slope of this line at each bias point corresponds to C_{gi} and the intersection is $2C_{gext\_inner}$. We also obtained C_{gi} by following an identical procedure in Type A and Type C devices with different L_G.

Figure 3-9 shows all of extracted C_{gi} in Type A, B, and C devices. C_{gi} rises up in the middle of a gate voltage range and begins to be saturated at high gate voltages in all three devices. Type B shows the highest values of C_{gi} in the comparison with Type A and C. Each one has a different threshold voltage because of Schottky barrier height variation from different gate metal, barrier thicknesses and channel composition as described in Table 3-1. In Ch. 4, we will compare these results with modeling results from our physical model and analyze the relative contributions of each capacitance component to the overall gate capacitance.

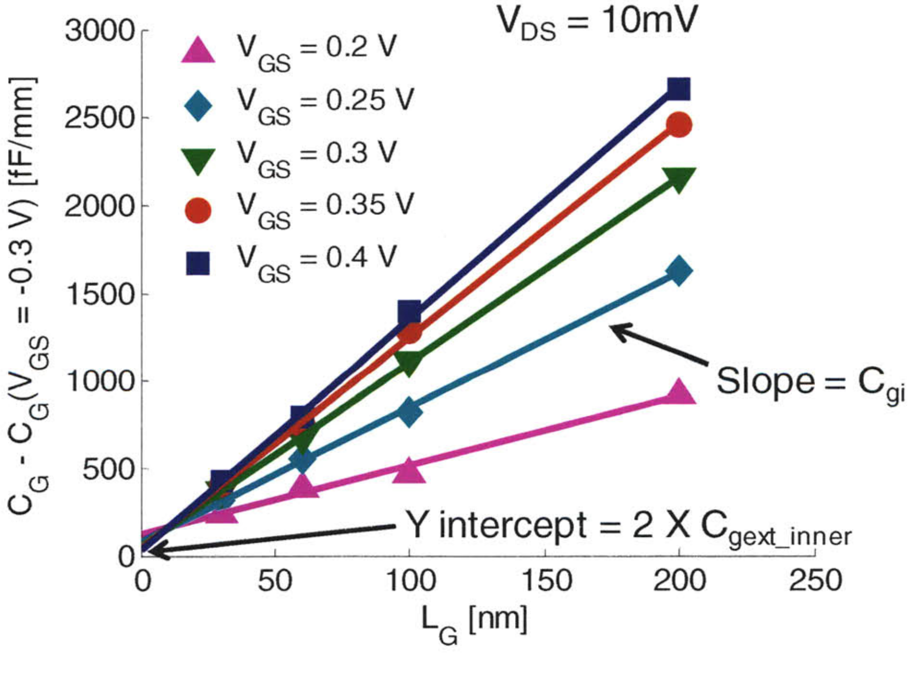

Figure 3-8. $C_G - C_G$ ($V_{GS} = -0.3$V) as a function of L_G at different V_{GS} for Type B devices. The slope of this line at each bias point corresponds to C_{gi}. The Y intercept is $2C_{gext\_inner}$ of Figure 3-4.

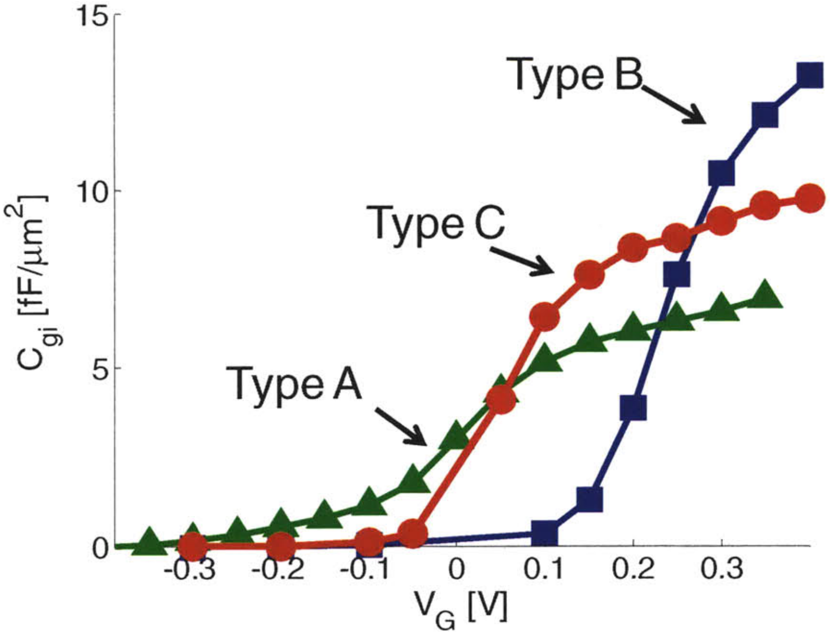

Figure 3-9. Intrinsic gate capacitance (C_{gi}) for Type A, B, and C heterostructures.

3.6. Summary

In this chapter, we describe a methodology to experimentally extract the gate capacitance from scaled-down InGaAs HEMTs. Devices with different gate lengths built on three different heterostructures are studied and their intrinsic gate capacitances are extracted. In the next chapter, these experimental intrinsic gate capacitances will be compared in detail with predictions from the physical model presented in Ch. 2. This will allow for a discussion on the relative contribution of the gate capacitance components to the overall capacitance.

Chapter 4. Comparison of Measurements with Physical Model

4.1. Introduction

In this chapter, we compare the experimental intrinsic gate capacitance of the HEMTs extracted in Ch.3 with the predictions of the gate capacitance model derived in Ch.2 for the three different heterostructures. We show that our model matches the measurements relatively well. From this, we are able to assess the relative contributions from the various capacitance components to the overall gate capacitance of the HEMTs.

4.2. Schottky barrier height and effective mass in our model

The experimental intrinsic gate capacitance presents a different threshold voltage for each type of heterostructure, as shown in Figure 3-9. This partially comes from the different Schottky barrier heights of gate metal stack, as listed in Table 3-1. Type A and C used a most common gate metal stack, Ti/Pt/Au, but Type B has a Pt-sinking gate metal stack, Pt/Ti/Mo/Au. In the Nextnano simulation code, we can control Schottky barrier height between metal and $In_{0.52}Al_{0.48}As$ barrier layer. In this work, we set Schottky barrier heights of 0.55 eV for Type A, 0.83 eV for Type B, and 0.63 eV for Type C which allows us to match the experimental threshold voltage. The experimental literature reports that Ti has a Schottky barrier of 0.65 eV on InAlAs layer [25] and that Pt-buried gate has a 0.83 eV on InAlAs layer [26]. Therefore, our Schottky barrier choices are reasonable.

As explained in Ch.2, our physical model uses one in-plane effective mass ($m_{\|}*$) to calculate C_Q, C_{cent}, and C_G expressions. Since our experimental HEMTs have a compound channel structure (InAs or $In_{0.7}Ga_{0.3}As$ core channel layer sandwiched with $In_{0.53}Ga_{0.47}As$ cladding layers), we computed an average effective mass for the three layer structure that weighs each region by its total electron concentration. The averaged effective masses are 0.031 m_o (m_o = electron mass) for Type A and B which is about 1.2 x InAs core channel layer $m_{\|}*$ and 0.040 m_o for Type C which is about 1.05 x $In_{0.7}Ga_{0.3}As$ core channel $m_{\|}*$. These small increments of $m_{\|}*$ from the core channel material effective mass come from the relatively small electron population in the cladding layers, as shown in Ch. 5. Using these $m_{\|}*$ values, our analytical model fits very well the Nextnano simulations for all three devices as shown in Figure 2-8. We are therefore in a position to compare the predictions of our physical model with the experimental results.

4.3. C_G in Type A (InAs channel, t_{ch}= 10 nm, t_{ins}= 10 nm)

Type A structure has 5 nm InAs core channel layer with t_{ch}= 10 nm and t_{ins} = 10 nm. Figure 4-1 shows the extracted experimental C_G and our modeled C_G and its components. The total gate capacitance predicted by the model (red line) is in excellent agreement with the measured values (blue squares). This gives us confidence over the calculation of the individual components of C_G. First of all, at t_{ins} = 10 nm, C_{ins} is 11.2 fF/μm^2. This is substantially bigger than the experimental gate capacitance, which has a maximum value of 7.0 fF/μm^2 or about 63% of C_{ins}. This sizable gate capacitance degradation mostly results from the finite inversion layer capacitance in this structure. As shown in Figure 4-1, C_{inv1} is quite comparable to C_{ins} in the strong inversion condition. In this structure, since C_{inv2} is negligible, the total inversion layer capacitance, C_{inv}, equals to C_{inv1}. In C_{inv1} , C_{Q1} is the dominant component due to the small effective mass of the InAs channel.

Figure 4-2 shows the evolution of the energy levels of the bottom of the 1st and 2nd subband with respect to the Fermi level (E_F). E_F only penetrates into the 1st subband energy in the operational gate voltage range, which suggests that only the 1st subband of the quantum well is populated with electrons. This is why C_{inv2} is negligible in Figure 4-1. In this structure, the 10 nm channel thickness results in an energy difference of about 0.25 eV between the 1st and 2nd subbands, and E_F cannot rise up this much in the operation range.

4.4. C_G in Type B (InAs channel, t_{ch}= 10 nm, t_{ins}= 4 nm)

The type B structure has a 5 nm InAs core channel layer with a total t_{ch} = 10 nm and t_{ins} = 4 nm. The difference from Type A is that t_{ins} is thinned down from 10 nm to 4 nm. Figure 4-3 shows the extracted experimental C_G and our modeled components for this structure. Our modeled total C_G is in moderate agreement with the experimental C_G. We will discuss the origin of the discrepancy that is observed later at the end of this chapter and in the next chapter. At t_{ins} = 4 nm, C_{ins} is 28.1 fF/μm^2 and becomes much bigger than the experimental gate capacitance. The maximum measured gate capacitance is 13.2 fF/μm^2 which is only 47% of C_{ins}. In this structure, C_{ins} is more than twice that of the Type A structure. As a result, C_{Q1} becomes even more of a dominant factor in the Type B structure. The channel layer structure and the thickness are the same to Type A, so C_{Q1} and C_{cent1} show similar features.

Figure 4-4 shows the evolution of the 1st and 2nd subband energy levels with respect to E_F. As in Type A, for the voltage range of interest here, E_F only penetrates into the 1st subband. Therefore, C_{inv2} is negligible in Figure 4-3. The energy difference between the 1st and 2nd subbands is also about 0.25 eV.

The fact that C_{ins} is substantially bigger than C_{inv} at t_{ins} = 4 nm indicates that barrier thickness scaling beyond 4 nm will not result in a gate capacitance enhancement anymore in this kind of heterostructure.

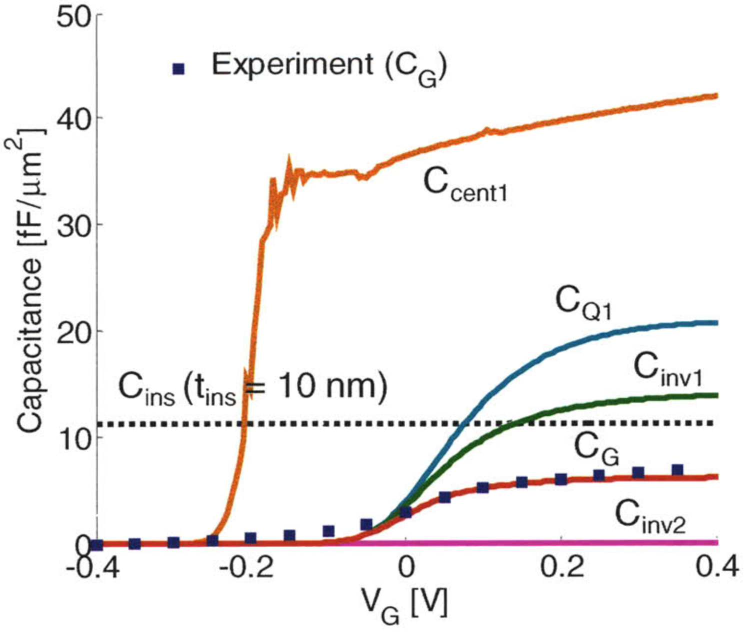

Figure 4-1. Experimental C_G and modeled C_G components of Type A (InAs channel, t_{ch} = 13 nm, t_{ins} = 4 nm) vs V_G.

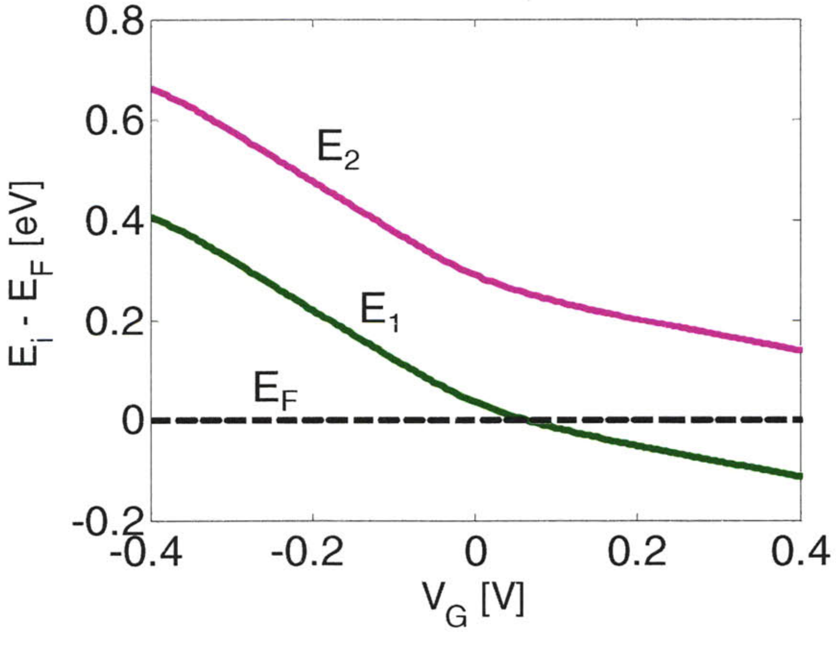

Figure 4-2. Evolution of the 1st and 2nd subband energy levels with respect to the Fermi level in the Type A structure.

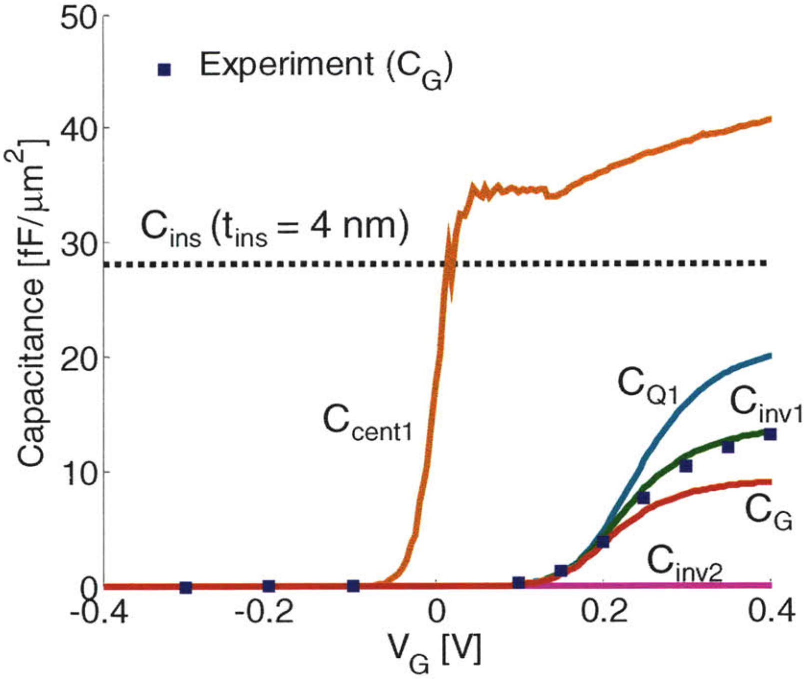

Figure 4-3. Experimental C_G and modeled C_G components of Type B (InAs channel, t_{ch} = 10 nm, t_{ins} = 4 nm) vs V_G.

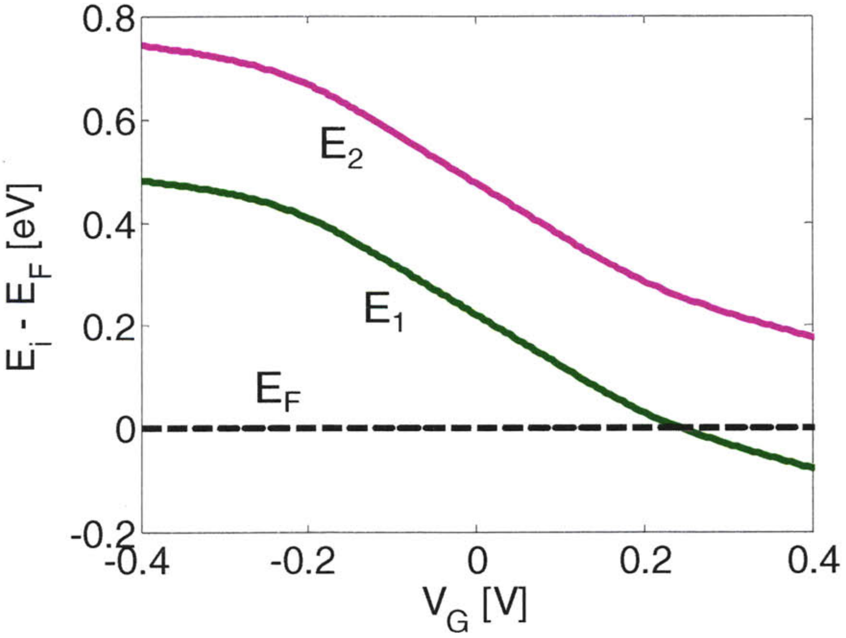

Figure 4-4. The evolution of the 1st and 2nd subband energy levels with respect to the Fermi level for the Type B structure. The 2nd subband contribution is negligible.

4.5. C_G in Type C ($In_{0.7}Ga_{0.3}As$ channel, t_{ch}= 13 nm, t_{ins}= 4 nm)

The type C heterostructure has an 8 nm $In_{0.7}Ga_{0.3}As$ core channel layer with a total t_{ch}= 13 nm and t_{ins} = 4 nm. The core channel layer has a 30% GaAs mole fraction which is characterized by an averaged m* larger than in the InAs channel cases. The channel layer is also thicker than in the Type A and B heterostructures. Figure 4-5 shows the extracted experimental C_G and our modeled C_G components. Our model matches the measurements quite well. C_{ins} is 28.1 fF/μm^2 which is bigger than experimental C_G with a maximum of 9.8 fF/μm^2 or 35% of C_{ins}. The overall gate capacitance is smaller than in the Type B heterostructure. This results from C_{cent1} which is smaller than in the Type A or B structures due to the thicker channel design which makes the centroid of charge distribution move further away from the InAlAs/InGaAs interface. On the contrary, an increase of the averaged m* in the channel layer due to higher m* from $In_{0.7}Ga_{0.3}As$ core layer contributes to an increase in C_{Q1}.

Figure 4-6 shows the evolution of the 1st and 2nd subband energy levels with respect to E_F as a function of V_G. In this structure, E_F penetrates into the 1st subband and approaches the 2nd subband. Therefore, there is some degree of electron population in the 2nd subband which contributes to the gate capacitance, as noted by the C_{inv2} line in Figure 4-5. The reason for this is that the thicker channel layer and the higher effective mass produces an energy splitting between the 1st and 2nd subbands of only 0.12 eV, in comparison with 0.25 eV in the Type A and B structures.

4.6. Key findings from the comparison of Type A, B and C

The agreement between modeled and experimental capacitance that is demonstrated in this chapter is reasonable although there are some discrepancies that we will discuss in the next chapter. In this section, we compare the results obtained from the comparative analysis of experimental and theoretical gate capacitance in Type A, B and C structures.

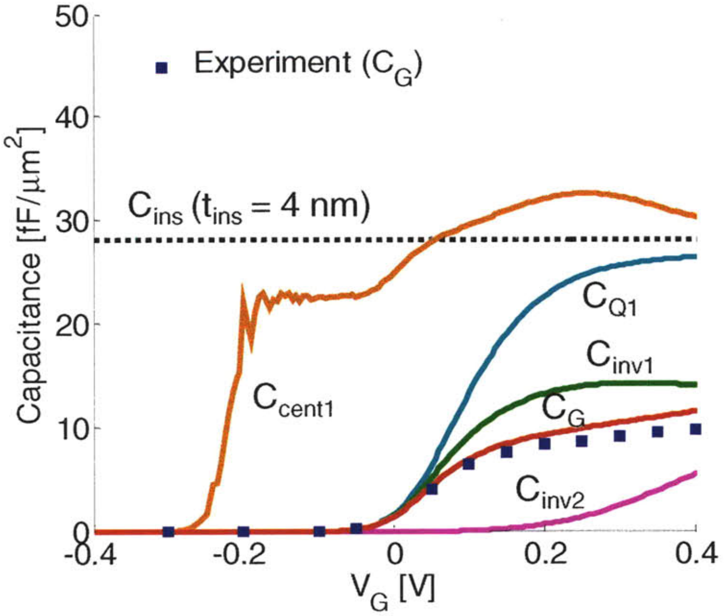

Figure 4-5. Experimental C_G and modeled C_G components of Type C ($In_{0.7}Ga_{0.3}As$ channel, t_{ch} = 13 nm, t_{ins} = 4 nm) vs V_G.

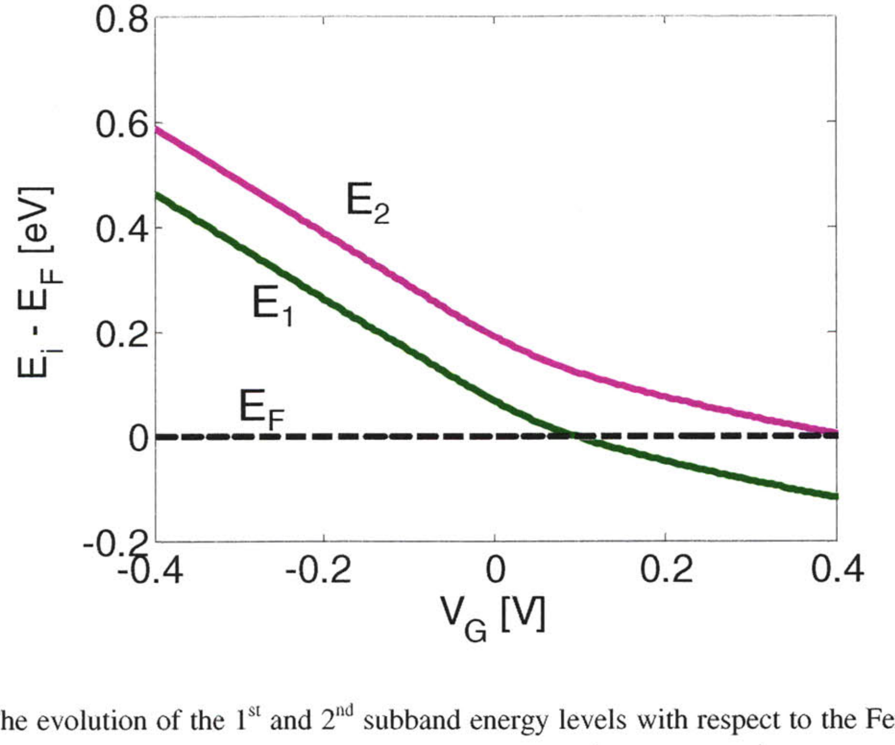

Figure 4-6. The evolution of the 1st and 2nd subband energy levels with respect to the Fermi level for the Type C structure. The 2nd subband is populated in a minor way.

First of all, in all three cases, the degradation in overall gate capacitance that comes from the finite inversion-layer capacitance is evident. The measured C_G in strong inversion is only between 35% and 62% of C_{ins}. The finite inversion-layer capacitance mostly results from a relatively small quantum capacitance. This is particularly the case for the InAs channels where the effective mass is the smallest.

An additional conclusion is that the 1st subband dominates the overall gate capacitance in the operational range of a scaled down HEMT. This is particularly the case of the InAs channel structures (Type A and B) which have significant more channel quantization due to a thinner channel and a lower effective mass. We only find a small C_{inv2} contribution to C_G in the Type C case which has a relatively thicker channel layer design and a higher effective mass. This indicates that in future scaled down devices which will be characterized by a very thin channel, it will not be possible to increase the inversion-layer capacitance by populating higher subbands in the channel.

A last conclusion is that C_{ccnt1} is also highly relevant to determining C_G. Comparing Type B and Type C in Figures 4-3 and 4-5, we see that C_{ccnt1} is significantly larger in the thin channel device (Type B) vs. the thicker channel device (Type C). This more than compensates for the lower quantum capacitance of the InAs channel and ends up with a higher overall value of C_G. This suggests that scaling down the channel thickness is an effective way to enhance gate capacitance. All of these conclusions represent good guidance to design future scaled down III-V MOSFETs.

4.7. Discrepancy in Type B

The agreement between model and experiments is worst for the Type B structure in Figure 4-3. There could be some experimental uncertainties such as measurement error in t_{ins} but, as shown in the next chapter, the discrepancy cannot come from this alone. Type B is the structure in which the quantum capacitance is most relevant. This suggests that we analyze in detail the assumptions that have been made in our calculation of C_Q.

In Nextnano simulations, we are limited to the use of spherical bands with a single well defined effective mass for each material which we have set to be equal to the bulk effective mass. In our model expression, we used an identical averaged in-plane effective mass ($m_{||}^*$) at all subband energy levels (E_i) which matches the Nextnano simulations. However, since the Type B structure has a very narrow quantum well channel, the effective mass can be larger than the bulk value and not constant at different energies. Compressive biaxial channel strain in the InAs channel layer (as a result of the lattice constant difference between InAs and the InP substrate) [27], channel quantization [28, 29], and non-parabolicity of the conduction band [30] are all predicted to increase $m_{\|}^{*}$ [31, 32] in the core channel layer. This should result in a large increase in the density of states [32] and the quantum capacitance limit. In addition, there is also some contribution of the channel cladding layer to the overall gate capacitance. In next chapter, we will discuss these issues in more detail and show improved fitting results by using a modified effective mass model.

4.8. Summary

In this chapter, we have compared experimental measurements of gate capacitance from all three heterostructures with our theoretical model. We found that the quantum capacitance acts as a limiting factor in the determination of the overall gate capacitance when the barrier thickness approaches a few nanometers. In addition, we figured out that the centroid capacitance is also highly relevant to determine the overall gate capacitance. For the structures that we have studied, only the 1st subband is substantially populated with electrons in the operational gate voltage range. A relatively big discrepancy between measurements and model results is observed in the Type B structure with a very narrow channel layer and insulator. We suggest that this is evidence for an increased in-plane effective mass which originates from compressive biaxial strain, non-parabolicity and strong quantization effects. More detailed studies about these effects will be given in Ch. 5.

Chapter 5. Discussion and Projections

5.1. Introduction

In this chapter, we discuss possible origins for the discrepancy between the measured and model gate capacitance observed in Type B structures. We suggest that the combination of non-parabolicity, quantization and compressive biaxial strain effects in the thin channel layer increases the in-plane effective mass. We have modeled this effect by artificially increasing the in-plane effective mass value for the InAs channel and this mitigated the discrepancy. In this chapter, we also explore the implications of our study for the gate capacitance and sheet-carrier concentration of a future 10 nm III-V MOSFET.

5.2. Possible sources of discrepancy in type B structure

The agreement between model and experiments is worst in Type B structures. This discrepancy may come from experimental uncertainty but there are other possible sources too. C_G in Type B is mostly determined by C_{Q1}. In our analytical model, we have used the bulk effective mass for the InAs channel core to model C_{Q1} and C_{cent1}. As explained in Ch. 2, our model uses input from Nextnano to track the movement of the band structure with respect to the Fermi level as V_G changes. Nexnano correctly uses different effective masses in the channel core and in the cladding layers but these are also the bulk values. If the real effective masses in these layers, especially in the 5 nm InAs channel layer, are different from the bulk values, the capacitance predictions of our model will be affected.

Actually, the combination of biaxial strain that is present in the InAs channel with the strong quantization and the non-parabolicity of the conduction band in this material are all expected to increase the effective mass in the InAs channel layer [27, 30, 33]. In fact, estimates of the InAs effective mass in the InAs channel for Type B structures have been carried out by N. Kharche in Prof. Klimeck's group at Purdue University using OMEN_FET, a tool that can account for all these effects. They estimated an effective mass for InAs in this structure that is as high as 0.050 m_0 (m_0 = electron mass), whereas the bulk value is 0.026 m_0.

In the next subsections, we discuss the possible origins of discrepancy in the gate capacitance of the Type B structures. We also present modified model results using an increased effective mass in the section 5.3.

5.2.1. Experimental uncertainty

There can be measurement error in the geometrical parameters of the devices such as the barrier thickness (t_{ins}), channel thickness (t_{ch}) and gate length (L_G). Especially small deviations in the measurement of t_{ins} result in a relatively large variation in the overall gate capacitance since it directly affects the insulator capacitance component. t_{ins} is obtained from TEM measurements with an uncertainty of about ±0.5 nm. This could contribute somehow to resolve the discrepancy, but as shown in detail in section 5.3, it cannot account for all of it.

5.2.2. Channel quantization

The quantum well channel structure changes the continuous bulk band states of the semiconductor channel into separate 2D subbands. The quantization energy of each subband is inversely proportional to the quantum well channel thickness. Confinement from the quantum well structure allows electrons to move only on the quantum well plane. If this quantization energy is relatively small compared to the bandgap of the channel, the effective mass of the bulk material at the conduction band edge can be used to describe the motion of electrons. However, as the quantum well width becomes narrower, the parabolic dispersion model cannot be used anymore especially for small band gap materials and the effective mass begins to be strongly dependent on the quantization energy [28]. In a quantum well structure, we can define two different kinds of electron effective mass. The *in-plane effective mass* ($m_{\|}^*$) which is relevant for transport parallel to the layers and defines the density of states [34]. The other effective mass is the *perpendicular effective mass* or *quantization effective mass* ($m_{\perp}^*$) which determines the electron quantization energies. As explained in section 2.2.2, in our model, $m_{\|}^*$ is considered to be equal to $\sqrt{m_x^* \cdot m_y^*}$ where m_x^*, m_y^* and m_z^* are defined in 3D bulk material (quantum well plane is assumed to be in x-y plane). In particular, III-V materials such as InAs, GaAs, or InP used in our study have an identical effective mass into the three independent directions in a 3D bulk structure, so $m_x^* = m_y^* = m_z^* = m^*$ and $m_{\|}^*$ in a 2D quantum channel well is simply considered to be equal to m^* in a 3D bulk. However, because of non-parabolicity, in a few nanometer size quantum well structure, $m_{\|}^*$ cannot be simply expressed as $\sqrt{m_x^* \cdot m_y^*}$ and strongly depends on the quantization energy [35].

Figure 5-1 describes this phenomenon very well. It shows $m_{\|}^*$ of the lowest subband in a lattice matched $In_{0.53}Ga_{0.47}As$/InP quantum well as a function of well width and quantization energy [28]. The solid lines come from their effective mass model calculation and the symbols are experimental findings. For wide well widths, $m_{\|}^*$ nearly equals the bulk effective mass ($m^* = 0.044\ m_o$) of $In_{0.53}Ga_{0.47}As$ but as the well width decreases below 10 nm, $m_{\|}^*$ begins to sharply increase and approaches the InP bulk effective mass ($m^* = 0.080\ m_o$). The amount of electron quantization energy strongly affects the increase of $m_{\|}^*$ in a narrow quantum well. In a InAs/InP quantum well structure with a 5 nm well width, the InAs $m_{\|}^*$ is estimated to be increased by about a factor of two [36]. This is due to the quantization effect discussed here but also with an additional contribution from the biaxial compressive strain explained in section 5.2.4.

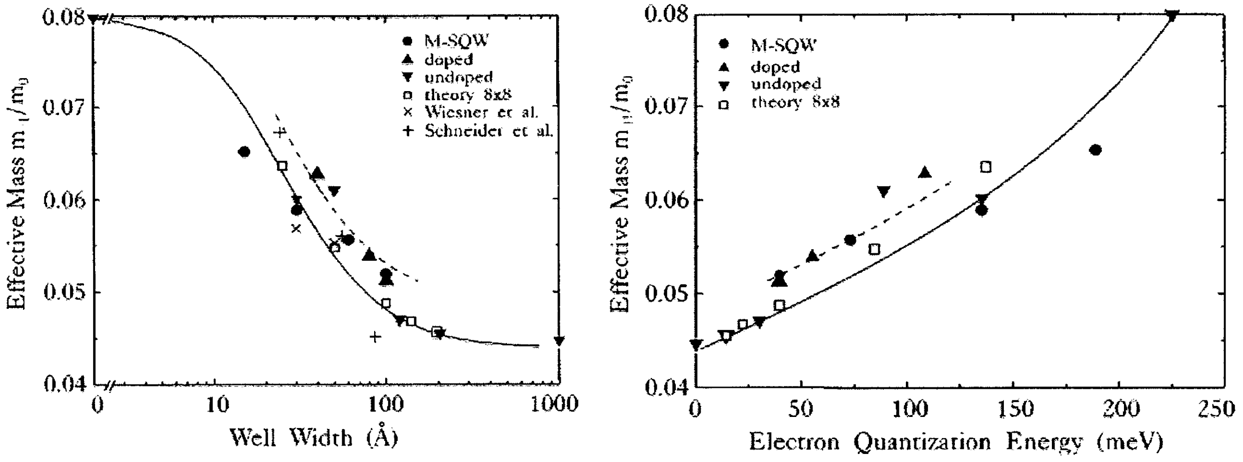

Figure 5-1. The in-plane effective mass ($m_{\parallel}$*) as a function of well width (a) and electron quantization energy (b) in $In_{0.53}Ga_{0.47}As/InP$ quantum well structures. Symbols represents reported experimental results and lines corresponds to calculated results from the model suggested by Wetzel [28]. [Figs. taken from Wetzel [28]].

5.2.3. Nonparabolicity of the conduction band

$m_{\parallel}^*$ in a 2D system is analytically defined like below.

$$\frac{1}{m_{\parallel}*} = \frac{1}{\hbar^2} \cdot \frac{d^2E}{dk_{\parallel}^2} \tag{1}$$

The electron kinetic energy at the bottom of the subband can be simply modeled as $\frac{\hbar^2 k_{\parallel}^2}{2m*}$ based on a spherical and parabolic conduction band model [32]. Then, according to equation (1), $m_{\parallel}^*$ in a 2D system becomes identical to the bulk effective mass. However, this assumption becomes less valid as the well width decreases. As the quantization energy increases by narrowing the quantum well, the subband levels locate far above the conduction band edge where the dispersion-relation shows non-parabolic characteristics. Therefore, under strong quantum confinement, the kinetic energy does not simply depend in a quadratic manner on $k_{\parallel}$ and nonparabolicity effects should be considered.

Figure 5-2 helps us to comprehend this effect. It shows the nonparabolicity of the E-k diagram in an $Al_xGa_{1-x}As$/GaAs quantum well. In a very narrow width quantum well, the curvature from the E-k diagram (dashed line) at the bottom of the subband shows some deviation from a purely parabolic curve (dotted line). This nonparabolicity property in the E-k curve is reflected in the increase of $m_{||}^*$ [31].

The nonparabolicity is also relevant to the band-filling effect. Higher sheet carrier concentration is expected to increase $m_{||}*$ and this increment is substantial [30]. Electrons fill up the E-k curve from the bottom of energy states and as the charge increases, more electrons must be located at higher energy states in the band. The curvature in this high energy region is smaller than at the bottom. Therefore, when a quantum well is filled with a large amount of electrons, $m_{||}^*$ increases.

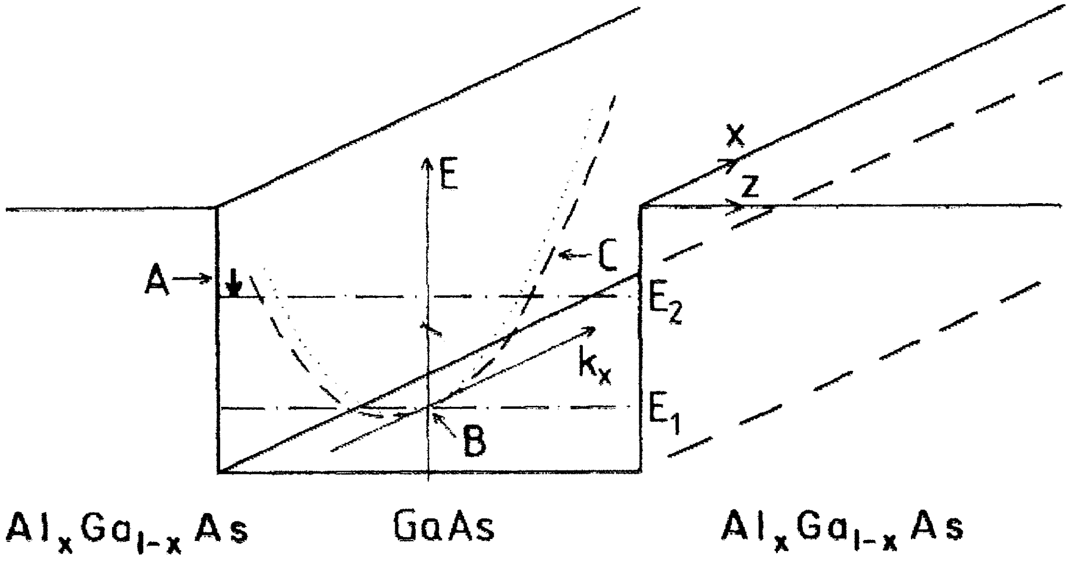

Figure 5-2. A sketch of nonparabolicity in E-k diagram inside a AlxGa1-xAs/GaAs quantum well. A indicates quantization energy shifts. B points out in-plane effective mass ($m_{||}*$) at the bottom of a subband energy level which is calculated from the deviated curvature (dashed line) compared with parabolic curvature (dotted line) at C. [Fig. taken from Ekenberg, [31]].

5.2.4. Biaxial channel strain

In our heterostructures, the channel material has a larger relaxed lattice constant than the InP substrate. As a result of this, the channel is under biaxial compressive strain [37]. This changes the crystal structure and the dispersion relation in the channel, which finally changes the effective mass [27, 38].

Figure 5-3 illustrates the changes of in-plane effective mass (marked as "par.") and perpendicular effective mass (marked as "perp.") with respect to biaxial strain for InGaAs with different GaAs mole fractions [27]. Symbols show results from k.p calculation and lines come from a model. This graph shows dramatic changes in the effective mass in response to applied biaxial strain. For example, according to this result, InAs channel (x=0) on an InP substrate yields about -0.031 biaxial strain and it brings up both of $m_{\parallel}^*$ and $m_{\perp}^*$ to more than $0.03m_0$.

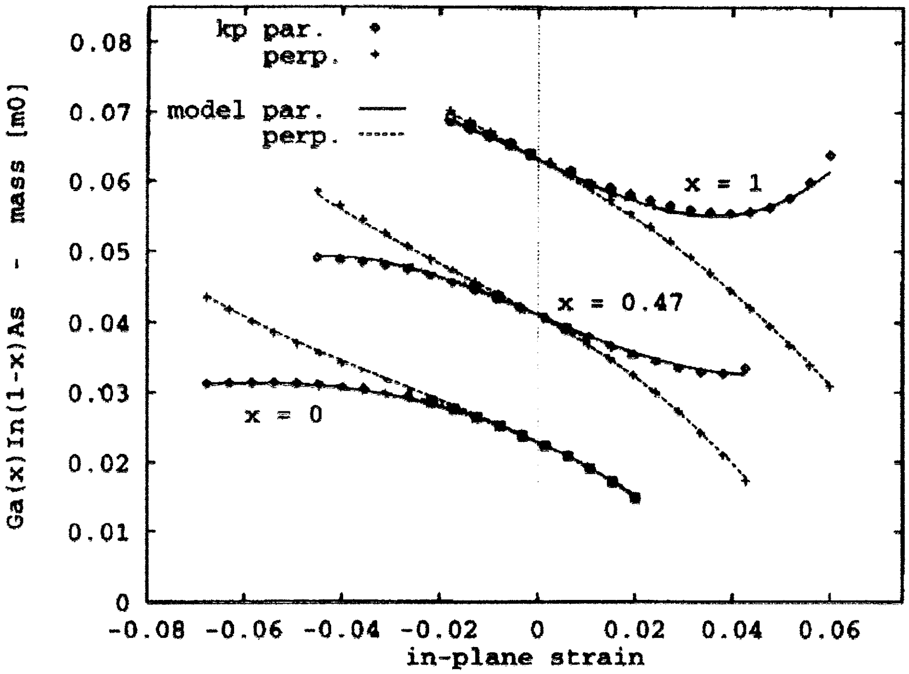

Figure 5-3. $In_xGa_{1-x}As$ in-plane (noted as par.) and perpendicular (noted as perp.) effective mass as a function of in-plane (biaxial) strain for different In mole fractions (x = 0, 0.47 and 1). [Fig. taken from Kopf [27]].

5.2.5. Role of channel cladding

In our devices, the channel structure consists of an InAs core layer and two lattice-matched cladding layers ($In_{0.53}Ga_{0.47}As$). Since the conduction band discontinuity between the core and cladding is not very high, electrons are not completely confined to the core channel layer. The electron wave function can spread into the cladding layers and result in a certain amount of electrons there. Figure 5-4 shows this phenomenon in the simulation results of a Type B structure. It shows that for $V_G - V_T = 0.22$ V, about 25% of the electrons reside in the cladding layer ($V_T = V_G$ at $N_S = 5 \times 10^{10}$ cm^{-2}). Since the cladding layer has a heavier effective mass (0.045 m_o) than the InAs core layer (0.026 m_o), an averaged $m_{||}^*$ (0.031 m_o) was used in our model which is slightly higher than the value of pure InAs. However, if the population of the core and cladding layer changes with V_G, then the averaged $m_{||}^*$ will change with V_G too. Therefore, we need to estimate the relative charge distribution inside the channel cladding layers for the entire operational V_G range and we need to verify that using a constant averaged $m_{||}^*$ for our model expression is a reasonable assumption. Furthermore, we also need to verify that most of charge stays in core channel layer when we change the InAs $m_{||}^*$ to account for the effects discussed in this chapter.

5.3. Modified model C_G results

We have explored the impact in our estimation of C_G from the uncertainty in t_{ins} and the impact of quantization, nonparabolicity and biaxial strain by artificially selecting a higher value of $m_{||}^*$ for the InAs layer. We have not done the same for the InGaAs cladding layers because we estimate the electron concentration there to be less than 25% of the total for the entire V_G range.

5.3.1. Channel cladding layer contribution to C_G

Nextnano just uses one bulk effective mass for each layer. It reads the effective mass information from a database.in file which has a list of most III-V material properties. However, we can artificially change the material parameter information in this database.in file. By changing the

$m_{\parallel}^{*}$ value for InAs, we can have a first-order examination of the impact of quantization, non-parabolicity and strain effects in C_G.

Before doing this, we first estimated how many electrons are located in the InAs core-channel layer for the entire V_G range. Figure 5-5 shows this for the Type B for two $m_{\parallel}^{*}$ values for InAs. Figure 5-5 (a) shows that the electron concentration in the core-layer is fairly constant with V_G and with a value of about 75% when we use the bulk InAs effective mass. As we increase $m_{\parallel}^{*}$ values of InAs up to 0.050 m_o (Figure 5-5 (b)) the proportion remains constant with V_G and it approaches 80%. This is because the density of states in the core-channel layer increases through increasing $m_{\parallel}^{*}$. We also tested different $m_{\parallel}^{*}$ values of InAs material around 0.050 m_o and found out that the electron concentration in the core-channel layer is always constant and in excess of 75%.

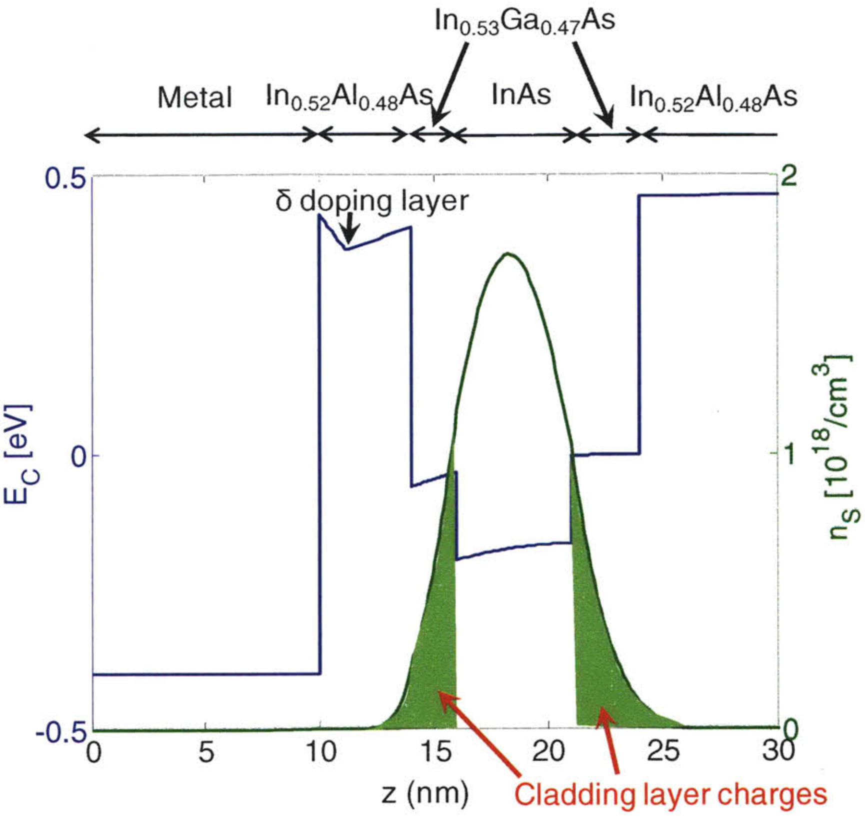

Figure 5-4. Conduction band (E_C) and electron density ($n_S(z)$) profile of Type B at $V_G - V_T = 0.22$ V. Green shadow area under the n_S curve indicates electrons in the cladding layers.

Therefore, we can conclude that most of charges are located in the core-channel layer and the proportion is constant regardless of bias. This confirms that the use of a constant averaged $m_{\|}^*$ for the entire V_G range is reasonable. In Nextnano simulation, the increase of InAs $m_{\|}^*$ up to 0.050 m_o also affects the cladding layer effective mass which increases from 0.045 m_o to 0.058 m_o. The averaged $m_{\|}^*$ value becomes 0.052 m_o. This value is nearly equal to the increased InAs $m_{\|}^*$. Thus, as InAs $m_{\|}^*$ increases, the electron concentration in the core layer becomes larger due to the increased DOS in the core layer and cladding layer contributes less. Therefore, we have explored the impact of changes of the InAs core channel $m_{\|}^*$ on C_G by artificially selecting a higher value of $m_{\|}^*$ for the InAs layer. The control of effective mass in the cladding layers is not precisely explored because of their small contribution to C_G.

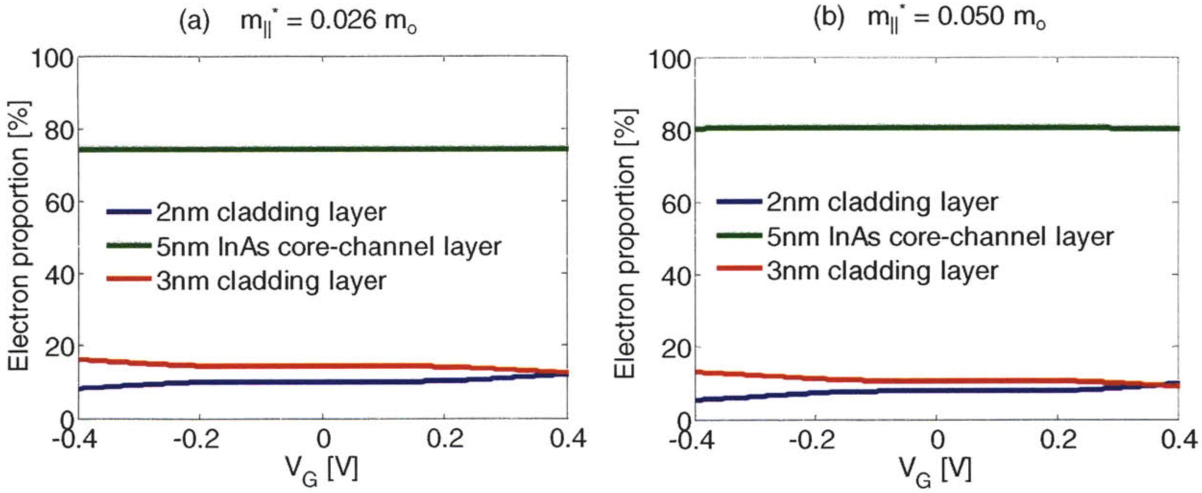

Figure 5-5. Electron concentration proportion [%] in each channel layer of Type B for InAs layer $m_{\|}^*$ = 0.026 m_o, 0.050 m_o.

5.3.2. Reduced discrepancy in Type B

Figure 5-6 shows that the agreement between experiments and the model for the InAs-channel devices improves when we increase $m_{\|}^*$ of InAs from bulk value (0.026m_e) to a value around 0.05 m_o with ±0.005m_o variation. This seems like a large increase but it is actually expected by theoretical and experimental studies of these effects [27, 30, 33], and simulation work by N.

Kharche at Purdue University. The uncertainty in t_{ins} is estimated to be about ±0.5 nm which translates in error bars in the calculated C_G of around ±0.6 fF/μm^2 at $V_G = 0.4$ V as shown in Figure 5-6. The change of InAs $m_{||}^*$ produces a larger impact on Type B than Type A structures, as shown in Figure 5-6. That is because the quantum capacitance is more dominant in Type B devices.

With the combination of thickness uncertainty and effective mass increase, the discrepancy between model and measurements is significantly reduced. However, there is still residual discrepancy in high V_G range. We suggests that this may come from band-filling effect [30]. As V_G increases, N_S keeps increasing and raising the $m_{||}^*$ value of InAs which ends up with higher C_G results. In fact, looking in detail at Figure 5-6, we see that the discrepancy between the measurements and the model increases as V_G increases, which is consistent with the band-filling effect. Since Nextnano can only handle a constant effective mass regardless of the applied voltage, this effect cannot be included in our modified model results.

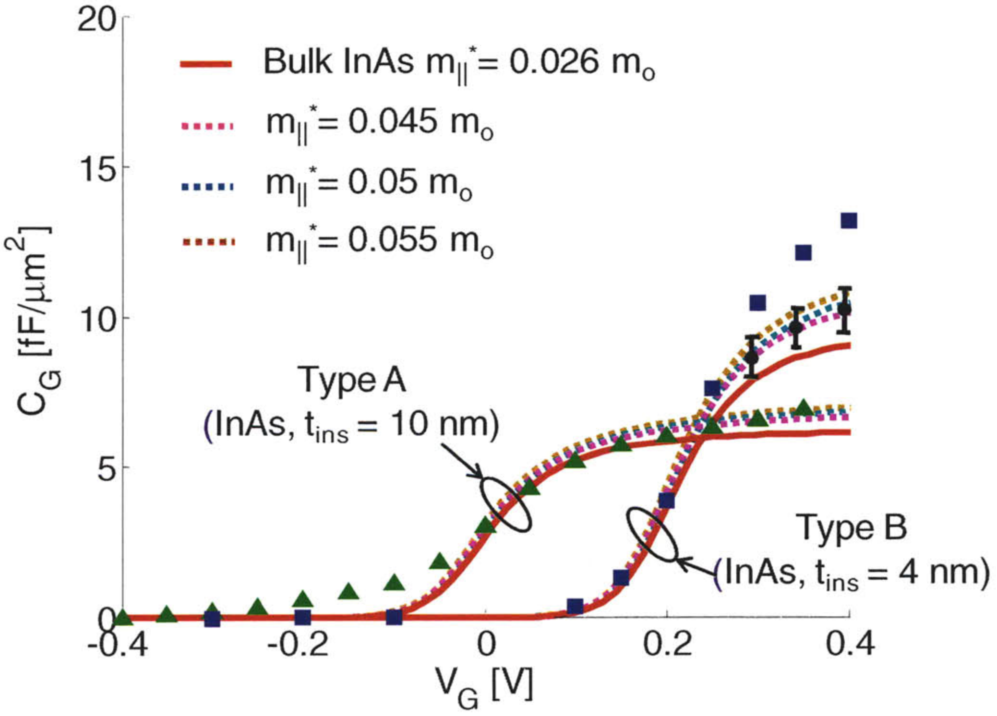

Figure 5-6. Gate capacitance as a function of gate voltage for the InAs-channel heterostructures (Type A and B). The dotted lines show calculations using heavier values of effective mass in the InAs layer. The error bars indicate the uncertainty in the calculated C_G that comes from an uncertainty of ±0.5 nm in t_{ins}.

5.4. Projection for 10 nm III-V MOSFET

Our gate capacitance model allows us to estimate C_G in future scaled down high-k dielectric III-V FETs. In future devices, the adoption of a high-k gate dielectric and the use of a very thin quantized channel with a low-effective mass material will establish the quantum capacitance of the first subband as the dominant term in C_G. Using our model, we can examine the implications of this.

First of all, we designed a 10 nm III-V MOSFET prototype by maintaining the electrostatic aspect ratio of current state-of-the-art 30 nm gate length HEMTs [14]. In these, t_{ins} is 4 nm and t_{ch} is 10 nm (the aspect ratio of t_{ch} to t_{ins} is 2.5). A 10 nm gate length device made out of the same materials with an identical electrostatic integrity should have t_{ins} = 1.3 nm and t_{ch} = 3 nm. If a high-k gate dielectric ($\varepsilon = 25\varepsilon_o$) is used as barrier material, t_{ins} can then be doubled. Figure 5-7 shows our 10 nm gate length III-V MOSFET prototype design which has 2.6 nm high-k ($\varepsilon = 25\varepsilon_o$) dielectric layer with t_{ins} = 2.6 nm and t_{ch} = 3 nm. There is a δ doping layer with 5×10^{12} cm^{-2} doping concentration under the channel layer. A low $m_{\parallel}^*$ III-V material will be used for the channel layer.

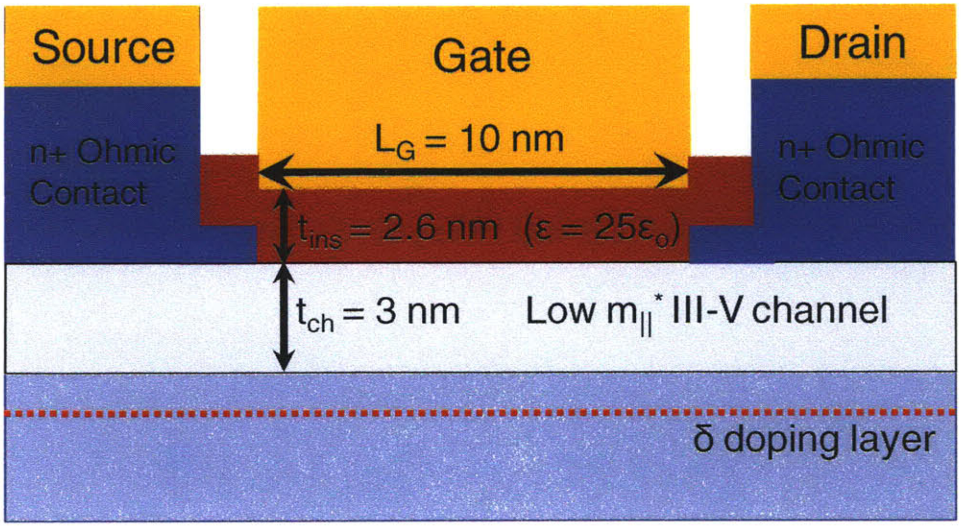

Figure 5-7. High-k dielectric 10 nm gate length III-V MOSFET prototype with t_{ch}= 3 nm and t_{ins} = 2.6 nm ($\varepsilon = 25\varepsilon_o$).

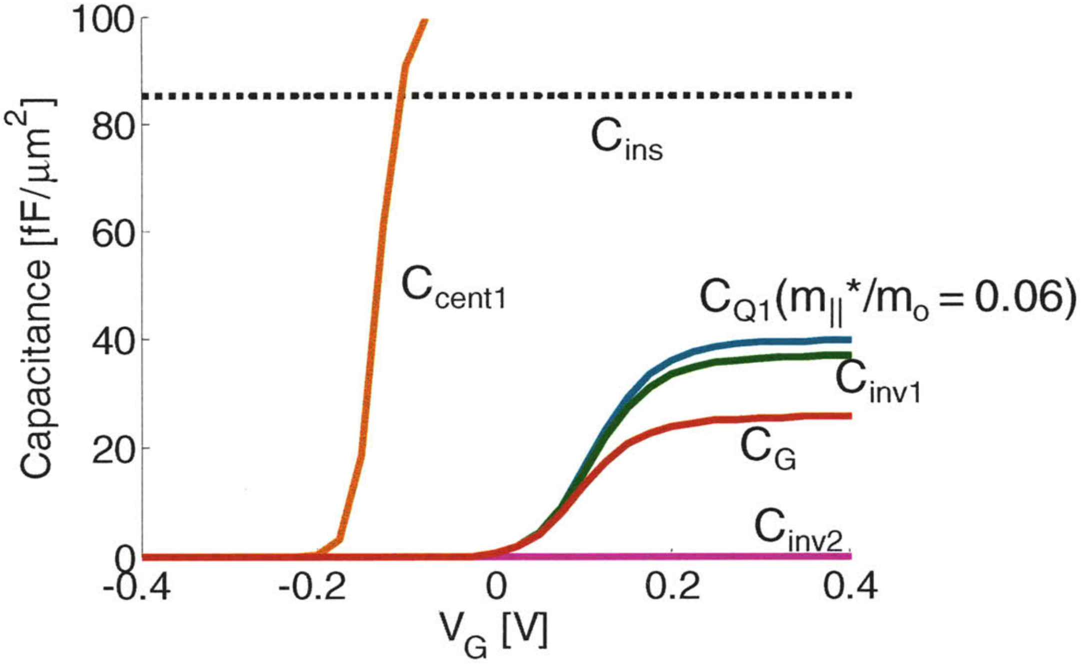

Figure 5-8. Modeled C_G and capacitance components in 10 nm III-V MOSFET prototype. 0.06 m_o is used for $m_{||}^*$ in the channel layer.

Next, we estimated how much C_G and N_S we can expect from this 10 nm III-V MOSFET structure in Figure 5-7 by using our model. Figure 5-8 shows our modeled C_G and capacitance component results in this prototype device. We used InAs as a channel layer material but changed $m_{||}^*$ to 0.06 m_o which is technically achievable by a 3 nm thin channel layer design [39]. Even though $m_{||}^*$ is much bigger than the bulk InAs m*, C_{ins} is more than double the value of C_{Q1} and C_{cent1} is much larger than C_{ins}. Thus, C_{inv1} and the overall C_G are mostly determined by C_{Q1}. C_{inv2} is negligible, which means that there is no 2nd subband charge contribution to C_G. This result indicates that in this deeply scaled down device, C_G is mostly controlled by C_{Q1} and the $m_{||}^*$ value in the channel layer.

Figure 5-9 shows the change of the modeled C_G curves in the 10 nm III-V MOSFET prototype structure for different $m_{||}^*$ values in the channel layer. Due to the dominance of C_{Q1} on C_G, the change of $m_{||}^*$ greatly affects the overall C_G. This eventually results in a high increase of N_S.

Figure 5-10 shows N_S as a function of V_G for different channel $m_{\|}^*$ values. V_T is defined as V_G at $N_S = 5 \times 10^{10}\ \mathrm{cm}^{-2}$. For $V_{DD} = 0.5$ V, we set V_T as $V_{DD}/3$ and estimated N_S for a gate overdrive of $V_{DD}-V_T = 0.33$ V. In this condition, N_S approaches the mid $10^{12}\ \mathrm{cm}^{-2}$ range (For example, $N_S = 4.1 \times 10^{12}\ \mathrm{cm}^{-2}$ for $m_{\|}^* = 0.05\ m_o$ and $4.6 \times 10^{12}\ \mathrm{cm}^{-2}$ for $m_{\|}^* = 0.06\ m_o$). Therefore, it is clear that in future 10 nm gate length III-V MOSFETs, an enhancement of channel effective mass is essential to attain N_S in the mid $10^{12}\ \mathrm{cm}^{-2}$ range. This seems eminently feasible through non-parabolicity, the strong quantum confinement expected from a very thin channel and by proper engineering of in-grown biaxial strain.

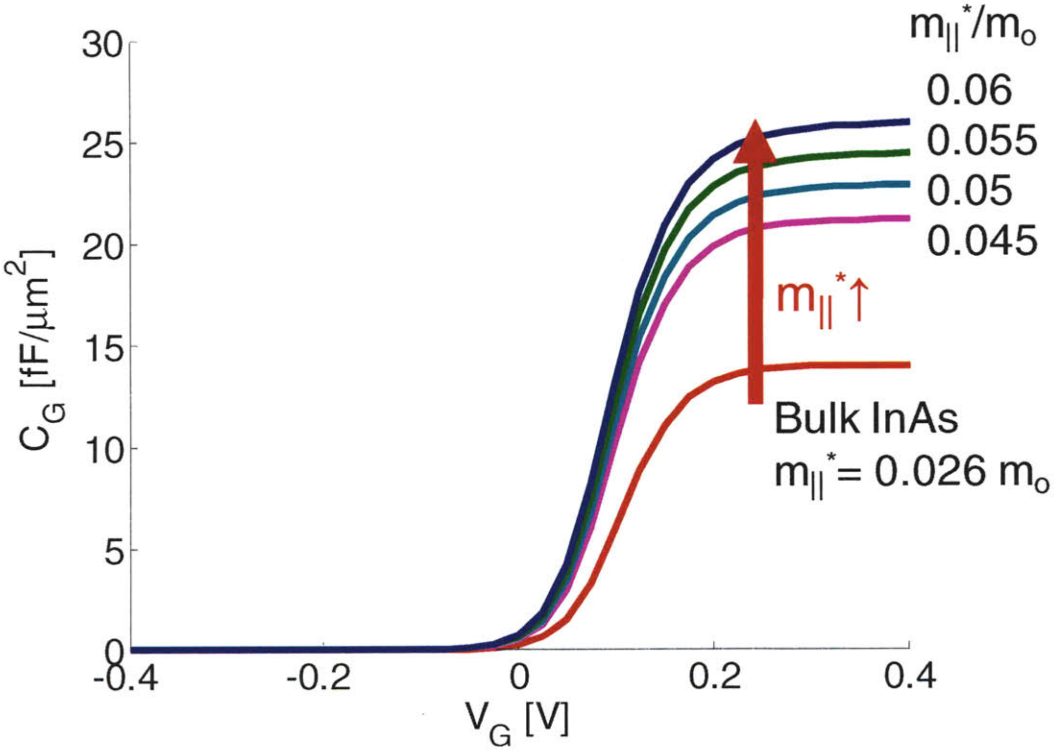

Figure 5-9. Gate capacitance of 10 nm III-V MOSFET prototype for different $m_{\|}^*$ values in the channel layer.

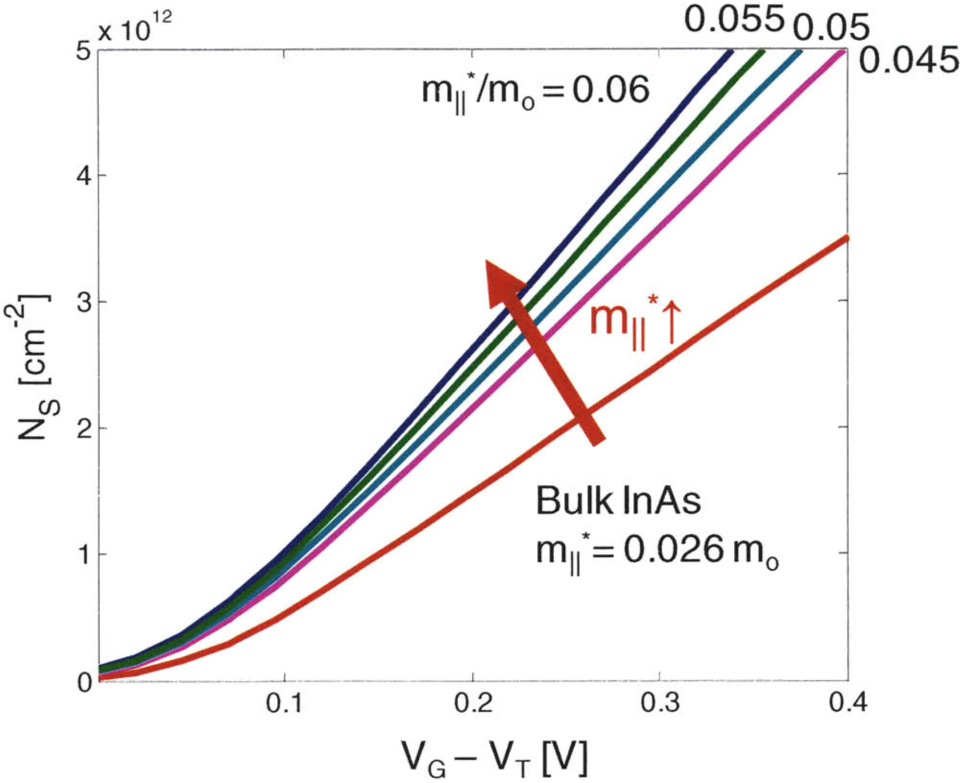

Figure 5-10. Sheet carrier concentration (N_S) of 10 nm III-V MOSFET prototype for different $m_{\|}$* values in the channel layer. V_T is defined as V_G at $N_S = 5 \times 10^{10}$ cm^{-2}.

5.5. Summary

We have discussed possible physical origins for the discrepancy observed in the measured and modeled gate capacitance of Type B structures. We suggest that the combination of non-parabolicity, quantization and strain effects contribute to an increase in the in-plane effective mass in a thin quantum well channel. This phenomenon is particularly relevant in deeply scaled down devices because of the dominance of the quantum capacitance in the overall gate capacitance. This effect can be used to solve the DOS bottleneck in future 10 nm gate length III-V MOSFETs by a thin channel layer design and by proper engineering of in-grown biaxial strain in the channel layer. This in-plane effective mass increase is estimated to be enough to provide the required sheet carrier concentration around mid 10^{12} cm^{-2} range for high performance logic operation at $V_{DD} = 0.5$ V.

Chapter 6. Conclusions and Suggestions

In this thesis, we have developed a simple quantitative model for the gate capacitance in III-V FETs that includes the density of states effect. We analyzed experimental gate capacitance on scaled down InGaAs and InAs HEMTs with different designs. Our model suggests that quantum capacitance will dominate in future scaled III-V MOSFETs. In this chapter, key findings of this gate capacitance analysis work are briefly summarized and suggestions for future work are provided.

6.1. Conclusions

In this work, we have built a physical gate capacitance model for III-V FETs that incorporates quantum capacitance and centroid capacitance in the channel. The inversion-layer capacitance is defined as a parallel combination of the inversion layer capacitance related to each occupied electron subband. These in turn are modeled as a series of the quantum capacitance and the centroid capacitance. The quantum capacitance originates in the penetration of the Fermi level inside the 2D subbands of a quantum well due to the finite density of states. Our model includes this density of states effect represented by the in-plane effective mass of the channel material. The centroid capacitance is related to the shape of the charge distribution in the inversion layer. In low effective mass III-V channels, both capacitance components can be relatively small. Our model is verified by a 1-D Poisson-Schrodinger Solver (Nextnano) simulation tool. Three model heterostructures with different barrier and channel thicknesses are tested for this, and our model shows very well-matched sheet carrier concentration and overall gate capacitance results to the Nextnano simulations.

The experimental InAs and InGaAs HEMTs studied in this work have three different types of device structures which include different (4 or 10 nm) barrier thickness and (10 or 13 nm) channel thickness. We extracted the gate capacitance of these scaled down HEMTs from S-parameter measurements. For each type, devices with different gate lengths ranging from 200 nm to 30 nm are fabricated and S-parameters are measured in the frequency range from 1 to 40 GHz. The intrinsic and parasitic capacitances for our experimental HEMT structures are modeled and determined from the S-parameter measurements.

We compared these experimental intrinsic gate capacitances with our physical model and showed very good agreement. Based on this match, the decomposition of the total gate capacitance into each gate capacitance model component allowed us to obtain physical insight into the key dependencies of the quantum capacitance and the centroid capacitance in scaled-down devices.

In all three cases studied, we observed evident degradation of the overall gate capacitance that comes from the finite inversion-layer capacitance. The measured gate capacitance in strong inversion is only between 35% and 62% of the insulator capacitance. The finite inversion-layer capacitance mostly results from a relatively small quantum capacitance. In addition, we also found that the 1st subband dominates the overall gate capacitance in the operational range of scaled down HEMTs. 10 or 13 nm thin and low effective mass channel layers produce strong channel quantization and allow for a very small 2nd subband contribution to the overall electron population in the channel.

Another conclusion is that the centroid capacitance is also a highly relevant component in the overall gate capacitance. A 10 nm channel thickness InAs device is observed to have a significantly larger centroid capacitance than a 13 nm channel thickness InGaAs device. This more than compensates for the lower quantum capacitance of the InAs channel and ends up with a higher overall value of gate capacitance.

The results obtained in this thesis help us to predict that in future scaled down devices with a very thin channel, an increase of the inversion-layer capacitance will not be possible by populating higher subbands in the channel but perhaps by increasing the quantum capacitance and the centroid capacitance of the 1st subband.

The discrepancy between the measured and model gate capacitance observed in InAs thin channel HEMT devices can be considered as evidence of increased in-plane effective mass in the channel layer. In the Nextnano simulations and in our model, we are limited to the use of spherical bands with a single well defined in-plane effective mass for each material in the channel which we have set equal to the bulk effective mass of the core channel. However, the in-plane effective mass can be greatly increased by band non-parabolicity, strong channel quantization, and compressive biaxial strain due to the lattice mismatch between channel and substrate. A modified physical model using a higher value of in-plane effective mass shows better results.

This suggests a solution to the worrying quantum capacitance limit from the density of states bottleneck in future III-V MOSFETs. The adoption of a high-k dielectric layer and a thinner channel layer design will make the insulator capacitance and the centroid capacitance much bigger and establish the quantum capacitance as the dominant component of the overall gate capacitance. However, the quantum capacitance can be greatly enhanced by proper engineering of in-grown biaxial strain and the combination of non-parabolicity and quantization effects in the thin channel layer design. Our model predicts that in future 10 nm III-V MOSFETs a sheet carrier concentration in the mid 10^{12} cm^{-2} range seems eminently feasible at $V_{DD} = 0.5$ V.

6.2. Suggestions for future work

We suggest a number of future studies that can continue the gate capacitance research presented here.

- Experimental and theoretical study of gate capacitance in InAs HEMTs with a thinner channel layer design below 10 nm. The in-plane effective mass of the InAs channel layer is expected to increase more under the 5 nm channel thickness.

- Develop a suitable high-k dielectric layer in III-V FETs and perform gate capacitance analysis. The interface-state density should be carefully considered as an additional capacitance component.
- Use a simulation tool beyond Nextnano which has an ability to accurately calculate band non-parabolicity, quantization and biaxial strain effects in III-V compound channel heterostructures. The band filling effect should be also explored through this simulator. A proper effective mass model including all these effects needs to be proposed.
- Carry out a study to establish an analytical expression which links the centroid capacitance expression to the averaged distance of charge distribution in the III-V quantum well channel structure.
- Investigate other logic parameters such as subthreshold slope and transconductance based on the gate capacitance model in the linear and saturation regimes with different temperature conditions.

Bibliography

[1] T. Skotnicki, J. A. Hutchby, K. Tsu-Jae, H. S. P. Wong, and F. Boeuf, "The end of CMOS scaling: toward the introduction of new materials and structural changes to improve MOSFET performance," *Circuits and Devices Magazine, IEEE,* vol. 21, pp. 16-26, 2005.

[2] R. Chau, S. Datta, and A. Majumdar, "Opportunities and challenges of III-V nanoelectronics for future high-speed, low-power logic applications," in *Compound Semiconductor Integrated Circuit Symposium, 2005. CSIC '05. IEEE*, 2005, p. 4 pp.

[3] D. Cheng, C. Liao, K. U. Cheng, and M. Feng, "Process Development and Characteristics of Nano III-V MOSFET," in *CS MANTECH*, 2008, p. 4 pp.

[4] M. Passlack, R. Droopad, K. Rajagopalan, J. Abrokwah, P. Zurcher, and P. Fejes, "High Mobility III-V Mosfet Technology," in *Compound Semiconductor Integrated Circuit Symposium, 2006. CSIC 2006. IEEE*, 2006, pp. 39-42.

[5] Y. Sun, S. J. Koester, E. W. Kiewra, J. P. de Souza, N. Ruiz, J. J. Bucchignano, A. Callegari, K. E. Rogel, and R. Germann, "Post-Si CMOS: III-V n-MOSFETs with High-k Gate Dielectrics," in *Compound Semiconductor Integrated Circuit Symposium*, May 2007.

[6] K. M. Lau, T. Chak Wah, L. Haiou, and Z. Zhenyu, "AlInAs/GaInAs mHEMTs on silicon substrates grown By MOCVD," in *Electron Devices Meeting, 2008. IEDM 2008. IEEE International*, 2008, pp. 1-4.

[7] M. V. Fischetti, T. P. O'Regan, N. Sudarshan, C. Sachs, J. Seonghoon, K. Jiseok, and Z. Yan, "Theoretical Study of Some Physical Aspects of Electronic Transport in nMOSFETs at the 10-nm Gate-Length," *Electron Devices, IEEE Transactions on,* vol. 54, pp. 2116-2136, 2007.

[8] M. V. Fischetti, L. Wangt, B. Yut, C. Sachs, P. M. Asbeckt, Y. Taurt, and M. Rodwell, "Simulation of Electron Transport in High-Mobility MOSFETs: Density of States Bottleneck and Source Starvation," in *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, 2007, pp. 109-112.

[9] S. Takagi and A. Toriumi, "Quantitative understanding of inversion-layer capacitance in Si MOSFET's," *Electron Devices, IEEE Transactions on,* vol. 42, pp. 2125-2130, 1995.

[10] H. S. Pal, K. D. Cantley, S. S. Ahmed, and M. S. Lundstrom, "Influence of Bandstructure and Channel Structure on the Inversion Layer Capacitance of Silicon and GaAs MOSFETs," *Electron Devices, IEEE Transactions on,* vol. 55, pp. 904-908, 2008.

[11] Y. Bo, W. Lingquan, Y. Yu, P. M. Asbeck, and T. Yuan, "Scaling of Nanowire Transistors," *Electron Devices, IEEE Transactions on,* vol. 55, pp. 2846-2858, 2008.

[12] J. Genoe, C. Van Hoof, W. Van Roy, J. H. Smet, K. Fobelets, R. P. Mertens, and G. Borghs, "Capacitances in double-barrier tunneling structures," *Electron Devices, IEEE Transactions on,* vol. 38, pp. 2006-2012, 1991.

[13] J. A. Lopez-Villanueva, P. Cartujo-Casinello, J. Banqueri, F. Gamiz, and S. Rodriguez, "Effects of the inversion layer centroid on MOSFET behavior," *Electron Devices, IEEE Transactions on,* vol. 44, pp. 1915-1922, 1997.

[14] D.-H. Kim and J. A. del Alamo, "30 nm E-mode InAs PHEMTs for THz and future logic applications," in *Electron Devices Meeting, 2008. IEDM 2008. IEEE International*, 2008, pp. 1-4.

[15] D.-H. Kim and J. A. del Alamo, "Scaling Behavior of In0.7Ga0.3As HEMTs for Logic," in *Electron Devices Meeting, 2006. IEDM '06. International*, 2006, pp. 1-4.

[16] D.-H. Kim and J. A. del Alamo, "30-nm InAs Pseudomorphic HEMTs on an InP Substrate With a Current-Gain Cutoff Frequency of 628 GHz," *Electron Device Letters, IEEE,* vol. 29, pp. 830-833, 2008.

[17] D.-H. Kim and J. A. del Alamo, "Logic Performance of 40 nm InAs HEMTs," in *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, 2007, pp. 629-632.

[18] B. Afzal, A. Zahabi, A. Amirabadi, Y. Koolivand, A. Afzali-Kusha, and M. E. Nokali, "Analytical model for C-V characteristic of fully depleted SOI-MOS capacitors," *Solid-State Electronics,* vol. 49, pp. 1262-1273, 2005.

[19] S. Luryi, "Quantum capacitance devices," *Applied Physics Letters,* vol. 52, pp. 501-503, 1988.

[20] C. R. Selvakumar, "Approximations to Fermi-Dirac integrals and their use in device analysis," *Proceedings of the IEEE,* vol. 70, pp. 516-518, 1982.

[21] T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Reviews of Modern Physics,* vol. 54, p. 437, 1982.

[22] F. Stern and W. E. Howard, "Properties of Semiconductor Surface Inversion Layers in the Electric Quantum Limit," *Physical Review,* vol. 163, p. 816, 1967.

[23] G. Dambrine, A. Cappy, F. Heliodore, and E. Playez, "A new method for determining the FET small-signal equivalent circuit," *Microwave Theory and Techniques, IEEE Transactions on,* vol. 36, pp. 1151-1159, 1988.

[24] D. Lovelace, J. Costa, and N. Camilleri, "Extracting small-signal model parameters of silicon MOSFET transistors," in *Microwave Symposium Digest, 1994., IEEE MTT-S International*, 1994, pp. 865-868 vol.2.

[25] L. P. Sadwick, C. W. Kim, K. L. Tan, and D. C. Streit, "Schottky barrier heights of n-type and p-type Al0.48In0.52As," *Electron Device Letters, IEEE,* vol. 12, pp. 626-628, 1991.

[26] K. J. Chen, T. Enoki, K. Maezawa, K. Arai, and M. Yamamoto, "High-performance InP-based enhancement-mode HEMTs using non-alloyed ohmic contacts and Pt-based buried-gate technologies," *Electron Devices, IEEE Transactions on,* vol. 43, pp. 252-257, 1996.

[27] C. Kopf, H. Kosina, and S. Selberherr, "Physical models for strained and relaxed GaInAs alloys: Band structure and low-field transport," *Solid-State Electronics,* vol. 41, pp. 1139-1152, 1997.

[28] C. Wetzel, R. Winkler, M. Drechsler, B. K. Meyer, J. Scriba, J. P. Kotthaus, and F. Scholz, "Electron effective mass and nonparabolicity in Ga0.47In0.53As/InP quantum wells," *Physical Review B,* vol. 53, p. 1038, 1996.

[29] C. Wetzel, A. L. Efros, A. Moll, B. K. Meyer, P. Omling, and P. Sobkowicz, "Dependence on quantum confinement of the in-plane effective mass in Ga0.47In0.53As/InP quantum wells," *Physical Review B,* vol. 45, p. 14052, 1992.

[30] U. Wiesner, J. Pillath, W. Bauhofer, A. Kohl, A. M. Kusters, S. Brittner, and K. Heime, "In-plane effective masses and quantum scattering times of electrons in narrow modulation-doped InGaAs/InP quantum wells," *Applied Physics Letters,* vol. 64, pp. 2520-2522, 1994.

[31] U. Ekenberg, "Nonparabolicity effects in a quantum well: Sublevel shift, parallel mass, and Landau levels," *Physical Review B,* vol. 40, p. 7714, 1989.

[32] U. Ekenberg, "Enhancement of nonparabolicity effects in a quantum well," *Physical Review B,* vol. 36, p. 6152, 1987.
[33] M. Luisier, N. Neophytou, N. Kharche, and G. Klimeck, "Full-band and atomistic simulation of realistic 40 nm InAs HEMT," in *Electron Devices Meeting, 2008. IEDM 2008. IEEE International*, 2008, pp. 1-4.
[34] R. Winkler and U. R?sler, "General approach to the envelope-function approximation based on a quadrature method," *Physical Review B,* vol. 48, p. 8918, 1993.
[35] G. Hendorfer, M. Seto, H. Ruckser, W. Jantsch, M. Helm, G. Brunthaler, W. Jost, H. Obloh, and D. J. As, "Enhancement of the in-plane effective mass of electrons in modulation-doped In(x)Ga(1-x)As quantum wells due to confinement effects," *Physical Review B,* vol. 48, p. 2328, 1993.
[36] B. R. Nag and S. Mukhopadhyay, "In-plane effective mass in narrow quantum wells of nonparabolic semiconductors," *Applied Physics Letters,* vol. 62, pp. 2416-2418, 1993.
[37] R. People, "Effects of coherency strain on the band gap of pseudomorphic In(x)Ga(1 - x)As on (001) InP," *Applied Physics Letters,* vol. 50, pp. 1604-1606, 1987.
[38] R. People and S. K. Sputz, "Band nonparabolicities in lattice-mismatch-strained bulk semiconductor layers," *Physical Review B,* vol. 41, p. 8431, 1990.
[39] "Private discussions with N.Kharche at Purdue University."