A Compact Transport and Charge Model for GaN-based High Electron Mobility Transistors for RF applications

by

Ujwal Radhakrishna

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

© Massachusetts Institute of Technology 2013. All rights reserved.

Author ....
Department of Electrical Engineering and Computer Science
May 22, 2013

Certified by ....
Dimitri A. Antoniadis
Ray and Maria Stata Professor of Electrical Engineering
Thesis Supervisor

Accepted by ....
Leslie Kolodziejski
Chairman, Department Committee on Graduate Theses

A Compact Transport and Charge Model for GaN-based High Electron Mobility Transistors for RF applications

by

Ujwal Radhakrishna

Submitted to the Department of Electrical Engineering and Computer Science
on May 22, 2013, in partial fulfillment of the
requirements for the degree of
Master of Science in Computer Science and Engineering

Abstract

Gallium Nitride (GaN)-based high electron mobility transistors (HEMTs) are rapidly emerging as front-runners in high-power mm-wave circuit applications. For circuit design with current devices and to allow sensible future performance projections from device engineering in such a rapidly evolving technology, compact device models are essential.

In this thesis, a physics-based compact model is developed for short channel GaN HEMTs. The model is based on the concept of virtual source (VS) transport originally developed for scaled silicon field effect transistors. Self-consistent current and charge expressions in the model require very few parameters. The parameters have straightforward physical meanings and can be extracted through independant measurements. The model is implemented in Verilog-A and is compatible with state of the art circuit simulators. The new model is calibrated and validated with experimental DC I-V and S-parameter measurements of fabricated devices. Using the model, a projection of cut-off frequency (f_T) of GaN-based HEMTs with scaling is performed to highlight performance bottlenecks.

Thesis Supervisor: Dimitri A. Antoniadis
Title: Ray and Maria Stata Professor of Electrical Engineering

Acknowledgments

मनश्चेन्न लग्नं गुरोरङ्घ्रिपद्मे ततः किं ततः किं ततः किं ततः किम्

I would like to express my gratitude to my guru Prof. Antoniadis. His remarkable insight in the field of semiconductors and many other things intrigue me every time. Hopefully one day many decades later, I will be able to achieve some degree of that expertise. I would like to express my sincere thanks to Prof. Palacios, whose efforts in GaN technology at MIT is the reason I could do research in this area. Meetings with him always help me see the bigger picture. I would also like to thank Dr. Lan Wei, the amazing post-doc in our group who made my first year here at MIT seem too easy. Her take on this subject made me love device physics even more.

Compact modeling of devices is meaningless without validation against device simulations or device measurements. I would like to thank Dong Seup Lee for fabricating for us, the best behaved state of the art devices. It was a pleasure modeling them. I would also like to thank Han Wang for providing the base TCAD simulation deck. I am very grateful to Jamie and Omair for helping me with measurements in our labs. I also thank Winston for the long and interesting conversations we have on various GaN related topics.

Semiconductor physics would have been just another topic of my interest, if my undergraduate mentors at IIT-Madras, Profs. Amitava Dasgupta, Nandita Dasgupta and Anjan Chakravorty had not guided me. I will be always grateful for the foundation that they gave me in this area.

I am very grateful for all my friends here at MIT and outside. Thank you Jaffas. Finally, I would like to deeply thank my appa, amma and tamma for always being there for me and for supporting my seemingly irrational decision of pursuing graduate study.

ಕಲಿತದ್ದು ಎಳ್ಳಷ್ಟು ಕಲಿಯಬೇಕಾದ್ದು ಬೆಟ್ಟದಷ್ಟು

Contents

1 Introduction **15**

1.1 Gallium Nitride technology: A brief history 15

1.2 GaN: Material properties and applications 16

1.3 Thesis outline and goals 19

2 GaN HEMT: Operating principle **21**

2.1 Basics of device operation 21

2.1.1 Generic HEMT structure 21

2.1.2 GaN HEMT operation : Charge 22

2.1.3 GaN HEMT operation : Transport 23

2.1.4 Physical effects unique to GaN 25

2.2 DC and RF figures of merit 27

2.2.1 DC parameters 27

2.2.2 RF parameters 28

2.3 Device under study 31

3 GaN compact models: Overview **33**

3.1 Angelov model 33

3.2 Curtice model 36

3.3 EEHEMT model 37

3.4 Comparison of GaN models 40

4 The MVSG-RF model **42**

4.1 Intrinsic transistor model . . . 43
4.1.1 Saturation regime (High V_{DSi}) . . . 44
4.1.2 Non-saturation regime . . . 46
4.1.3 GaN specific effects . . . 47
4.2 Access regions . . . 48
4.2.1 Long channel access transistor model . . . 50
4.2.2 The implicit-gate . . . 52
4.3 Gate diode model . . . 54
4.4 Channel charge model . . . 55

5 Experimental verification of MVS-G-RF model 57
5.1 Validation of access region model . . . 57
5.2 Validation of MVS-G-RF DC model . . . 59
5.3 Validation of charge model . . . 62
5.3.1 S-parameter measurements . . . 62
5.3.2 Extraction of fringing capacitances . . . 63
5.3.3 Gate capacitance: model vs. measurements . . . 65
5.4 Unity current gain frequency (f_T) . . . 66
5.5 Technology projections: g_m and f_T . . . 68
5.5.1 g_m projection . . . 68
5.5.2 f_T projection . . . 69

6 Summary and future work 71

A Model parameter extraction 73
A.1 Device parameters . . . 74
A.2 C_{inv} extraction . . . 74
A.3 Extraction of V_{To}, S, n_d and δ . . . 75
A.4 Extraction of v_{inj}, β and θ_v, θ_μ, η_v and η_μ . . . 76

B Model parameters 79
B.1 Model parameters for TLMs . . . 79

B.2 Model parameters for GaN HEMTs 80

List of Figures

1-1 Number of publications (based on INSPEC search) and milestones of GaN research [1]. GaN research in both optoelectronics and electronics progressed rapidly after demonstrating the growth of high quality GaN and conductivity control of both p- and n-type GaN in the early 1990s. 16

1-2 Table showing material properties. GaN shows a combination of high electron mobility, electron velocity and breakdown field [2] 17

1-3 Projection of the role of GaN on Si platform in IR's market [3]. The FoM ('efficiency × density/ cost') is sustained in future through GaN based technologies. 18

1-4 R_{on} vs. BV tradeoff for Si, GaN and SiC. For a given BV GaN shows significantly lower R_{on} [3] 19

2-1 A rough schematic of a typical GaAs based HEMT heterostructure . 22

2-2 (a) Device structure schematic (b) Band diagram of heterostructure showing different types of charges 23

2-3 Velocity-field profile of GaN HEMTs; MC simulation and measurements 24

2-4 Velocity-field profile assumed in MVS-G-RF model 25

2-5 Structure schematic showing different regions of interest (a) red-intrinsic transistor (b) green- access regions 26

2-6 Schematic showing the reason for dynamic R_{on} effects. Electrons from the gate compensate the positive donor states creating depletion region in the 2DEG at drain edge of the gate thereby increasing R_{on}. 27

2-7 Simplified circuit for RF FoM extraction 29

2-8 BV vs. f_T trade off for different technologies as given [4]. GaN based devices have the highest BV with medium to high f_T. 30

2-9 Schematic of the device structure fabricated by Prof. Palacios' group [5] 31

3-1 Equivalent circuit of Angelov model [6] 34

3-2 Typical g_m characteristics for different material systems (after,[6]) . . 34

3-3 Equivalent circuit of the original version of Curtice model (after [7]) . 36

3-4 Different regions of EEHEMT model based on g_m [8] 38

3-5 Equivalent circuit of EEHEMT model[8] 40

3-6 Comparison of available GaN compact models[9] 41

4-1 Subcircuit model of MVS-G-RF model showing implicit-gate access regions and Schottky-gate diodes along with the intrinsic transistor. . 42

4-2 Band profile of Intrinsic transistor in saturation under drain bias. . . 43

4-3 An example for VS-point charge Q_{ixo} described by the model showing simplified expressions in strong and weak inversion. 44

4-4 An example for F_{sat} function showing its asymptotic forms in linear and saturation regimes. 46

4-5 Access regions modeled as implicit-gate transistors. The dashed-lines in the sub-circuit elements denote the absence of actual physical gate terminal in these regions. 49

5-1 Comparison of current-voltage characteristics of TLMs: Model vs. measurements. Parameters except L were left unchanged in the implicit-gate transistor model. 58

5-2 Comparison of current-voltage characteristics of 105 nm gate length GaN HEMT: Model vs. measurements. 59

5-3 Comparison of current-voltage characteristics of 42 nm gate length GaN HEMT: Model vs. measurements. 60

5-4 Comparison of current-voltage characteristics of 42 nm gate length GaN HEMT: Model with linear access resistors vs. measurements. Model mismatch illustrates the significance of access regions on device behavior. 61
5-5 Equivalent circuit assumed for gate capacitance extraction from S-parameters (after [10]). 62
5-6 Schematic showing different capacitances in on and off states. The channel screens C_{if} in on-state. 63
5-7 C_{if} and C_{of} extraction procedure from gate capacitance measurements. 64
5-8 Gate capacitance comparison for 105-nm and 42-nm devices: model vs. measurements. 65
5-9 Extraction of f_T from S-parameter measurements. 66
5-10 f_T values with scaling of gate length: model vs. device measurements. Good match is achieved for both lot1 and lot2 devices. 67
5-11 Technology projections on g_m. The projections show significant improvement in peak g_m with access region scaling and reduction of R_c. 68
5-12 Technology projections on f_T. A significant boost in f_T is possible through technology advances such as reduction of SCE and access region scaling. 69

A-1 Flowchart showing parameter extraction sequence of important MVS-G-RF model parameters. 73
A-2 Extraction of C_{inv} from gate capacitance vs. gate length measurements. 75
A-3 Extraction of V_{to}, S and δ from subthreshold regime of transfer curves. 76
A-4 Extraction of v_{inj}, β , η_v and η_μ from output characteristics. 77
A-5 Extraction of gate leakage model parameters from gate current characteristics. 78

Chapter 1

Introduction

1.1 Gallium Nitride technology: A brief history

The first Gallium Nitride (GaN) material was produced by passing ammonia over hot Gallium by Jusa and Hahn in 1938. Since then numerous research breakthroughs and progress have been made in GaN based electronics [11]. Originally, nitride semiconductors were considered as suitable candidates for optoelectronics due to their unique properties such as direct tunable bandgap from 6.2eV (AlN) to 0.7eV (InN), piezoelectricity and polarization and so on. Large area GaN was grown by hydride vapor phase epitaxy (HVPE) directly on sapphire by Maruska and Tietjen in 1969 [12] and in 1993 high brightness blue light emitting diodes (LEDs) was developed by Nichia. Since then this material system has become primary choice for blue LEDs, blue LASER diodes and other optoelectronic devices. Figure 1-1 shows some of the major milestones in GaN research [1].

Today, GaN-based transistors are starting to become key components of power amplifiers (PAs) impacting high-power transmitter design at microwave frequencies [13]. This has opened up a new arena for nitride semiconductors in the area of electronics. Gallium-Nitride-based high electron mobility transistors (GaN HEMTs) are proving to be attractive candidates for high voltage (HV) and high frequency (HF) applications. As the market for cellular, personal communications services, and broadband access are expanding and fifth-generation (5G) mobile systems are coming

closer to reality, GaN-based RF and microwave power amplifiers are beginning to be the focus of attention.

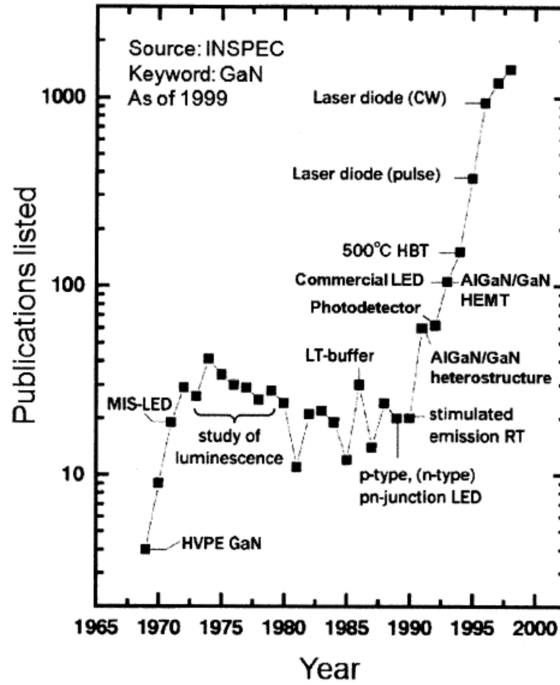

Figure 1-1: Number of publications (based on INSPEC search) and milestones of GaN research [1]. GaN research in both optoelectronics and electronics progressed rapidly after demonstrating the growth of high quality GaN and conductivity control of both p- and n-type GaN in the early 1990s.

1.2 GaN: Material properties and applications

For mm-wave applications, a variety of power amplifier technologies are vying for market share including Si-LDMOS (Lateral-diffused MOS) and Bipolar transistors, GaAs MESFETs, GaAs (or GaAs-InGaP) heterojunction bipolar transistors (HBTs), SiC (Silicon Carbide) MESFETs and GaN HEMTs [2]. The competitive advantage of GaN in this area is due to its superior and unique properties such as high electron density ($\approx 1 \times 10^{13} cm^{-2}$), high electron mobility in two dimensional electron gas (2DEG)($\approx 1500 cm^2/Vs$), good thermal conductivity ($\approx 1.5 W/cm.K$) and high breakdown field ($\approx 3.0 MV/cm$). These properties are compared with other materials

and summarized in Fig.1-2. According to a recent survey by Strategies Unlimited, the total GaN Electronic Device market is expected to reach USD 500 M by the end of this decade [2]. RF and microwave applications are likely to be the largest share of the GaN device market. The targets for GaN HEMT are both military and commercial applications. The former include RADARs (ship-board, airborne and ground) and high performance space electronics. The latter include: Base station transmitters, C-band Satcom, Ku-K band VSAT and broadband satellites, LMDS and digital radio [2].

Semiconductor		Si	AlGaAs/InGaAs	InAlAs/InGaAs	SiC	AlGaN/GaN
Characteristic	unit					
Bandgap	eV	1.1	1.42	1.35	3.26	3.49
Electron mobility at 300K	cm2/Vs	1500	8500	5400	700	1500-2200
Saturated (Peak) electron velocity	(10^7) cm/s	1.0 (1.0)	1.3 (2.1)	1.0 (2.3)	2.0 (2.0)	1.3 (2.1)
Critical breakdown field	V/cm	0.3	0.4	0.5	3	3

Figure 1-2: Table showing material properties. GaN shows a combination of high electron mobility, electron velocity and breakdown field [2]

In addition to the RF market, properties unique to GaN material systems such as high breakdown field and low on-resistance are opening up new avenues in power conversion market. Already several applications between 20 V to 1200 V in the power electronic marketplace are being designed using GaN on Si platform. These applications include switch mode power supplies (SMPS), DC-DC converters, Power ICs (PIC) etc. As the silicon-based power FET is approaching a performance plateau, further enhancements are incremental and costs of advancements are becoming prohibitively high. Thus, to meet new requirements and challenges, novel materials and transistor structures are needed. Even though silicon carbide (SiC) FETs have emerged and have been undergoing refinements for the past 10 years to address these issues, they suffer from significant cost premiums due to limited-quality material supply, as well as the intrinsic cost structure of the material. At this critical juncture, GaN HEMTs find a unique opportunity to meet these demands. Figure. 1-3 shows the role of GaN projected by International Rectifiers [3]. The FoM defined in the figure is 'efficiency $\times$ density/ cost' which is appearing to currently saturate. GaN

based technologies provide the advantage of higher density with greater efficiency needed to continue the improvement in the FoM.

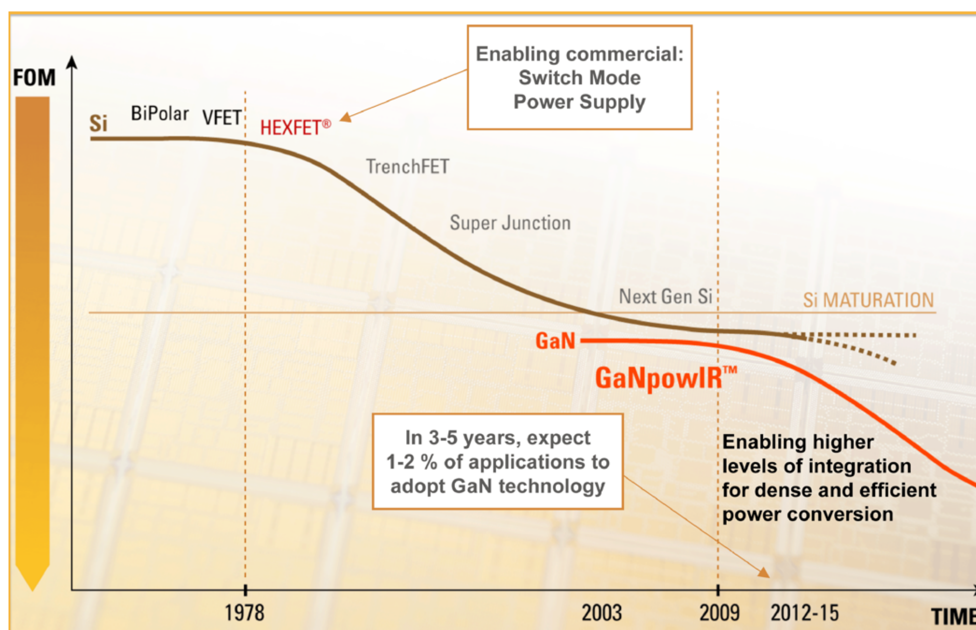

Figure 1-3: Projection of the role of GaN on Si platform in IR's market [3]. The FoM ('efficiency × density/ cost') is sustained in future through GaN based technologies.

From Fig. 1-3 it is clear that at present, GaN is about to make significant strides in power electronic industry. As mentioned earlier there are two main reasons for this: namely, significantly lower specific on-resistance (R_{on}) and high breakdown voltage (BV). Figure. 1-4 shows this trade-off for silicon, SiC and GaN. For the same BV, GaN has 10 times lower R_{on} than SiC. This improves the figure of merit of GaN devices. For the same BV voltage, GaN power devices will have lower footprint or the footprint could be traded for higher BV and lower R_{on}.

Since GaN HEMTs are poised to take the lead in the above mentioned applications, it becomes necessary to understand the operating principle of the device. Once the underlying physics of device operation is well understood, simple physics-based compact models can be developed for circuit design. The central objective of this thesis is to develop such a compact model for highly scaled GaN HEMTs for RF circuit design. The goals of the thesis are briefly explained in the next section.

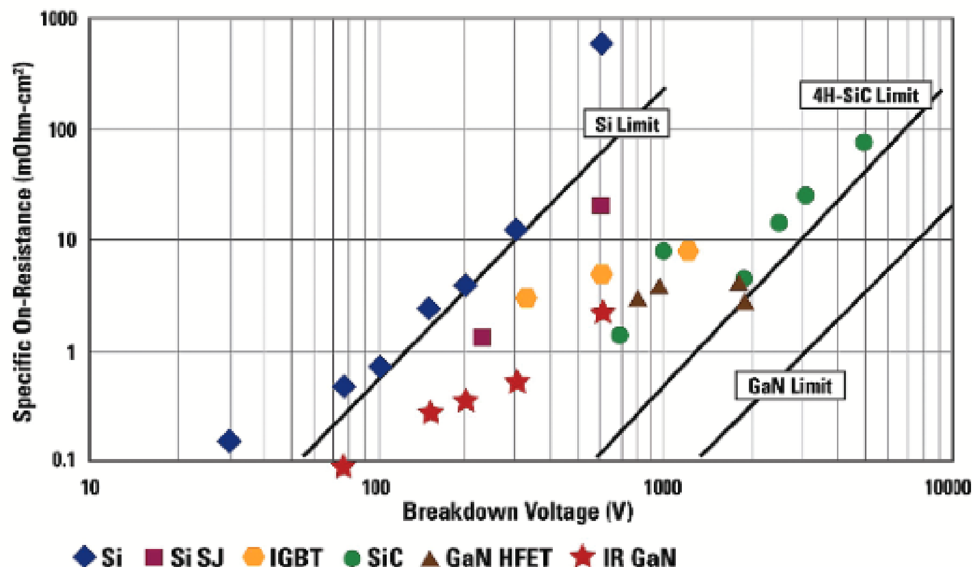

Figure 1-4: R_{on} vs. BV tradeoff for Si, GaN and SiC. For a given BV GaN shows significantly lower R_{on} [3]

1.3 Thesis outline and goals

This thesis has two main objectives. Firstly, to understand the operation of GaN HEMTs under various external bias conditions; both voltage and temperature. In addition, to understand the essential physics in different spatial regions of the device. Secondly, to develop compact device models which can capture significant physics aspects while using simple analytical expressions suited for circuit simulations. We propose such a compact model for GaN HEMTs for RF applications which we call the MIT Virtual Source GaNFET-RF (MVS-G-RF) model.

The thesis is organized as follows:

- In chapter 2, brief operating principle of GaN HEMTs is described. Important parameters and figures of merit for DC and RF conditions will be explained. In addition a brief summary of device fabrication flow for the devices fabricated at MIT is discussed.

- Chapter 3 deals with a detailed literature review of available GaN compact models. Comparison of these models with the MVS-G-RF model is done. Model

simplicity and accuracy, physical effects captured, C-continuity, self-heating effects, dimension scalability, trapping and other GaN specific effects: These are some of the properties which serve as baseline for comparison.

- Details of MVS-G-RF model are explained in chapter 4. These include models for the intrinsic transistor, access regions, device self-heating, gate leakage.

- In Chapter 5, comparison of the model with measurements of fabricated devices is carried out. In addition to DC measurements, S-parameter measurements are used to study dynamic device behavior. Along with a description of measurement procedure, performance projection based on the model is demonstrated in this chapter.

- Finally in Chapter 6, summary and conclusions are presented. Further work that is needed in this area is also discussed.

Chapter 2

GaN HEMT: Operating principle

In this chapter, a brief introduction to the physics of operation of GaN HEMTs will be given. Some physical effects such as polarization, piezoelectricity and charge trapping that are unique to GaN-based HEMTs will also be dealt. Important DC and RF parameters will be discussed and finally, the process flow that is followed for device fabrication will be described.

2.1 Basics of device operation

2.1.1 Generic HEMT structure

A High electron mobility transistor (HEMT) is a heterojunction field-effect device. The heterostructure formed between two materials with different crystal structures and bandgaps aids in the formation of a quantum well at the interface. Although similar band confinement is achieved through inversion in MOSFETs, the key difference of a HEMT structure is that the two dimensional electron gas at this interface is highly mobile. Selective doping of heterostructure can be done to achieve an undoped channel. A rough schematic of a typical HEMT structure is shown in Fig. 2-3 which shows the band structure along the heterostructure. The spatial separation of doped layer (n^+ -AlGaAs in this case) from the mobile electrons in the channel (GaAs in this case) enables high carrier mobility due to reduced ion impurity scattering.

The first report on HEMT was published by Takashi Mimura in 1979 [14]. Many of the first generation HEMT devices belonged to the AlGaAs/GaAs or InAlAs/InGaAs family.

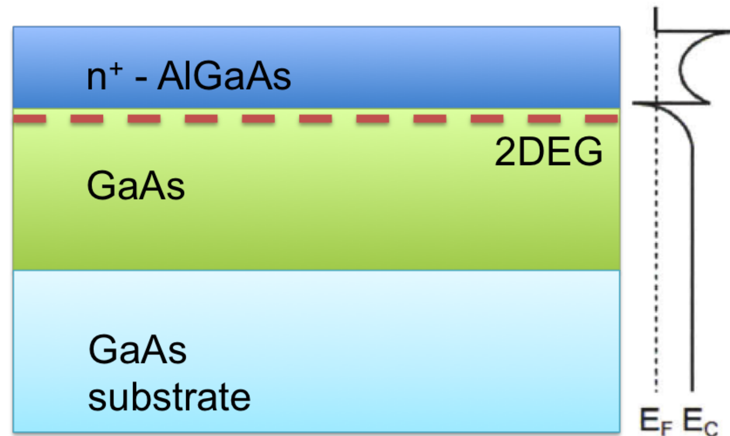

Figure 2-1: A rough schematic of a typical GaAs based HEMT heterostructure

2.1.2 GaN HEMT operation : Charge

The first demonstration of HEMT on GaN system was by Asif Khan et. al in 1993. [15]. The hetero-junction in these devices is formed between AlGaN / InGaN/ InAlGaN and GaN, with different lattice constants. A schematic of the device structure is shown in Fig. 2-2a. The difference in electron affinity (χ) and band gap (E_g) results in the formation of a nearly triangular potential well at the interface on the GaN side where electrons can be confined resulting in a two-dimensional electron gas (2DEG) as shown in Fig. 2-2b. The cause for the 2DEG in GaN HEMTs is different from other HEMT devices. The polar nature of the AlGaN/GaN system results in spontaneous polarization and in addition, the difference in lattice constants of the two layers results in piezoelectric polarization. In GaN HEMTs, the 2DEG is not induced by doping but instead by donor-like surface states on the AlGaN layer facilitated by spontaneous and piezoelectric electric field in the AlGaN layer. The 2DEG density in GaN-based HEMTs is therefore a strong function of the barrier layer (AlGaN) thickness.

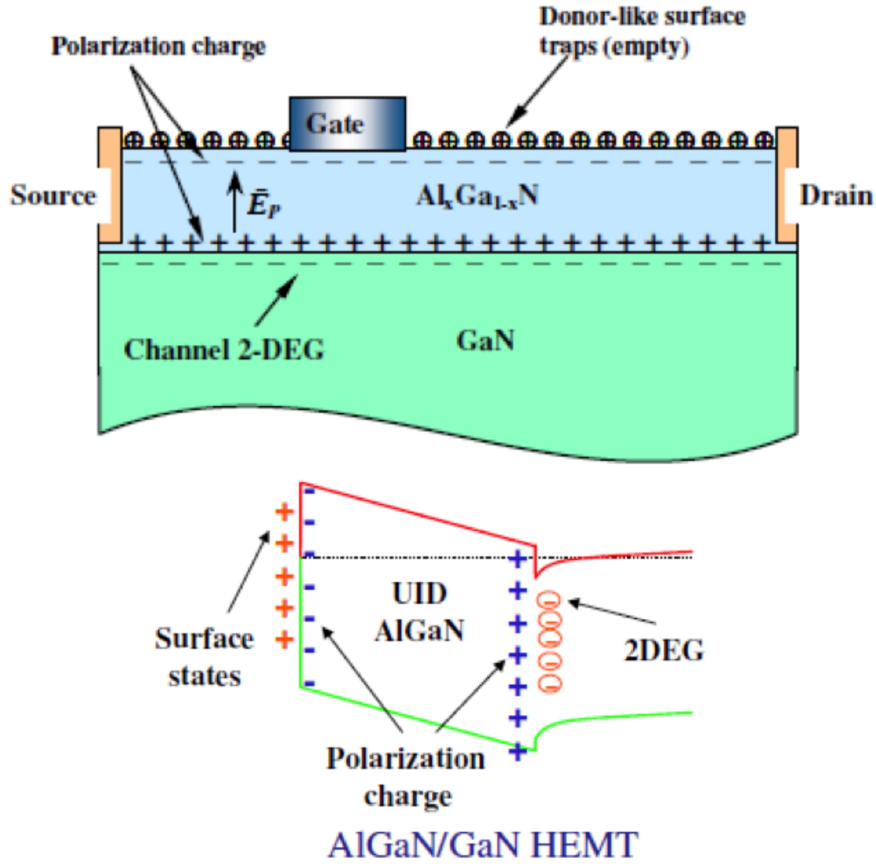

Figure 2-2: (a) Device structure schematic (b) Band diagram of heterostructure showing different types of charges

Operation of the device requires a gate electrode to control the 2DEG density and hence modulate the drain to source current (I_{DS}). Since the 2DEG exists due to the heterostructure, a negative gate voltage is required to deplete it under the gate and hence cut-off current flow. Therefore, GaN HEMTs without special gate stack engineering are normally-on (depletion-mode, D-mode) devices. With this understanding of the origin of charge in the device, next we look at the carrier drift velocity.

2.1.3 GaN HEMT operation : Transport

Confinement of electrons in the quantum well and absence of dopants in the channel result in high mobility and peak velocity. There are several reports in literature on

the velocity-field dependence of electrons in GaN systems [16]. Figure 2-3 shows the Monte-Carlo simulation results of velocity-field characteristic in GaN as given in [17]. According to these simulations, the velocity saturation field in GaN is about 250 KV/cm (not shown in the figure). The velocity does not show simple saturation as observed in Si but shows the 'hump and dip' characteristic similar to other III-V systems. The drift velocity in AlGaN/GaN HEMTs peaks at approximately 3×10^7 cm/s at an electric field of about 160 kV/cm and then decreases. At even higher electric fields, the drift velocity saturates to a value of $\approx 1 \times 10^7$ cm/s. This negative differential mobility is caused due to inter-valley scattering according to simulation studies[18].

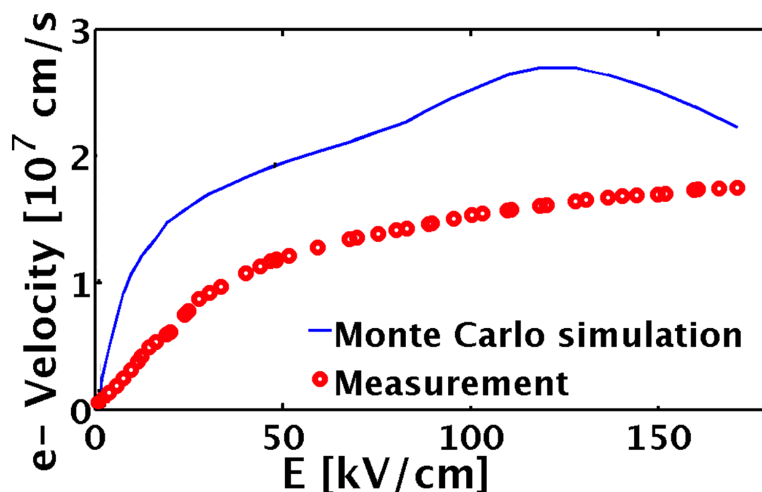

Figure 2-3: Velocity-field profile of GaN HEMTs; MC simulation and measurements

The velocity vs. field profile can be extracted from measurements. The velocity-field profile that is obtained through measurements does not show any negative differential mobility. Figure 2-3 also shows the velocity-field profile that is extracted from IV measurements of transmission line method structures (TLMs). By calculating the 2DEG charge from Hall-probe measurement and accounting for voltage drops in contact resistance, the velocity-field profile shown in red in the figure was obtained. Clearly no sign of 'hump and dip' characteristic can be seen. Possible explanation for this could be self-heating in DC measurements which could potentially mask the 'hump and dip'. However, pulsed IV measurements done in [16], where most of the

self-heating is removed, also does not exhibit negative differential mobility.

In this work, a simple velocity saturation model for carriers in 2DEG is assumed to get a closed-form expression for current. The saturation velocity in our model lies in between the peak-velocity and saturation velocity of MC simulations. This is depicted in Figure 2-4. The error between the assumed and actual profile in the low field region of interest is quite small. For most of the RF applications, the bias conditions are such that the device operates in the 'region of interest' shown in Fig. 2-4.

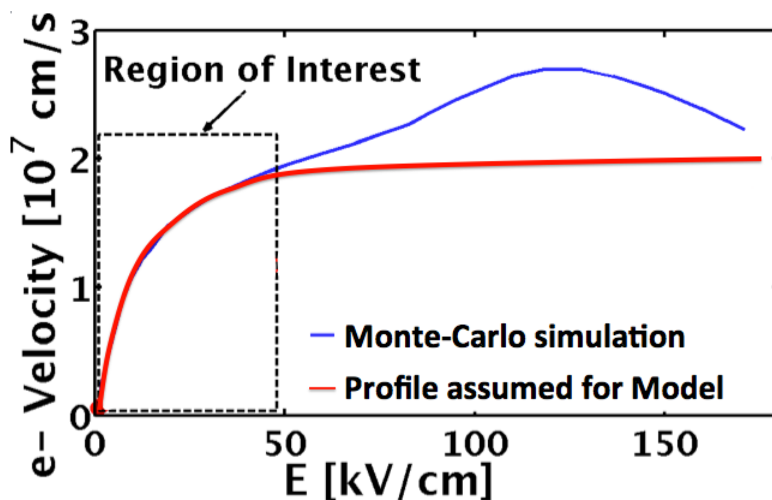

Figure 2-4: Velocity-field profile assumed in MVS-G-RF model

2.1.4 Physical effects unique to GaN

While charge and transport in GaN HEMTs are similar to HEMTs of other material systems, some properties are unique to GaN HEMTs. They are briefly described in this section.

Access regions

GaN HEMTs usually have Schottky gates and the source and drain are not self-aligned to avoid gate leakage. Even in cases where devices have gate dielectric, e.g. the devices fabricated at MIT currently do not have self-alignment. This results in

source and drain access regions which are essentially un-gated heterostructure regions. This is shown in the schematic of Fig. 2-5

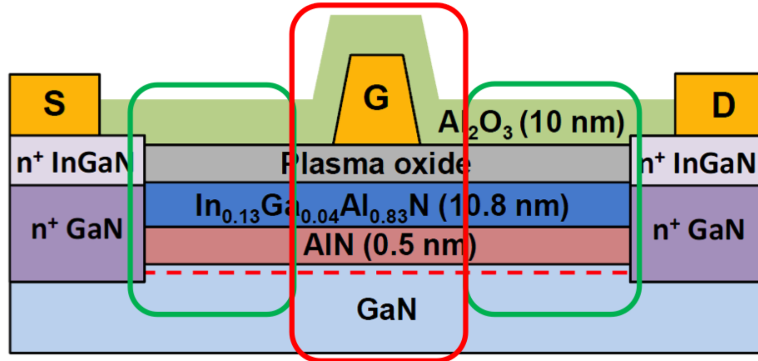

Figure 2-5: Structure schematic showing different regions of interest (a) red-intrinsic transistor (b) green- access regions

Access regions are nothing but resistive regions under low bias conditions. Under high bias conditions, velocity saturation, self-heating and pinch-off in these regions introduce non-linearity in access region behavior. The MVS-G-RF model deals with these effects as will be explained in chapter 4.

Trapping effects

Charge trapping, current collapse or knee walk-out is a much discussed nonideality of GaN HEMTs [19] - [20]. Simply put, current collapse is the increase in on-resistance and reduction in on-current during pulsed switching conditions. The effect is severe when the gate is biased to large negative voltages and the drain is biased to large positive voltages before switching. Several mechanisms have been proposed for this effect. Some of them include charging of virtual gate [21], gate-bias-induced nonuniform strain [22],and hot electron injection and trapping in the buffer[23]. The theories are still controversial but the explanation of virtual gate charging appears most convincing. Fig. 2-6 depicts a rough schematic of this scenario.

Since the 2DEG in the channel of a GaN HEMT is created by donor type surface states, any impact on the surface states reflects in the 2DEG charge density. In the virtual gate charging scheme, the electrons from the gate compensate some of the

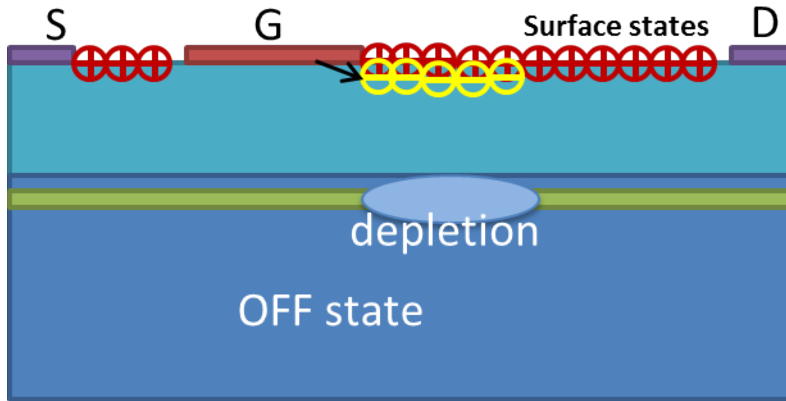

Figure 2-6: Schematic showing the reason for dynamic R_{on} effects. Electrons from the gate compensate the positive donor states creating depletion region in the 2DEG at drain edge of the gate thereby increasing R_{on}.

surface states in the drain access region adjacent to the gate electrode. This results in the creation of a depletion layer in the channel next to the gate in the drain access region. This region is therefore like a virtual gated region which increases the drain access region resistance causing current collapse. This trapping effect is still to be incorporated in the model.

In addition, effects such as gate leakage, device self-heating which are included in the model will be explained in chapter 4. In the next section, some of the important DC and RF performance figures will be described.

2.2 DC and RF figures of merit

2.2.1 DC parameters

The important DC parameters of GaN HEMTs for mm-wave applications are transconductance (g_m) and on-resistance (R_{on}). Highly scaled gate-length devices are used for RF applications, in which transport in the channel is mostly velocity saturated. The intrinsic saturation transconductance is given as below:

$$g_{msat,int} = \frac{\partial I_{Dsat}}{\partial V_{GS,int}} = WC_g v_{sat} \tag{2.1}$$

where

$$I_{Dsat} = WQ_{sat}v_{sat} = WC_g \left(V_{GS,int} - V_T\right) v_{sat} \tag{2.2}$$

However this is the intrinsic transconductance ($g_{msat,int}$). The actual transconductance (g_{msat}) is much lower due to the source access region voltage drop.

$$I_{Dsat} = WQ_{sat}v_{sat} = WC_g \left(V_{GS,ext} - R_S I_{Dsat} - V_T\right) v_{sat} \tag{2.3}$$

$$g_{msat,ext} = \frac{\partial I_{Dsat}}{\partial V_{GS,ext}} = \frac{g_{msat,int}}{(1 + g_{msat,int}Rs)} \tag{2.4}$$

The non-linearity of the source access region also introduces non-linearity in gm in addition to reducing it. Transconductance non-linearity introduces harmonic distortions in RF circuit design [24]. The MVS-G model can capture these effects as will be shown later.

The on-resistance (R_{on}) has significant contribution from the source and drain contact resistance. With recent developments, in ohmic-contact formation using Ti/Al/Ni/Au metal stacks, very low contact resistance ($\approx 0.3\ \Omega$ .mm) [25] is achievable.

2.2.2 RF parameters

At microwave frequencies of device operation (300 MHz to 300 GHz), standard circuit theory fails as signal wavelength approaches device dimensions. Phase shifts in signal voltages across the device can be accounted-for by assuming two-port network theory model with a simple small signal equivalent circuit for the device. This enables the extraction of RF performance parameters such as unity current gain cutoff frequency (f_T), and Unity power gain cutoff frequency (f_{max}). A detailed extraction procedure of f_T and f_{max} will be given in chapter 5. The key dependencies of these two RF figures-of-merit can be calculated by assuming a simple small signal equivalent circuit as shown in Fig. 2-7 [10].

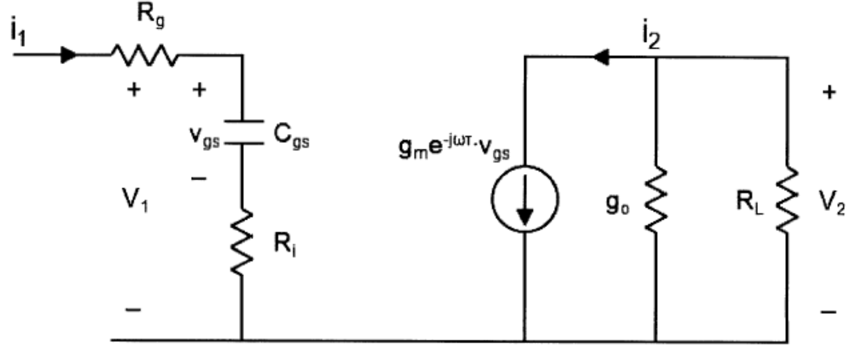

Figure 2-7: Simplified circuit for RF FoM extraction

The current gain (h_{21}) for this circuit is given as (for these calculations R_i is ignored)

$$h_{21} = \frac{i_2}{i_1} = \frac{g_m}{j\omega C_{gs}} \tag{2.5}$$

The unity current gain frequency (f_T) is then given by

$$f_T = \frac{g_m}{2\pi C_{gs}} \tag{2.6}$$

Similarly voltage gain (A_v) can be computed as

$$A_v = \frac{v_2}{v_1} = \frac{g_m(R_o||R_L)}{\sqrt{1+\omega^2 C_{gs}^2 R_g^2}} \approx \frac{g_m(R_o||R_L)}{\omega C_{gs} R_g} \tag{2.7}$$

The power gain then becomes

$$A_i A_v = \frac{i_2 v_2}{i_1 v_1} \approx \frac{g_m}{\omega C_{gs}} \frac{g_m(R_o||R_L)}{\omega C_{gs} R_g} \tag{2.8}$$

For the load matched case with $R_o = R_L$, maximum power delivered to the load is given as

$$G_p = A_i A_v/2 \approx \frac{g_m}{\omega C_{gs}} \frac{g_m(R_o)}{4\omega C_{gs} R_g} \tag{2.9}$$

The unity power gain frequency (f_{max}) is then given by

$$f_{max} = \frac{f_T}{2}\sqrt{\frac{R_o}{R_g}} \tag{2.10}$$

From above equations it is clear that for maximizing f_T and f_{max}, device scaling is a straightforward solution. In addition to scaling, device engineering to reduce device parasitics and gate resistance also improves the RF FoMs. The f_T for GaN FETs are much higher than in Si FETs for the same breakdown voltage. A study of breakdown voltage vs. f_T done in [4] for various technologies is reproduced here in Fig. 2-8. GaN technology can serve the needs of high BV ($> 100V$) and medium-high f_T (30-300 GHz) applications.

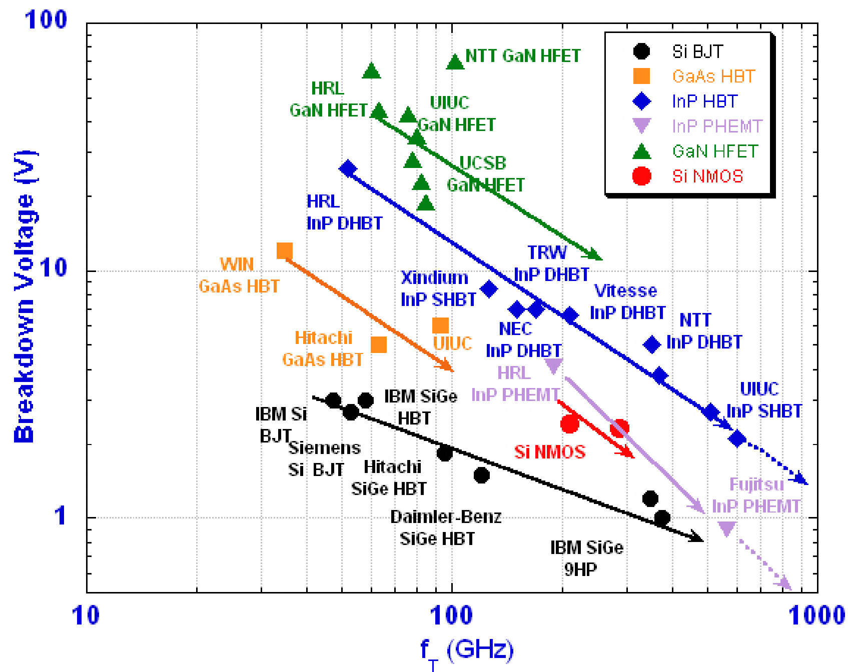

Figure 2-8: BV vs. f_T trade off for different technologies as given [4]. GaN based devices have the highest BV with medium to high f_T.

The figures-of-merit discussed in this section serve as baselines for technology innovations in GaN HEMT fabrication. The MVS-G-RF model can be used to project many of the above FoM and technology bottlenecks affecting their improvement. This will be explained in chapter 5. In the next section, a brief outline of the process flow for the device under study will be given.

2.3 Device under study

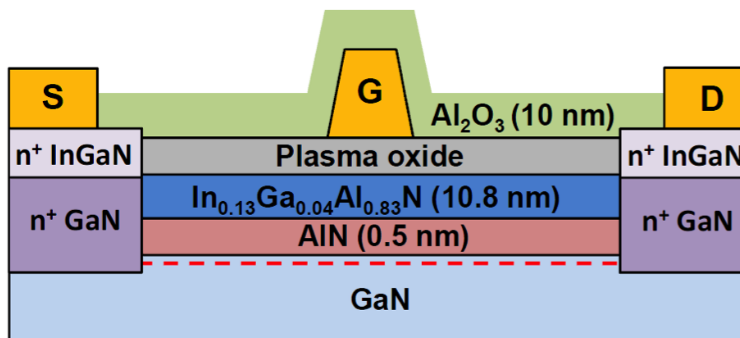

Figure 2-9: Schematic of the device structure fabricated by Prof. Palacios' group [5]

The devices required for this model development were fabricated by group members of Prof. Palacios at MIT. The RF devices studied have structures as depicted in Fig. 2-9. The heterostructure of the fabricated devices consist of a 10.8 nm $In_{0.13}Al_{0.83}Ga_{0.04}N$ layer near-lattice-matched to GaN, and 0.5 nm AlN as top barrier with 1.8 μm GaN buffer. Devices have typical gate lengths between 25 nm and 120 nm and total on resistance of 0.8 $\Omega - mm$. The total source to drain distance (L_{SD}) is approximately 1 μm [5]. Hall measurements after surface passivation with a 10 nm Al_2O_3 dielectric shows a 2-D charge density (2DEG) of $1.5 \times 10^{13} cm^{-2}$, mobility of 1600 $cm^2V^{-1}s^{-1}$, and sheet resistance of 210 Ω/sq.

A brief outline of the fabrication process flow is described here. A detailed description can be found in [26]. The lattice-matched heterostructure is grown on a SiC substrate by metal-organic chemical vapor deposition. The device fabrication process begins with mesa isolation using Cl_2/BCl_3-plasma-based dry etching. Then, ohmic contacts are formed by depositing a Ti/Al/Ni/Au metal stack, followed by annealing at 850^oC for 30 s in a N_2 atmosphere. After the ohmic anneal, the surface of the InAlN is subjected to an oxygen plasma treatment at 800 W by using a Branson IPC plasma asher. Electron-beam lithography is used to define rectangular gate electrodes in a single layer of polymethyl methacrylate. The gate metalization consists of a Ni

(10 nm)/Au (40 nm) metal stack. After gate metalization, the oxide layer on the access region, which is formed by the oxygen treatment, is removed by wet etching with buffered oxide etch. Finally, the devices are passivated with a 10-nm Al_2O_3 layer deposited by atomic layer deposition.

In this chapter, a brief overview of the device operating principle, important FoM and a short summary of fabrication process flow was given. With this basic understanding of the device, we look at the current status of compact modeling efforts of the device in the next chapter.

Chapter 3

GaN compact models: Overview

A literature review of the current status of GaN HEMT compact models is the topic of this chapter. As GaN technology matures in terms of process and fabrication technology, the focus of GaN research is shifting to modeling and circuit design. In this scenario, there is increasing demand for accurate non-linear models. A detailed survey of available GaN models is already done in [9]. In addition to the usual requirements of compact models such as accuracy, numerical simplicity, ability to model both static and dynamic behavior, GaN models are expected to capture many behavioral nuances described in chapter 2. Many potential candidate models for this purpose are available among which the Curtice model [7], the Angelov model [6], the EEHEMT (originally Eesof GaAs HEMT) [27] are well known.

The organization of this chapter is as follows: First a brief description of each of the non-linear models described is discussed. This is followed by a comparison of the models in terms of the effects that they can capture and their computational complexity. The MVSG-RF model performance with respect to these models is also studied.

3.1 Angelov model

The Angelov model (Chalmers U. model) is a popular empirical non-linear IV model. The model was first proposed in 1992 by Prof. Iltcho Angelov of Chalmers university,

Sweden. It is an equivalent circuit type model suitable for GaAs, GaN, SiC and even CMOS technologies. Shown in Fig. 3-1 is the equivalent circuit that is implemented by the model as given in [6].

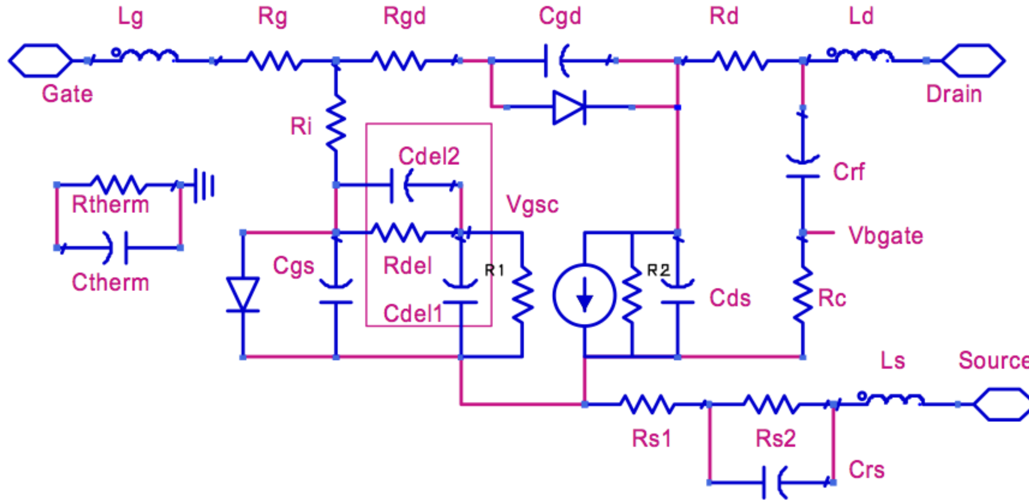

Figure 3-1: Equivalent circuit of Angelov model [6]

While many of the elements in the circuit shown in Fig.3-1 capture parasitics at device nodes, the core model equations govern currents I_d, I_g and capacitances C_{gs}, C_{gs} and C_{ds}. The approach followed in Angelov model is to capture the non-linearity in transconductance (g_m) observed in GaN, GaAs and other material technologies. This is shown in Fig. 3-2. As can be seen from the figure, the g_m plots exhibit the so called 'bell shaped' characteristic. Capturing this shape is the key idea around which Angelov model seems to have been built.

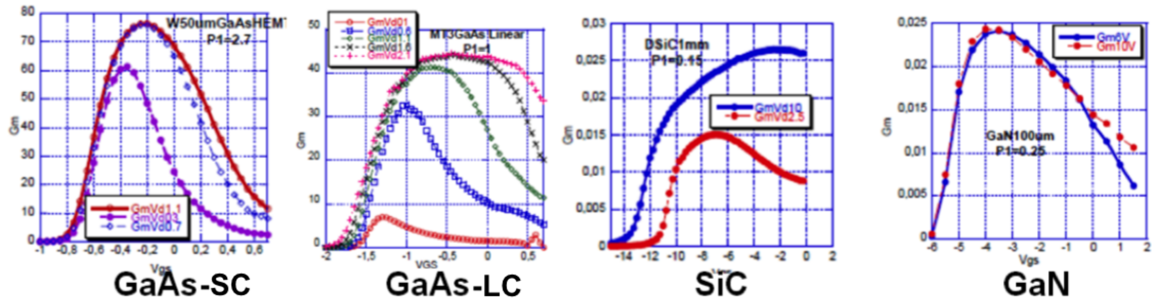

Figure 3-2: Typical g_m characteristics for different material systems (after,[6])

The formulation for g_m (ignoring V_{ds} dependence) in the model is given by

$$g_m = g_{mpk}\left(1 - tanh^2[p_{1m}\left(V_{gs} - V_k\right)]\right) \tag{3.1}$$

where g_{mpk}, p_{1m} and V_k are fitting parameters. The reason for using tanh() formulation is to ensure that current expressions have infinite derivatives. Clearly the core model is empirical in nature and is not physics based. The current then becomes,

$$I_{ds} = I_{pks}\left(1 + tanh(\psi_p)\right) tanh(\alpha V_{ds})\left(1 + \lambda V_{ds}\right) \tag{3.2}$$

Here, I_{pks}, α and λ are fitting parameters. The non-linearity and hence harmonics are captured by powered dependence of I_{ds} on V_{gs} as

$$\psi_p = P_{1m}\left(V_{gs} - V_{pk0}\right) + P_2\left(V_{gs} - V_{pks}\right)^2 + P_3\left(V_{gs} - V_{pksm}\right)^3 \tag{3.3}$$

More such empirical terms (with additional fitting parameters: P_{1m}, P_2, P_3, V_{pk0}, V_{pks} and V_{pksm}) are added as and when device non-idealities such as DIBL and other short channel effects (SCE) are required to be captured. In addition to currents, charge model expressions have similar form with their own set of parameters. A typical capacitance expression in the model is:

$$C_{gs} = C_{gsp} + C_{gso}\left(1 + tanh(\psi_1)\right)\left(1 + tanh(\psi_2)\right) \tag{3.4}$$

with more parameters C_{gsp}, C_{gso} and the non-linearity achieved through

$$\psi_1 = P_{10} + P_{11}V_{gs} + P_{111}V_{ds}; \psi_2 = P_{20} + P_{21}V_{ds} \tag{3.5}$$

Clearly the current and charge models are similar in form but not self-consistent. The model also has formulation for gate leakage which is physics-based. It includes electro-thermal models for self-heating. In addition the access regions are modeled as empirical non-linear elements. Many of the parameters can be extracted from measurements and most of them are just fitting parameters without physical interpretation.

According to [9], Angelov model has about 90 parameters. This empirical nature of the model with large fitting parameters increases the model accuracy if the parameters are correctly extracted. The model does indeed seem to capture the non-linear behavior and resulting harmonics in power circuits [6]. However the model is non-speculative since it is not physics based. Any changes to the device dimensions or process technology will mean that the model has to re-fitted to the characteristics. This increases the optimization time. In addition, it makes the model non-scalable with regard to geometry, which is an important criterion expected of compact models. In the next section we look at another well-known GaN HEMT model; the Curtice model.

3.2 Curtice model

There are several variants of Curtice model in literature such as CFET, CHEMT and so on. The original version of the model can be traced back to the paper published in 1985 by Prof. Walter Curtice of university of Michigan [7]. The model was originally proposed for GaAs FETs but has been extended to other material systems including recently to GaN HEMTs.

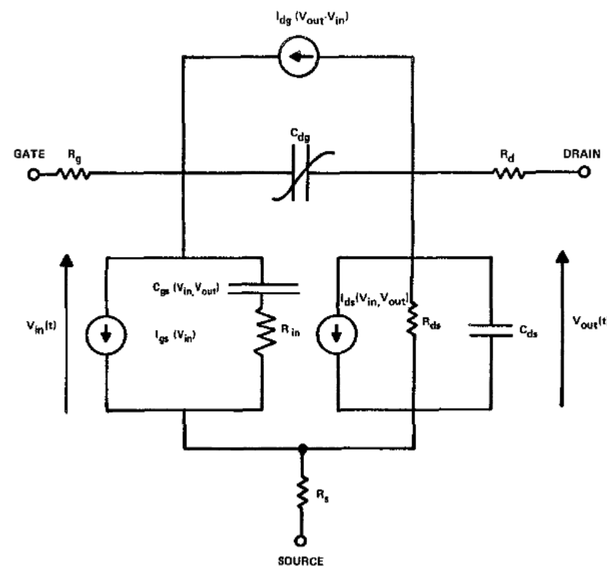

Figure 3-3: Equivalent circuit of the original version of Curtice model (after [7])

Figure 3-3 is the equivalent circuit of Curtice model as depicted in [7]. Once again the terminal currents I_{ds}, I_{dg} and I_{gs} along with capacitances C_{gs} and C_{gd} are modeled through non-linear semi-empirical functions dependent on terminal voltages. The currents have a polynomial dependence on the input voltage as shown

$$I_{ds} = \left(A_0 + A_1 V_1 + A_2 V_1^2 + A_3 V_1^3\right) tanh(\gamma V_{out}(t)); \tag{3.6}$$

$$V_1 = V_{in}(t - \tau)[1 + \beta(V_{out}^o - V_{out}(t)] \tag{3.7}$$

Here β is the current factor while the remaining parameters are fitting parameters. In a similar fashion, the capacitances are also non-linear functions of voltages. The small signal equivalent circuit elements such as g_m, g_{ds} etc. which are obtained from differentiating currents have to be tweaked to account for access region effects [7]. As given in [9], the Curtice model in its Curtice3 form is not geometry scalable and has 59 parameters. It does not model self-heating effects. However the compact modeling council (CMC) variant of the model has 55 parameters with geometry scalability and electro thermal effects incorporated.

3.3 EEHEMT model

The Eesof HEMT model originally proposed for GaAs HEMTs is an extension of Curtice model [27]. It is an empirical model that was developed by Agilent technologies for the express purpose of fitting measured electrical behavior of HEMTs. The model includes: isothermal I_{ds} model fits, flexible g_m formulation, self-heating correction for I_{ds}, charge model that fits measured capacitance values, dispersion model that permits fitting of high-frequency conductances and DC characteristics, breakdown model that describes I_{gd} as a function of both V_{gs} and V_{ds}.

A thorough description of the model details is given in [8]. Here important equations governing I_{ds} and C_{gs} will be given. The EEHEMT model is defined for different regions of device operation based on V_{gs}. The different model equations with their

own parameters are then 'stitched' together to get the complete model. The different regions are based on g_m plot and is shown in Fig. 3-4.

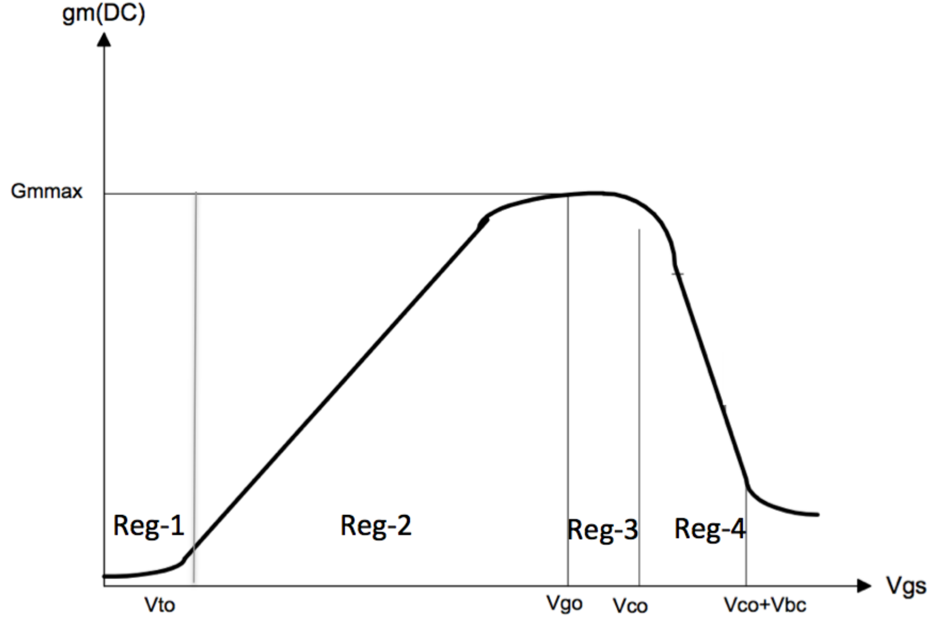

Figure 3-4: Different regions of EEHEMT model based on g_m [8]

The first region (Reg-1) is below threshold voltage, the second region (Reg-2) is the linear g_m region between maximum transconductance and threshold voltage, the remaining regions (Regs -3 and 4) are after maximum transconductance and include the compressed g_m effects at large V_{gs}.

In Reg-3 and Reg-4 the I_{ds} ($V_{gs} > V_g$) expressions are given by

$$V_g = V_{ch} + \frac{V_{go} - V_{ch}}{1 + \gamma(V_{dso} - V_{ds})} \tag{3.8}$$

$$I_{ds} = gm_{max}[V_x - \frac{V_{go} + V_{to}}{2} + V_{ch}] \tag{3.9}$$

$$V_x = (V_{gs} - V_{ch})(1 + \gamma(V_{dso} - V_{ds})) \tag{3.10}$$

In Reg-2 where $V_t < V_{gs} < V_g$, the current equation is given by

$$I_{ds} = \frac{gm_{max}}{2}[V_x - (V_{to} - V_{ch}) + \frac{V_{to} - V_{go}}{\pi} sin\left(\pi \frac{V_x - (V_{go} - V_{ch})}{V_{to} - V_{go}}\right)] \quad (3.11)$$

In Reg-1, below threshold, the I_{ds} is equal to zero. The threshold voltage itself accounts for drain bias induced effects as below

$$V_t = V_{ch} + \frac{V_{to} - V_{ch}}{1 + \gamma(V_{dso} - V_{ds})} \quad (3.12)$$

In addition to this, the EEHEMT model separates the DC and AC behavior for a simpler model extraction. The small signal models have their own add-ons in equations. This brings a certain degree of independence to fitting exercise but increases parameter count and model complexity. Further details in this regard can be found in [8].

The charge model in EEHEMT includes two gate capacitances connected between gate to source and gate to drain and output capacitance between drain and source. The gate charge model is given by the expressions below. As can be seen the form of charge and current are not self-consistent. This behavior of EEHEMT model is common to all empirical models.

$$Q_g = \left(\frac{C_{11o} - C_{11th}}{2} g(V_j) + C_{11th}(V_j - V_{infl})\right)(1 + \lambda(V_o - V_{dso})) - C_{12sat}V_o \quad (3.13)$$

$$g(V_j) = V_j - V_{infl} + \frac{\Delta gs}{3} ln\left(cosh\left(\frac{3}{\Delta gs}(V_j - V_{infl})\right)\right) \quad (3.14)$$

$$V_j = 0.5(2V_{gs} - V_{ds} + V_o); V_o = \sqrt{V_{ds}^2 + \Delta_{ds}^2} \quad (3.15)$$

Similar expressions exist for other two terminal charges and currents. In addition, there are model add-ons to include the gate current and breakdown effects [8]. The equivalent circuit of the EEHEMT model looks as shown in Fig. 3-5.

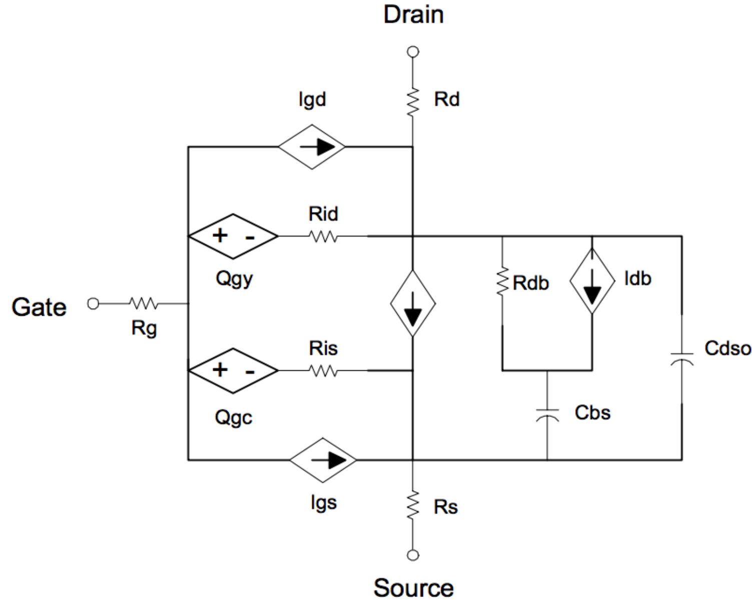

Figure 3-5: Equivalent circuit of EEHEMT model[8]

The EEHEMT model has about 71 parameters with geometry scalability. With this overview of the well-known GaN HEMT compact models, in the next section a comparison of these models with MVSG-RF model is given.

3.4 Comparison of GaN models

A comparison of existing models is tabulated in Fig. 3-6 as given in [9]. It also includes models used for GaAs FETs, LDMOS etc. According to the table, the least number of parameters, 48, corresponds to the CFET model. All models have substantial number of parameters. Larger parameter count means longer parameter extraction times. Larger parameter set does not necessarily imply correct physics. In fact as stated in [9] models with more parameters tend to be empirical. The validity of parameters reduces as parameter count increases which could result in poor performance optimization routines. Some of the models such as Angelov and Curtice3 in the table also are incapable of geometry scaling. Several models do not capture electro-thermal effects.

The MVSG-RF model performance in terms of the above criteria is quite good.

FET Models	Approx. Number of Parameters	Electrothermal (Rth-Cth) Model	Geometry Scalability Built-In	Original Device Context
Curtice3	59	No	No	GaAs MESFET
Motorola Electrothermal (MET)	62	Yes	Yes	LD MOSFET
CMC (Curtice/ Modelithics/Cree)	55	Yes	Yes	LD MOSFET
BSIMSOI3	191	Yes	Yes	SOI MOSFET
CFET	48	Yes	Yes	HEMT
EEHEMT	71	No	Yes	HEMT
Angelov	80	Yes	No	HEMT/MESFET
Angelov GaN	90	Yes	No	HEMT
Auriga	100	Yes	Yes	HEMT

Figure 3-6: Comparison of available GaN compact models[9]

Firstly, since the model is physics based, most of its parameters have physical meanings with easy extraction procedures. In addition, the current and charge expressions are self-consistent with same parameters. Unlike models such as EEHEMT model, no separate model equations are necessary for DC, large or small signal operation. The number of parameters required by MVSG-RF model is about 35 as will be shown in next chapter. This is much less than the all the models listed above including the CFET model. The model is geometry scalable and includes self-heating effects as well. Enough attention is not paid to the access regions in all the above GaN models. These regions play an important role in device behavior. MVSG-RF model captures the physics of carrier transport in these regions.

In conclusion, the review of GaN models carried out in this chapter reveals that all the existing models are empirical with large number of non-physical parameters. They do not yet capture all the GaN HEMT behavioral nuances. In this context, MVSG-RF could prove to be an attractive candidate for GaN compact modeling. In the next chapter, details of the governing equations of MVSG-RF model will be presented.

Chapter 4

The MVSG-RF model

In this chapter, the core model equations of the MIT Virtual Source GaNFET - Radio Frequency model are discussed in detail. The Key model components are best described with the help of the sub-circuit shown in Fig.4-1.

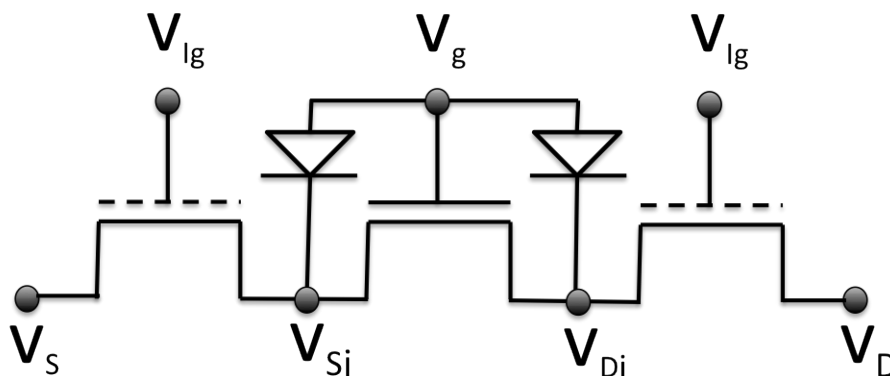

Figure 4-1: Subcircuit model of MVS-G-RF model showing implicit-gate access regions and Schottky-gate diodes along with the intrinsic transistor.

The model equations for the current in the intrinsic transistor region below the gate electrode are described in the first section. In addition, the model current equations for the non-linear access regions which are modeled as implicit-gate transistors is given in the second section. This is followed by the charge model governing channel charges. Gate current (leakage) is also modeled by two Schottky-diodes as shown in the figure, and is presented at the end of this chapter.

4.1 Intrinsic transistor model

The intrinsic transistor model in MVSG-RF is based on the Virtual-Source (VS) model developed at MIT [28] which was originally developed for highly scaled Si FETs with near-ballistic (quasi-ballistic) mode of transport. The model is based on the concept of "top-of-the-barrier-transport".

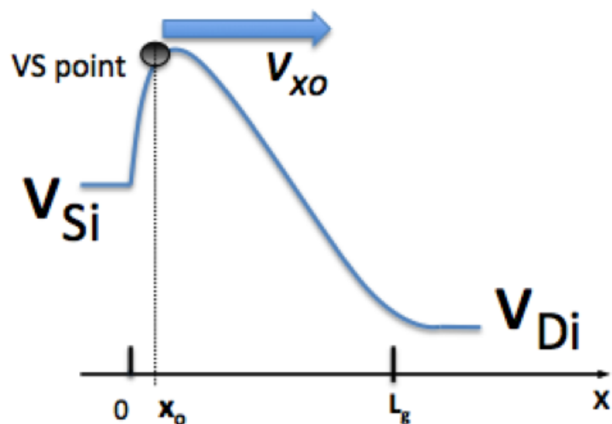

Figure 4-2: Band profile of Intrinsic transistor in saturation under drain bias.

The physical picture of carrier transport in the intrinsic transistor region of scaled HEMTs with gate lengths in the order of 10-100 nanometers can be explained with the band diagram of Fig. 4-2 shown in saturation regime. Carriers are injected into the source access region from source terminal which is a thermal-equilibrium reservoir. The carriers reaching the channel, then face the potential energy barrier in the channel whose height is modulated by the gate voltage. In highly scaled HEMTs such as RF-GaN-HEMTs, this barrier peaks at the beginning of the channel at $x = x_o$ as shown. Beyond this point, for sufficiently high drain bias, the lateral field in the channel is large enough to invalidate the application of gradual channel approximation (GCA). However, the density of carriers at the top of the barrier can be set by 1D-MOS electrostatics by the application of GCA.

4.1.1 Saturation regime (High V_{DSi})

In the charge-sheet approximation, the drain current normalized by width (I_D/W) can be described by the product of the local charge areal density times the local carrier velocity anywhere in the channel. It is particularly useful to write this expression (4.1) at the location of the top of the energy barrier ($x = x_o$) where the channel charge density Q_{ixo}, is easiest to model by applying GCA:

$$I_D = W Q_{ixo} v_{xo} \tag{4.1}$$

The average carrier velocity at this point is v_{xo}. The maximum possible value of v_{xo} is approximately the uni-directional thermal velocity and occurs when all the positive velocity carriers at the beginning of the channel are injected from the thermal equilibrium source [29]. Backscattering over a very short distance into the channel determines how close to this upper limit the device operates. In saturation, v_{xo} has been empirically found to be a bias independent parameter [28], however for GaN HEMTs it appears to show dependence on vertical field (or on Q_{ixo}).

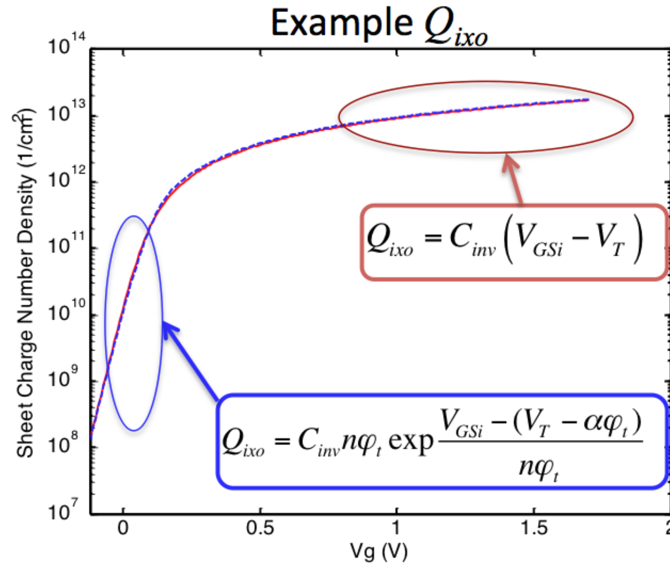

Figure 4-3: An example for VS-point charge Q_{ixo} described by the model showing simplified expressions in strong and weak inversion.

The virtual source charge density is given by the empirical form as in [28]. This

expression allows for a continuous expression for the inversion charge density at the virtual source from weak to strong inversion. The basic form of the expression was first proposed by Wright [30] with the αF_f term introduced in [28] as follows:

$$Q_{ixo} = C_{inv} n\phi_t ln\left(1 + exp\left(\frac{V_{GSi} - (V_T - \alpha\phi_T F_f)}{n\phi_T}\right)\right) \tag{4.2}$$

Fig.4-3 depicts the way in which (4.2) captures the charge behavior in both strong and weak inversion along with the asymptotic forms of (4.2) in the two regimes. C_{inv} is the effective gate-to-channel capacitance per unit area in strong inversion (we borrow the term inversion from MOSFETs to denote the state of the GaN channel charge density), ϕ_t is the thermal voltage, V_{GSi} is the internal gate-source voltage, corrected for the voltage drop on the source access region, n is the sub-threshold coefficient, which is related to the sub-threshold swing (S) as:

$$n = \frac{S}{\phi_t ln10} + n_d V_{DSi} \tag{4.3}$$

where modest punch-through is modeled using the parameter n_d. V_T is the threshold voltage corrected for drain-induced barrier lowering, DIBL, a well-known short channel effect as given by:

$$V_T = V_{To} - \delta V_{DSi} \tag{4.4}$$

Here δ is the DIBL parameter and V_{To} is the threshold voltage at $V_{DSi} \approx 0$. The term following V_T in (4.2) allows for the requirement of different values of threshold voltage in strong and weak inversion (a shift of V_T by $\alpha\phi_t = 3.5\phi_t$). The function F_f is a Fermi function that allows for a smooth transition between the two values of V_T and is centered at the point halfway between them [28]:

$$F_f = \frac{1}{\left(1 + exp\left(\frac{V_{GSi} - (V_T - \alpha\phi_T/2)}{\alpha\phi_T}\right)\right)} \tag{4.5}$$

4.1.2 Non-saturation regime

The description of current with v_{xo} as a bias-independent parameter is valid only for saturation regime of device operation where the carrier velocity is independent of lateral field. To account for the non-saturation region, the velocity v_{xo} in (4.1) is multiplied with the empirical saturation function F_{sat} [28]. F_{sat} increases smoothly from 0, at $V_{DSi} = 0$, to 1, at $V_{DSi} >> V_{DSAT}$, where V_{DSAT} is the saturation voltage. F_{sat} is similar to the Caughey and Thomas [31] formulation for carrier velocity saturation.

$$F_{sat} = \frac{V_{DSi}/V_{DSAT}}{\left(1 + (V_{DSi}/V_{DSAT})^{\beta}\right)^{1/\beta}} \tag{4.6}$$

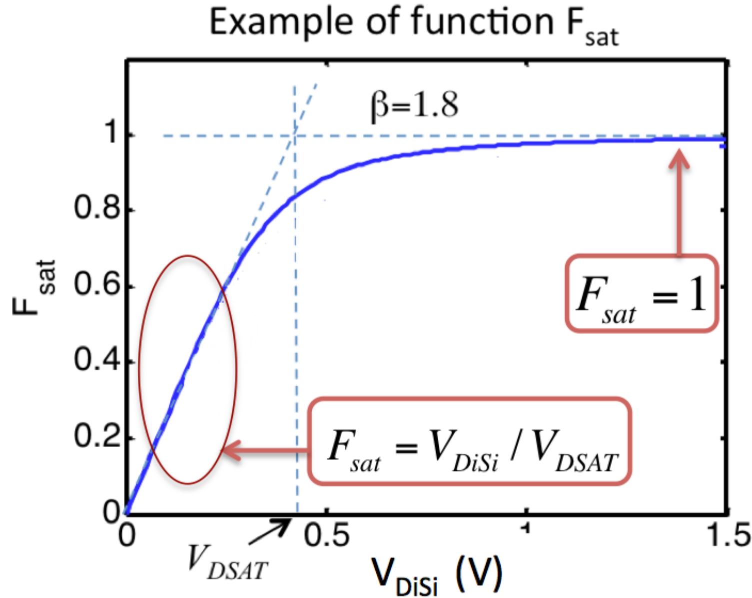

Figure 4-4: An example for F_{sat} function showing its asymptotic forms in linear and saturation regimes.

An example depiction of F_{sat} is given in Fig.4-4. It shows how V_{DSAT} influences the linear-to-saturation transition for a particular value of β. The current from linear to saturation is then given by the single expression

$$I_D = WQ_{ixo}v_{xo}F_{sat} \tag{4.7}$$

While β is a fitting parameter, it has been found over many fits to experimental devices, from Si MOSFETs to GaN HEMTs that β=1.0-3.0 provides excellent fit. Finally, to allow a smooth transition between the strong- and weak-inversion values of the saturation voltage V_{DSAT}, a generalized form of the saturation voltage is introduced by employing again F_f.

$$V_{DSAT} = V_{DSATS}(1 - F_f) + \phi_t F_f \tag{4.8}$$

$$V_{DSATS} = \frac{v_{xo} L_g}{\mu} \tag{4.9}$$

Here μ is the carrier mobility and L_g is the gate length. Note that Eqs. (4.3)-(4.5) are heuristic in nature [28] but the model has been validated successfully against several scaled Si and III-V devices [32]. This model for intrinsic transistor current must also include certain GaN-specific effects described in the next sub-section

4.1.3 GaN specific effects

The velocity at the virtual source (v_{xo}) in the model is described in detail in this section. For highly scaled Si FETs in the near-ballistic regime, the velocity, v_{xo}, at the VS point is proportional to the unidirectional thermal velocity [28], (ballistic velocity) v_T. This velocity, in the degenerate case (i.e. low channel charge density at the VS point), is independent of bias, while in the degenerate case theoretically it increases with channel charge density. However, through many fits of the MVSG Model to experimental data it is found that v_{xo} is best left independent of charge density. For transistors with slightly longer gate length (compared to mean free path), in the quasi-ballistic transport regime, backscattering reduces the velocity from the ballistic thermal velocity limit ($v_{xo} = Bv_T$, where B($<=$ 1) is the backscattering coefficient) but the velocity can still be approximated to be independent of bias.

In GaN HEMTs, small electron mean free path means that transport might not be ballistic even at highly scaled gate lengths. However model fits to experimental devices seem to suggest that for gate lengths of the order of few hundred nanometers,

carrier injection velocity at the VS point is independent of lateral field. Hence the VS model can be extended to model GaN HEMTs with a different interpretation of v_{xo}. In addition, in GaN HEMTs electron velocity decreases as charge density increases due to strong electron-optical phonon interaction (scattering) and this is modeled empirically using Q_{ixo}. Including this and the temperature dependence due to self-heating, v_{xo} is modeled as [33]-[34]:

$$v_{xo} = \frac{v_{inj}}{\left(1 + \theta_v \frac{Q_{ixo}}{C_{inv}}\right)} (1 - \eta_v I_D V_{DS}) \tag{4.10}$$

where v_{inj} is the bias independent injection velocity at low channel charge. The term in the denominator accounts for carrier scattering and the last term in the numerator accounts for velocity reduction due to self-heating. θ_v and η_v are fitting parameters. Carrier mobility also decreases due to scattering and self-heating and is modeled as [33]-[34]:

$$\mu = \frac{\mu_o}{\left(1 + \theta_\mu \frac{Q_{ixo}}{C_{inv}}\right) \left(1 + \eta_\mu \frac{I_D V_{DS}}{T_o}\right)^\epsilon} \tag{4.11}$$

Here θ_μ, η_μ and ϵ are fitting parameters and T_o is a reference temperature.

4.2 Access regions

Access regions in GaN HEMTs are important because state-of-the-art GaN HEMTs are not self-aligned, and these regions affect device performance. The impact of access regions on device performance is highlighted in the next chapter. In this section, details of the model equations for these regions are given. In the MVS-G-RF model, the behavior of access regions is captured by modeling them as implicit-gate transistors. Figure 4-5 shows the two sub-circuit access region elements of Fig.4-1 that will be described in this section.

If access region lengths are comparable to scaled gate lengths (few hundred nanometers), then the intrinsic transistor model described in the previous section could be used to model access transistors with suitable modifications to ensure field continu-

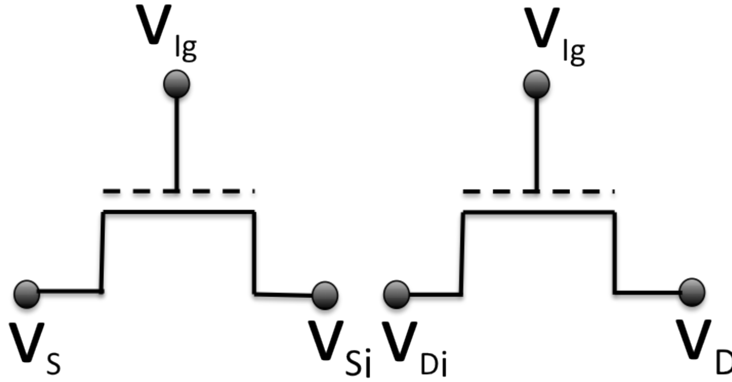

Figure 4-5: Access regions modeled as implicit-gate transistors. The dashed-lines in the sub-circuit elements denote the absence of actual physical gate terminal in these regions.

ity between different regions and with correct implicit-gate overdrive. The devices fabricated at MIT have typical access region lengths in the order of few microns. In these length scales which are much larger than carrier mean free path (carrier mean free path in GaN is about few tens of Angstroms), carrier transport in access regions is drift-diffusion based. The access region transistor model in long channel regime is described in the next sub-section.

The sub-circuit model of Fig. 4-1 shows that the intrinsic transistor and access region transistors are connected in series, which implies that at the intrinsic nodes Di and Si, the carriers undergo complete thermalization by losing kinetic energy. However, in these devices, carriers reaching the intrinsic transistor (at Si) have non-zero kinetic energy which is gained while traveling through the source access region and carriers reaching the intrinsic drain node (Di) have sufficient kinetic energy which is gained in the channel. This physical effect is ignored in this work and carriers reaching intrinsic transistor and drain access transistor are assumed to have no kinetic energy. The assumption of thermalized carriers at intrinsic source node (Si) is reasonable since the field in source access region is relatively low and does not impart significant energy to carriers at Si. The assumption might not be reasonable for carriers at Di as the field in the channel is quite high at large V_{DSi} and carriers have significant

kinetic energy. Field continuity between different regions needs to be incorporated so that these effects are captured correctly, and is being investigated.

4.2.1 Long channel access transistor model

In the long channel regime, the velocity at the virtual source (v_{xo}) in saturation is no longer a bias independent quantity but is dependent on the local lateral field as $v_{xo} = \mu E_{xo}$. The current is evaluated using the following procedure:

$$I_D = WQ_i(x)v(x) = WQ_i(x)\mu E(x) = WQ_i(x)\mu\frac{d\psi(x)}{dx} \tag{4.12}$$

Here $Q_i(x)$ is the channel charge at location x in the channel, $\psi(x)$ is the potential at that position and μ is the carrier mobility as described in (4.10). To account for carrier velocity saturation effects, (4.10) is modified as below:

$$I_D = WQ_i(x)\mu\frac{\frac{d\psi(x)}{dx}}{\left(1+\left(\frac{\frac{d\psi(x)}{dx}}{\frac{v_{sat}}{\mu}}\right)^{\beta}\right)^{1/\beta}} \tag{4.13}$$

Here v_{sat} is the carrier saturation velocity and the form of carrier velocity saturation is the Caughey-Thomas form [31]. To enable a charge-based rather than potential-based formulation for current, a procedure similar to [35] is followed by relating $\psi(x)$ and $Q_i(x)$ as

$$\frac{dQ_i(x)}{d\psi(x)} = C_{invi} \tag{4.14}$$

Here C_{invi} is the channel to gate capacitance which in our case is the implicit-gate areal capacitance. Eq. (4.13) is valid at all x provided GCA holds, which is reasonable for large channel length. Substituting for $\psi(x)$ in (4.12) using (4.13) and integration of (4.12) can be done from x=0 to x=L . Here L is the access region length i.e. for source access region, $L = L_{GS}$, and for drain access region, $L = L_{GD}$. Substituting $Q_i(0) = Q_{is}$ and $Q_i(L) = Q_{id}$ along with employing current continuity we get the following expression for current in long channel access transistor:

$$I_D = W\frac{\mu}{2C_{invi}L}\frac{Q_{is}^2 - Q_{id}^2}{\left(1+\left(\frac{Q_{is}-Q_{id}}{\frac{C_{invi}vL}{\mu}}\right)^{\beta}\right)^{1/\beta}} \tag{4.15}$$

Here v is the carrier velocity combining strong and weak inversion regimes, as discussed below. In order to make the current expression look similar to that of VS model expression in (4.1), the current expression can be re-formatted as

$$I_D = W\frac{v}{2C_{invi}V_{DSAT}}\frac{Q_{is}^2 - Q_{id}^2}{\left(1+\left(\frac{Q_{is}-Q_{id}}{C_{invi}V_{DSAT}}\right)^{\beta}\right)^{1/\beta}} = Wv\frac{Q_{is}+Q_{id}}{2}F_{vsat} \tag{4.16}$$

where F_{vsat} is given by

$$F_{vsat} = \frac{\frac{Q_{is}-Q_{id}}{C_{invi}V_{DSAT}}}{\left(1+\left(\frac{Q_{is}-Q_{id}}{C_{invi}V_{DSAT}}\right)^{\beta}\right)^{1/\beta}} \tag{4.17}$$

F_{vsat} governs the transition from non-velocity saturation to velocity saturation regime of carrier transport as channel length is scaled down. F_{vsat} has a form similar to the Caughey-Thomas form [31] which was employed for F_{sat} in the intrinsic transistor model (4.4). V_{DSAT} and v are saturation voltage and saturation velocity respectively and have the form below to account for both strong and weak inversion.

$$V_{DSAT} = \frac{v_{sat}L}{\mu}(1-F_f) + 2n\phi_T F_f \tag{4.18}$$

$$v = v_{sat}(1-F_f) + 2\phi_T\frac{\mu}{L}F_f \tag{4.19}$$

where ϕ_T is the thermal voltage, n is the sub-threshold factor related to sub-threshold slope by $S = n\phi_T ln(10)$. The charges themselves have a form close to that of the charges for intrinsic transistor. The source access region charges are given as:

$$Q_{is} = C_{invi}2n\phi_t ln\left(1+exp\left(\frac{V_{IgS}-(V_{Ti}-\alpha\phi_T F_f)}{2n\phi_T}\right)\right) \tag{4.20}$$

$$Q_{id} = C_{invi} 2n\phi_t ln\left(1 + exp\left(\frac{V_{IgS} - V_{DSAT}F_{sat} - (V_{Ti} - \alpha\phi_T F_f)}{2n\phi_T}\right)\right) \quad (4.21)$$

The key distinction of the charge expressions here from that in (4.2) is that the charge expressions have additional factors of 2 to ensure correct sub-threshold current behavior, but this creates inconsistency in sub-threshold slope for charges in weak inversion. This issue needs further study but since capacitances in sub-threshold regime is dominated by parasitic capacitances, the discrepancy in access region capacitances can be ignored.

The linear-to-saturation transition in source-access-region-current is achieved through $V_{DSAT}F_{sat}$ in Q_{id} as given by

$$V_{DSAT}F_{sat} = V_{DSAT}\frac{\frac{V_{SSi}}{V_{DSAT}}}{\left(1 + \left(\frac{V_{SiS}}{V_{DSAT}}\right)^{\beta}\right)^{1/\beta}} = \frac{V_{SiS}}{\left(1 + \left(\frac{V_{SiS}}{V_{DSAT}}\right)^{\beta}\right)^{1/\beta}} \quad (4.22)$$

Here α and F_f are as defined in (4.2)-(4.3). $V_{SiS} = V_{Si} - V_S$, $V_{IgS} = V_{Ig} - V_S$ with the node voltages as depicted in Fig.4-5. V_{Ti} is the implicit-gate transistor threshold voltage which is explained in detail in next sub-section. Similarly the charges for drain access region are given by same expressions as (4.19)-(4.21) with V_{SiS} replaced with $V_{DDi} = V_D - V_{Di}$ and V_{IgS} replaced with $V_{IgDi} = V_{Ig} - V_{Di}$ which are drain access region node voltages as depicted in Fig.4-5.

4.2.2 The implicit-gate

Access regions are ungated regions of GaN HEMTs as shown in Fig. 2-7. Since there is no physical gate in these regions, modeling them as implicit-gate transistors warrants further explanation. The access region model in MVS-G-RF model is not that of actual gated transistors but that of implicit-gate transistors. The vertical field that terminates at the 2DEG in the access regions could originate from somewhere in the device (donor states at AlGaN surface, metal layers running above etc.). In the limiting case one could assume, that the field originates at infinity in which

case the implicit-gate is at infinity. The location of the origin of this vertical field is the implicit-gate electrode. Since it is difficult to determine a location for the implicit-gate, the areal implicit-gate-capacitance (C_{invi}) is a fitting parameter in the MVS-G-RF model.

The value of the implicit gate voltage (V_{Ig}) is linked to the 2DEG sheet charge density in the access regions. Under very low drain bias conditions, the access regions are nothing but linear resistors. The current through the access regions under these conditions is given by

$$I_{access} = \frac{V_{access}}{R_{access}} = \frac{V_{access}}{\frac{R_{sh}L}{W}} \tag{4.23}$$

Where I_{access} is the access region current, V_{access} is the voltage drop across the access region and R_{sh} is the access region sheet resistance (the analysis holds for both source and drain access regions).

From the implicit-gate transistor model, the access region current at such low voltages is given from (4.14) as

$$I_{access} = \frac{W}{L}\mu C_{invi}(V_{Ig} - V_{Ti})V_{access} \tag{4.24}$$

Equating the two forms of currents we get the following expression for gate-overdrive voltage:

$$(V_{Ig} - V_{Ti}) = \frac{1}{R_{sh}\mu C_{invi}} \tag{4.25}$$

The implicit-gate overdrive can be computed from μ, R_{sh} and C_{invi}. This overdrive is used in the Q_{is} and Q_{id} charge expressions of source and drain access regions. C_{invi} is the only additional parameter needed for the access regions in the MVS-G-RF model. The access region model described in this section captures the observed non-linear behavior of these regions [16]. By modeling as implicit-gate transistors, the non-linear behavior is captured as pinch-off and velocity saturation. The model is validated against gateless transmission line method structures (TLMs) and is discussed in the next chapter.

4.3 Gate diode model

A simple Schottky-diode model is used for the gate diodes. The MVS-G-RF model adopts the procedure similar to the Angelov model [6] to include gate breakdown effects in the diode model.

The gate leakage currents through gate-source and gate-drain diodes are given by

$$I_{GSi} = I_S \, exp\left(\frac{-\phi_B}{\eta\phi_T}\right)\left(exp\left(\frac{V_{GSi}}{\eta\phi_T}\right) - 1\right) \tag{4.26}$$

$$I_{GDi} = I_S \, exp\left(\frac{-\phi_B}{\eta\phi_T}\right)\left(exp\left(\frac{V_{GDi}}{\eta\phi_T}\right) - 1\right) \tag{4.27}$$

Here $I_S = WA^*T^2$ is the pre-exponent diode current term derived from the thermionic emission model. A^* is the Richardson constant, ϕ_B is the Schottky barrier height which for GaN/AlGaN systems is typically about 1.2 eV, η is the diode ideality factor. For MVSS-G-RF model fits, I_S and ϕ_B are extractable parameters from gate current measurements. The gate breakdown effect is incorporated by suitably modifying the above equations as in [6].

$$I_{GSi,bd} = I_{GSi}\left(1 + K_{bd} \, exp\frac{V_{SiG} - V_{bd}}{\phi_{bd}}\right) \tag{4.28}$$

$$I_{GDi,bd} = I_{GDi}\left(1 + K_{bd} \, exp\frac{V_{DiG} - V_{bd}}{\phi_{bd}}\right) \tag{4.29}$$

This expression accounts for reverse gate breakdown. V_{bd} is the reverse breakdown voltage, K_{bd} and ϕ_{bd} are fitting parameters.

In addition to terminal currents, terminal charge models are necessary for dynamic circuit simulations. The channel charge model description in MVS-G-RF model is given in next section.

4.4 Channel charge model

In order to account for dynamical behavior, the model terminal charges must produce the full matrix of capacitive components and their voltage dependences. The intrinsic channel charge model under ballistic/quasi-ballistic regime is implemented as in [36]. The channel charge partitioning following the Ward-Dutton scheme [37] is done to get source and drain side components.

$$Q_s = W \int_{x=0}^{x=L_g} \left(1 - \frac{x}{L_g}\right) Q_i(x) dx \tag{4.30}$$

$$Q_d = W \int_{x=0}^{x=L_g} \left(\frac{x}{L_g}\right) Q_i(x) dx \tag{4.31}$$

The areal channel charge ($Q_i(x)$) is not known at all positions except at the VS point ($x = x_o$). The charge at the VS point is given by (4.2) and the velocity at this point is $v_x o$. Current continuity can be employed to express $Q_i(x)$ as a function of carrier velocity as follows

$$Q_i(x)v(x) = Q_{ixo}v_{xo} \rightarrow Q_i(x) = Q_{ixo}\frac{v_{xo}}{v(x)} \tag{4.32}$$

The carrier velocity at any location is computed by the application of energy conservation for carriers.

$$\frac{1}{2}m_e v(x)^2 = \frac{1}{2}m_e v_{xo}^2 + \gamma q V(x) \tag{4.33}$$

where $V(x)$ is the channel potential and γ is the parameter which accounts for energy loss due to carrier scattering. For V(x), a linear or parabolic profile can be assumed to get closed form expressions for terminal channel charge contributions. The resulting error in gate capacitances is shown to be low (5 percent) in [36]. With the linear potential assumption, we get the following expressions for velocity and charge in saturation regime:

$$v(x) = \sqrt{v_{xo}^2 + \frac{2\gamma q x V_{DiSi}}{m_e L_g}} \tag{4.34}$$

$$Q_s = WL_gQ_{ixo}\left(\frac{(4k+4)\sqrt{k+1}-(6k+4)}{3k^2}\right) \tag{4.35}$$

$$Q_d = WL_gQ_{ixo}\left(\frac{(4k-4)\sqrt{k+1}+4}{3k^2}\right) \tag{4.36}$$

where $k = \frac{2\gamma qxV_{DiSi}}{m_e v_{xo}^2}$. In addition to these channel charges, fringing capacitances also contribute to the total gate charge. The origin of inner fringing (C_{if}) and outer fringing capacitances (C_{if}) are explained in detail in the next chapter and here the charges due to these are given.

$$Q_{s,of} = WC_{of}V_{GSi} \tag{4.37}$$

$$Q_{d,of} = WC_{of}V_{GDi} \tag{4.38}$$

$$Q_{s,if} = WC_{if}\left(V_{GSi} - n\phi_T\ ln\left(1 + exp\frac{V_{GSi}-V_T}{n\phi_T}\right)\right) \tag{4.39}$$

$$Q_{d,if} = WC_{if}\left(V_{GDi} - n\phi_T\ ln\left(1 + exp\frac{V_{GDi}-V_T}{n\phi_T}\right)\right) \tag{4.40}$$

While C_{of} affects gate charges $Q_{s,of}$ and $Q_{d,of}$ in both strong inversion, C_{if} affects gate charge through $Q_{s,if}$ and $Q_{d,if}$ only in weak inversion. The screening of C_{if} in strong inversion is achieved through a screening function as shown in (4.39)-(4.40). The gate capacitances are appropriate derivatives of these charges.

$$C_{gs} = -\frac{\partial(Q_s + Q_{s,if} + Q_{s,of})}{\partial V_g} \tag{4.41}$$

$$C_{gd} = -\frac{\partial(Q_d + Q_{d,if} + Q_{d,of})}{\partial V_g} \tag{4.42}$$

$$C_{gg} = C_{gs} + C_{gd} \tag{4.43}$$

Chapter 5

Experimental verification of MVS-G-RF model

The MVS-G-RF model proposed in the previous chapter is validated here against device measurements. The chapter begins with the validation of the access region model against gateless structures. This is followed by validation of the model for transistor currents and their derivatives against measurements. S-parameter derived gate capacitance measurements are then compared against charge-derived model capacitances. The last section of the chapter deals with device performance projection in terms of transconductance and unity-current gain frequency.

5.1 Validation of access region model

The implicit-gate transistor model for access regions is tested using transmission line method (TLM) structures. TLMs are ungated resistive structures of different lengths that are typically used to extract sheet resistance and contact resistance. Ideally TLMs are expected to behave like resistors with linear current-voltage characteristics. However, the characteristics obtained from measurements show non-linear behavior with currents displaying quasi-saturation at relatively low voltages. Reports in literature attribute this behavior to device self-heating [38]. The devices fabricated at MIT are on SiC substrate and self-heating is expected to be low (if not absent) at bias

conditions at which current quasi-saturation occurs. With the implicit-gate model, the non-linear behavior of access regions is attributed to pinch-off and or velocity saturation in these regions.

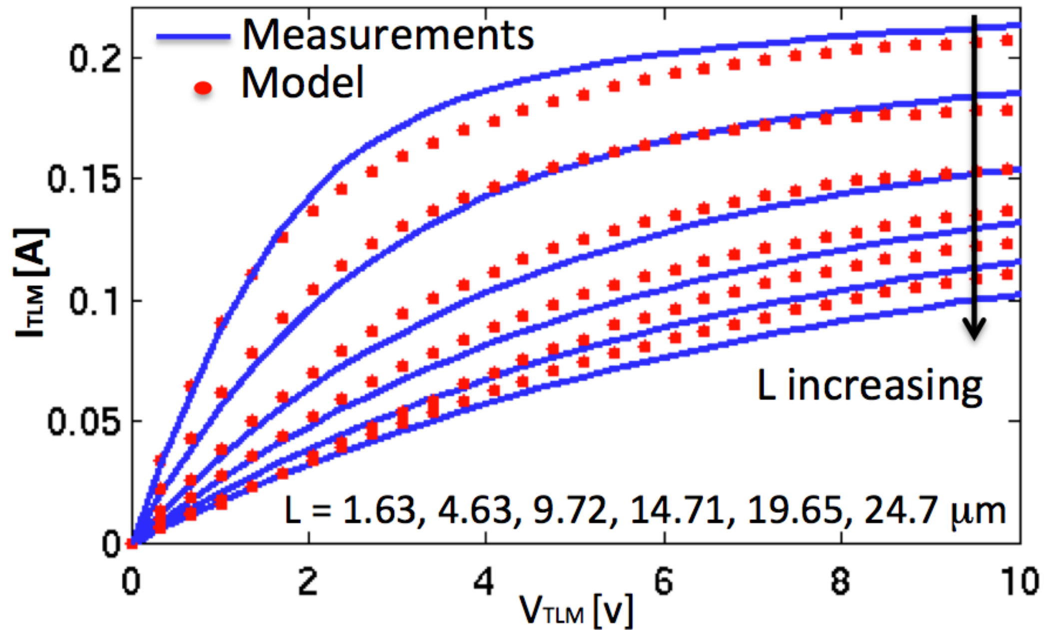

Figure 5-1: Comparison of current-voltage characteristics of TLMs: Model vs. measurements. Parameters except L were left unchanged in the implicit-gate transistor model.

TLMs are measured using an Agilent 4155 DC prober. These structures are on the same die as active gated transistors. The current voltage characteristics that are obtained for different length structures are shown in Fig.5-1. Also shown in the figure are currents predicted from the model. Good match is observed between model and measurements. The key point to note here is that for the model, no parameter other than TLM length (L) is changed to get the fits. The extraction procedure for key model parameters is given in Appendix A and the list of parameters used in the model is given in Appendix B. This ensures model scalability and confirms the physics of the model.

5.2 Validation of MVS-G-RF DC model

In this section, the comparison of terminal currents and their derivatives obtained from the model is validated against experimental devices. To ensure model scalability, two devices on the same die (as that of TLM structures) but with different gate lengths are measured. First, shown in Fig. 5-2 is the output, transfer, output conductance and transconductance of 105 nm gate length device. The model and measurements show good match.

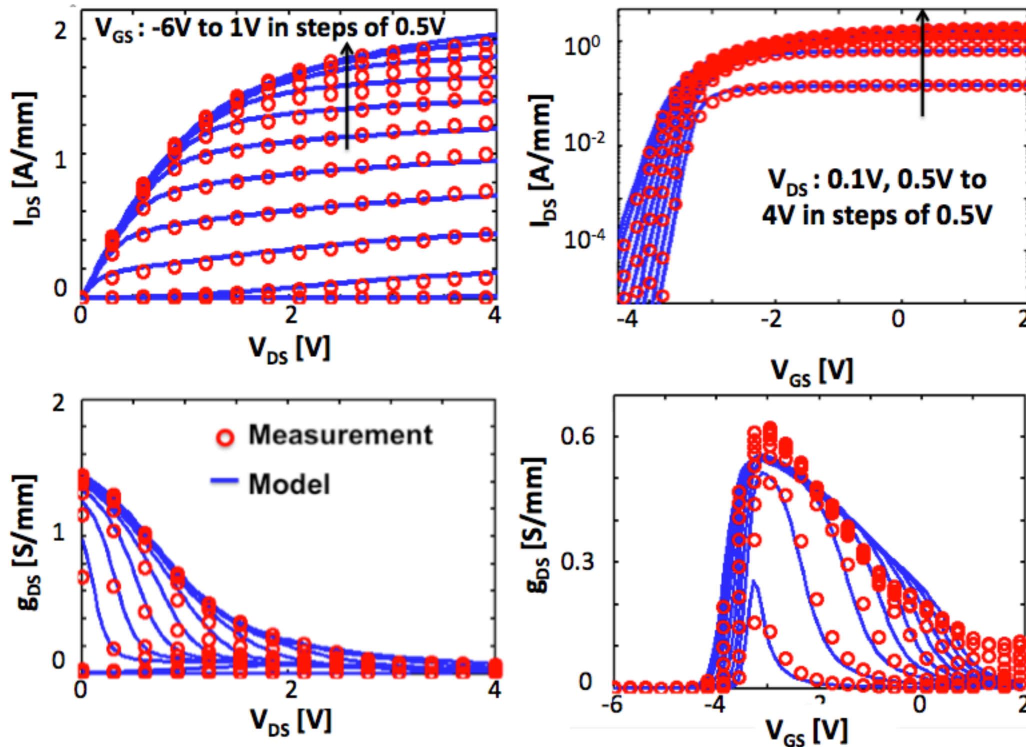

Figure 5-2: Comparison of current-voltage characteristics of 105 nm gate length GaN HEMT: Model vs. measurements.

The output characteristics show an upper bound due to the non-linear behavior of access regions. This is captured by the model. In addition, the transconductance plot is highly non-linear. Presently this is both due to access regions and velocity reduction in the channel due to carrier scattering. The list of parameter values used in the model is given in Appendix B and the extraction procedure of these parameters is highlighted in Appendix A.

The model is also compared against a 42 nm gate length device. This is shown in Fig. 5-3. With exception to parameters that govern electrostatics (DIBL, SS, n), the parameters used for this fit are the same as those used for the 105 nm gate length device (refer Appendix. B). The 42 nm gate length device shows significant short channel effects such as DIBL and modest punch-through. The model fits well with both devices indicating correctness of physics.

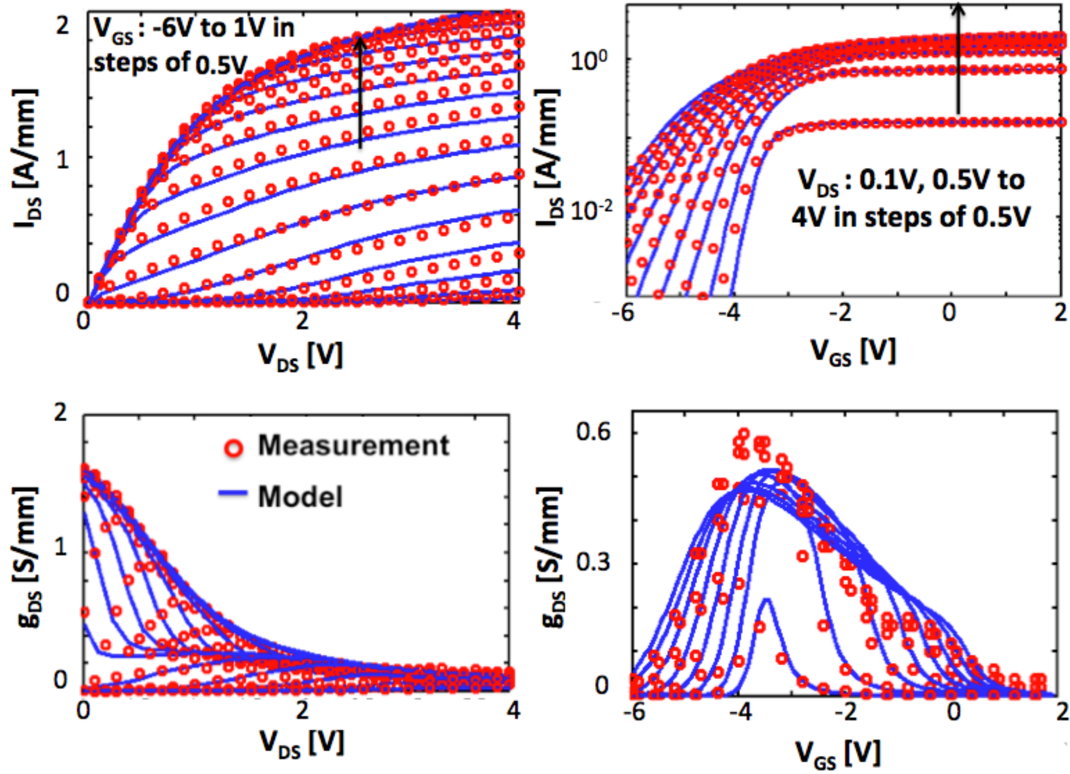

Figure 5-3: Comparison of current-voltage characteristics of 42 nm gate length GaN HEMT: Model vs. measurements.

The importance of access region non-linear behavior on device characteristics can be best understood if the model with access regions modeled as linear resistors is compared against device measurements as given in Fig. 5-4. The mismatch in output characteristics at high V_{GS} with the model predicting much higher on-currents is due to incorrect access region behavior. Pinch-off and velocity saturation in access regions limit the maximum current in these devices. In addition, these regions also increase the non-linearity in g_m as can be seen from the figure.

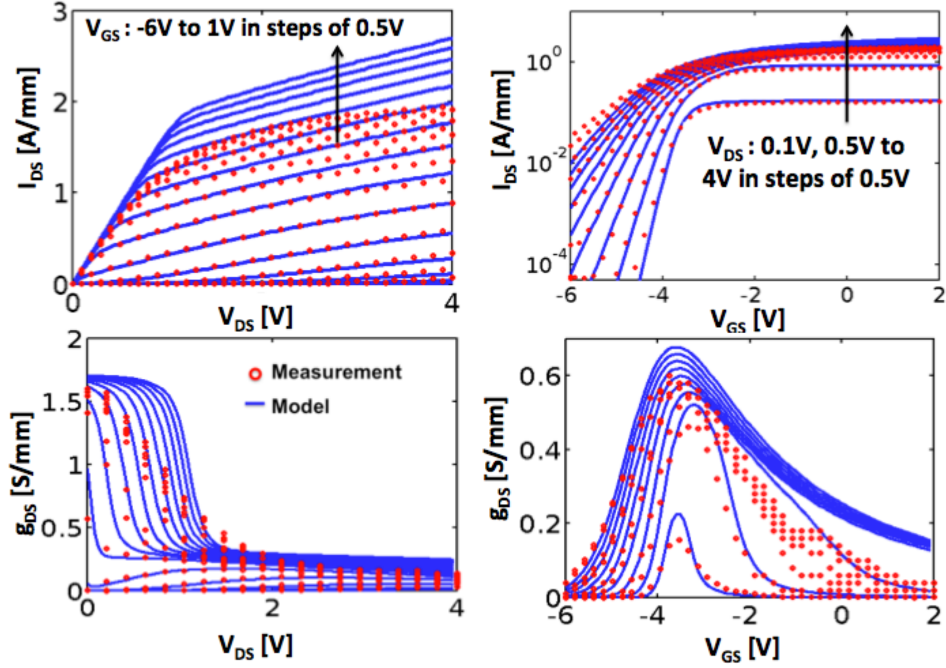

Figure 5-4: Comparison of current-voltage characteristics of 42 nm gate length GaN HEMT: Model with linear access resistors vs. measurements. Model mismatch illustrates the significance of access regions on device behavior.

In our previous work in [38], the access region non-linearity was attributed to velocity saturation along with self-heating and the regions were modeled as semi-empirical non-linear resistors instead of implicit-gate transistors. Since the devices are on SiC substrate which is a good thermal conductor, self-heating is not the main cause for the non-linear behavior and the incorrect assumption lead to the use of high scattering coefficient for velocity (θ_v) to get the currents from the model to match measurements. Implementing access regions as implicit-gate transistor model results in good match between model and measurements (Figs. 5-2 -5-1) with much lower scattering coefficient θ_v than earlier. θ_v has reduced to 0.05 (in the model with implicit-gate transistors) from earlier value of 0.3 (in the model with non-linear access resistors) indicating that while velocity reduction due to electron-phonon interaction could be present, it is not the main cause for g_m non-linearity which is due to pinch-off and velocity saturation in the access regions.

5.3 Validation of charge model

5.3.1 S-parameter measurements

In addition to DC currents, gate capacitance from the model can be compared against measurements. Since the devices are highly scaled with small widths ($25\mu m$), the channel capacitances are masked by parasitic components such as pad and fringing capacitances. Therefore proper de-embedding is necessary to get meaningful gate capacitance values. Since the values of gate capacitances are of the order of fF, S-parameter measurements are used to obtain them.

The S-parameters of the transistors are characterized from 100 MHz to 45 GHz by using an Agilent N5250C network analyzer. The system is calibrated with an off-wafer line-reflect-match calibration standard. The measured S-parameters are de-embedded using on-wafer open and short test structures as described in [10]. The equivalent circuit assumed to extract gate-capacitances from the de-embedded S-parameters is given in Fig.5-5.

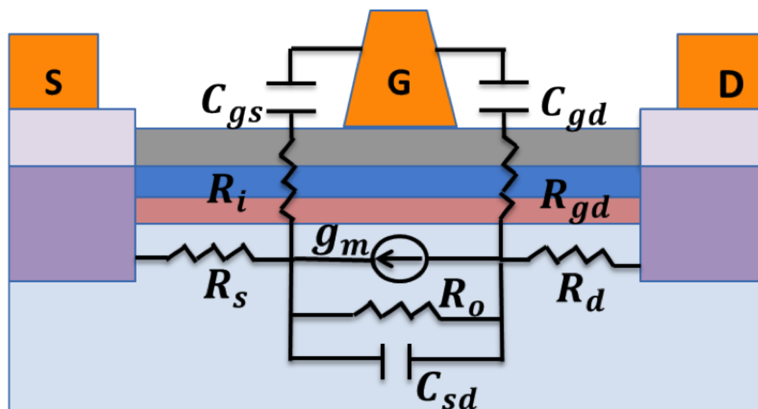

Figure 5-5: Equivalent circuit assumed for gate capacitance extraction from S-parameters (after [10]).

The de-embedded S-parameters (with effects of pad capacitances and terminal resistance and inductance parasitics removed) are converted to Y-parameters. The key circuit elements can be then extracted from the resulting Y-parameters as given by [10]:

$$C_{gs} = \frac{Im(Y_{11}) + Im(Y_{12})}{\omega} \left(1 + \left(\frac{Re(Y_{11}) + Re(Y_{12})}{Im(Y_{11}) + Im(Y_{12})}\right)^2\right) \tag{5.1}$$

$$C_{gd} = -\frac{Im(Y_{12})}{\omega} \left(1 + \left(\frac{Re(Y_{12})}{Im(Y_{12})}\right)^2\right) \tag{5.2}$$

$$C_{gg} = C_{gs} + C_{gd} \tag{5.3}$$

$$C_{ds} = \frac{Im(Y_{22}) + Im(Y_{12})}{\omega} \tag{5.4}$$

$$g_o = Re(Y_{22}) + Re(Y_{12}) \tag{5.5}$$

$$g_m = \sqrt{(Re(Y_{21}) - Re(Y_{12}))^2 + (Im(Y_{21}) - Im(Y_{12}))^2} \tag{5.6}$$

5.3.2 Extraction of fringing capacitances

The gate capacitances that are obtained from S-parameters include fringing capacitances in addition to gate-to-channel capacitances. These must be extracted and included in the model for proper description of the dynamic device behavior. There are two types of fringing capacitances as shown in fig.5-6:

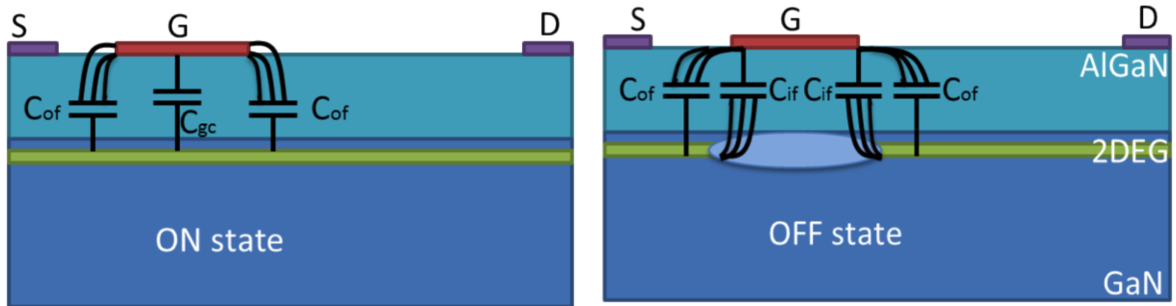

Figure 5-6: Schematic showing different capacitances in on and off states. The channel screens C_{if} in on-state.

(i). Outer fringing capacitances (C_{of}): These are capacitances that are due to the fringing fields from the gate metal to the 2DEG in the access regions. The fringing fields and the capacitance exist in both strong and weak inversion. In the MVS-G-RF model, the outer fringing capacitance is a fixed parameter added to the gate capacitance.

(ii) Inner fringing capacitances (C_{if}): These are capacitances present only in the sub-threshold (off-state) when there is no significant charge in the channel. The inner fringing capacitance is due to the inner fringing field from the gate metal to the access region-2DEG through the channel region. The inner fringing fields are screened by channel-2DEG in the on-state. To mimic this behavior the MVS-G-RF model uses screening functions in inner fringing charges as given in (4.39)-(4.40).

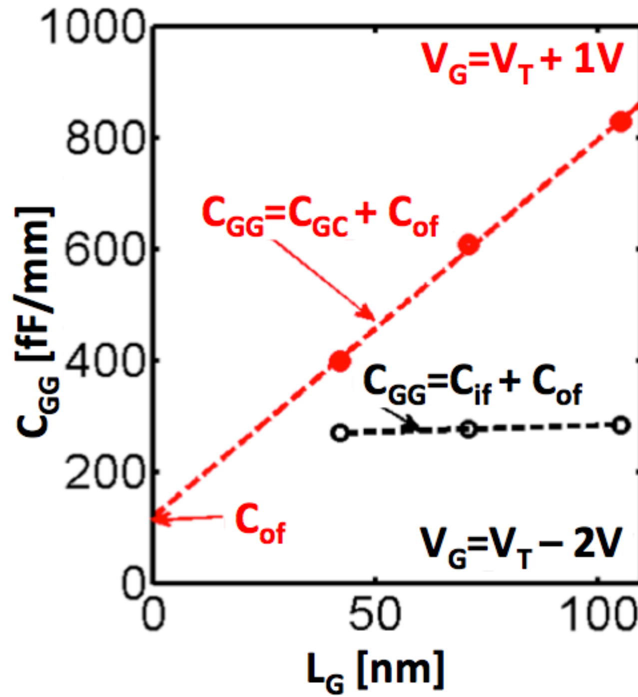

Figure 5-7: C_{if} and C_{of} extraction procedure from gate capacitance measurements.

The measured gate capacitances are plotted as a function of gate length as shown in Fig.5-7, to extract the values of the fringing capacitances for the devices under study. In the on-state ($V_{gs} = V_T + 1V$), the measured gate capacitance is the sum of channel (C_{gc}) and outer fringing capacitance (C_{of}). The intercept gives the value

of C_{of} and from the slope, the areal gate capacitance (C_{inv}) can be known. In the off-state ($V_{gs} = V_T - 2V$), the measured capacitance is the sum of outer (C_{of}) and inner fringing capacitance (C_{if}) and with the extracted value of C_{of}, the value of C_{if} can be known. The gate capacitance validation is then done.

5.3.3 Gate capacitance: model vs. measurements

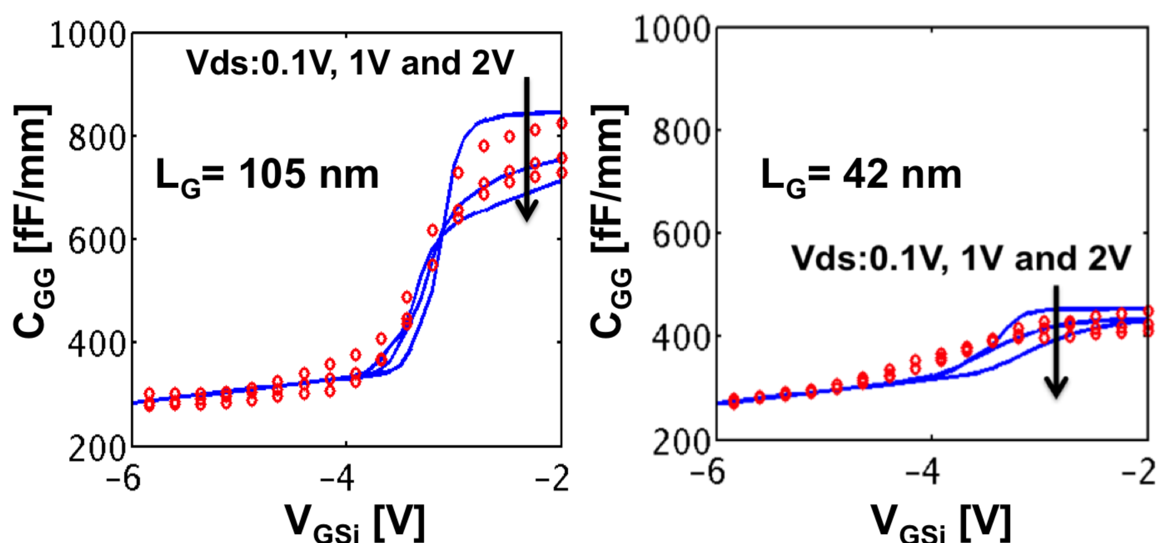

Figure 5-8: Gate capacitance comparison for 105-nm and 42-nm devices: model vs. measurements.

The measured gate capacitance (C_{gg} of (5.3)) is compared against model gate capacitance (C_{gg} of (4.40) with added fringing capacitance). This is shown in Fig.5-8 for 105-nm and 42-nm gate length devices. The model gives a reasonable match with measurements confirming the validity of the gate charge model. Since the capacitance in the on-state is dominated by the channel capacitance, which scales with L_g, the gate capacitance for the 42-nm device is lower than that of the 105-nm device. The off capacitance in this case was dominated by C_{of}. The list of parameters used for these fits are tabulated in Appendix B. The parameter γ used for capacitance fits turns out to be 0.2 which means that there is significant carrier scattering in the channel as

indicated by the low carrier mean free paths. Once the currents and charges of the device are modeled, performance metrics such as unity-current-gain-frequency (f_T) and transconductance (g_m) can be measured and compared against the model. This is given in the following sections.

5.4 Unity current gain frequency (f_T)

f_T is one of the important figures of merit of scaled GaN HEMTs used in RF applications. f_T is obtained from measurements by plotting the magnitude of current gain (h_{21}) as a function of frequency and extrapolating to 0 dB as shown in Fig.5-9. These devices were optimized for maximum f_T. The 105-nm device has f_T of about 152 GHz and the 42-nm device has f_T of about 210 GHz. From the equivalent circuit model of Fig.5-5, f_T can be estimated using the expression:

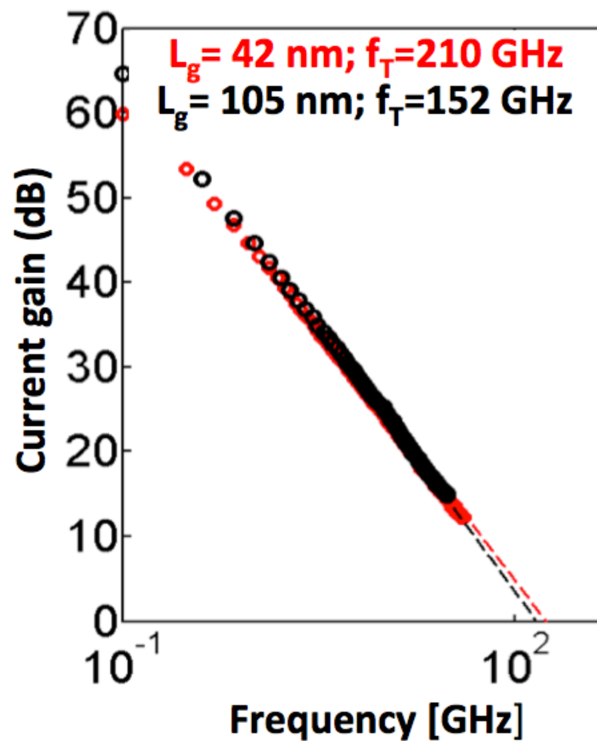

Figure 5-9: Extraction of f_T from S-parameter measurements.

$$f_T = \frac{1}{2\pi} \frac{g_{m,int}}{C_{gs} + C_{gd} + g_{m,int}\left(R_S + R_D\right)\left(1 + \frac{g_{o,int}}{g_{m,int}}\left(1 + \frac{C_{gs}}{C_{gd}}\right)\right)} \tag{5.7}$$

where $g_{m,int}$ is the intrinsic transconductance obtained after correcting for access region voltage drops, $g_{o,int}$ is the intrinsic output conductance, C_{gs} and C_{gd} are the gate capacitances described in previous sections. R_S and R_D are the access region and contact resistances. At bias conditions corresponding to peak f_T (in this case, $V_{GS} = -3V, V_{DS} = 4V$), the access regions behave as linear resistances. The f_T from model is compared against measured f_T for the two devices (lot 1) in Fig.5-10. In addition, a second set of devices (lot 2) with slightly lower fringing capacitances are also used for f_T comparison. These f_T values are also given along with modeled f_T values as a function of gate length. Peak f_T of about 300 GHz is achieved in this technology by gate length scaling to 23 nm. The measured and modeled f_T give a good match thus setting baseline for making technology projections in terms of f_T.

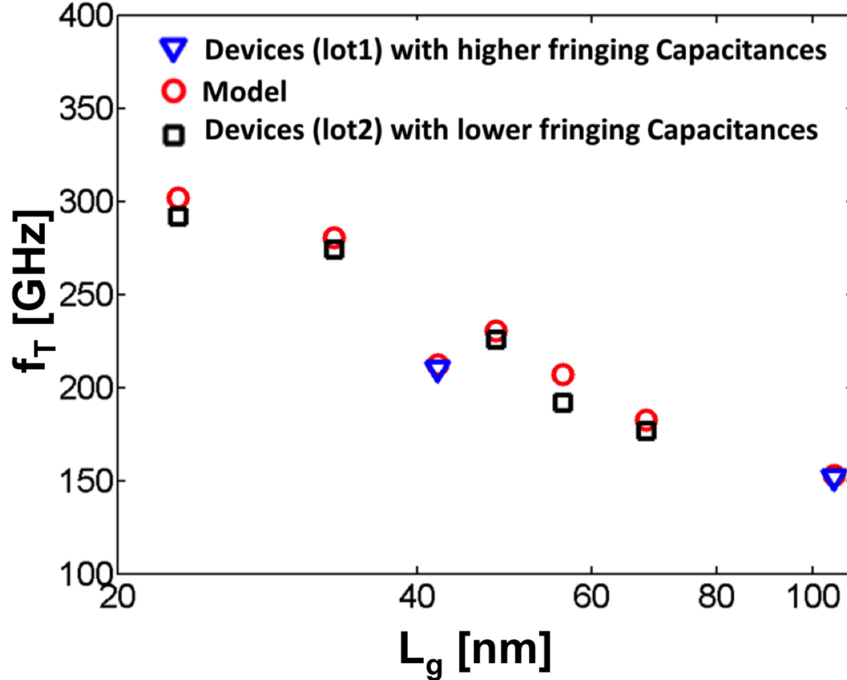

Figure 5-10: f_T values with scaling of gate length: model vs. device measurements. Good match is achieved for both lot1 and lot2 devices.

5.5 Technology projections: g_m and f_T

5.5.1 g_m projection

The transconductance (g_m) is an important device performance metric for RF-GaN HEMTs. This parameter relates to the possible signal-gain from the device and the device's use in an amplifier. The nature of g_m also affects the output signal quality. In Fig.5-11, the impact of access regions and contact resistances on peak g_m in the device is described. The value of g_m has direct bearing on the value of f_T and maximum speed of operation of the device.

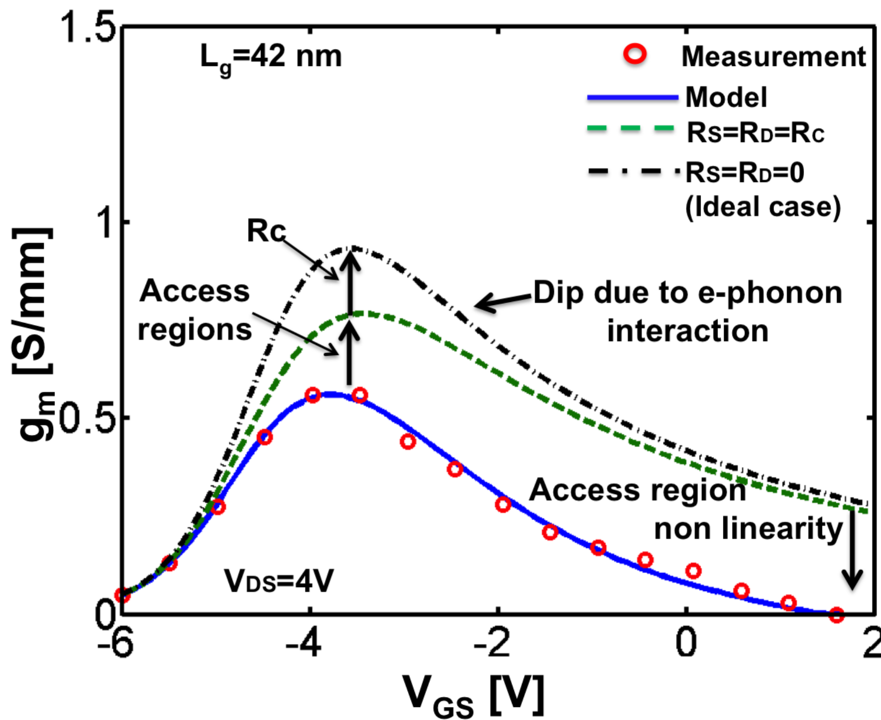

Figure 5-11: Technology projections on g_m. The projections show significant improvement in peak g_m with access region scaling and reduction of R_c.

Contact and access region resistances have significant impact on peak transconductance ($g_{m,peak}$) as shown in Fig. 5-10. The model shows 66 percent reduction in peak g_m (blue solid line) compared to peak intrinsic, $g_{m,int}$ (black dot-dash line). Thus without any other changes, reduction of contact resistance and gate self-alignment

would be sufficient to significantly increase g_m. However, sharp drop in g_m above certain V_G is due to both non-linear source access region behavior and due to Q_{ixo}-dependent v_{xo}, which could be due to interface scattering and optical phonon interaction [39]-[40]. The latter is a GaN material-specific phenomenon and it might not be possible to eliminate this through device scaling [40].

5.5.2 f_T projection

The maximum speed of operation of the device is characterized by the parameter f_T and is an important FoM for RF applications. From the model, it is possible to study the impact of the various technology innovations that could result in improved f_T in the device without having to resort to L_g scaling. This is explained in of Fig. 5-12.

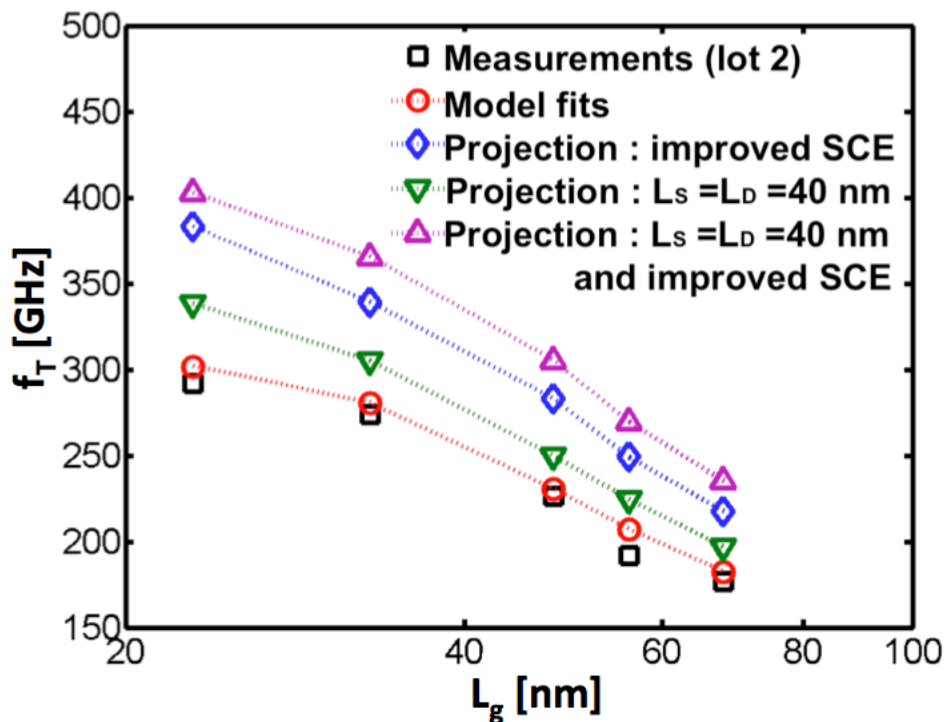

Figure 5-12: Technology projections on f_T. A significant boost in f_T is possible through technology advances such as reduction of SCE and access region scaling.

As shown in Fig. 5-12, scaling-down source access region (L_S) and drain access region (L_D) and suppressing short channel effects (SCE) can both effectively improve

f_T. Scaling-down L_S and L_D reduces R_S and R_D, which essentially decreases parasitic delay and thus improves f_T. Suppressing SCE, on the other hand, enhances f_T by increasing $g_{m,int}$ while reducing g_o. With $L_S + L_D = 1\mu m$, and poor SCE (DIBL = 380 mV/V and SS=260 mV/dec) at $L_g = 23\ nm$, the base-line technology (lot 2) provides f_T of 290 GHz. Scaling-down L_S and L_D to 40 nm, which can be potentially achieved by adopting self-aligned structure [41], would help in boosting f_T to 340 GHz. Recent technology advancements such as use of back barriers have resulted in suppression of SCE [25]. Using the model and assuming improved DIBL=150 mV/V and SS=150 mV/dec, f_T can be increased to 380 GHz. Devices combining the two improvements, are expected to show f_Ts above 400 GHz at $L_g = 23nm$.

The above discussions illustrate that the physics behind the MVS-G-RF model enables us to make technology projections to evaluate possible ways to improve device performance. Key results with regard to the work on the MVS-G-RF model were presented in this chapter. Model validation for DC and RF device characteristics was followed by two examples on possible technology projections using the MVS-G-RF model. In the concluding chapter that follows, summary of the work done in this thesis is given along with possible ways to improve the current version of the model. In addition, further work needed to incorporate other GaN-specific effects is discussed.

Chapter 6

Summary and future work

The MVS-G-RF model developed in this work consists of a compact model for terminal currents and charges for scaled GaN HEMTs. It describes the access region behavior along with the intrinsic transistor in these devices and is able to predict device metrics such as g_m and f_T accurately. In addition, it could be used to make technology projections and to highlight bottlenecks to device performance. The model is implemented in Verilog-A and is therefore suitable for circuit simulations.

The following points highlight further work that is needed to improve the model.

- The model in its current form extracts threshold voltage and electrostatic parameters such as SS, DIBL etc. and uses them as parameters. This limits the use of the model when required to predict device performance of scaled gate length devices. To enable this, we need model equations for electrostatic parameters if geometry and other device specific information are given.

- The self-heating model that is included in its present form is semi-empirical and is a static model. The thermal coefficients for physical quantities such as mobility and VS-velocity need to be validated. In addition, an RC thermal network that accounts for the thermal resistance (R_{th}) and thermal capacitance (C_{th}) are needed to capture dynamic heating effects and also so that these thermal properties can be independently calculated from device geometry and substrate information rather than being left as fitting parameters.

- The Schottky gate leakage model needs to be validated and also a gate leakage model in the presence of gate dielectric needs to be developed. This would enable estimation of f_{max} in addition to f_T.

- The parasitic terminal capacitances, inductances and resistances must be included in the model to account for the magnitude and phase shifts of input/output signals between the device and terminals. The parasitic Kondoh model [42] is currently being tried in this regard.

- A key issue with GaN HEMTs is that of charge trapping. This was mentioned in chapter 2. In the current form MVS-G-RF model does not account for charge trapping and the associated current collapse. This has to be included to enable meaningful RF circuit simulations. Pulsed IV measurements are necessary to study this effect.

- The model can be extended to long channel regime to model GaN HEMTs suited for high voltage applications. This version of the model is christened MVS-G-HV model and is currently being developed to model HV switching devices along with options to include field plates.

Appendix A

Model parameter extraction

The following flowchart shows the extraction procedure and extraction sequence of important parameters of MVS-G-RF model.

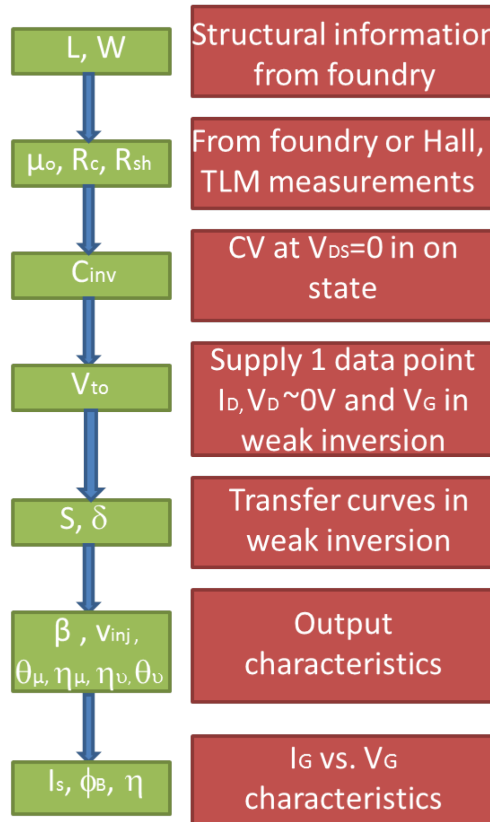

Figure A-1: Flowchart showing parameter extraction sequence of important MVS-G-RF model parameters.

This is not the exhaustive list of parameters but the most significant ones. Most of the other parameters are either fitting parameters or constants for GaN HEMT.

A.1 Device parameters

Geometry and structural information are best provided by foundry. For the model, geometry parameters that are needed are: Gate length (L_g), Source access region (L_{gs}), drain access region length (L_{gd}), Device width (W). In addition, parameters related to field plate such as field plate length, inter-layer dielectric thickness etc. might be necessary for future modeling. Other additional useful parameters required by the model are: Low field mobility (μ_o), Contact resistance (R_c) and sheet resistance (R_{sh}), 2DEG density. If these are not provided, additional measurements might be needed to extract them. μ_o and R_{sh} extraction would require Hall measurements and special hall structures. R_c can be extracted by measuring resistances of TLM structures of different lengths and extracting the offset at $L_g = 0$. Correct extraction of these parameters can also be verified from R_{on} match in output characteristics.

A.2 C_{inv} extraction

C_{inv} is an important model parameter in MVS-G-RF model. Its accurate extraction is essential for correct modeling. C_{inv} must be extracted from CV measurements rather than from analytical calculation. Accurate analytical calculation must also include quantum correction as charge centroid in 2DEG is shifted away from the interface. This needs dedicated calculations/simulation, which from a compact modeling perspective might not be critical as long as we can directly measure the capacitance. To get C_{inv}, two-terminal CV (with $V_{DS} = 0$) of devices with different gate length but identical widths and access region lengths must be measured. The gate to channel capacitance in strong inversion scales as a function of L_g. From the slope we get the value of C_{inv} (areal gate capacitance) and the intercept gives the parasitic capacitance. Parasitic capacitance includes outer fringing capacitance (C_{of}) and pad

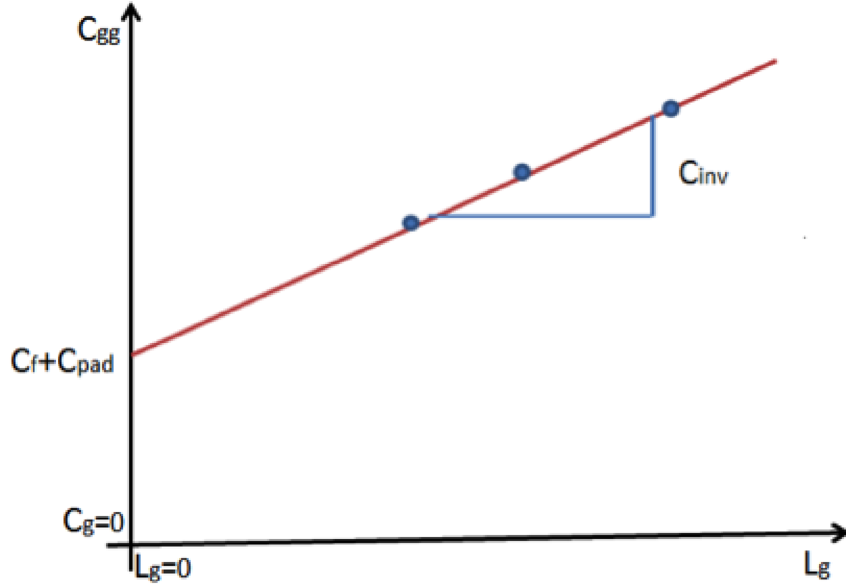

Figure A-2: Extraction of C_{inv} from gate capacitance vs. gate length measurements.

parasitic. To remove the pad parasitic, we need CV of open test structures.

A.3 Extraction of V_{To}, S, n_d and δ

V_T computation requires knowledge of piezoelectric charges at the interface, heterostructure composition and thickness. Again the model is simplistic in the sense that V_{To} needed for the model can be extracted using a simple Matlab routine. The input to the routine is one data point (V_G, I_D, V_D) in weak inversion (just beyond strong to weak inversion transition) at low V_D ($\approx 0V$ where DIBL has negligible impact). Alternately V_{To} can be approximated as V_G at the same (V_G, I_D, V_D) point on transfer curve. Sub threshold slope (S) is obtained from the slope of the transfer curve on log scale in weak inversion regime at low V_D ($\approx 0V$). Low V_D is preferred to avoid shift of S due to modest punch through (a-dependence on V_{DS}) in the device. The parameter extracted must make sense for the L_g of the device. The punchthrough factor (n_d) is a fitting parameter to account for increasing S with V_{DS}. This is fitted to get the 'spread' in $I_D - V_G$ in sub-threshold regime as shown in Fig. A-3. DIBL is extracted from the lateral shift of I_D (due to shift of V_T) as V_D is increased in the

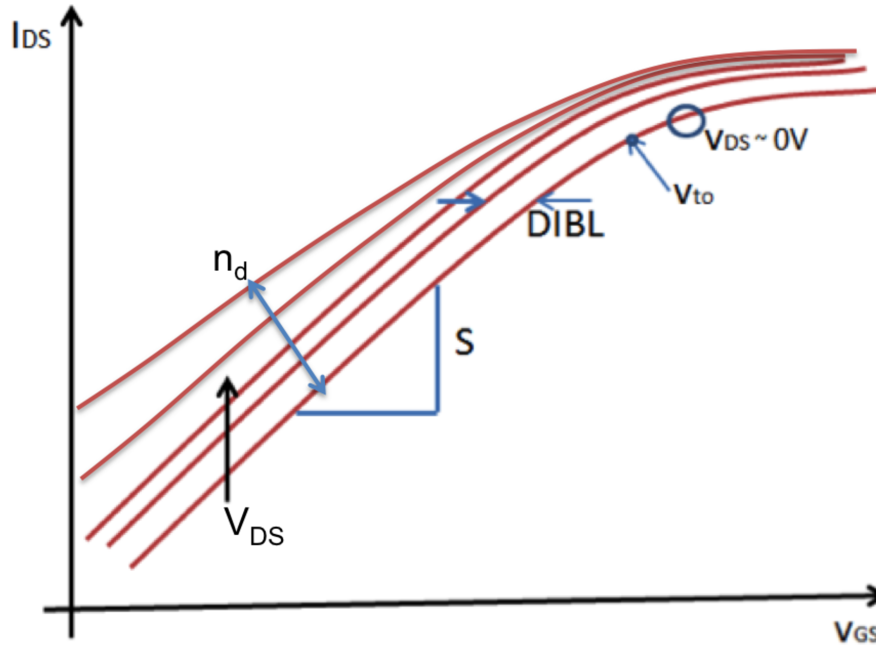

Figure A-3: Extraction of V_{to}, S and δ from subthreshold regime of transfer curves.

transfer curves in weak inversion. δ is multiplied by intrinsic drain voltage (V_{Di}) in the expression for threshold voltage in the model.

A.4 Extraction of v_{inj}, β and θ_v, θ_μ, η_v and η_μ

v_{inj}, β and θ_v, θ_μ, η_v and η_μ can be extracted from output characteristics. v_{inj} can be fitted to get accurate match with saturation current level at low V_{GS} where other effects such as self-heating and scattering are minimum. v_{inj} extracted must lie between the bracket of peak electron velocity ($2.5 \times 10^7 cm/s$) and saturation velocity ($1.3 \times 10^7 cm/s$) depending on gate length. The transition from linear to saturation current in output characteristics is governed by F_{sat} which has a parameter β. β should ideally lie between 1.5-3 for these type of HEMTs depending on saturation. A larger value of β is needed for more abrupt kind of saturation.

η_v and η_μ are self-heating parameters affecting μ and velocity respectively and can be extracted from slope of output curves in saturation at large V_G where self-heating

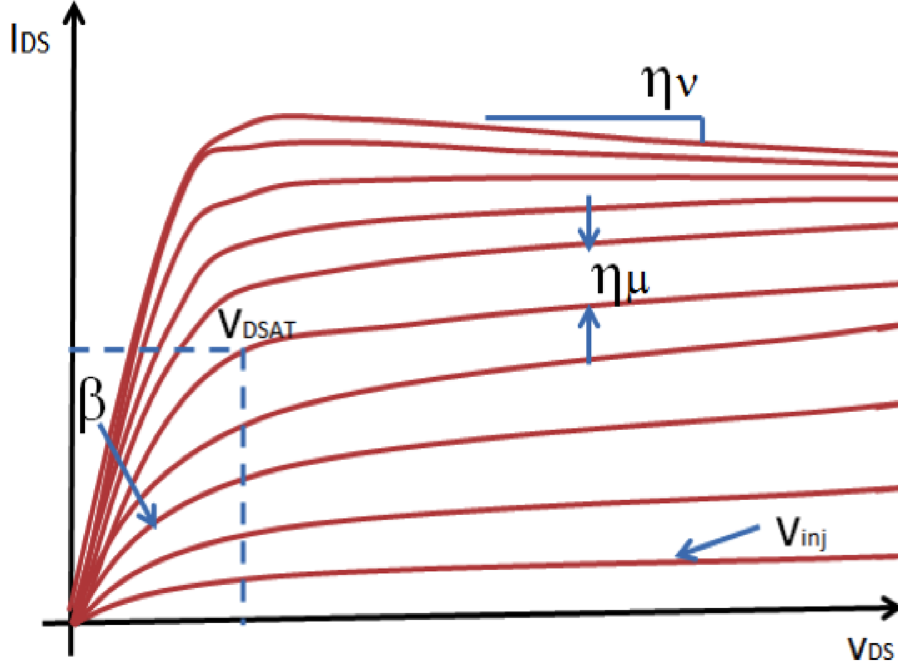

Figure A-4: Extraction of v_{inj}, β , η_v and η_μ from output characteristics.

is dominant. η_v is directly responsible for the negative slope of the output curves and can be extracted from fitting. η_μ reduces mobility and hence increases V_{DSAT}. θ_v and θ_μ are also fitting parameters which affect V_{DSAT} and can be extracted from V_{DSAT} at lower V_{GS} when self-heating has not yet kicked in. Thus by fitting to get correct linear-to-saturation transition voltages at larger V_{GS} we can get value of η_μ. Typically both η_v and η_μ are not fitting parameters. η_v and η_μ are directly dependent on thermal resistance (R_{th}) of the thermal network and thermal coefficient of mobility and velocity respectively. Since R_{th} characterization through TCAD and evaluation of thermal coefficients through multi-temperature measurements has not been done yet, η_v and η_μ have been reduced to fitting parameters for now.

Extraction of I_S, ϕ_B and η

The gate diode current model is not yet verified but the extraction procedure is given here. The important gate-leakage parameters can be extracted from $I_G - V_G$ characteristics at different V_{DS}. I_S is the reverse saturation current of the diode and

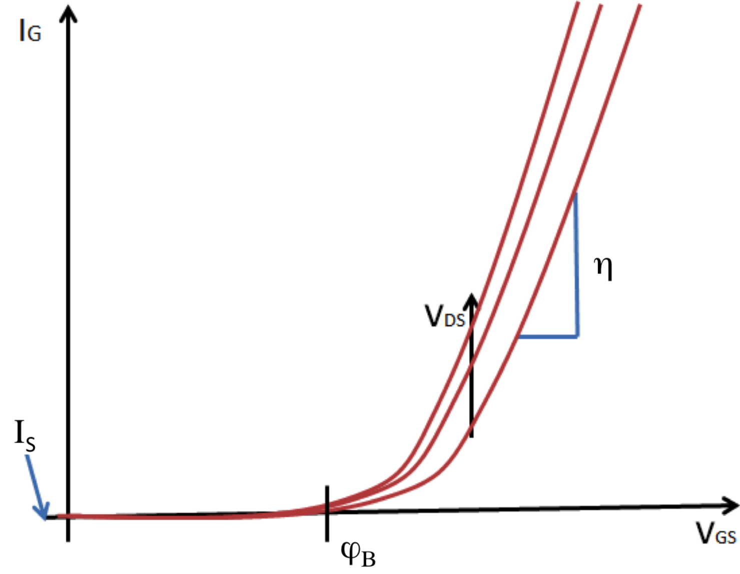

Figure A-5: Extraction of gate leakage model parameters from gate current characteristics.

can be extracted from reverse bias region of $log|I_G| - V_G$ characteristics at $V_D = 0$. η (ideality factor) and ϕ_B (Schottky barrier height) are material dependent properties and are fixed if the heterostructure composition is fixed. Typically ϕ_B for AlGaN-GaN heterostructure systems is about 1.2eV. η can be extracted from the slope of the linear region of the forward-biased characteristics of gate current plot shown in Fig. A-5. accurate extraction of ϕ_B can be done from multi-temperature measurements of reverse saturation current (I_S).

Appendix B

Model parameters

B.1 Model parameters for TLMs

Parameters	Values	Description
Geometric parameters		
L (μm)	1.63, 4.63, 9.72, 14.71, 19.65, 24.7	TLM length
Parameters from measurements and device characteristics		
R_{sh} ($\Omega/sq.$)	205	Sheet resistance of access regions
R_c ($\Omega - mm$)	0.12	Contact resistance (per-side)
μ_o (cm^2/Vs)	1650	Low-field mobility
Fitting parameters (Generic VS model)		
α	3.5	Sub-threshold transition parameter
δ ((V/V))	0	Drain-induced-barrier-lowering
S (V/Dec)	0.1	Sub-threshold swing
Fitting parameters (MVS-G-RF model specific)		
v_{xo} (cm/s)	1.6e7	Velocity at VS point
β	1.5	Lateral-field-saturation-parameter
ϵ	2.3	Mobility self-heating parameter
β	1.5	Lateral-field-saturation-parameter
C_{invi} (F/cm^2)	1.5e-7	Areal implicit-gate capacitance

B.2 Model parameters for GaN HEMTs

Parameters	105 nm	42 nm	Description
Geometric parameters			
L_g (nm)	42	105	Channel length
L_{gs} (μm)	0.56	0.46	Source access region length
L_{gd} (μm)	0.43	0.46	Drain access region length
Parameters from measurements and device characteristics			
C_{inv} (F/cm^2)	6.5e-7	6.5e-7	Areal gate capacitance
R_{sh} $(\Omega/sq.)$	205	205	Sheet resistance of access regions
R_c $(\Omega - mm)$	0.24	0.24	Contact resistance (per-side)
μ_o (cm^2/Vs)	1650	1650	Low-field mobility
V_{To} (V)	-3.3	-3.0	Threshold voltage for $V_D \approx 0\ V$
C_{of} (fF/mm)	140	140	Outer fringing capacitance
C_{if} (fF/mm)	130	144	Inner fringing capacitance
Fitting parameters (Generic VS model)			
α	3.5	3.5	Sub-threshold transition parameter
δ $((V/V))$	0.31	0.12	Drain-induced-barrier-lowering
S (V/Dec)	0.26	0.14	Sub-threshold swing
n_d	3.3	0.4	Punch through factor
Fitting parameters (MVS-G-RF model specific)			
v_{xo} (cm/s)	1.35e7	1.4e7	Velocity at VS point
β	1.5	1.5	Lateral-field-saturation-parameter
ϵ	2.3	2.3	Mobility self-heating parameter
β	1.5	1.5	Lateral-field-saturation-parameter
η_μ	0.1	0.1	Mobility self-heating parameter
η_v	3e-3	1e-3	Velocity self-heating parameter
θ_μ	0.0	0.0	Vertical field dependent mobility parameter
θ_v	0.05	0.05	Vertical field dependent velocity parameter
γ	0.19	0.19	Energy loss parameter for capacitance
C_{invi} (F/cm^2)	1.5e-7	1.5e-7	Areal implicit-gate capacitance

Bibliography

[1] I. Akasaki, "Nitride semiconductors-impact on the future world," *Journal of Crystal Growth*, vol. 237-239, Part 2, no. 0, pp. 905 – 911, 2002. The thirteenth international conference on Crystal Growth in conjunction with the eleventh international conference on Vapor Growth and Epitaxy.

[2] U. Mishra, P. Parikh, and Y.-F. Wu, "AlGaN/GaN HEMTs-an overview of device operation and applications," *Proceedings of the IEEE*, vol. 90, pp. 1022 – 1031, Jun 2002.

[3] M. A. Briere, "Gan-based power devices offer game-changing potential in power-conversion electronics," tech. rep., International Rectifier, December 30 2008.

[4] M. Feng, S.-C. Shen, D. C. Caruth, and J.-J. Huang, "Device technologies for RF front-end circuits in next-generation wireless communications," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 354–375, 2004.

[5] D. S. Lee, et al., "High frequency performance of GaN HEMTs at cryogenic temperatures," *Electron Devices, IEEE Transactions on*, vol. 51, pp. 1432 – 1440, Sept. 2004.

[6] I. Angelov, H. Zirath, and N. Rosman, "A new empirical nonlinear model for HEMT and MESFET devices," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 40, no. 12, pp. 2258–2266, 1992.

[7] W. R. Curtice and M. Ettenberg, "A nonlinear gaas fet model for use in the design of output circuits for power amplifiers," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 33, no. 12, pp. 1383–1394, 1985.

[8] C. William, *Small and large signal modeling of mm-wave HEMT devices.* PhD thesis, University of South Florida, 2003.

[9] L. Dunleavy, C. Baylis, W. Curtice, and R. Connick, "Modeling GaN: Powerful but challenging," *Microwave Magazine, IEEE*, vol. 11, no. 6, pp. 82–96, 2010.

[10] J. W. Chung, *Millimeter-wave GaN high electron mobility transistors and their integration with silicon electronics.* PhD thesis, Massachusetts Institute of Technology, 2011.

[11] E. Cho, *Gallium Nitride Based Heterostructure Growth and Application to Electronic Devices and Gas Sensors*. ProQuest, 2009.

[12] H. Maruska and J. Tietjen, "The preparation and properties of vapor-deposited single-crystal-line GaN," *Applied Physics Letters*, vol. 15, no. 10, pp. 327–329, 1969.

[13] L. Dunleavy, C. Baylis, W. Curtice, and R. Connick, "Modeling GaN: Powerful but challenging," *Microwave Magazine, IEEE*, vol. 11, pp. 82 –96, Oct. 2010.

[14] T. Mimura, S. Hiyamizu, T. Fujii, and K. Nanbu, "A new field-effect transistor with selectively doped GaAs/n-$Al_xGa_{1-x}As$ heterojunctions," *Japanese journal of applied physics*, vol. 19, pp. L225–L227, 1980.

[15] M. Asif Khan, A. Bhattarai, J. Kuznia, and D. Olson, "High electron mobility transistor based on a gan-al x ga 1-x n heterojunction," *Applied Physics Letters*, vol. 63, no. 9, pp. 1214–1215, 1993.

[16] T. Palacios, S. Rajan, A. Chakraborty, S. Heikman, S. Keller, S. DenBaars, and U. Mishra, "Influence of the dynamic access resistance in the gm and ft linearity of AlGaN/GaN HEMTs," *Electron Devices, IEEE Transactions on*, vol. 52, pp. 2117 – 2123, Oct. 2005.

[17] M. Singh and J. Singh, "Design of high electron mobility devices with composite nitride channels," *Journal of applied physics*, vol. 94, no. 4, pp. 2498–2506, 2003.

[18] B. Gelmont, K. Kim, and M. Shur, "Monte carlo simulation of electron transport in gallium nitride," *Journal of applied physics*, vol. 74, no. 3, pp. 1818–1821, 1993.

[19] G. Meneghesso, F. Rampazzo, P. Kordos, G. Verzellesi, and E. Zanoni, "Current collapse and high-electric-field reliability of unpassivated GaN/AlGaN/GaN HEMTs," *Electron Devices, IEEE Transactions on*, vol. 53, no. 12, pp. 2932–2941, 2006.

[20] H. Hasegawa, T. Inagaki, S. Ootomo, and T. Hashizume, "Mechanisms of current collapse and gate leakage currents in AlGaN/GaN heterostructure field effect transistors," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 21, no. 4, pp. 1844–1855, 2003.

[21] R. Vetury, N. Q. Zhang, S. Keller, and U. K. Mishra, "The impact of surface states on the DC and RF characteristics of AlGaN/GaN HFETs," *Electron Devices, IEEE Transactions on*, vol. 48, no. 3, pp. 560–566, 2001.

[22] G. Simin, A. Koudymov, A. Tarakji, X. Hu, J. Yang, M. A. Khan, M. Shur, and R. Gaska, "Induced strain mechanism of current collapse in AlGaN/GaN heterostructure field-effect transistors," *Applied Physics Letters*, vol. 79, no. 16, pp. 2651–2653, 2001.

[23] S. C. Binari, K. Ikossi, J. A. Roussos, W. Kruppa, D. Park, H. B. Dietrich, D. D. Koleske, A. E. Wickenden, and R. L. Henry, "Trapping effects and microwave power performance in AlGaN/GaN HEMTs," *Electron Devices, IEEE Transactions on*, vol. 48, no. 3, pp. 465–471, 2001.

[24] I. Khalil, *Intermodulation distrotion in GaN HEMT*. PhD thesis, Technical Universitat Berlin, Berlin, July 2009.

[25] D. S. Lee, X. Gao, S. Guo, and T. Palacios, "InAlN/GaN HEMTs with AlGaN back barriers," *Electron Device Letters, IEEE*, vol. 32, no. 5, pp. 617–619, 2011.

[26] D. S. Lee, J. Chung, H. Wang, X. Gao, S. Guo, P. Fay, and T. Palacios, "245-GHz InAlN/GaN HEMTs with oxygen plasma treatment," *Electron Device Letters, IEEE*, vol. 32, pp. 755 –757, June 2011.

[27] V. Zomorrodian, Y. Pei, U. K. Mishra, and R. A. York, "A scalable EE_HEMT based large signal model for multi-finger AlGaN/GaN HEMTs for linear and non-linear circuit design," *physica status solidi (c)*, vol. 7, no. 10, pp. 2450–2454, 2010.

[28] A. Khakifirooz, O. M. Nayfeh, and D. Antoniadis, "A simple semiempirical short-channel MOSFET current–voltage model continuous across all regions of operation and employing only physical parameters," *Electron Devices, IEEE Transactions on*, vol. 56, no. 8, pp. 1674–1680, 2009.

[29] M. Lundstrom and Z. Ren, "Essential physics of carrier transport in nanoscale MOSFETs," *Electron Devices, IEEE Transactions on*, vol. 49, no. 1, pp. 133–141, 2002.

[30] G. Wright, "Threshold modelling of MOSFETs for CAD of CMOS-VLSI," *Electronics Letters*, vol. 21, no. 6, pp. 223–224, 1985.

[31] D. Caughey and R. Thomas, "Carrier mobilities in silicon empirically related to doping and field," *Proceedings of the IEEE*, vol. 55, no. 12, pp. 2192–2193, 1967.

[32] D.-H. Kim, J. Del Alamo, D. Antoniadis, and B. Brar, "Extraction of virtual-source injection velocity in sub-100 nm III–V HFETs," in *Electron Devices Meeting (IEDM), 2009 IEEE International*, pp. 1–4, IEEE, 2009.

[33] Y. Chang, Y. Zhang, Y. Zhang, and K. Tong, "A thermal model for static current characteristics of AlGaN/ GaN high electron mobility transistors including self-heating effect," *Journal of applied physics*, vol. 99, p. 044501, 2006.

[34] R. van Langevelde and F. M. Klaassen, "Effect of gate-field dependent mobility degradation on distortion analysis in mosfets," *Electron Devices, IEEE Transactions on*, vol. 44, no. 11, pp. 2044–2052, 1997.

[35] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*, vol. 2. Oxford University Press New York, 1999.

[36] L. Wei, O. Mysore, and D. Antoniadis, "Virtual-source-based self-consistent current and charge FET models: From ballistic to drift-diffusion velocity-saturation operation," *Electron Devices, IEEE Transactions on*, vol. 59, no. 5, pp. 1263–1271, 2012.

[37] D. Ward and R. Dutton, "A charge-oriented model for MOS transistor capacitances," *Solid-State Circuits, IEEE Journal of*, vol. 13, no. 5, pp. 703–708, 1978.

[38] U. Radhakrishna, L. Wei, D.-S. Lee, T. Palacios, and D. Antoniadis, "Physics-based GaN HEMT transport and charge model: Experimental verification and performance projection," in *Electron Devices Meeting (IEDM), 2012 IEEE International*, pp. 13–6, IEEE, 2012.

[39] A. Matulionis, J. Liberis, I. Matulioniene, M. Ramonas, and E. Sermuksnis, "Ultrafast removal of LO-mode heat from a GaN-based two-dimensional channel," *Proceedings of the IEEE*, vol. 98, pp. 1118 –1126, July 2010.

[40] T. Fang, R. Wang, H. Xing, S. Rajan, and D. Jena, "Effect of optical phonon scattering on the performance of GaN transistors," *Electron Device Letters, IEEE*, vol. 33, pp. 709 –711, May 2012.

[41] K. Shinohara, D. Regan, A. Corrion, D. Brown, S. Burnham, P. Willadsen, I. Alvarado-Rodriguez, M. Cunningham, C. Butler, A. Schmitz, S. Kim, B. Holden, D. Chang, V. Lee, A. Ohoka, P. Asbeck, and M. Micovic, "Deeply-scaled self-aligned-gate GaN DH-HEMTs with ultrahigh cutoff frequency," in *Electron Devices Meeting (IEDM), 2011 IEEE International*, pp. 19.1.1 –19.1.4, Dec. 2011.

[42] H. Kondoh, "An accurate FET modelling from measured S-parameters," in *Microwave Symposium Digest, 1986 IEEE MTT-S International* (IEEE, ed.), pp. 377–380, 1986.