

MODELING AND OPTIMAL DESIGN OF NANOSCALE DOUBLE-GATE CMOS DEVICES  
AND TECHNOLOGY

By  
SHISHIR AGRAWAL

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY  
UNIVERSITY OF FLORIDA

2009

© 2009 Shishir Agrawal

To my parents and sister whose unyielding support and sacrifice made this dissertation possible

## ACKNOWLEDGEMENTS

I would like to express a deep sense of gratitude to my advisor, Dr. Jerry Fossum, for his excellent guidance and support throughout the course of this work. His knowledge, experience, and enthusiasm for the area of semiconductor devices has been invaluable in making this work a reality. I would like to thank Dr. Scott Thompson, Dr. Jing Guo, and Dr. Selman Hershfield for agreeing to be on my committee as well as for helpful discussions during my research work. I would like to acknowledge Freescale Semiconductor and Sematech Inc. for providing us with the FinFET data used in this work. I especially thank Leo Mathew at Freescale and Casey Smith at Sematech for patiently answering our questions and offering invaluable suggestions about the data. I have been fortunate to have the company of my colleagues, Siddharth Chouksey, Zhichao Lu, Zhenming Zhou, Dabraaj Sarkar, Weimin Zhang, Murshed Chowdhury, and Sueng-Hwan Kim. I thank all of them for help with technical issues as well as for their good companionship through these years. This work would not have been possible without the unyielding support and sacrifice of my parents and sister. I am deeply indebted to them for the confidence they placed in me during difficult times of my life. This work is dedicated to them. Last, but not the least, thanks go out to my wife, Shweta, for adding so much happiness and cheer to my years in graduate school.

## TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGEMENTS .....	4
LIST OF TABLES .....	8
LIST OF FIGURES .....	10
LIST OF ABBREVIATIONS .....	14
ABSTRACT .....	15
 CHAPTER	
1 INTRODUCTION .....	17
2 A NOVEL ITFET DESIGN FOR ENHANCED SCALABILITY .....	21
2-1 Introduction .....	21
2-2 Conventional ITFET Design .....	22
2-3 Scaled ITFET Design .....	23
2-4 Scalability .....	24
2-5 Design Considerations .....	25
2-6 Summary .....	26
3 ON THE SUITABILITY OF A HIGH-K GATE DIELECTRIC IN NANOSCALE FINFET CMOS TECHNOLOGY .....	39
3-1 Introduction .....	39
3-2 Device Design .....	40
3-3 Simulation Setup .....	42
3-4 Performance Projections .....	44
3-5 Scalability .....	46
3-6 Summary and Discussion .....	47
4 A PHYSICAL MODEL FOR FRINGE CAPACITANCE IN DOUBLE-GATE MOSFETS WITH NON-ABRUPT SOURCE/DRAIN JUNCTIONS .....	56
4-1 Introduction .....	56
4-2 Characterization of C <sub>if</sub> .....	57
4-3 Characterization of L <sub>u</sub> and C <sub>of</sub> .....	60
4-4 Verification and Discussion .....	61
4-5 Model Implementation in UFDG (Ver. 3.8) .....	62
4-6 Significance of C <sub>f</sub> for Optimal DG CMOS Design .....	63
4-7 Summary and Discussion .....	65

5	A REVERSE-ENGINEERING METHODOLOGY TO EXTRACT THE SOURCE/ DRAIN DOPING PROFILE OF NANOSCALE FINFETS FOR OPTIMAL SOURCE/ DRAIN PROCESS DESIGN.....	75
	5-1 Introduction .....	75
	5-2 Description of Reverse-Engineering Methodology .....	76
	5-3 Calibration of UFDG to C-V Characteristics (Step 1) .....	77
	5-4 Calibration of UFDG to I-V Characteristics (Step 1).....	78
	5-5 Determination of SDE Doping Profile (Step 2).....	80
	5-6 Carrier Mobility Degradation in Short-Channel FinFETs.....	81
	5-7 Device Redesign for Better Ion/Ioff Tradeoff.....	81
	5-8 Summary.....	84
6	PHYSICAL INSIGHTS ON CARRIER MOBILITIES IN SHORT-CHANNEL FINFETS.....	108
	6-1 Introduction .....	108
	6-2 Significance of High Mobility for FinFET CMOS Technology.....	109
	6-3 Literature Survey .....	109
	6-4 UFDG Calibration of Long- and Short-Channel FinFETs .....	110
	6-5 Possible Explanations of the Mobility Degradation .....	112
	6-6 Comparison of a “Pragmatic” FinFET Technology to Bulk Technology.....	114
	6-7 Summary.....	117
7	SUMMARY AND FUTURE WORK.....	124
	7-1 Summary.....	124
	7-2 Future Work.....	126
<b>APPENDIX</b>		
A	USING TAURUS AND UFDG FOR UTB DEVICE SIMULATION.....	129
	A-1 Introduction.....	129
	A-2 A Brief Description of Taurus and UFDG Mobility Models.....	130
	A-3 Calibration of Weak-Inversion I-V Characteristics .....	131
	A-4 Calibration of Strong-Inversion I-V Characteristics.....	132
	A-5 Additional Note .....	134
B	UFDG CODE REVISIONS FOR HIGH-K DEVICE SIMULATION.....	143
	B-1 Introduction.....	143
	B-2 Code Revision to Account for Gate Dielectric Material.....	143
	B-3 Code Revisions for Fringe Capacitance Model .....	144

LIST OF REFERENCES.....	150
BIOGRAPHICAL SKETCH.....	155

## LIST OF TABLES

<u>Table</u>	<u>page</u>
2-1 Structure of conventional ITFET structure at 45nm node .....	27
2-2 Characteristics of conventional ITFET structure at 45nm node .....	27
2-3 Structure of scalable ITFET structure at 45nm node .....	28
2-4 Characteristics of scalable ITFET structure at 45nm node .....	28
2-5 Structure of scalable ITFET structure at 22nm node .....	29
2-6 Characteristics of scalable ITFET structure at 22nm node .....	29
3-1 Taurus-predicted weak-inversion characteristics of pragmatic device for different values of fin thickness .....	49
3-2 Physical parameters of pragmatic and high-k device .....	49
3-3 UFDG/Spice3-predicted propagation delays of unloaded CMOS ROs comprising pragmatic and high-k device .....	50
4-1 Physical parameters of device structures used to verify fringe capacitance model .....	66
4-2 Effective extension length of device structures used to verify fringe capacitance model .....	66
4-3 UFDG/Spice3-predicted propagation delays of unloaded CMOS ROs comprising Device III with different spacer materials .....	66
5-1 Approximate device and process information available for Sematech FinFETs .....	85
5-2 Structural and physical parameters of Sematech FinFETs after calibration to UFDG .....	86
5-3 Source/drain doping profile for Sematech n-channel FinFETs estimated using reverse-engineering methodology .....	86
6-1 Summary of mobility parameters of n-channel FinFETs calibrated to UFDG .....	118
6-2 Summary of mobility parameters of p-channel FinFETs calibrated to UFDG .....	118
6-3 Extracted mobility due to undetermined defects ( $\mu_N$ ) for $L_g=32\text{nm}$ as a function of temperature and inversion charge density .....	119

6-4	Possible causes and solutions to the problem of carrier mobility degradation in short- $L_g$ FinFETs .....	119
A-1	A summary of the mobility models to be used for calibration between Taurus and UFDG.....	135

## LIST OF FIGURES

<u>Figure</u>	<u>page</u>
2-1 Cross-sectional diagram of ITFET structure.....	30
2-2 Taurus-predicted weak-inversion constituent currents of conventional ITFET .....	31
2-3 Taurus-predicted strong-inversion constituent currents of conventional ITFET .....	32
2-4 Taurus-predicted weak-inversion constituent currents of scalable ITFET with $L_g = 18\text{nm}$ , SOI width, $t_{\text{Si}} = 10\text{nm}$ .....	33
2-5 Taurus-predicted weak-inversion constituent currents of scalable ITFET with $L_g = 18\text{nm}$ , SOI width, $t_{\text{Si}} = 11\text{nm}$ .....	34
2-6 Taurus-predicted weak-inversion constituent currents of scalable ITFET with $L_g = 18\text{nm}$ , SOI width, $t_{\text{Si}} = 12\text{nm}$ .....	35
2-7 Taurus-predicted strong-inversion constituent currents of scalable ITFET with $L_g = 18\text{nm}$ , SOI width, $t_{\text{Si}} = 10\text{nm}$ .....	36
2-8 Taurus-predicted weak-inversion constituent currents of scalable ITFET with $L_g = 9\text{nm}$ .....	37
2-9 Taurus-predicted strong-inversion constituent currents of scalable ITFET with $L_g = 9\text{nm}$ .....	38
3-1 Taurus-predicted weak-inversion $I_{\text{DS}}-V_{\text{GS}}$ characteristics of n-channel $L_g = 18\text{nm}$ pragmatic and high-k device.....	51
3-2 UFDG-predicted $I_{\text{DS}}-V_{\text{GS}}$ characteristics of n-channel $L_g = 18\text{nm}$ pragmatic and high-k device with all physical models on.....	52
3-3 UFDG-predicted $C_{\text{G}}-V_{\text{GS}}$ characteristics of pragmatic and high-k device with QM model on and off .....	53
3-4 UFDG-predicted $I_{\text{DS}}-V_{\text{GS}}$ characteristics of n-channel $L_g = 18\text{nm}$ pragmatic and high-k device with QM model off, $R_{\text{S/D}} = 0$ .....	54
3-5 UFDG-predicted $I_{\text{SD}}-V_{\text{GS}}$ characteristics of p-channel $L_g = 18\text{nm}$ pragmatic and high-k device with all physical models on.....	55
4-1 Diagram of actual and simplified DG MOSFET structure .....	67
4-2 Diagram of device structure portraying inner and outer fringing fields .....	68

4-3	Taurus and model-predicted weak-inversion fringe capacitance vs. spacer permittivity.....	69
4-4	Taurus and model-predicted strong-inversion fringe capacitance vs. spacer permittivity.....	70
4-5	Taurus and model-predicted weak-inversion fringe capacitance vs. gate height.....	71
4-6	Taurus and model-predicted strong-inversion fringe capacitance vs. gate height .....	72
4-7	Taurus and model-predicted weak-inversion fringe capacitance vs. silicon width.....	73
4-8	UFDG-predicted $C_G$ - $V_{GS}$ characteristics of Device III with spacer permittivity, $k_{sp} = 1, 7.5$ .....	74
5-1	Flowchart demonstrating reverse-engineering methodology to extract source/drain doping profile .....	87
5-2	Calibration of n-channel $N_{inv}$ - $V_{GS}$ data to UFDG.....	88
5-3	Calibration of p-channel $N_{inv}$ - $V_{GS}$ data to UFDG.....	89
5-4	Calibration of UFDG to n-channel $L_g = 1\mu\text{m}$ $g_m/I_{DS}^2$ - $V_{GS}$ data .....	90
5-5	Calibration of UFDG to n-channel $L_g = 1\mu\text{m}$ $I_{DS}$ - $V_{GS}$ data .....	91
5-6	Channel resistance ( $R_{ch}$ ) of n-devices plotted vs. gate length, $L_g$ .....	92
5-7	Calibration of UFDG to n-channel $L_g = 75\text{nm}$ $g_m/I_{DS}^2$ - $V_{GS}$ data .....	93
5-8	Calibration of UFDG to n-channel $L_g = 75\text{nm}$ $I_{DS}$ - $V_{GS}$ data.....	94
5-9	Calibration of UFDG to n-channel $L_g = 32\text{nm}$ $I_{DS}$ - $V_{GS}$ data (weak-inversion).....	95
5-10	Calibration of UFDG to n-channel $L_g = 32\text{nm}$ $g_m/I_{DS}^2$ - $V_{GS}$ data .....	96
5-11	Calibration of UFDG to n-channel $L_g = 32\text{nm}$ $I_{DS}$ - $V_{GS}$ data (strong-inversion) .....	97
5-12	Calibration of UFDG to p-channel $L_g = 10\mu\text{m}$ $\mu_{eff}$ - $N_{inv}$ data .....	98
5-13	Calibration of UFDG to p-channel $L_g = 1\mu\text{m}$ $I_{SD}$ - $V_{GS}$ data .....	99
5-14	Channel resistance ( $R_{ch}$ ) of p-devices plotted vs. gate length, $L_g$ .....	100
5-15	Calibration of UFDG to p-channel $L_g = 75\text{nm}$ $I_{SD}$ - $V_{GS}$ data.....	101
5-16	Calibration of UFDG to p-channel $L_g = 32\text{nm}$ (linear-scale) $I_{SD}$ - $V_{GS}$ data.....	102

5-17	Calibration of UFDG to p-channel $L_g = 32\text{nm}$ (log-scale) $I_{SD}-V_{GS}$ data.....	103
5-18	Determination of $N_{SD}(y)$ by matching UFDG- and Taurus-predicted characteristics.....	104
5-19	UFDG-predicted $\mu_{\text{eff}}-N_{\text{inv}}$ for n-channel FinFETs of various $L_g$ .....	105
5-20	Taurus-predicted weak-inversion characteristics for redesigned device.....	106
5-21	Extracted S/D doping profile, $N_{SD}(y)$ , for n-channel $L_g = 32\text{nm}$ device and projected $N_{SD}(y)$ for redesigned device.....	107
6-1	UFDG calibration of n-channel $L_g = 32\text{nm}$ FinFET and projection with ideal mobility parameters.....	120
6-2	UFDG calibration of p-channel $L_g = 32\text{nm}$ FinFET and projection with ideal mobility parameters.....	121
6-3	UFDG-predicted $\mu_{\text{eff}}-N_{\text{inv}}$ for $L_g = 1\mu\text{m}$ FinFET at different temperatures.....	122
6-4	UFDG-predicted $\mu_{\text{eff}}-N_{\text{inv}}$ for $L_g = 32\text{nm}$ FinFET at different temperatures.....	123
7-1	A flowchart describing possible future study for FinFET source/drain design .....	128
A-1	A comparison of the Taurus “constant mobility” and “velocity saturation” models in weak-inversion for an abrupt junction device .....	136
A-2	Electric potential as a function of distance for abrupt junction device in weak inversion.....	137
A-3	Calibration between Taurus- and UFDG-predicted $I_{DS}-V_{GS}$ characteristics for an abrupt junction device in weak inversion.....	138
A-4	Electric potential as a function of distance for abrupt junction device in strong inversion.....	139
A-5	Calibration between Taurus- and UFDG-predicted $I_{DS}-V_{GS}$ characteristics for an abrupt junction device in strong inversion .....	140
A-6	Calibration between Taurus- and UFDG-predicted $I_{DS}-V_{DS}$ characteristics for an abrupt junction device in strong inversion.....	141
A-7	Electric potential as a function of distance for underlapped device in strong inversion.....	142
B-1	Basic two-plate model of fringe capacitance .....	146

B-2	A schematic diagram of the G-S/D underlap structure along with fringe capacitance components .....	147
B-3	A schematic diagram of underlap structure for $C_{if}$ derivation in device with $SiO_2$ dielectric.....	148
B-4	A schematic diagram of underlap structure for $C_{if}$ derivation in device with high-k dielectric.....	149

## LIST OF ABBREVIATIONS

MOSFET	Metal-Oxide-Semiconductor Field-Effect Transistor
CMOS	Complementary MOS
SOI	Silicon-on-Insulator
FD	Fully Depleted
ITFET	Inverted-T FET
UTB	Ultra-Thin Body
UFDG	University of Florida Double-Gate
SCE	Short-Channel Effect
DIBL	Drain-Induced Barrier Lowering
QM	Quantum-Mechanical
SDE	Source/Drain Extension
SG	Single Gate
S	Subthreshold Slope

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

MODELING AND OPTIMAL DESIGN OF NANOSCALE DOUBLE-GATE CMOS DEVICES  
AND TECHNOLOGY

By

Shishir Agrawal

May 2009

Chair: Jerry G. Fossum  
Major: Electrical and Computer Engineering

This dissertation seeks to model underlying physical effects, and gain insights into the optimal design of nanoscale Double-Gate (DG) MOSFETs, particularly the quasi-planar FinFET structure. Our work includes design of a novel FinFET-based device structure (scalable ITFET), evaluation of the suitability of high-k gate dielectrics for nanoscale FinFETs, a model for parasitic fringe capacitance in DG FinFETs, and insights on optimal source/drain (S/D) process design and carrier mobility in short-channel FinFETs.

The ITFET is a recently proposed hybrid device which consists of a FinFET and planar FD/SOI MOSFET combined into a single device with a common gate. The advantage of the ITFET is that it enhances the on-state current per pitch by ~100%; its disadvantage is poor scalability. We propose a novel ITFET design which is scalable to near the end of the ITRS roadmap while still giving reasonable enhancement in on-state current,  $I_{on}$  (~20-35%).

We perform a realistic assessment of the performance and scalability advantage of high-k dielectrics in FinFET-CMOS technology. Our results indicate that a high-k dielectric actually undermines circuit performance while giving limited improvement in scalability. We conclude that high-k gate dielectrics are not suitable for nanoscale FinFETs.

We present a physical model for fringe capacitance ( $C_f$ ) in DG MOSFETs with non-abrupt S/D junctions. We model  $C_f$  in terms of the device structure and short-channel effects (SCEs). The model is implemented in our physical/process based compact model, UFDG, and will enable quasi-predictive device/circuit simulations.

In undoped UTB FinFETs, the lateral S/D doping profile,  $N_{SD}(y)$ , defines the tradeoff between SCEs and parasitic resistance,  $R_{S/D}$ , via gate-source/drain underlap. We demonstrate a reverse-engineering methodology to extract  $N_{SD}(y)$  from FinFET  $C_G$ - $V_{GS}$  and  $I_{DS}$ - $V_{GS}$  data. The extracted  $N_{SD}(y)$  is then used to redesign the S/D process to effect a better tradeoff between SCEs and  $R_{S/D}$ .

Finally, we discuss the degradation of mobility in short-channel FinFETs possibly due to S/D defects/dopants. We explore possible causes of the phenomenon and make device processing suggestions to help mitigate the effect.

## CHAPTER 1 INTRODUCTION

The Double-Gate (DG) MOSFET is considered one of the most promising technological options to replace conventional (i.e, bulk-Si and PD/SOI) technologies [1]. The primary advantage of the DG MOSFET is excellent control of short-channel effects (SCEs) [2], which makes it potentially scalable to the end of the SIA ITRS roadmap [1]. Also, since the body of the DG MOSFET can be left undoped, random variation in threshold voltage and other device characteristics due to process variations can be greatly reduced [3]. The undoped body also implies high carrier mobility without the need for mobility enhancement techniques [4]. The research described herein seeks to model physical effects and gain insights into the optimal design of DG MOSFETs.

In recent years, FinFET technology has been considered to be most feasible for the implementation of the DG MOSFET [5][6]. Since it is a quasi-planar technology, it is most compatible with today's standard processing techniques. One of the recent innovations in FinFET technology is the ITFET structure [7][8]. The ITFET structure consists of a planar FD/SOI MOSFET fabricated on the thick insulating BOX around the FinFET. The FD/SOI MOSFET and the FinFET have a common gate, effectively forming a hybrid device. The advantage of this device is enhanced drive current per pitch (~100%), as well as to allow an "analog" device width as opposed to the FinFET "digital" width. The disadvantage of the ITFET is reduced scalability since the SCEs are governed by the single-gate FD/SOI MOSFET. Thus, while this 'conventional' ITFET will be useful in applications where high current drive is required, like I/O and interconnect drivers, it will not be useful in extremely scaled CMOS circuits. In Chapter 2, we discuss a novel method of enhancing the scalability of the ITFET structure, while still ensuring reasonable improvement in drive current (~20-35%). We use the Taurus device simulator [9] to

project the scalability and on-state current of the novel structure. Appendix A discusses the physical models used in the simulations. This ‘scalable’ ITFET can be used in scaled CMOS circuits which require device ratioing, like SRAM cells. We stress, however, that the above-discussed ITFET structures may not be appropriate for applications where considerations of device scalability and gate capacitance optimization (in addition to on-state current) are important. For these applications, the ITFET can be manufactured with the SOI ungated, thus rendering the electrical characteristics of the device the same as the FinFET. The SOI on the BOX lends greater mechanical stability to the fin, making it possible to fabricate thinner, taller fins.

Since the FinFET is a novel device structure, optimal device design decisions may be different, often counterintuitive, from what we expect in conventional technology. In Chapter 3, we examine the impact of high-k gate dielectrics on the device and circuit performances of FinFET CMOS technology via physics-based device/circuit simulations. DG FinFETs are designed with high-k at the HP-45nm node of the 2005 SIA ITRS [1] (gate length,  $L_g = 18\text{nm}$ ), and are compared with a *pragmatic* design in which the traditional  $\text{SiO}_2/\text{SiON}$  gate dielectric is retained, and kept relatively thick to avoid excessive gate tunneling current. We design the devices using the Taurus device simulator [9] and use the simulation results to calibrate our physical/process-based compact model, UFDG [10], which we use for quasi-predictive performance projections. To simulate devices with high-k gate dielectric using UFDG, we make some changes in the model which are discussed in Appendix B. Whereas it is presumed that a high-k dielectric will significantly enhance FinFET-CMOS scalability and performance, we show that there are heretofore unacknowledged compromising effects associated with it that undermine this enhancement. In fact, our results show that a high-k gate dielectric actually undermines speed performance while giving little improvement in scalability relative to the pragmatic design, while the latter can be scaled, with good performance, to the end of the ITRS. We conclude that a high-

k dielectric is not worthwhile given the reduced circuit performance and the technological challenges of integrating a high-k dielectric into the process flow. We believe, based on our study, that the pragmatic approach to nanoscale DG CMOS should be taken, with good performance expected, to the end of the roadmap.

Nanoscale DG MOSFETs are typically designed with gate-source/drain (G-S/D) underlap to effect a good tradeoff between SCEs, parasitic S/D resistance ( $R_{S/D}$ ), and fringe capacitance ( $C_f$ ). Underlapped DG MOSFETs are typically undoped; thus S/D dopants diffuse into the extensions during the S/D anneal process causing a non-abrupt S/D lateral doping profile,  $N_{SD}(y)$ . In Chapter 4, we present, for the first time, a physical model of  $C_f$  for underlapped DG MOSFETs with (realistic) non-abrupt  $N_{SD}(y)$ .  $C_f$  consists of two components; the inner fringe capacitance,  $C_{if}$ , and the outer fringe capacitance,  $C_{of}$ . We relate  $C_{if}$  and  $C_{of}$  to the device SCEs as governed by the effective G-S/D underlap,  $L_{eSD}$ . The model is verified by numerical simulations of DG MOSFETs with varying device parameters including the spacer dielectric constant. Our model for  $C_f$  is implemented in our physical/process-based compact model, UFDG [10]. Since the model is quasi-predictive, it will be useful in trading off  $R_{S/D}$  and  $L_{eSD}$  for optimal speed performance.

In Chapter 5, we discuss the problem of S/D process design for nanoscale FinFETs with G-S/D underlap. G-S/D underlap implies an increase in the weak-inversion effective channel length,  $L_{eff(weak)} = L_g + 2L_{eSD}$ , while not significantly affecting the effective channel length in strong inversion,  $L_{eff(strong)} \cong L_g$  [11].  $L_{eSD}$  is determined primarily by the spacer extension length,  $L_{ext}$ , and  $N_{SD}(y)$ . An increase in  $L_{eSD}$  implies better SCEs and also lower  $C_f$ . However, it also implies increased  $R_{S/D}$ , which results in reduced on-state current ( $I_{on}$ ). Thus, the S/D process implies a tradeoff between SCEs and  $R_{S/D}$ . It has also been shown [12] that it may be possible to engineer the threshold voltage,  $V_t$ , of a device by allowing controlled densities of S/D dopants to diffuse into the channel. To better understand the above design tradeoffs, we seek to relate the S/D

process to  $N_{SD}(y)$ . We therefore design and demonstrate a reverse-engineering methodology to extract  $N_{SD}(y)$  from fabricated FinFET  $C_G-V_{GS}$  and  $I_{DS}-V_{GS}$  data. We also show how the extracted  $N_{SD}(y)$  can be used to redesign the S/D process to effect a better tradeoff between SCEs and  $I_{on}$ .

In Chapter 6, we discuss the problem of carrier mobility degradation in short- $L_g$  FinFETs, which may be related to the S/D processing. In our work, we have observed a severe, unexplained degradation of carrier mobility in  $L_g < \sim 100\text{nm}$  FinFETs. This effect has been reported in the literature as well [13][14][15][16]. Several possible causes for this degradation have been proposed, from Coulomb/remote Coulomb scattering due to S/D dopants to scattering due to “neutral defects.” We try and gain more insight into this phenomenon with a thorough literature review as well as mobility measurements in short- and long- $L_g$  FinFETs. Based on these insights, we make device processing suggestions which may lead to an improvement in mobility for short  $L_g$  FinFETs.

In Chapter 7, we summarize our work and also discuss needed future work in the area of optimal FinFET design.

## CHAPTER 2 A NOVEL ITFET DESIGN FOR ENHANCED SCALABILITY

### 2-1 Introduction

In FinFET technology, the vertical silicon fin is very thin, and it is possible to fabricate only one fin per lithographic pitch ( $P$ ). Thus, there is “wasted area” on the thick BOX between two fins. The ITFET [7][8] is a recently proposed hybrid device which consists of a Double-Gate (DG) FinFET and a single-gate (SG) planar FD/SOI MOSFET (FDFET), with all the gates common. The idea behind the ITFET (Inverted-T FET) is to integrate the planar FDFET with the FinFET to enhance the attained drive current per pitch, as well as to allow an “analog” device width, as opposed to the FinFET “digital” width. Further, we have discovered [8] that the SOI on the BOX lends mechanical stability to the fin, which makes it possible to fabricate taller, thinner fins, yielding even more current per pitch. Fig. 2-1 shows a cross-sectional illustration of the ITFET structure. The sole disadvantage of the ITFET is that short-channel effects (SCEs) are governed by the SG FDFET [17], and the limited SOI-UTB (ultra-thin body) thickness thus limits the device scalability [7].

In this chapter, we first examine the design and performance of a conventional ITFET at the 45nm node of the SIA ITRS [1], with a midgap metal gate on both constituent devices. We then propose a novel design to improve the scalability of the ITFET. The idea is to use different gate materials on the constituent devices, employing, in a doable technology, a higher work function for the FDFET to increase its threshold voltage ( $V_{tFD}$ ). Such design alleviates the influence of FDFET SCEs on the off-state current ( $I_{off}$ ) of the ITFET. We demonstrate the design via numerical device simulations, showing that the optimal ITFET is scalable to near the end of the SIA ITRS [1] where it still yields significant enhancement of on-state current ( $I_{on}$ ).

## 2-2 Conventional ITFET Design

We design a conventional ITFET (i.e., one with the same gate on both fin and FDFET) for the 45nm technology node [1] with gate length  $L_g = 18\text{nm}$ . We assume the novel processing discussed in [7] whereby the fin gate extends down to the BOX to prevent prodigious leakage current in the fin base (Fig. 2-1). This extension of the fin gate to the BOX is relatively thin, but it reduces the width of the planar FDFET by twice the fin-base gate thickness ( $t_{gb}$ ), and yields an ITFET that essentially comprises the two constituent devices in parallel. We thus model this ITFET, per pitch, using the 2-D simulator Taurus [9], by simulating each constituent device, per unit width, and summing their predicted current-voltage characteristics multiplied by the respective widths; the fin height ( $h_{Si}$ ) for the FinFET and  $(P - w_{Si} - 2t_{gb})$  for the FDFET, where  $w_{Si}$  is the fin-UTB width.

We design a near optimal  $L_g = 18\text{nm}$  FinFET which includes gate-source/drain (G-S/D) underlap [11] with an undoped UTB and midgap gate. The S/D extension length is  $L_{ext} = 18\text{ nm}$ , and the lateral doping straggle in the extension is  $\sigma_L = 9.6\text{nm}$ , giving an effective underlap ( $L_{eSD}$ ) of about 3.5nm. We choose  $w_{Si} = 12\text{nm}$  and gate oxide thickness  $t_{ox} = 1.2\text{nm}$  for adequate SCE control (DIBL = 100mV/V and  $S = 78\text{mV/dec}$ ). We assume a fin aspect ratio of  $R_f = h_{Si}/w_{Si} = 4$ , which means  $h_{Si} = 48\text{nm}$ . The series resistance ( $R_{S/D}$ ) for both the fin and FDFET is assumed to be  $80\ \Omega\text{-}\mu\text{m}$ . At the assumed technology node,  $P = 90\text{nm}$  [1], and if we assume  $t_{gb} = 3\text{nm}$ , the FDFET width is 72nm. We note that the ITFET fabrication process results in the FDFET gate length being about 20-30% longer than the gate length of the FinFET. This is due to the limited depth of focus of the lithographic tool. We therefore assume in our nominal design that  $L_g \cong 22\text{nm}$  for the FDFET. The underlap associated with the FDFET is the same as that of the FinFET. To determine the needed SOI width,  $t_{Si}$ , we use the insight that for a well tempered FDFET,  $t_{Si} \cong L_{eff(weak)}/5$  [17], where  $L_{eff(weak)}$  is the effective channel length in weak inversion. With

$L_{\text{eff(weak)}} = L_g + L_{eS} + L_{eD} \cong 29\text{nm}$ , using Taurus to simulate different values of SOI width, we arrive at  $t_{\text{Si}} = 6\text{nm}$ . The design is summarized in Table 2-1 with performance predictions in Table 2-2. No QM effects are turned on in Taurus and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V}\cdot\text{s}$ ) is used. The carrier velocity saturation effect is turned off for weak-inversion projections and turned on for strong-inversion projections (Appendix A). Fig. 2-2 shows Taurus-predicted ITFET constituent currents in weak inversion while Fig. 2-3 portrays constituent currents in strong inversion. We note that the enhancement in drive current is  $\sim 110\%$  as compared to a conventional FinFET. We also note that the  $I_{\text{DS}}\text{-}V_{\text{GS}}$  characteristic is smooth without any hump as may be expected from the hybrid device. We believe that the reason for this is that both devices have pretty much the same threshold voltage. We note, however, that this device is not very scalable since the SCEs are controlled by the single gate FDFET. This is discussed in more detail in Sec. 2-4.

### 2-3 Scaled ITFET Design

For improved n-channel ITFET scalability with undoped UTB, we propose to use a midgap metal gate on the FinFET and a  $p^+$ -polysilicon gate on the FDFET. We note that this design is doable because the FinFET is gated using a thin layer of midgap metal and then capped off with polysilicon [18], which can be n-type or p-type. Therefore, it is possible to deposit  $p^+$ -poly on the entire device, thus defining a high work function gate on the FDFET. The high threshold voltage of the FDFET ( $V_{\text{tFD}}$ ) lets us get away with poor SCEs without impacting the  $I_{\text{off}}$ . This implies more device scalability. The disadvantage of this design is that the high  $V_{\text{tFD}}$  yields a smaller enhancement in current drive compared to that of the ITFET with a common metal gate, as described in Sec. 2-2.

Via simulation, we first design and check a so-designed ITFET at the 45nm node (i.e.,  $L_g = 18\text{nm}$  for the FinFET) of the SIA roadmap [1]. The device design assumed is the same as for the

conventional ITFET described in Sec. 2-2 (Table 2-3) with the exception of  $t_{Si}$ . As discussed above, we can afford a much thicker  $t_{Si}$  because of the high  $V_{tFD}$ . We let  $t_{Si}$  vary (10, 11, and 12nm), and simulate the designed ITFET. Taurus-predicted weak-inversion  $I_{DS}-V_{GS}$  characteristics are shown in Fig 2-4 to 2-6. We observe that for  $t_{Si} = 10$  or 11nm,  $I_{off}$  of the FinFET dominates the total  $I_{off}$  of the ITFET, despite the poor subthreshold characteristics of the FDFET. However, for  $t_{Si} = 12$ nm,  $I_{off}$  of the FDFET dominates, and the ITFET shows worse subthreshold characteristics. To allow for process variations, a reasonable design should then be  $t_{Si} = 10-11$ nm. Fig. 2-7 depicts the constituent strong-inversion  $I_{DS}-V_{GS}$  characteristics of the scalable ITFET with  $t_{Si} = 10$ nm and specific performance parameters are given in Table 2-4. We observe from Fig. 2-7 that the enhancement in  $I_{on}$  over the FinFET is predicted to be  $\sim 23\%$ . We note that in the device-on condition ( $V_{GS} = V_{DS} = 1.0V$ ), the constituent FDFET is biased near moderate inversion because of the high  $V_{tFD}$ , which explains the relatively low  $I_{on}$  enhancement. We note again, the absence of a hump in the  $I_{DS}-V_{GS}$  characteristics. We believe this can be explained by the poor subthreshold slope of the constituent FDFET, which causes the FDFET component of the current to increase slower than the FinFET component. This bias condition also implies that small uncertainties in the device structure and/or underlying physics could translate to large errors in the results. For example, an error in the quantization modeling could mean a large uncertainty in the  $I_{on}$  enhancement.

#### 2-4 Scalability

The scalability of both the conventional and novel ITFET designs will be ultimately limited by how thin  $t_{Si}$  can be made, as noted in [7]. It has been shown [19] that for  $t_{Si} < 4$ nm,  $V_{tFD}$  increases dramatically with decreasing  $t_{Si}$  because of the energy quantization due to the structural confinement of carriers in the thin body. We therefore base our scalability estimate on the fact that it is not possible to realistically design devices with  $t_{Si} < \cong 5$ nm. It is evident that our

conventional ITFET (with  $t_{Si} = 6\text{nm}$  for  $L_g = 18\text{nm}$ ) cannot be scaled far beyond this node. The advantage of this device, however, is that the improvement in current drive is  $\cong 110\%$  per pitch and it will be possible to tailor the width of the FDFET in an analog fashion for optimized layout area.

For the novel ITFET design, scalability will be limited to the gate length at which  $I_{off}(\text{FDFET}) < I_{off}(\text{FinFET})$  for  $t_{Si} = 5\text{nm}$ . We design a  $L_g = 9\text{nm}$  FinFET with  $w_{Si} = 5\text{nm}$ ,  $t_{ox} = 1.2\text{nm}$  and S/D doping profile which gives us reasonable  $L_{eSD} = 2.5\text{nm}$  (Table 2-5).  $R_f = 4$  implies  $h_{Si} = 20\text{nm}$ . Correspondingly we design a FDFET with  $L_{gFD} \cong 1.2L_{gFF} = 11\text{nm}$ . Taurus simulations indicate that adequate SCE control of the FDFET is implied by  $t_{Si} = 5\text{nm}$ , which is at the scaling limit. The  $L_g = 9\text{nm}$  FinFET corresponds to the 22nm node of the SIA roadmap [1] (i.e.,  $P = 44\text{nm}$ ). Thus, the width of the FDFET =  $P - w_{Si} - 2t_{gb} = 33\text{nm}$ . The constituent currents of the scaled ITFET are depicted in Fig. 2-8 for weak inversion and Fig. 2-9 for strong inversion. We find that the  $I_{on}$  enhancement at  $V_{GS} = V_{DS} = V_{DD}$  is 36%. Therefore, we see that our design is scalable to near the end of the SIA roadmap with reasonable  $I_{on}$  enhancement. We conclude that our novel ITFET design is scalable with a modest improvement ( $\cong 20\text{-}35\%$ ) in  $I_{on}$  at every node upto the 22nm node.

## 2-5 Design Considerations

In summary, there are three possible ITFET designs and an appropriate one can be chosen based on device application considerations at every technology node. The first design has a midgap gate on both constituent devices with extremely high current drive per unit device width as well as an analog width for specific drive ratio requirements. It utilizes a gate length longer than that suggested at the particular node to control SCEs. It will be a good choice for interconnect drivers where layout area is less crucial. The second design uses a  $p^+$  poly gate on the FDFET and a midgap gate on the FinFET. It gives us a modest improvement in current over the FinFET while utilizing the same layout area. This design may be well utilized in extremely scaled

CMOS circuitry where additional current or modest device ratioing is desirable. A possible application might be pull-down transistors in SRAM cells. A third possible design is an ITFET structure with the FDFET ungated. As discussed earlier, the SOI on the BOX lends mechanical stability to the fin and makes it possible to fabricate taller, thinner fins. Thus, even with the SOI ungated, we can achieve improved current per pitch as compared to a conventional FinFET structure with lower  $R_f$ . This could be the design most useful for the design of logic devices where the optimization of gate capacitance with on-state current is crucial.

## 2-6 Summary

Using numerical device simulations, we have examined different ITFET designs, their scalability and possible applications. We found that the conventional ITFET gives a huge (~110%) enhancement in on-state current as compared to the FinFET but has limited scalability. We then showed that a novel ITFET, designed with a  $p^+$ -poly gate on the constituent planar SG FD/SOI MOSFET and a midgap gate on the DG FinFET, can possibly be scaled to near the end of the SIA roadmap [1], still yielding significant enhancement of current drive over that of the FinFET. We demonstrated via numerical simulations that a viable ITFET can be designed upto the 22nm node ( $L_g = 9\text{nm}$ ) with ~20-35% enhancement in on-state current. A third ITFET structure is one in which the constituent FDFET is left ungated, so that the electrical characteristics of the device are the same as the FinFET. The added mechanical stability provided to the ITFET fin by the SOI covering the pitch area is a huge advantage over the conventional FinFET, making it possible to fabricate thinner, taller fins, and thus attain more current per pitch.

Table 2-1. Summary of the conventional ITFET structure at the 45nm node. The constituent devices are depicted.

FinFET	FDfET
$w_{Si} = 12\text{nm}$	$t_{Si} = 6\text{nm}$
$\Phi_G = 4.71\text{V}$ (midgap metal)	$\Phi_G = 4.71\text{V}$ (midgap metal)
$h_{Si} = 48\text{nm}$	$W = 72\text{nm}$
$R_{S/D} = 80\Omega\text{-}\mu\text{m}$	$R_{S/D} = 80\Omega\text{-}\mu\text{m}$
$L_{gFF} = 18\text{nm}$	$L_{gFD} = 22\text{nm}$
$L_{ext} = 18\text{nm}$	$L_{ext} = 18\text{nm}$
$\sigma_L = 9.6\text{nm}$	$\sigma_L = 9.6\text{nm}$

Table 2-2. Taurus-predicted characteristics for the conventional ITFET (design in Table 2-1) at the 45nm node

DIBL (mV/V)	S (mV/dec)	$I_{on}$ (A per pitch)	$I_{off}$ (A per pitch)	$I_{on}$ enhancement (%)
110	78	$1.21 \times 10^{-4}$	$4 \times 10^{-9}$	109

Table 2-3. Summary of the scalable ITFET structure at the 45nm node. The constituent devices are depicted.

FinFET	FDFET
$w_{Si} = 12\text{nm}$	$t_{Si} = 10\text{nm}$
$\Phi_G = 4.71\text{V}$ (midgap metal)	$\Phi_G = 5.26\text{V}$ (heavily doped $p^+$ poly)
$h_{Si} = 48\text{nm}$	$W = 72\text{nm}$
$R_{S/D} = 80\Omega\text{-}\mu\text{m}$	$R_{S/D} = 80\Omega\text{-}\mu\text{m}$
$L_{gFF} = 18\text{nm}$	$L_{gFD} = 22\text{nm}$
$t_{ox} = 1.2\text{nm}$	$t_{ox} = 1.2\text{nm}$
$L_{ext} = 18\text{nm}$	$L_{ext} = 18\text{nm}$
$\sigma_L = 9.6\text{nm}$	$\sigma_L = 9.6\text{nm}$

Table 2-4. Taurus-predicted characteristics for the scalable ITFET (design in Table 2-3) at the 45nm node

DIBL (mV/V)	S (mV/dec)	$I_{on}$ (A per pitch)	$I_{off}$ (A per pitch)	$I_{on}$ enhancement (%)
100	78	$7.1 \times 10^{-5}$	$2 \times 10^{-9}$	23

Table 2-5. Summary of the scalable ITFET structure at its scaling limit (22nm node). The constituent devices are depicted.

FinFET	FDfET
$w_{Si} = 5\text{nm}$	$t_{Si} = 5\text{nm}$
$\Phi_G = 4.71\text{V}$ (midgap metal)	$\Phi_G = 5.26\text{V}$ (heavily doped p+ poly)
$h_{Si} = 20\text{nm}$	$W = 33\text{nm}$
$R_{S/D} = 80\Omega\text{-}\mu\text{m}$	$R_{S/D} = 80\Omega\text{-}\mu\text{m}$
$L_{gFF} = 9\text{nm}$	$L_{gFD} = 11\text{nm}$
$t_{ox} = 1.2\text{nm}$	$t_{ox} = 1.2\text{nm}$
$L_{ext} = 9\text{nm}$	$L_{ext} = 9\text{nm}$
$\sigma_L = 5\text{nm}$	$\sigma_L = 5\text{nm}$

Table 2-6. Taurus-predicted characteristics of a scalable ITFET (design in Table 2-5) at the 22nm node

DIBL (mV/V)	S (mV/dec)	$I_{on}$ (A per pitch)	$I_{off}$ (A per pitch)	$I_{on}$ enhancement (%)
100	78	$3.07 \times 10^{-5}$	$1 \times 10^{-9}$	36

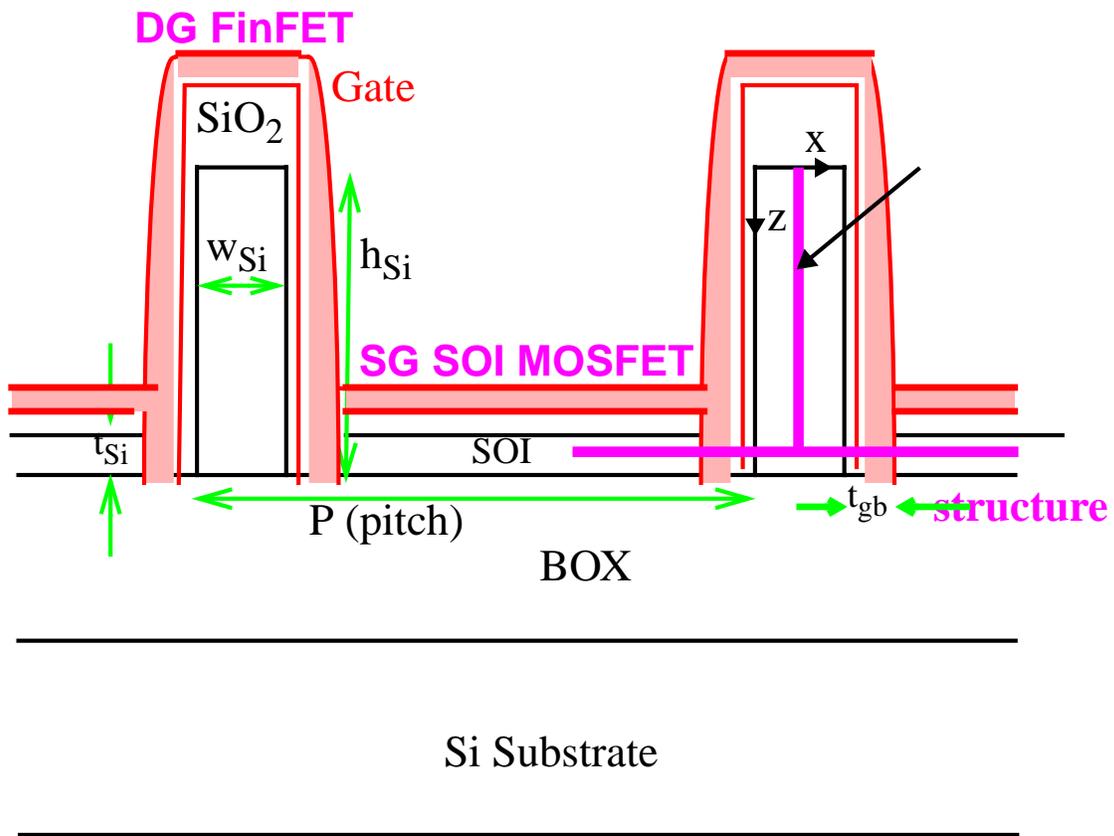


Figure 2-1. Cross-section of the ITFET structure, which shows the single gate FD/SOI MOSFET sandwiched between two double-gated fins, with all gates common.

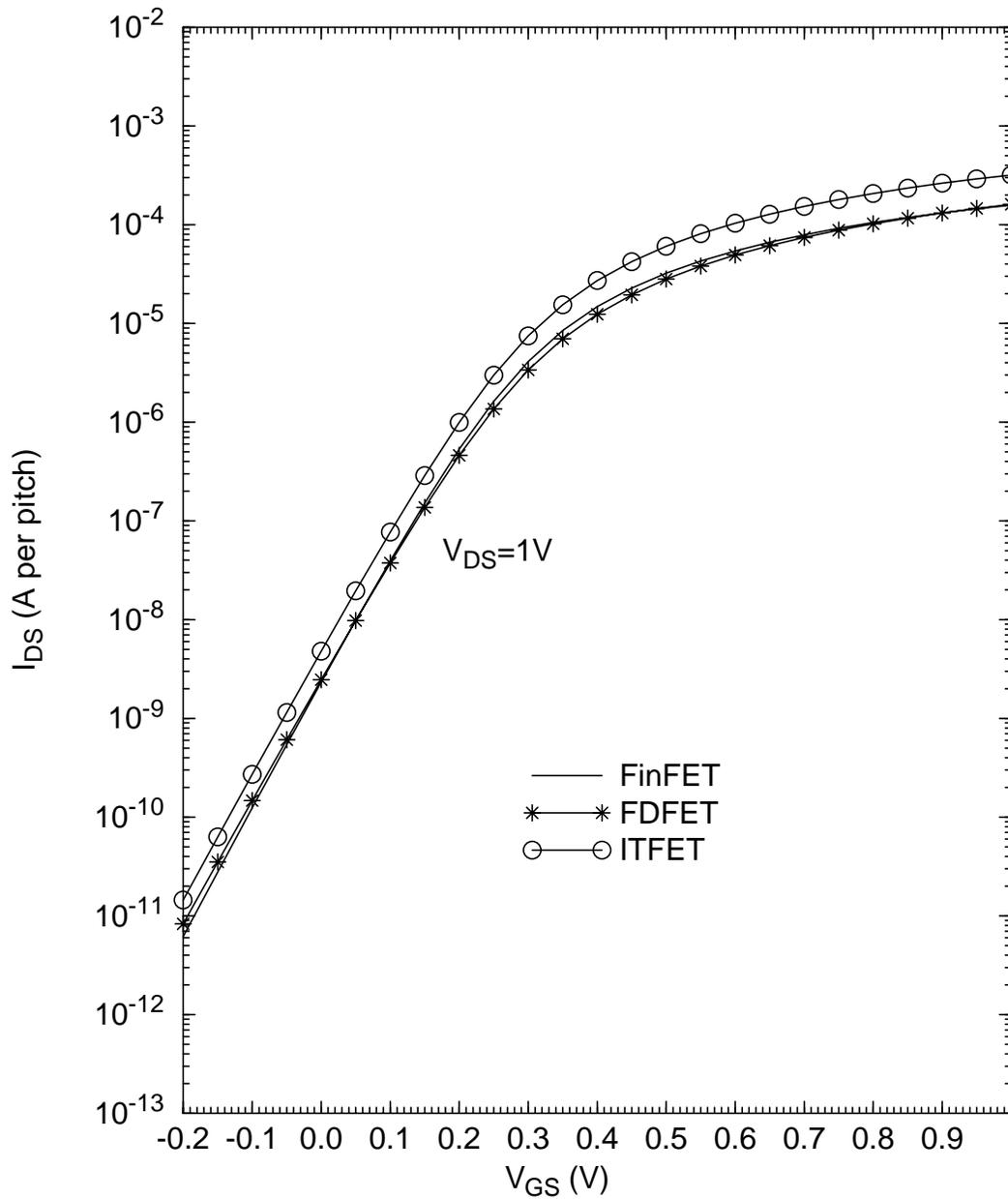


Figure 2-2. Taurus-predicted weak-inversion constituent currents of conventional ITFET with  $L_g = 18\text{nm}$  (Table 2-1 design). No QM effects are turned on, and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V}\cdot\text{s}$ ) without velocity saturation is used.

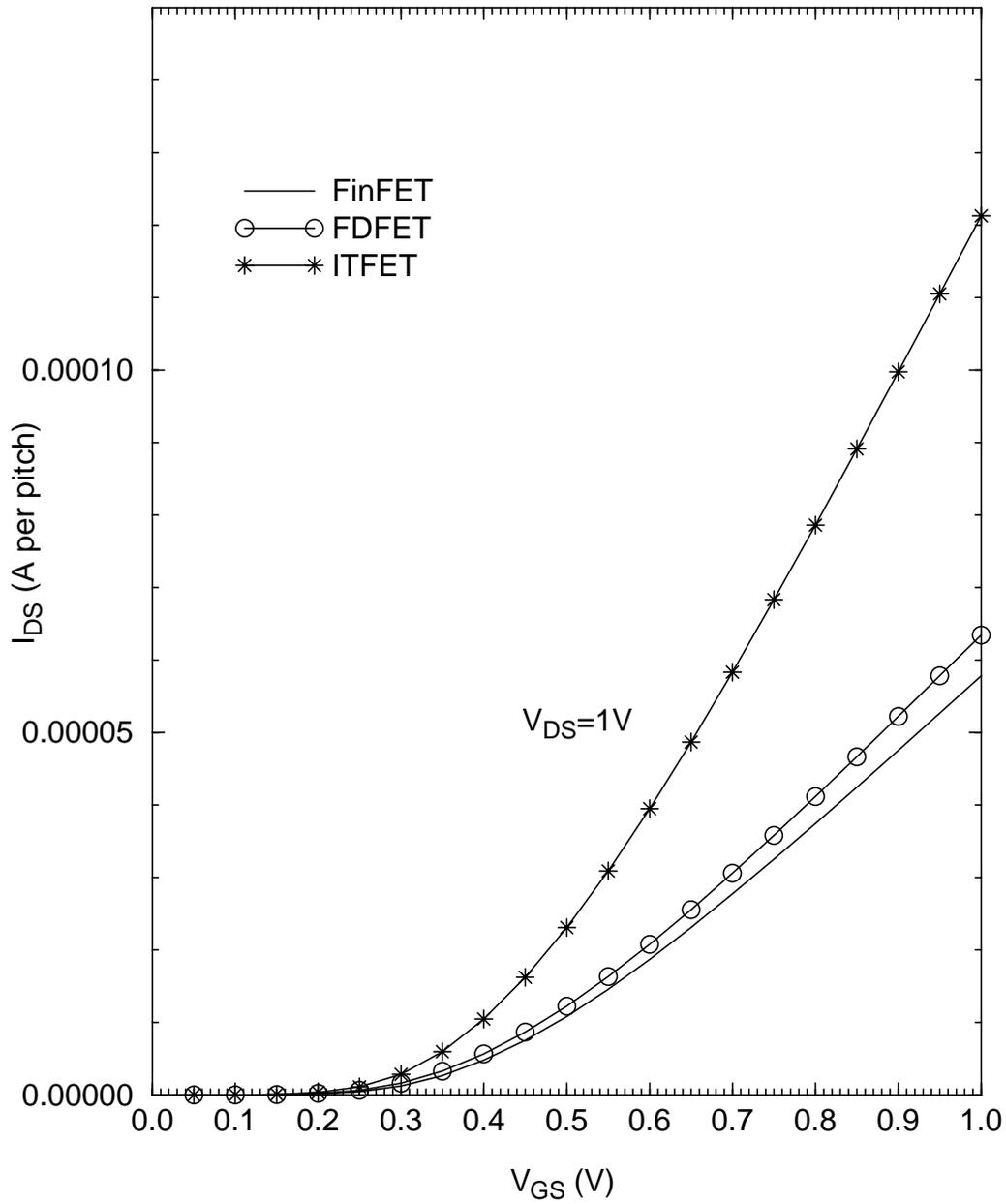


Figure 2-3. Taurus-predicted strong-inversion constituent currents of conventional ITFET with  $L_g = 18\text{nm}$  (Table 2-1 design). No QM effects are turned on, and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V}\cdot\text{s}$ ) with velocity saturation is used.

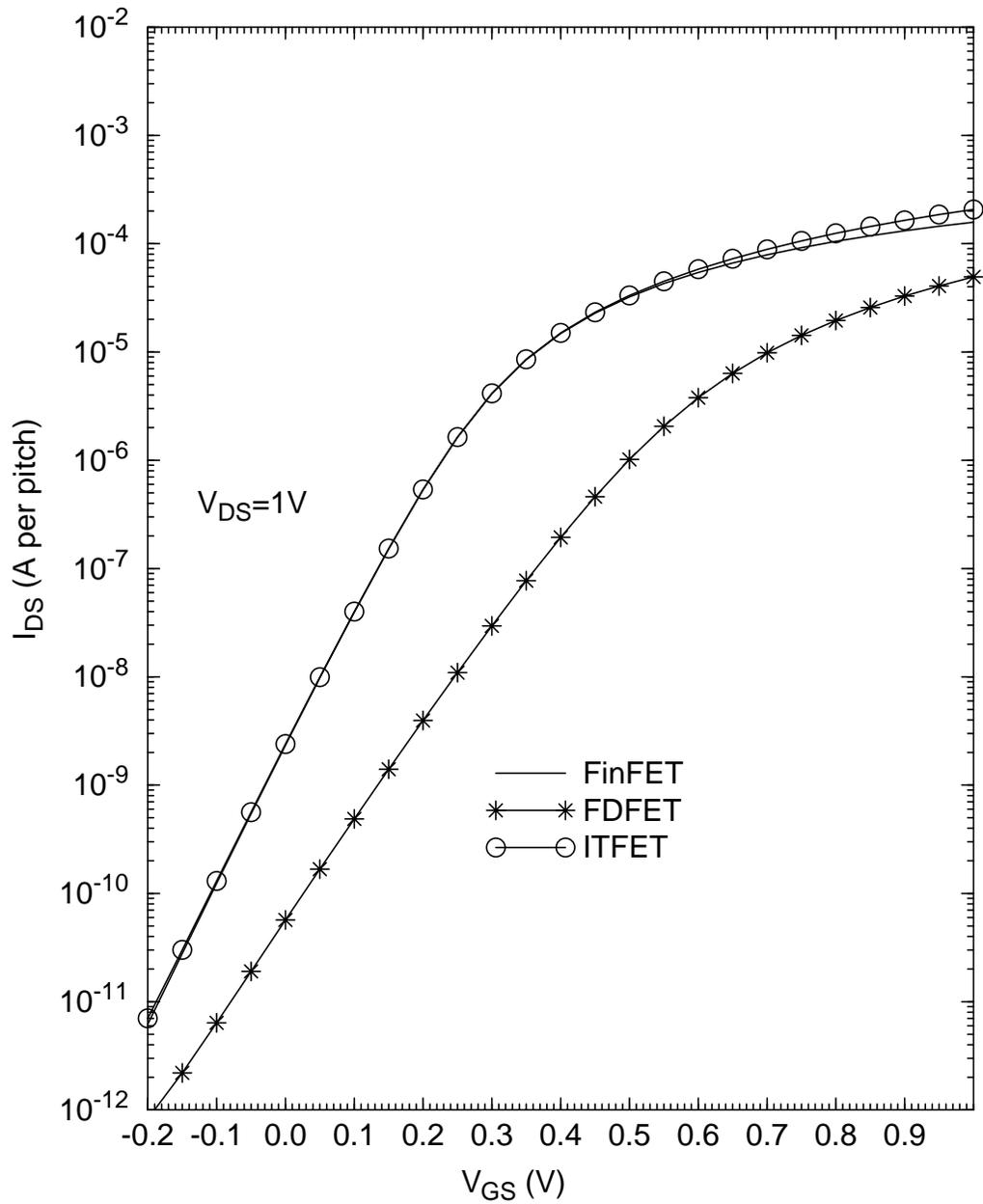


Figure 2-4. Taurus-predicted weak-inversion constituent currents of scalable ITFET with  $L_g = 18\text{nm}$ ,  $t_{Si} = 10\text{nm}$  (Table 2-3 design). No QM effects are turned on, and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V}\cdot\text{s}$ ) without velocity saturation is used.

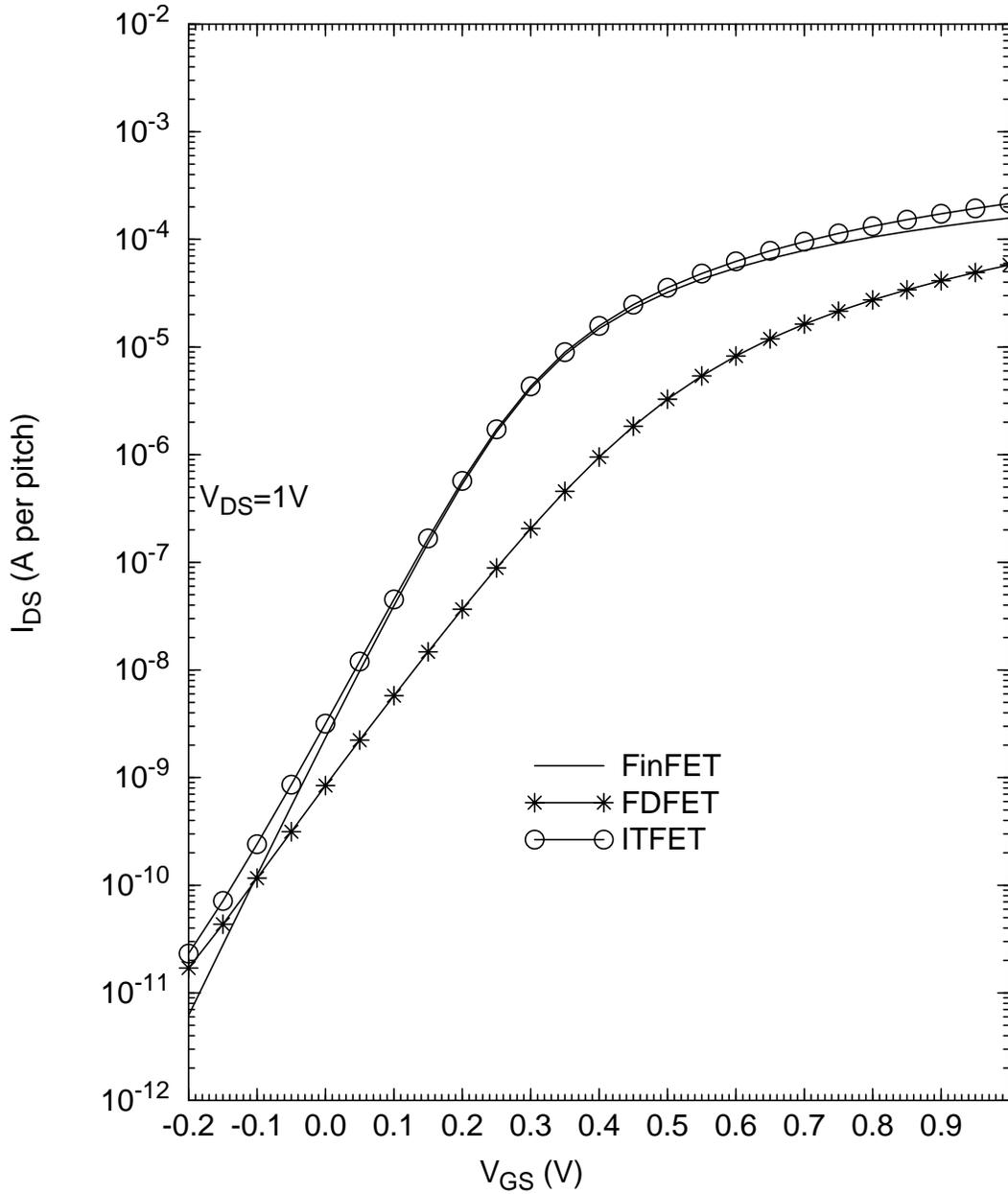


Figure 2-5. Taurus-predicted weak-inversion constituent currents of scalable ITFET with  $L_g = 18\text{nm}$ ,  $t_{Si} = 11\text{nm}$  (Table 2-3 design). No QM effects are turned on, and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V}\cdot\text{s}$ ) without velocity saturation is used.

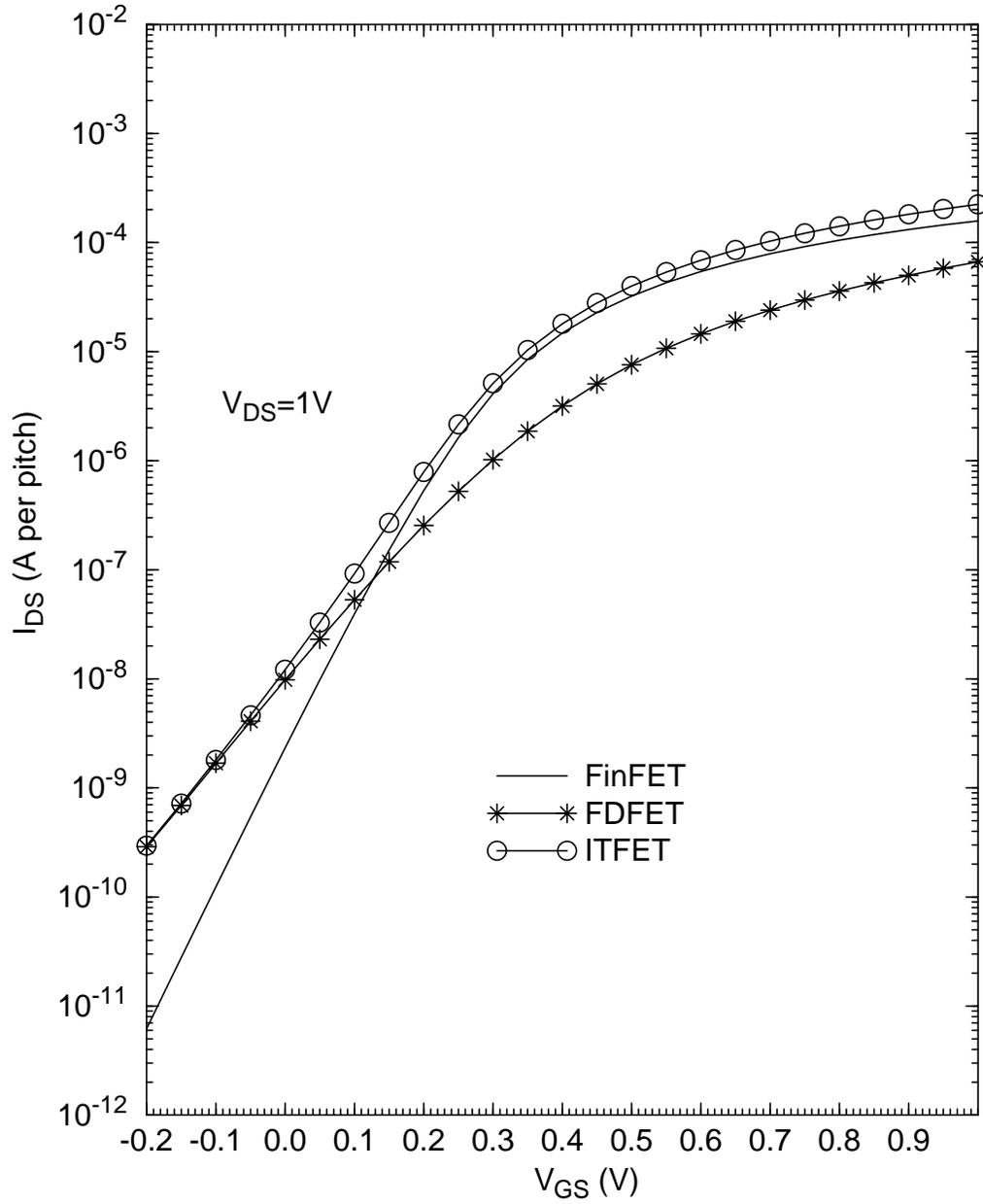


Figure 2-6. Taurus-predicted weak-inversion constituent currents of scalable ITFET with  $L_g = 18\text{nm}$ ,  $t_{Si} = 12\text{nm}$  (Table 2-3 design). No QM effects are turned on, and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V}\cdot\text{s}$ ) without velocity saturation is used.

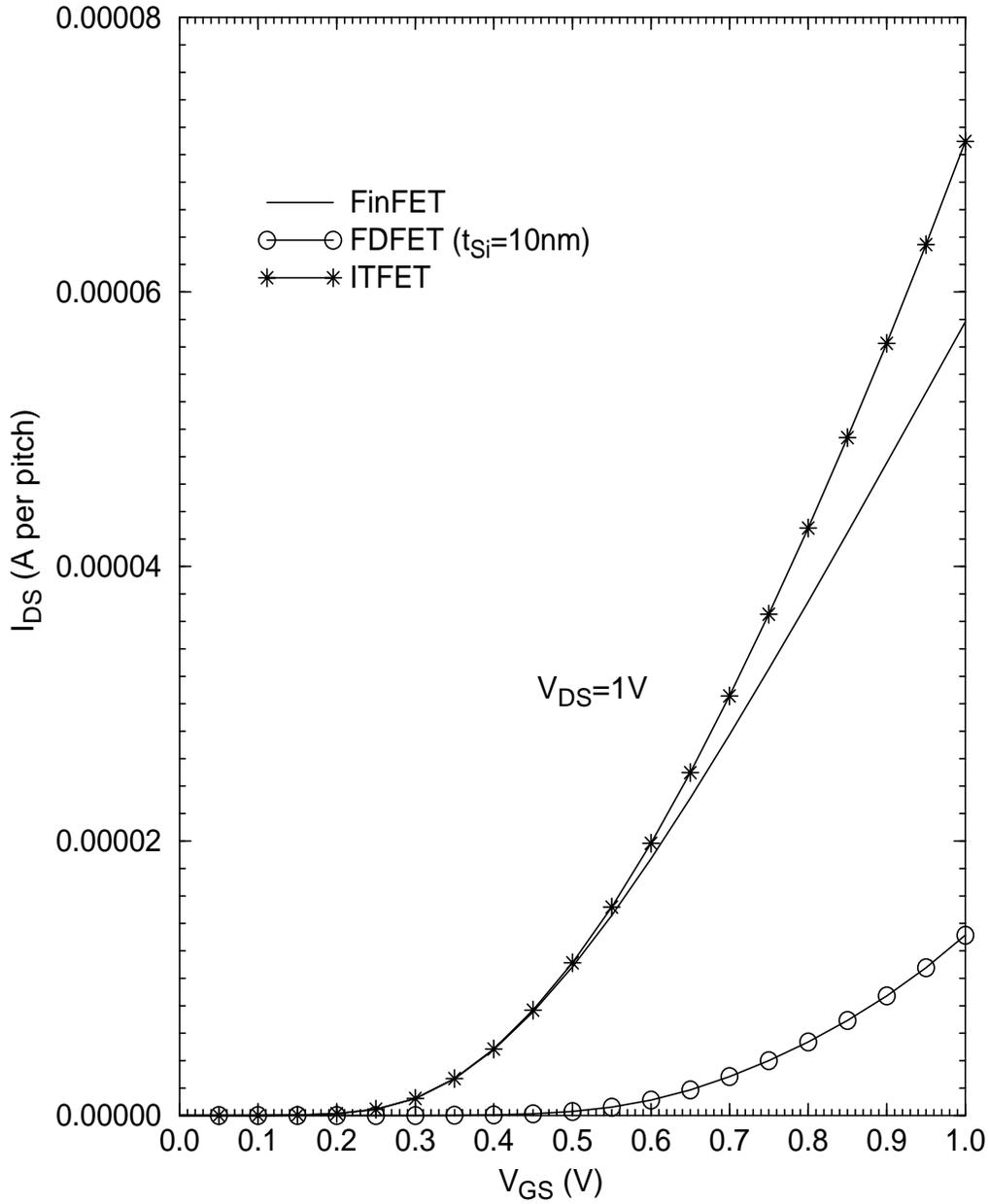


Figure 2-7. Taurus-predicted strong-inversion constituent currents of scalable ITFET with  $L_g = 18nm$ ,  $t_{Si} = 10nm$  (Table 2-3 design). No QM effects are turned on, and a constant electron mobility ( $\mu_n = 300cm^2/V\cdot s$ ) with velocity saturation is used.

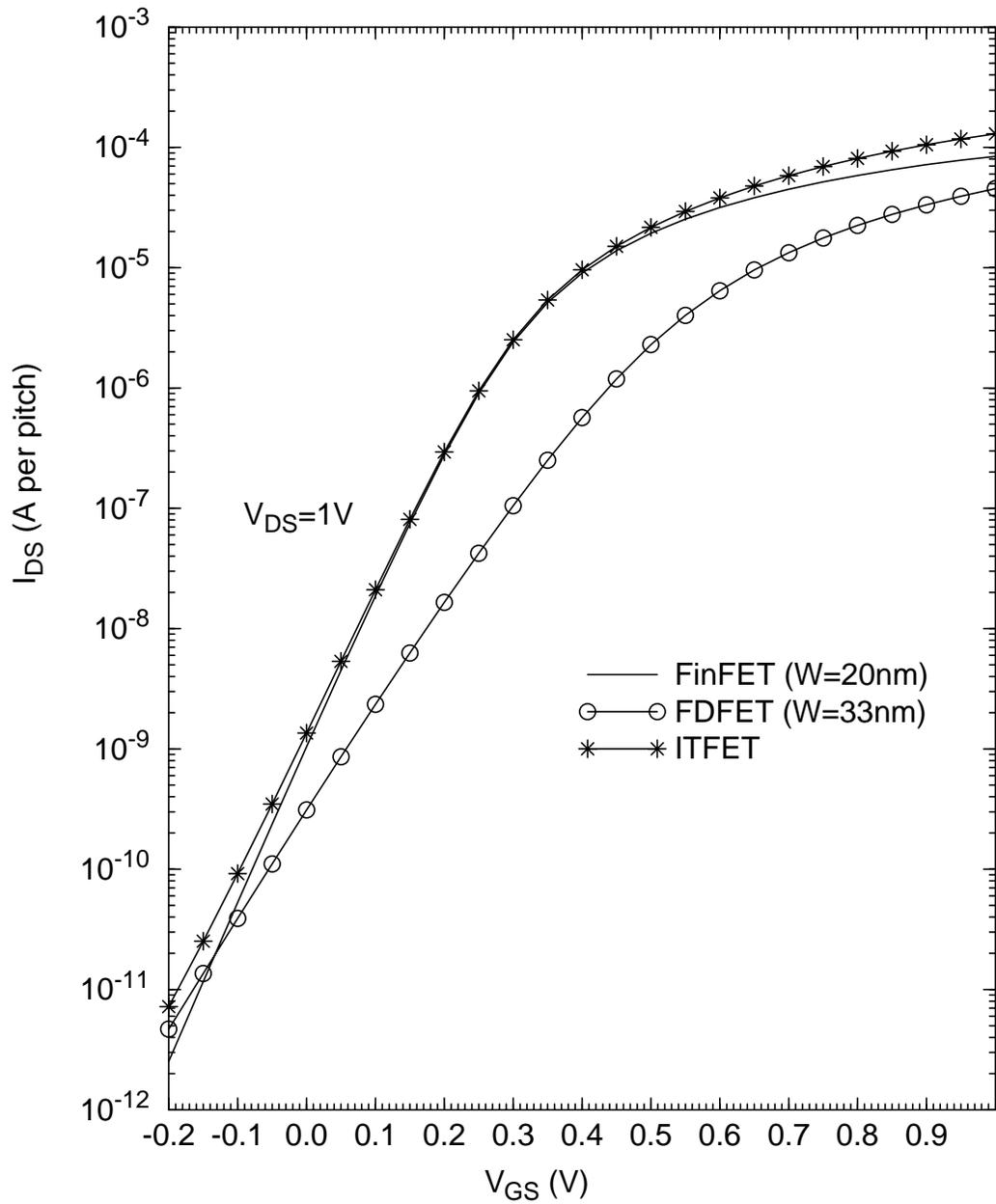


Figure 2-8. Taurus-predicted weak-inversion constituent currents of scalable ITFET with  $L_g = 9\text{nm}$  (Table 2-5 design). No QM effects are turned on, and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V}\cdot\text{s}$ ) without velocity saturation is used.

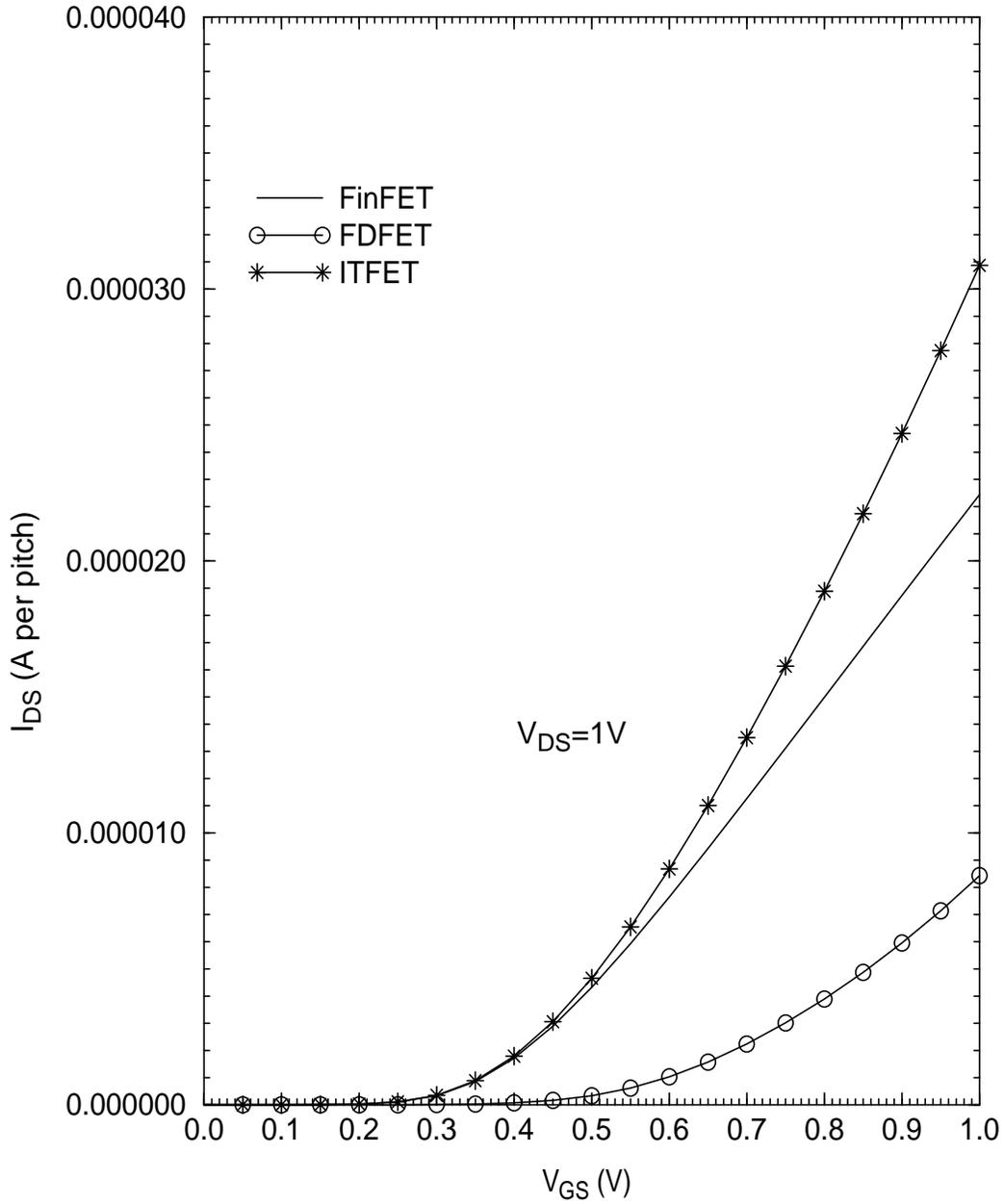


Figure 2-9. Taurus-predicted strong-inversion constituent currents of scalable ITFET with  $L_g = 9\text{nm}$  (Table 2-5 design). No QM effects are turned on, and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V}\cdot\text{s}$ ) with velocity saturation is used.

CHAPTER 3  
ON THE SUITABILITY OF A HIGH-K GATE DIELECTRIC IN NANOSCALE FINFET  
CMOS TECHNOLOGY

**3-1 Introduction**

In the nanoscale Double-Gate (DG) FinFET, control of short-channel effects (SCEs) is effected by the two, coupled gates on either side of a thin Si fin. As the gate length ( $L_g$ ) is scaled down, the fin-body ( $t_{Si}$ ) can be thinned further to maintain the SCE control, *without having to thin the gate dielectric* (EOT, or  $t_{ox}$ ) [2]. Such scaling is possible for the DG FinFET because of its zero body capacitance, which implies ideal subthreshold slope in the absence of SCEs, independent of  $t_{ox}$ . Further, bulk-inversion [20] in the undoped fin-body tends to render the gate capacitance (and hence inversion-charge density and on-state current) less dependent on a non-scaled  $t_{ox}$ . This *pragmatic* DG-FinFET CMOS design [2] could eliminate the need to replace the traditional SiON gate dielectric with a thicker high-k material [21][22] to avoid excessive gate tunneling current in the scaled device. However, a high-k gate dielectric could offer additional advantages, and the widespread notion is that it can significantly enhance CMOS performance as well as scalability, including DG-FinFET CMOS [23]. The possible high-k advantages include thinner effective oxide thickness (EOT), which implies higher gate capacitance ( $C_G$ ) and on-state current ( $I_{on}$ ), as well as improved scalability. Also, the larger physical thickness ( $t_{hk}$ ) of the high-k dielectric reduces the parasitic gate-source/drain (G-S/D) outer-fringe capacitance [24][25]. However, there are disadvantages as well. The larger  $t_{hk}$  results in the field-induced barrier lowering (FIBL) effect, which significantly degrades SCE control [25][26]. The channel mobility tends to be significantly degraded [27] due to both the poor quality of the Si-high-k dielectric interface and, fundamentally, the long-range scattering from optical phonons inherently present in high-k insulators [28]. Further, the integration of a high-k dielectric into the CMOS process presents formidable technological challenges [21][22].

Manoj and Rao [25] did a numerical simulation-based study of the impact of high-k dielectrics on nanoscale FinFET design and performance. They compared  $L_g = 32\text{nm}$  devices having different gate dielectric constants ( $k$ ). They maintained the same off-state current ( $I_{\text{off}}$ ) in each device by reducing  $t_{\text{Si}}$  with increasing  $k$  to suppress FIBL. They did CMOS ring-oscillator simulations, giving good insights, and reported a modest performance enhancement for “optimum”  $k \sim 20$ , relative to counterpart CMOS with  $\text{SiO}_2$  gate dielectric. However, their optimal high-k device needed  $t_{\text{Si}} = 6\text{nm}$ , as opposed to  $t_{\text{Si}} = 11\text{nm}$  for their  $\text{SiO}_2$  device. Using such an ultra-thin Si fin is not realistic at an  $L_g = 32\text{nm}$  technology node. Also, the design would not be scalable, irrespective of the technology, since  $t_{\text{Si}}$  cannot be thinned much below  $\cong 5\text{nm}$  due to the fundamental quantization effect on threshold voltage [19].

In this chapter, we do a more realistic assessment of high-k nanoscale DG FinFETs, comparing their projected CMOS performance with that of pragmatic DG-FinFET CMOS with  $\text{SiO}_2$  (or  $\text{SiON}$ ) gate dielectric. We use our process/physics-based (predictive) compact model for DG MOSFETs (UFDG [10]) in Spice3, supported by a numerical device simulator (Taurus [9]). We first design the devices at the HP-45nm technology node defined by the 2005 SIA ITRS [1], with  $L_g = 18\text{nm}$ , and project performances, noting heretofore unacknowledged compromising effects associated with the high-k dielectric in addition to those mentioned above. We then compare the scalability of the two technologies.

### 3-2 Device Design

We first design two  $L_g = 18\text{nm}$  DG FinFET structures using Taurus. Device I uses a relatively thick, pragmatic  $\text{SiO}_2$  gate dielectric and Device II uses a high-k gate dielectric. For Device I, we let  $t_{\text{ox}} = 1.2\text{nm}$ , thick enough to avoid excessive gate tunneling current ( $< 100\text{A}/\text{cm}^2$  [29]) in the HP device as specified in the ITRS [1]. To optimize the design, we incorporate a G-S/D underlap [11] via the S/D-extension lateral doping-density profile, assumed to be gaussian:

$N_{SD}(y) = N_{SD0} \exp(-y^2/\sigma_L^2)$ . The peak doping density at the edge of the S/D contact region is assumed to be  $N_{SD0} = 1 \times 10^{20} \text{cm}^{-3}$ . For an assumed S/D extension length ( $L_{ext}$ ) of 18nm, a reasonable doping straggle  $\sigma_L = 9.6\text{nm}$  yields, according to Taurus and UFDG simulations [11], an effective underlap  $L_{eSD} = 3.5\text{nm}$ , or a (weak-inversion) effective channel length  $L_{eff(weak)} = L_g + 2(3.5)\text{nm} = 25\text{nm}$ . We consider this  $N_{SD}(y)$  to be near-optimal, as well as reasonable, as explained in the following section. The gate work function is assumed to be midgap (e.g., TiN [18]), with the FinFET body/channel left undoped. We specify  $t_{Si}$  so as to obtain acceptable SCE performance,  $DIBL \cong 100 \text{ mV/V}$  and  $S < 90\text{mV}$ . Table 3-1 shows Taurus-predicted SCEs at room temperature, and  $I_{off}$  for (n-channel) Device I with different  $t_{Si}$ . No QM effects are turned on in Taurus and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V-s}$ ) is used. The carrier velocity-saturation effect is turned off (Appendix A). We choose  $t_{Si} = 12\text{nm}$  to meet the SCE specifications, getting  $I_{off} = 49\text{nA}/\mu\text{m}$  (per fin height). We stress, however, that  $I_{off}$  can be further lowered by thinning  $t_{Si}$  slightly, as shown in the table; and sensitivity to typical  $t_{Si}$  variations is acceptable [30].

The parasitic S/D series resistance,  $R_{S/D}$ , has two main components: the extension resistance  $R_{ext}$  and the contact resistance  $R_{con}$ . We estimate  $R_{ext}$  by numerical evaluation of

$$R_{ext} = \int_0^{L_{ext}} \frac{1}{q N_{SD}(y) \mu(N_{SD}) t_{Si}} dy, \quad (3-1)$$

where the carrier mobility  $\mu$  depends on  $N_{SD}$  via the classical Arora mobility model [31]. For the nFinFET,  $R_{ext}$  is estimated to be  $46\Omega\text{-}\mu\text{m}$ . Since the ITRS at 45nm targets  $R_{S/D} = 125\Omega\text{-}\mu\text{m}$  [1], we then assume a reasonable  $R_{con} = 79\Omega\text{-}\mu\text{m}$ . For the pFinFET, Eq. 3-1 yields  $R_{ext} = 77\Omega\text{-}\mu\text{m}$ . Then, again assuming  $R_{con} = 79\Omega\text{-}\mu\text{m}$ , we let  $R_{S/D} = 156\Omega\text{-}\mu\text{m}$ . The pragmatic Device I design is summarized in Table 3-2.

Device II, with a high-k dielectric, has almost the same structure as Device I. We assume an EOT of 0.7nm, in accordance with the HP-45nm node of the ITRS [1]. The gate dielectric constant is assumed to be  $k = 25$ , corresponding approximately to  $\text{HfO}_2$  [21], and hence  $t_{\text{hk}} = 4.5\text{nm}$ . With the thin EOT, the fin thickness can be relaxed to  $t_{\text{Si}} = 13.2\text{nm}$  for the same (Taurus-predicted)  $I_{\text{off}}$  as Device I; we do this to make a fair comparison of the two FinFET designs. With the thicker  $t_{\text{Si}}$ , Eq. 3-1 yields slightly lower  $R_{\text{ext}} = 42\Omega\text{-}\mu\text{m}/70\Omega\text{-}\mu\text{m}$  for the n/p-channel Device II, and so  $R_{\text{S/D}} = 121\Omega\text{-}\mu\text{m}/149\Omega\text{-}\mu\text{m}$ . The high-k Device II structure is summarized in Table 3-2 also. Taurus-predicted  $I_{\text{DS}}\text{-}V_{\text{GS}}$  curves of the n-channel versions of Devices I and II are plotted in Fig. 3-1 for low and high  $V_{\text{DS}}$ . Note the identical  $I_{\text{off}} = 49\text{nA}/\mu\text{m}$  for both devices, with  $I_{\text{on}}$  (for  $V_{\text{DD}} = 1.0\text{V}$ ) of Device II being only slightly higher than that of Device I. We note that these predictions are equivocal due to inadequate physical modeling in Taurus, which we discuss in the next section.

### 3-3 Simulation Setup

We use Taurus to simulate the basic FinFET structures (Table 3-2), and then use the results to calibrate our process/physics-based compact model, UFDG [10], which we use in Spice3 for CMOS performance projections. We are careful to avoid errors that may arise due to different physical models in Taurus and UFDG. So, we first turn-off the quantization models, and use a constant carrier mobility in both Taurus and UFDG (Appendix A) to achieve a good calibration in the weak-inversion region. This is necessary because the transport and quantization modeling in Taurus is, we believe, not well-suited to ultra-thin-body (UTB) devices. The calibration involves matching Taurus-predicted subthreshold characteristics by tuning in UFDG the effective G-S/D overlap ( $L_{\text{eSD}}$ ) [11]. It also involves matching Taurus-predicted weak-inversion  $C_{\text{G}}\text{-}V_{\text{GS}}$  characteristics by tuning two UFDG parameters related to the parasitic G-S/D fringe capacitance (FIF=0.8, QIF=5) [24]. Corresponding to the  $N_{\text{SD}}(y)$  assumed, we get from the UFDG calibration

$L_{\text{eSD}} = 3.5\text{nm}$ , as noted previously. This value of underlap has been shown to give near-optimal ring-oscillator (RO) intrinsic propagation delay for 18nm DG-FinFET CMOS [24]. Since UFDG assumes that the gate dielectric is  $\text{SiO}_2$ , we made changes in the model to account for the high-k dielectric in Device II. They include modifying the dielectric constant of the gate dielectric and changes to the parasitic capacitance model (Appendix B). In addition, since UFDG does not account for the FIBL effect [26], we use a small increase in the UFDG  $t_{\text{Si}}$  (from 13.2nm to 14nm) to match the Taurus-predicted subthreshold characteristics for Device II. This tuning of  $t_{\text{Si}}$  does not affect the predicted strong-inversion characteristics of the high-k device.

We then use UFDG, with its physical transport and quantization modeling, to perform quasi-predictive simulations of the DG-FinFET CMOS devices and ROs. UFDG includes QM-based models for UTB effective carrier mobility ( $\mu_{\text{eff}}$ ) [32], quantization-defined threshold-voltage shift ( $\Delta V_t$ ) [19] and inversion-layer capacitance ( $C_{\text{inv}}$ ) [33], carrier-velocity overshoot [34], and ballistic-limit current [4]. Typical values for the two  $\mu_{\text{eff}}$  parameters, derived from the model verification [32], are used. For electrons in the n-channel Device I with {110} fin surfaces, these parameters ( $\text{UO} = 1100\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.8$ ) yield  $\mu_{\text{n(eff)}} = 298\text{cm}^2/\text{V}\cdot\text{s}$  in the on-condition ( $V_{\text{DD}} = 1.0\text{V}$ ); and for holes ( $\text{UO} = 190\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 1$ ), they yield  $\mu_{\text{p(eff)}} = 118\text{cm}^2/\text{V}\cdot\text{s}$ . These mobilities are consistent with measured ones in actual undoped DG FinFETs [4]. For Device II, we neglect for now any mobility degradation due to the high-k dielectric [27][28]. For convenience in the discussion of the performance projections in the next section, we collectively refer to  $\mu_{\text{eff}}$  [32],  $\Delta V_t$  [19], and  $C_{\text{inv}}$  [33] as the QM model in UFDG. When the QM model is turned off, the quantization effects on  $V_t$  and  $C_{\text{inv}}$  are not modeled, and  $\mu_{\text{eff}}$  is replaced by a mobility ( $\mu_{\text{n}} = 298\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\mu_{\text{p}} = 118\text{cm}^2/\text{V}\cdot\text{s}$ ) with no dependence on  $t_{\text{Si}}$  nor transverse electric field ( $E_x$ ). However, dependence on the longitudinal electric field ( $E_y$ ) through the velocity-saturation/overshoot/ballistic-limit effects [34][4] is retained.

### 3-4 Performance Projections

The  $I_{DS}-V_{GS}$  characteristics predicted by UFDG for the n-channel Devices I and II are shown in Fig. 3-2. We note that the predicted enhancement in  $I_{on}$  in the high-k Device II, relative to the pragmatic Device I, is only 10%, which is not inconsistent with the Taurus predictions in Fig. 3-1. This seems surprisingly low since EOT is 58% (0.7/1.2) thinner than  $t_{ox}$  of Device I, making the equivalent oxide capacitance 72% higher. Examining our predictions, we find that the reason is two-fold. First,  $E_x$  in Device II is higher, at  $V_{GS} = 1.0V$ , because of the thinner EOT. This leads to additional  $C_G$ , or  $C_{inv}$ , reduction due to quantization [33] [35] (i.e., the electrical confinement of the inversion electrons) and hence a compromising of  $I_{on}$ . Fig. 3-3 shows the UFDG-predicted low- $V_{DS}$  intrinsic (without the parasitic G-S/D fringe capacitance)  $C_G-V_{GS}$  curves for Devices I and II, with the QM model turned on and off. In the on-condition,  $C_G$  for Device II is 71% higher than that for Device I when the QM model is turned off. However, when the QM model is turned on,  $C_G$  is only 40% higher. The undermining quantization effect is, therefore, exacerbated in Device II. Second,  $R_{S/D}$  in Device II is more detrimental. Due to the ohmic drop across  $R_S$ , the effective gate bias ( $V_{GS(eff)} = V_{GS} - I_{on}R_S$ ) in the on-condition is lower in Device II because of the higher  $I_{on}$ . The enhancement in  $I_{on}$  is thus compromised further. These two effects are reflected well by comparing Fig. 3-2 with Fig. 3-4, where we show predicted  $I_{DS}-V_{GS}$  characteristics of Devices I and II with the QM model turned off and  $R_{S/D} = 0$ . Note here that the predicted  $I_{on}$  enhancement ( $\cong 60\%$ ) is much higher than that in Fig. 3-2, and more in accord with the EOT/ $t_{ox}$  ratio. The enhancement is still less than the noted 72% because of the significant bulk (strong) inversion [20] in undoped DG MOSFETs, which implies that the intrinsic  $C_G$  is less than the gate oxide capacitance even when QM effects are not taken into account.

The UFDG-predicted  $I_{DS}-V_{GS}$  characteristics for the p-channel Devices I and II are shown in Fig. 3-5. We note here that there is almost no improvement in  $I_{on}$  afforded by the high-k

dielectric. We believe the additional compromising of the high-k enhancement in the p-channel Device II is related to the  $E_x$  dependence of hole mobility. In the nFinFET, because of the high electron mobility (Table 3-2),  $I_{on}$  is (ballistic-) limited by the injection velocity ( $v_{inj}$ ) at the source [4]:  $I_{on}/W = Q_{iS}v_{inj}$ , where  $Q_{iS}$  is the inversion charge density at the source which is enhanced by the high-k dielectric. However, in the pFinFET, the hole mobility is much lower (Table 3-2), and  $v_{inj}$  does not limit  $I_{on}$ . This means that  $I_{on}$  is more dependent on  $\mu_{p(eff)}$ , which is degraded ( $\sim 10\%$ ) more in Device II than in Device I because of the higher  $E_x$ . The high-k enhancement in  $I_{on}$  is thus further compromised in the pFinFET.

Since the increase in  $C_G$  yielded by the high-k dielectric (Fig. 3-3) is more than the increase in  $I_{on}$  (Fig. 3-2), the CV/I speed metric of Device II is actually degraded relative to pragmatic Device I. Whereas CV/I is not necessarily a reliable metric for undoped DG FinFETs [37], its implication here is borne out by results of UFDG/Spice3 unloaded CMOS-RO simulations. We consider three cases (all with equal nFinFET and pFinFET heights (i.e., device widths) because of the comparable currents in Figs. 3-2 and 3-5). Case I is the realistic case, with the devices in Table 3-2 simulated with the QM model on and with the noted  $R_{S/D}$ . The simulations for this case predict, as revealed in Table 3-3, that the RO with Device II actually has a 6% longer propagation delay ( $\tau_{pd}$ ) compared to the RO with Device I. We note also that since  $I_{on}$  of Device II is not significantly improved relative to Device I, this negative relative speed-projection result will not improve much with reasonable load capacitance. Case II is an ideal case, with the QM model turned off and  $R_{S/D} = 0$  (as in Fig. 3-4). The simulation results for this case, also given in Table 3-3, show  $\tau_{pd}$  of the Device II RO to be 19% shorter than that of the Device I RO. And, the much higher  $I_{on}$  of Device II for this case (Fig. 3-4) implies even better high-k performance enhancement with load capacitance. So, ideally, the high-k gate dielectric yields the anticipated enhancement, but, realistically, this enhancement is more than negated by the QM

effects and the finite  $R_{S/D}$ . To gain more insight on these compromising effects, we consider Case III in which the QM model is turned on, but with  $R_{S/D} = 0$ . The simulation results for this case, in Table 3-3, show that  $\tau_{pd}$  is virtually the same for the ROs with Devices I and II. Thus, considering all cases, we find that the QM and  $R_{S/D}$  effects both contribute significantly to the compromising of the anticipated performance advantage afforded by the high-k dielectric. We stress that we neglected any mobility degradation [27][28] due to the high-k dielectric in Device II. We note that this degradation cannot be completely suppressed, even in an “ideal” technology, because of the fundamental long-range scattering from the optical phonons present in high-k insulators [28]. Thus, the performance of the high-k FinFETs relative to the pragmatic ones will be even worse than what we are projecting.

### 3-5 Scalability

While speed performance is an important consideration, the primary benefit of a high-k gate dielectric is generally presumed to be enhanced CMOS scalability enabled by thin EOT without excessive gate tunneling current. As we noted in Sec. 3-2, we were able to relax  $t_{Si}$  from 12nm to 13.2nm for the same  $I_{off}$  in the high-k Device II. If we assume that a minimum  $t_{Si}$  correlates with a particular technology node, then this relaxation of  $t_{Si}$  afforded by the high k (i.e., by thinner EOT) translates to some improvement in device scalability. To quantify the improvement, and describe the actual impact of high k on scalability, we thinned  $t_{Si}$  in Device II back to 12nm (as for Device I), and then shortened  $L_g$  (keeping the same G-S/D underlap) to match, via Taurus simulations, the SCEs of the original Device II. We found that  $L_g$  can be shortened to 16.5nm (from the original 18nm) for the same  $I_{off}$  and similar SCEs. Thus, we get only about an 8%  $L_g$ -scaling boost, relative to the pragmatic Device I, from the high-k gate dielectric. The limited enhancement in scalability can be attributed to the FIBL effect [25][26], which undermines the SCE control afforded by the thin EOT and hence limits the allowed

increase in  $t_{Si}$ , or shortening of  $L_g$ , for the same  $I_{off}$ . This modest scaling benefit is really not worthwhile when we consider the reduced performance, as characterized in Sec. 3-4, as well as the formidable technological challenges of integrating a high-k dielectric into the process flow. In fact, additional simulations show that the same SCE control can be achieved with Device I scaled to 16.5nm by merely thinning  $t_{Si}$  from 12nm to 10.8nm. Such scaling, with the better Device I performance, seems to be the logical, pragmatic choice.

### 3-6 Summary and Discussion

Via physics-based device/circuit simulations, we have studied the impact of a high-k gate dielectric on the performance and scalability of nanoscale DG-FinFET CMOS. We designed a high-k FinFET (Device II), and compared it with a *pragmatic* FinFET (Device I) having thicker (> EOT), conventional SiO<sub>2</sub> gate dielectric at the HP-45nm node ( $L_g = 18\text{nm}$ ) of the 2005 SIA ITRS. We found that the high-k dielectric actually degrades CMOS-speed performance due to two heretofore unacknowledged compromising effects of the high k. One is additional  $C_G$  reduction due to quantization because of higher transverse electric field resulting from thinner EOT; the  $I_{on}$  enhancement expected for the thinner EOT is thus compromised. The second is additional loss of effective gate bias due to  $R_S$  because of the higher  $I_{on}$ ; the enhancement in  $I_{on}$  is thus compromised more. Further, in the pFinFET, hole mobility is degraded due to the higher transverse field, compromising  $I_{on}$  even more. We stress that we have not considered any degradation in channel mobility resulting from the integration of the high-k dielectric [27][28], which will cause further loss of performance relative to the pragmatic DG FinFET. Further, high-k dielectrics also suffer from reliability issues which make their integration into the process flow very challenging [36].

Our study focused on the HP-45nm node. However, our conclusions are applicable to LOP and LSTP applications [1] as well, for which thicker SiON for Device I [2] can be used to

adequately suppress gate current [29], with the fin thickness scaled to meet the ITRS  $I_{\text{off}}$  requirement. Further, the conclusions remain valid for scaling DG-FinFET CMOS to the end of the ITRS, which is enabled by well-designed G-S/D underlap [11][30] in the pragmatic device [2].

Our simulation-based study did imply a small improvement in the scalability of DG-FinFET CMOS due to the thinner EOT. However, given the reduced performance and the technological challenges of integrating a high-k dielectric into the process flow [36], we conclude that it is not worthwhile. We believe, also based on our study, that the pragmatic approach to nanoscale FinFET CMOS should be taken, with good performance expected to the end of the ITRS.

Table 3-1. Taurus-predicted weak-inversion characteristics of the pragmatic 18nm DG nFinFET (Device I) at room temperature for different values of Si-fin thickness. The off-state current is given per Si-fin height. No QM effects are turned on in Taurus and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V-s}$ ) is used. The carrier velocity saturation effect is turned off (Appendix A).

$t_{\text{Si}}$ (nm)	DIBL (mV/V)	S (mV)	$I_{\text{off}}$ (nA/ $\mu\text{m}$ ) at $V_{\text{DS}}=1.0\text{V}$
10	78	74	4
11	89	76	14
12	100	78	49
13	122	83	160

Table 3-2. Physical parameters for the pragmatic Device I and the high-k Device II. The parameters are the same for the nFinFET and pFinFET unless noted otherwise.

Parameter	Device I	Device II
Gate length ( $L_g$ )	18nm	18nm
Fin thickness ( $t_{\text{Si}}$ )	12nm	13.2nm
Effective gate oxide thickness (EOT)	1.2nm	0.7nm
Gate dielectric constant (k)	3.9	25
Gate work function	midgap	midgap
S/D-extension length ( $L_{\text{ext}}$ )	18nm	18nm
S/D-extension doping straggle ( $\sigma_L$ )	9.6nm	9.6nm
S/D contact resistance ( $R_{\text{con}}$ )	$79\Omega-\mu\text{m}$	$79\Omega-\mu\text{m}$
S/D-extension resistance ( $R_{\text{ext}}$ ) for nFinFET	$46\Omega-\mu\text{m}$	$42\Omega-\mu\text{m}$
S/D-extension resistance ( $R_{\text{ext}}$ ) for pFinFET	$77\Omega-\mu\text{m}$	$70\Omega-\mu\text{m}$
S/D resistance ( $R_{\text{S/D}}=R_{\text{ext}}+R_{\text{con}}$ ) for nFinFET	$125\Omega-\mu\text{m}$	$121\Omega-\mu\text{m}$
S/D resistance ( $R_{\text{S/D}}=R_{\text{ext}}+R_{\text{con}}$ ) for pFinFET	$156\Omega-\mu\text{m}$	$149\Omega-\mu\text{m}$
Effective electron mobility in on-condition ( $\mu_{\text{neff}}$ ) for nFinFET	$298\text{cm}^2/\text{V-s}$	$260\text{cm}^2/\text{V-s}$
Effective hole mobility in on-condition ( $\mu_{\text{peff}}$ ) for pFinFET	$118\text{cm}^2/\text{V-s}$	$107\text{cm}^2/\text{V-s}$

Table 3-3. UFDG/Spice3-predicted propagation delays of unloaded CMOS ring oscillators comprising pragmatic-Device I and high-k-Device II 18nm DG FinFETs, for three cases as noted.

Case	Device I $\tau_{pd}$ (ps)	Device II $\tau_{pd}$ (ps)	High-k $\tau_{pd}$ benefit (%)
I. Realistic: QM model on, actual $R_{S/D}$ (Table 3-2)	1.94	2.06	-6
II. Ideal: QM model off, $R_{S/D} = 0$	0.73	0.59	+19
III. Quasi-Ideal: QM model on, $R_{S/D} = 0$	1.08	1.07	+1

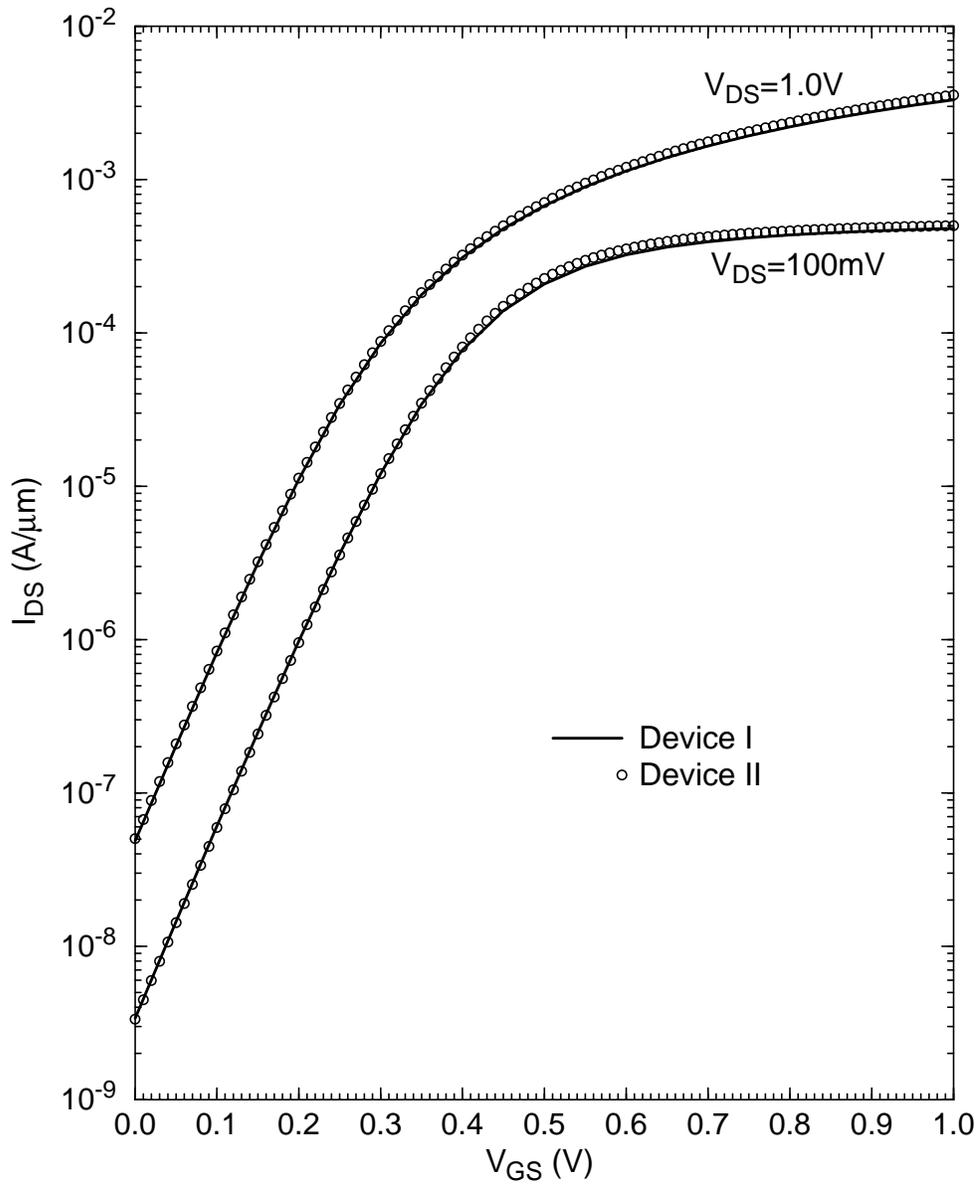


Figure 3-1. Taurus-predicted  $I_{DS}$ - $V_{GS}$  characteristics (per fin height) of n-channel 18nm DG FinFETs: pragmatic Device I and high-k Device II. No QM effects are turned on, and a constant electron mobility ( $\mu_n = 300\text{cm}^2/\text{V}\cdot\text{s}$ ) without velocity saturation is used

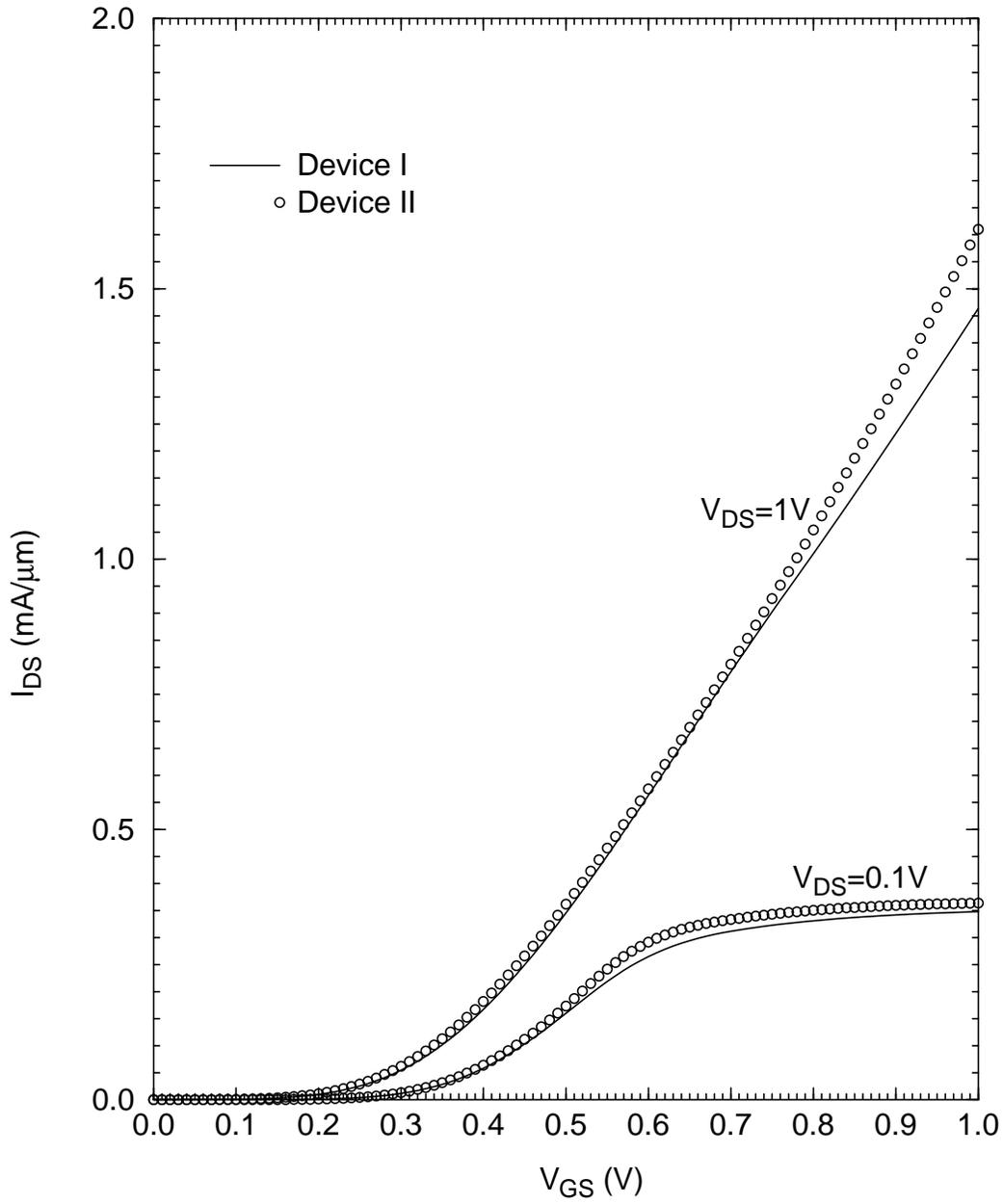


Figure 3-2. UFDG-predicted  $I_{DS}$ - $V_{GS}$  characteristics (per fin height) of n-channel pragmatic Device I and high-k Device II. The QM model is turned on and  $R_{S/D} = 125\Omega\text{-}\mu\text{m}/121\Omega\text{-}\mu\text{m}$  for Device I/II.

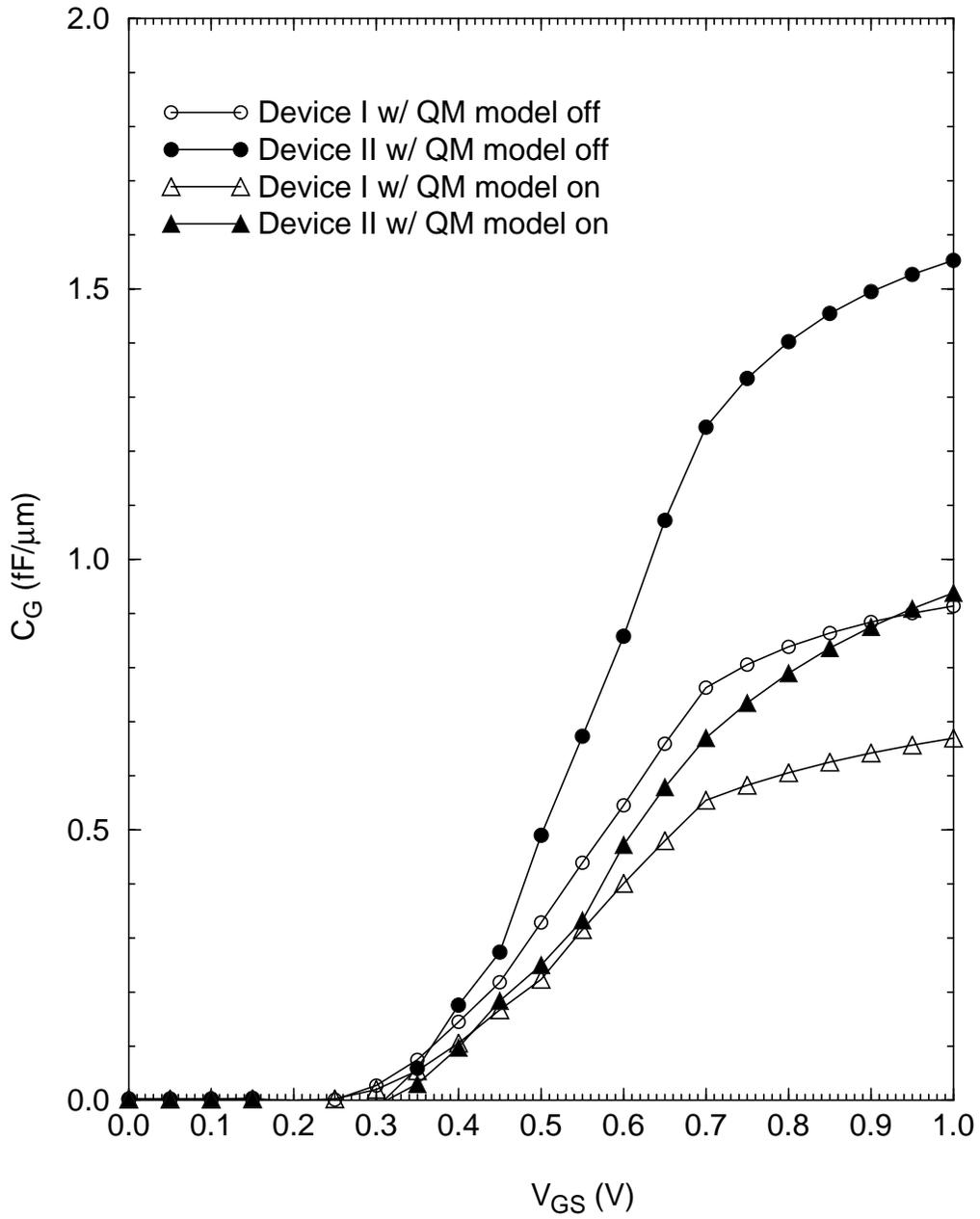


Figure 3-3. UFDG-predicted low- $V_{DS}$  intrinsic  $C_{GS}$ - $V_{GS}$  characteristics (per fin height) of n-channel pragmatic Device I and high-k Device II, with the QM model turned on and off.

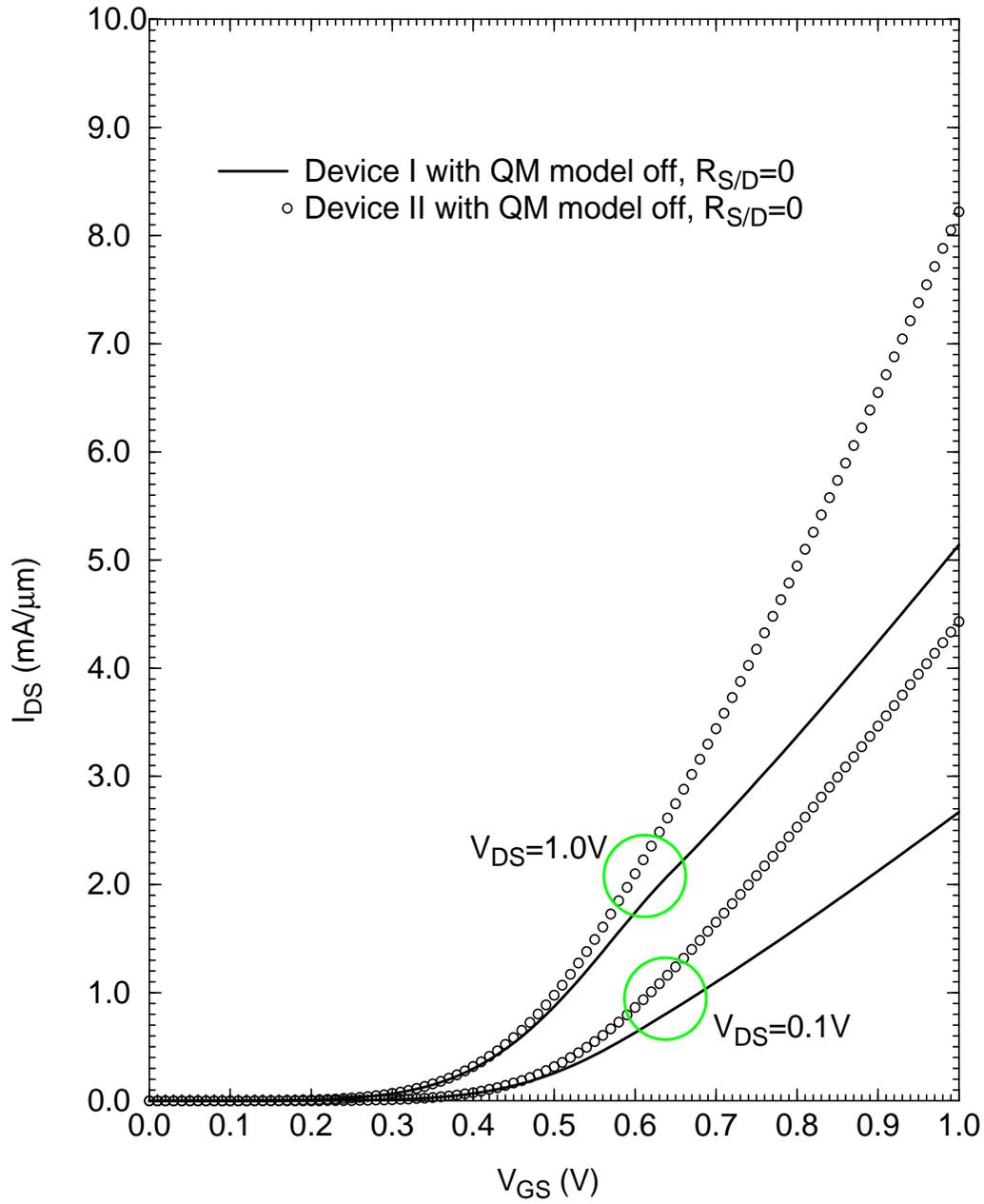


Figure 3-4. UFDG-predicted  $I_{DS}$ - $V_{GS}$  characteristics (per fin height) of n-channel pragmatic Device I and high-k Device II. The QM model is turned off and  $R_{S/D} = 0$ .

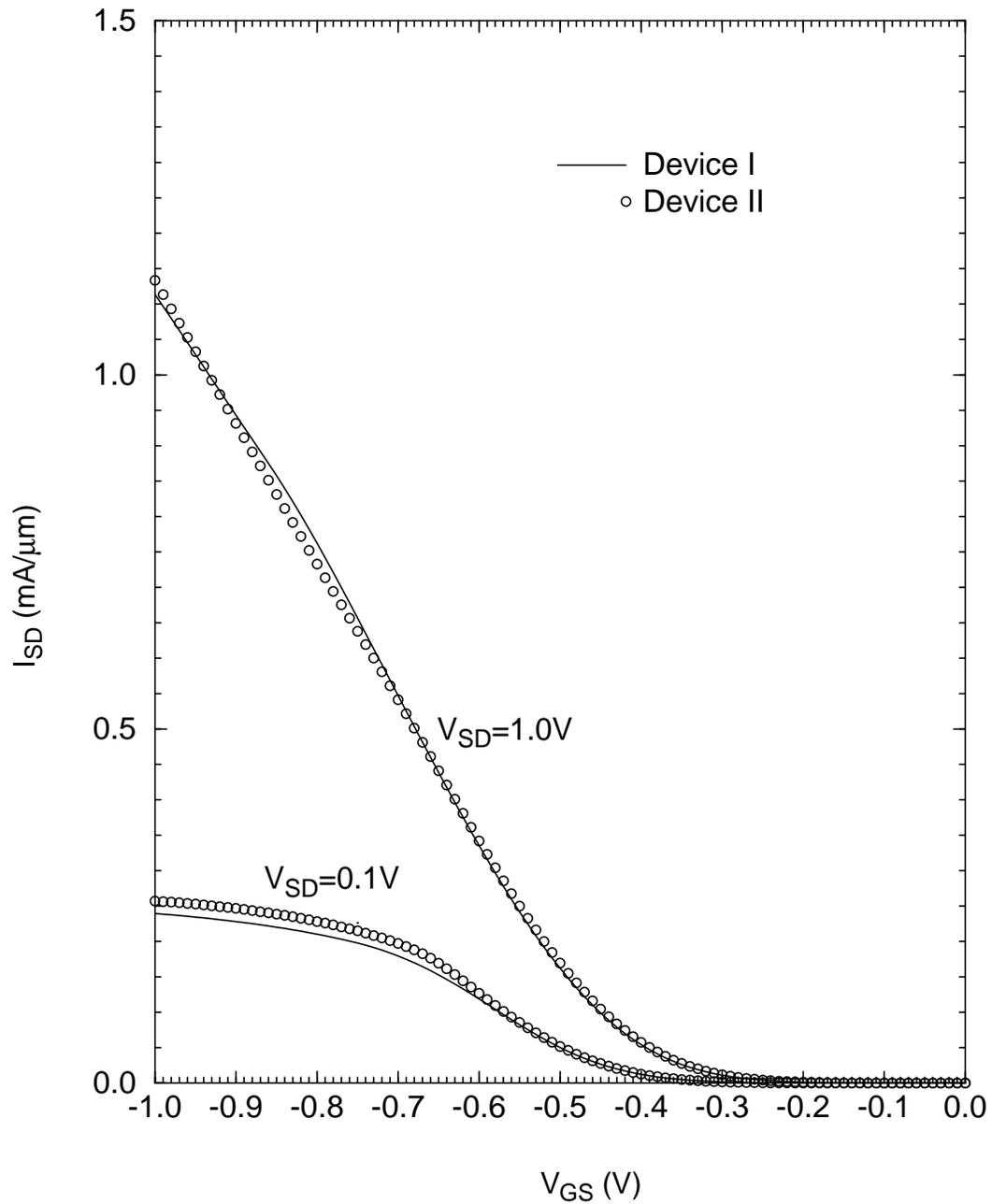


Figure 3-5. UFDG-predicted  $I_{SD}$ - $V_{GS}$  characteristics (per fin height) of p-channel pragmatic Device I and high-k Device II. The QM model is turned on and  $R_{S/D} = 156\Omega\text{-}\mu\text{m}/149\Omega\text{-}\mu\text{m}$  for Device I/II.

CHAPTER 4  
A PHYSICAL MODEL FOR FRINGE CAPACITANCE IN DOUBLE-GATE MOSFETS WITH  
NON-ABRUPT SOURCE/DRAIN JUNCTIONS

**4-1 Introduction**

Double-Gate (DG) MOSFETs typically have undoped bodies; they are therefore designed with reasonably long extensions ( $L_{\text{ext}}$ ) to prevent substantial density of source/drain (S/D) dopants in the channel. S/D dopants diffuse into the extension during the S/D implant/anneal process thereby defining the S/D lateral doping profile,  $N_{\text{SD}}(y)$ .  $N_{\text{SD}}(y)$  can be reasonably assumed to be gaussian;  $N_{\text{SD}}(y) = N_{\text{SD0}}\exp(-(y-L_{\text{ext}})^2/\sigma_L^2)$ , where  $y=0$  is chosen to be the edge of the gate (Fig. 4-1A). Due to the relatively light doping density near the gate edges, the gate can modulate portions of the S/D extension via the inner and outer fringing fields (Fig. 4-2). This encroachment of the gate field effect into the S/D extension causes a lengthening of the effective channel length in weak inversion,  $L_{\text{eff(weak)}}$ , which defines short-channel effects (SCEs) [11]. The SCEs of an underlapped device (Fig. 4-1A) with gate length,  $L_g$ , can be characterized in terms of a simplified device structure of gate length,  $L_{\text{eff(weak)}} = L_g + 2L_{\text{eSD}}$  (Fig. 4-1B).  $L_{\text{eSD}}$  is governed by  $N_{\text{SD}}(y)$  and the spacer dielectric constant,  $k_{\text{sp}}$ .  $L_{\text{eSD}}$  is not modeled directly; it is defined and derived by matching the SCEs of the actual device in Fig. 4-1A to those defined by the simplified device structure in Fig. 4-1B. Shorter  $\sigma_L$  implies fewer dopants in the S/D extension, which causes the inner fringing field to penetrate further into the extension thereby lengthening  $L_{\text{eSD}}$ . Similarly, increasing  $k_{\text{sp}}$  causes stronger coupling between the gate and S/D extension causing a lengthening of  $L_{\text{eSD}}$ . In strong inversion, the inversion charge in the body screens out the inner fringing field via the short Debye length,  $L_D \propto (1/n)^{0.5}$ , where  $n$  is the inversion charge density. Thus, the effective channel length in strong inversion (which defines on-state current) is

$$L_{\text{eff(strong)}} \cong L_g.$$

The G-S/D parasitic fringe capacitance,  $C_f$ , of nanoscale DG MOSFETs is a significant CMOS speed limiter [24].  $C_f$  consists of two components: the inner fringe capacitance,  $C_{if}$ , and the outer fringe capacitance,  $C_{of}$  as reflected in Fig. 4-2. It has been related to  $L_{eSD}$  via modeling that assumes abrupt S/D junctions [24]. Such  $C_f$  modeling provides good qualitative insight, but is inaccurate because it approximates  $N_{SD}(y)$  as an abrupt junction at  $y=L_{eSD}$  in weak inversion. The inaccuracy is due to the fact that  $L_{eSD}$  is the effective *gated* extension which defines SCEs in the simplified structure (Fig. 4-1B). Thus,  $L_{eSD}$  is different from the length that relates to the depletion of the S/D extension,  $y=L_u$ , in the actual structure (Fig. 4-1A), which defines  $C_f$  [24]. Accurate models for  $C_{of}$  in terms of  $L_u$  are presented in [38] and [39], but they do not provide any insight on how  $L_u$  is defined in a realistic device structure with finite  $\sigma_L$ . In this chapter, we characterize  $C_{if}$  and  $L_u$  in terms of  $L_{eSD}$  and the device structure. Then, we characterize  $C_{of}$  in terms of  $L_u$  using the modeling in [38]. The result is then a complete accounting for the effects of the crucial G-S/D underlap in nanoscale DG MOSFETs. We present models for  $C_{if}$  and  $L_u$  in Sec. 4-2 and 4-3 respectively. In Sec. 4-4, we verify our modeling by comparing predictions with numerical simulation results for varying device parameters, UTB thickness ( $t_{Si}$ ), gate oxide thickness ( $t_{ox}$ ), gate height ( $t_g$ ), and  $k_{sp}$ . In Sec. 4-5, we implement the model for  $C_f$  in our physical, process-based model, UFDG. In Sec. 4-6, we perform ring oscillator (RO) simulations using UFDG/Spice3 to study the significance of  $k_{sp}$  in defining CMOS circuit speed.

#### 4-2 Characterization of $C_{if}$

Fig. 4-1A depicts the actual DG MOSFET device structure. The spacer width defines the S/D extension length,  $L_{ext}$ . Lateral S/D dopant diffusion defines a gaussian S/D profile,  $N_{SD}(y)$ , in the extension region. The origin of the co-ordinate system is taken to be at the edge of the gate-oxide as shown. We assume  $V_{DS} = 0$  to simplify the analysis, since numerical simulation results show that  $C_f$  is only weakly dependent on  $V_{DS}$ . Our  $C_f$  modeling applies to the drain as well as to

the source. Charge in the extension region, which goes to zero (neutrality) for increasing y, can be related to  $C_{if}$  and  $C_{of}$  as follows:

$$(V_{GS} - \Phi_{GS})(C_{if} + C_{of}) = -q \int_{y=0}^{\infty} \int_{x=0}^{t_{Si}/2} [N_{SD}(y) - n(x, y)] dx dy \quad (4-1)$$

where  $\Phi_{GS}$  is the workfunction difference between the gate and S/D extension, and  $n(x, y)$  is the electron density in the extension. Further, we can write Poisson's equation in the extension as:

$$\frac{dE_x}{dx} + \frac{dE_y}{dy} = \frac{q}{\epsilon_{si}} [N_{SD}(y) - n(x, y)] \quad (4-2)$$

We make a reasonable assumption, based on numerical simulations, that  $dE_x/dx$  is constant in x at given y;  $dE_x/dx \cong -E_{sx}(y)/(t_{Si}/2)$  where  $E_{sx}$  is the electric field at the extension-spacer interface in the x-direction. Now, we integrate Eq. 4-2 along x and y directions to get:

$$\int_{y=0}^{\infty} \int_{x=0}^{t_{Si}/2} \frac{dE_y}{dy} dx dy = \frac{q}{\epsilon_{si}} \int_{y=0}^{\infty} \int_{x=0}^{t_{Si}/2} [N_{SD}(y) - n(x, y)] dx dy + \int_{y=0}^{\infty} E_{sx}(y) dy \quad (4-3)$$

We make the reasonable assumption that  $C_{of}$  is associated only with the x-component of the electric field,  $E_{sx}$ , (Fig. 4-2) which implies:

$$\int_{y=0}^{\infty} \epsilon_{si} E_{sx}(y) dy = C_{of} (V_{GS} - \Phi_{GS}) \quad (4-4)$$

Combining Eq. 4-1, 4-3, and 4-4 we get:

$$\int_{x=0}^{t_{Si}/2} (E_y(x, \infty) - E_y(x, 0)) dx = \frac{-C_{if}}{\epsilon_{si}} (V_{GS} - \Phi_{GS}) \quad (4-5)$$

With  $E_y(x, \infty) = 0$  and letting  $\int_{x=0}^{t_{Si}/2} E_y(x, y) dx = \overline{E_Y}(y) \frac{t_{Si}}{2}$ , we differentiate Eq. 4-5 to get

a simple but physical expression for  $C_{if}$ :

$$\frac{d\overline{E_Y}(y=0)}{dV_{GS}} = \frac{2C_{if}}{\epsilon_{si}t_{Si}} . \quad (4-6)$$

To characterize  $C_{if}$  in weak inversion ( $C_{ifw}$ ), we recognize that if the SCEs of the actual device (Fig. 4-1A) are comparable to the simplified device structure (Fig. 4-1B), the potential distributions within  $L_g$  of the two structures should also be comparable. We therefore use the 2D Poisson solution in weak inversion of the simplified structure to approximate the derivative in Eq. 4-6. Recognizing that  $y=0$  in the actual device structure maps to  $y=L_{eSD}$  in the simplified structure and using Yeh's analysis for the latter [40], we get:

$$\frac{d\overline{E_Y}(y=0)}{dV_{GS}} \approx \frac{dE_y(t_{Si}/2, L_{eSD})(simplified)}{dV_{GS}} = \frac{\exp(-L_{eSD}/\lambda)}{\lambda} \quad (4-7)$$

where

$$\lambda = \frac{t_{Si}}{2} \sqrt{0.5 \left( 1 + \frac{4\epsilon_{si}t_{ox}}{\epsilon_{ox}t_{Si}} \right)}$$

From Eq. 4-6 and Eq. 4-7, we now have

$$C_{ifw} = \frac{\epsilon_{si}t_{Si} \exp(-L_{eSD}/\lambda)}{2\lambda} . \quad (4-8)$$

Eq. 4-8 is a compact relation for  $C_{ifw}$  in terms of the device structural parameters and  $L_{eSD}$ . It physically relates  $C_{ifw}$  to the G-S/D underlap that governs the SCEs. In strong inversion,  $E_y$  gets screened out by free carriers in the body. This implies that the LHS of Eq. 4-6 is approximately zero, and thus  $C_{if}$  in strong inversion goes to  $C_{ifs} = 0$ .

### 4-3 Characterization of $L_u$ and $C_{of}$

To characterize  $C_{of}$ , we first characterize  $L_u$ , which is the depleted length of the S/D extension region; beyond  $L_u$ , towards the S/D, there is significant free-carrier density, and so  $L_u$  defines  $C_{of}$  [38]. To characterize  $L_u$ , we assume the simplified picture of the inner and outer fringing field illustrated in Fig. 4-2. We assume that the electric field in  $0 < y < L_u$ , which defines  $C_{if}$ , is predominantly y-directional ( $E_y$ ) while the field in  $y > L_u$ , which defines  $C_{of}$ , is x-directional (as was assumed to get Eq. 4-4). In effect, we assume that the depletion charge in  $0 < y < L_u$  is related exclusively to  $C_{if}$ . Then, using an average  $\bar{N}_{SD}$  for  $N_{SD}(y)$  in  $0 < y < L_u$ , we can write:

$$(V_{GS} - \Phi_{GS})C_{if} \approx -q\bar{N}_{SD}L_u \frac{t_{Si}}{2} . \quad (4-9)$$

Also, using the 1-D Poisson's equation for  $0 < y < L_u$  we have,

$$\Psi(L_u) - \Psi(0) = \Psi_{bi} - \Psi(0) = \frac{q}{2\varepsilon_{si}}\bar{N}_{SD}L_u^2 \quad (4-10)$$

where  $\Psi_{bi}$  is the built-in potential of the S/D junction. Combining Eq. 4-9 and Eq. 4-10 to eliminate  $\bar{N}_{SD}$  we get

$$L_u = \frac{-\varepsilon_{si}t_{si}(\Psi_{bi} - \Psi(0))}{(V_{GS} - \Phi_{GS})C_{if}} . \quad (4-11)$$

To evaluate  $L_u$  in weak inversion, we let  $C_{if} = C_{ifw}$  in Eq. 4-11 and estimate  $\Psi(0)$  using the weak-inversion solution [40] of the simplified device structure (as in Sec. 4-2).

$$\Psi(0) \approx \Psi(t_{si}/2, L_{eSD})(simplified) \quad (4-12)$$

$$= \frac{(\Psi_{bi} - V_{GS} + \Phi_{MS})}{\sinh\left(\frac{L_{eff}}{\lambda}\right)} \left[ \sinh\left(\frac{L_{eSD}}{\lambda}\right) + \sinh\left(\frac{L_{eff} - L_{eSD}}{\lambda}\right) \right] + (V_{GS} - \Phi_{MS})$$

Here,  $\Phi_{MS}$  is the workfunction difference between the gate and body. The  $V_{GS}$  dependence of  $L_u$

(via Eq. 4-11 and Eq. 4-12) is relatively weak. We therefore choose a nominal value of  $V_{GS}$  for the purpose of evaluating  $L_u$ . For an undoped DG MOSFET with midgap gates, we choose  $V_{GS}=0$ . Thus, noting that  $\Phi_{GS}\approx\psi_{bi}$  for undoped bodies, we get from Eq. 4-11:

$$L_u = \frac{\epsilon_{si} t_{si}}{C_{ifw}} \left( 1 - \frac{\Psi(0)}{\Psi_{bi}} \right) . \quad (4-13)$$

Eq. 4-13 with the modeling in [38] analytically gives  $C_{ofw}$  in terms of  $L_{eSD}$  and the device structure. In strong-inversion,  $C_{ifs}=0$ ; thus Eq. 4-9 implies that  $L_u=0$ . Then, [38] defines  $C_{ofs}$ .

#### 4-4 Verification and Discussion

To evaluate our model, we use Taurus [9] to do 2-D simulations of DG MOSFETs with different structures, and compare  $C_f$  predicted by Taurus to our model-predicted  $C_f = C_{if} + C_{of}$ . The device structures assumed are summarized in Table 4-1. The S/D doping profile is assumed to be gaussian,  $N_{SD}(y) = N_{SD0} \exp(-(y-L_{ext})^2/\sigma_L^2)$  in Fig. 4-1A;  $N_{SD0}$  is the peak S/D doping density and  $\sigma_L$  is the lateral doping straggle. To evaluate  $L_{eSD}$ , we tune the Taurus-predicted SCEs of a given device to our physical/process-based model, UFDG [10]. UFDG models the weak-inversion region of a device via an analytical solution [40] for the 2D poisson equation in the UTB of the simplified device structure in Fig. 4-1B. Table 4-2 gives  $L_{eSD}$  evaluated for all four device structures in Table 4-1 with different  $k_{sp}$ . Increasing  $L_{ext}$  or reducing  $\sigma_L$  implies fewer dopants near the gate edge. Fewer dopants near the gate edge implies fewer free carriers, due to which the inner fringing field can penetrate deeper into the extension region thereby increasing  $L_{eSD}$ . Also, as  $k_{sp}$  increases,  $L_{eSD}$  increases, indicating better control of the potential in a portion of the extension through the spacer.

Fig. 4-3 shows Taurus- and model-predicted  $C_f$  in weak inversion,  $C_{fw}$  vs.  $k_{sp}$  for all four device structures. We observe that there is good agreement between the numerical results and the model. We also find that for increasing  $k_{sp}$  and also for increasing  $L_{eSD}$ , the model becomes less

accurate. We believe this is because our assumption (Fig. 4-2) that  $C_{if}$  and  $C_{of}$  are defined exclusively by  $E_y$  (Eq. 4-9) and  $E_x$  (Eq. 4-4), respectively, tends to break down. For increasing  $k_{sp}$  and  $L_{eSD}$ , the outer fringing field tends to have a significant  $E_y$ -component which contributes to  $L_u$ . Since Eq. 4-9 does not take this into account, it tends to underpredict  $C_{ofw}$ . However, for reasonable DG MOSFET design where  $k_{sp} = 1-7.5$  and  $L_{eSD} = 1-4nm$ , the model predictions are quite good ( $< \sim 10\%$  error). Fig. 4-4 displays Taurus- and model-predicted  $C_f$  in strong inversion,  $C_{fs}$  vs.  $k_{sp}$  for Device III. We observe that the model is reasonable and predicts  $C_{fs}$  with  $< \sim 15\%$  error. Further, we evaluate the model for different values of the device structural parameters,  $t_g$ , and  $t_{Si}$ . Fig. 4-5 and 4-6 display  $C_{fw}$  and  $C_{fs}$  respectively vs.  $t_g$  for Device III; Fig. 4-7 displays  $C_{fw}$  vs.  $t_{Si}$  for Device III. On the whole, we observe that the model predicts the trends with respect to the device structure reasonably well with errors  $< \sim 15\%$ .

#### 4-5 Model Implementation in UFDG (Ver. 3.8)

We implement our analytical model for  $C_f$  in our physical/process-based model, UFDG [10]. UFDG is a compact Poisson-Schrodinger solver for undoped UTB devices with regional analyses for the weak- [40] and strong-inversion [41] regions linked by a 2-D cubic spline for the moderate-inversion region. G-S/D underlap is modeled via different effective channel lengths for the weak-inversion ( $L_{eff(weak)} = L_g + L_{eS} + L_{eD}$ ) [11] and strong-inversion ( $L_{eff(strong)} = L_g$ ) [41] solutions. UFDG physically accounts for QM-based effects: increase in threshold voltage due to structural confinement [19] and inversion-layer capacitance [33]. UFDG also models transport phenomena via a QM-based carrier mobility model [32], carrier temperature-dependent velocity overshoot [34], and ballistic-limited current [4]. The above modeling makes UFDG quasi-predictive and hence useful for performance projections and design of nanoscale DG CMOS technology.

The model implementation is done as follows. A flag (CFF) is used to turn the fringe capacitance model on (CFF = 1,2) and off (CFF = 0). When CFF = 1, the current model for  $C_f$  [24] is retained and when CFF = 2, the above-described model is used. The above-described model is for symmetrical Double-Gate (SDG) devices and should only be used as such. It can be extended for asymmetrical Double-Gate (ADG) and FD/SOI devices in the future. The implementation is facilitated by the regional analyses mentioned above, with the model for  $C_{fw} = C_{ifw} + C_{ofw}$  implemented in the weak-inversion solution and for  $C_{fs} = C_{ifs} + C_{ofs}$  implemented in the strong-inversion solution. The terminal charges defined by  $C_{fw}$  and  $C_{fs}$  are added to the intrinsic gate and source/drain charges in the weak- and strong-inversion solutions respectively. The 2-D spline interpolates the terminal charges in the moderate-inversion region. Currently, UFDG (ver. 3.7) implicitly assumes that the spacer dielectric is  $\text{SiO}_2$  ( $k_{sp} = 3.9$ ). To generalize the model, we add a parameter for the spacer permittivity (KSP). KSP is used to define  $C_{ofw}$  and  $C_{ofs}$ . In the weak-inversion solution,  $C_{ifw}$  is computed using Eq. 4-8 and  $L_u$  is computed using Eq. 4-13. Equations (5), (7) and (12) in [38] define  $C_{ofw}$  in terms of  $L_u$ ,  $t_g$ ,  $t_{ox}$ ,  $k_{sp}$ , and the S/D width,  $L_{SD}$ . Since  $C_{ofw}$  is a very weak function of  $L_{SD}$ , we hard-code the value of  $L_{SD} = 50\text{nm}$ . In the strong-inversion model for  $C_f$ ,  $C_{ifs}$  is assumed to be zero consistent with the insight in Sec. 4-2.  $C_{ofs}$  is defined in terms of  $L_u = 0$ ,  $t_g$ ,  $t_{ox}$ ,  $k_{sp}$ , and  $L_{SD} = 50\text{nm}$  using Eqs. (5), (7), and (12) in [38]. Fig. 4-8 displays UFDG-predicted  $C_G$ - $V_{GS}$  characteristics at low and high  $V_{DS}$  for Device III (Table 4-3) with  $k_{sp} = 1$  and  $k_{sp} = 7.5$ . We note the significance of  $C_f$  in defining  $C_G(V_{GS})$  which underscores the need to reduce  $C_f$  for optimal CMOS speed performance.

#### 4-6 Significance of $C_f$ for Optimal DG CMOS Design

It has been shown [24] that  $C_f$  is a significant speed limiter in nanoscale DG CMOS technology.  $C_f$  can be reduced by using a low-k dielectric spacer, preferably an air-spacer [42]. To study the significance of spacer material on optimal speed performance of nanoscale FinFET

CMOS, we use UFDG to perform quasi-predictive ring oscillator (RO) simulations of Device III with different  $k_{sp}$ . We assume realistic mobility parameters in UFDG ( $UO = 1100\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.8$  for nFinFET;  $UO = 190\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 1$  for pFinFET) which correspond to actual measured mobilities in undoped UTB FinFETs [32]. These parameters imply electron mobility,  $\mu_{N(\text{eff})} = 298\text{cm}^2/\text{V}\cdot\text{s}$  and hole mobility,  $\mu_{P(\text{eff})} = 118\text{cm}^2/\text{V}\cdot\text{s}$  in the on-condition ( $V_{DD}=1\text{V}$ ). While increasing  $k_{sp}$  leads to increased  $C_f$ , it also leads to reduced  $R_{S/D}$  via modulation of the carrier density in the S/D extension region by the gate-field through the spacer. Taurus simulations of Device III with different  $k_{sp}$  show that  $I_{on}$  increases by 7% and 12% when  $k_{sp}$  is increased from  $k_{sp} = 1$  to 3.9 and 7.5 respectively. To account for this effect in our RO simulations, we tune  $R_{S/D}$  in UFDG to reflect the increased  $I_{on}$  for higher  $k_{sp}$ . We assume  $R_{S/D} = 125\Omega\text{-}\mu\text{m}$  [1] for  $k_{sp} = 1$ . For the same percentage change in  $I_{on}$  with increasing  $k_{sp}$ , we set  $R_{S/D} = 105\Omega\text{-}\mu\text{m}$  for  $k_{sp} = 3.9$  and  $R_{S/D} = 90\Omega\text{-}\mu\text{m}$  for  $k_{sp} = 7.5$ . We also assume equal nFinFET and pFinFET widths since the UFDG-predicted  $I_{on}$  for both devices is comparable. Table 4-3, which displays the per stage delay  $t_{pd}$  vs.  $k_{sp}$ , shows the significance of reducing  $k_{sp}$  to improve circuit performance. We note that  $t_{pd}$  reduces by ~25% when the spacer material is changed from  $\text{Si}_3\text{N}_4$  to  $\text{SiO}_2$  and by ~40% when it is changed to air ( $k_{sp} = 1$ ). We stress that  $C_f$  does not scale with  $L_g$ ; the importance of  $C_f$  in defining circuit speed would increase with scaling  $L_g$ . Also, the gate to S/D contact capacitance [43] has not been considered in our discussion; this would increase with reducing gate pitch and become increasingly important in defining circuit speed.

$C_f$  can also be reduced by increasing  $L_{eSD}$  via appropriate  $N_{SD}(y)$  design as is evident from the device designs in Table 4-1 (which also implies reduced SCEs). However, such design also increases the  $R_{S/D}$  [11], implying lower  $I_{on}$  and thereby defining a design tradeoff.  $N_{SD}(y)$  needs to be carefully engineered to effect an optimal trade-off between SCEs,  $C_f$ , and  $I_{on}$ . We examine this problem in Chapter 5 where we develop and demonstrate a reverse-engineering

methodology to relate  $N_{SD}(y)$  to the S/D process. Based on the extracted  $N_{SD}(y)$ , the S/D process can be redesigned to effect a better device design trade-off.

#### **4-7 Summary and Discussion**

We have presented, for the first time, a physical model for parasitic fringe capacitance ( $C_f$ ) in DG MOSFETs with (realistic) non-abrupt source/drain junctions. We evaluated our model for different device structures and found good agreement with numerical simulation results. The model is implemented in our physical/process-based model, UFDG. We performed ring-oscillator simulations using UFDG in Spice3 to demonstrate the significance of  $C_f$  in defining CMOS circuit speed. Our model will be helpful in DG MOSFET design to trade-off SCEs, on-state current and parasitic capacitance for optimal CMOS performance.

Table 4-1. Physical parameters for the Device Structures I-IV

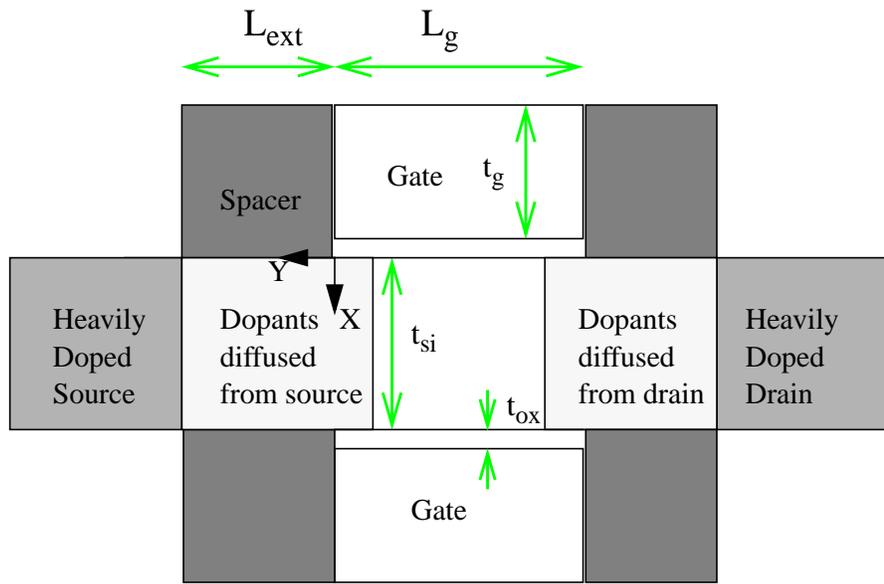
Parameter	Device I	Device II	Device III	Device IV
Gate length ( $L_g$ ) (nm)	18	18	18	18
Fin thickness( $t_{Si}$ ) (nm)	10	10	10	10
Effective gate oxide thickness (EOT) (nm)	1.2	1.2	1.2	1.2
Gate height ( $t_g$ ) (nm)	15	15	15	15
Gate work function	midgap	midgap	midgap	midgap
S/D peak doping density ( $N_{SD0}$ )( $cm^{-3}$ )	$1 \times 10^{20}$	$1 \times 10^{20}$	$1 \times 10^{20}$	$1 \times 10^{20}$
S/D-extension length ( $L_{ext}$ )(nm)	12	12	18	18
S/D-extension doping straggle ( $\sigma_L$ ) (nm)	8	7.07	9.6	7.77

Table 4-2. Effective gated extension length,  $L_{eSD}$  obtained for Devices I-IV for different values of spacer permittivity,  $k_{sp}$ , and material.  $L_{eSD}$  was obtained by calibrating the SCEs of a given device structure to UFDG.

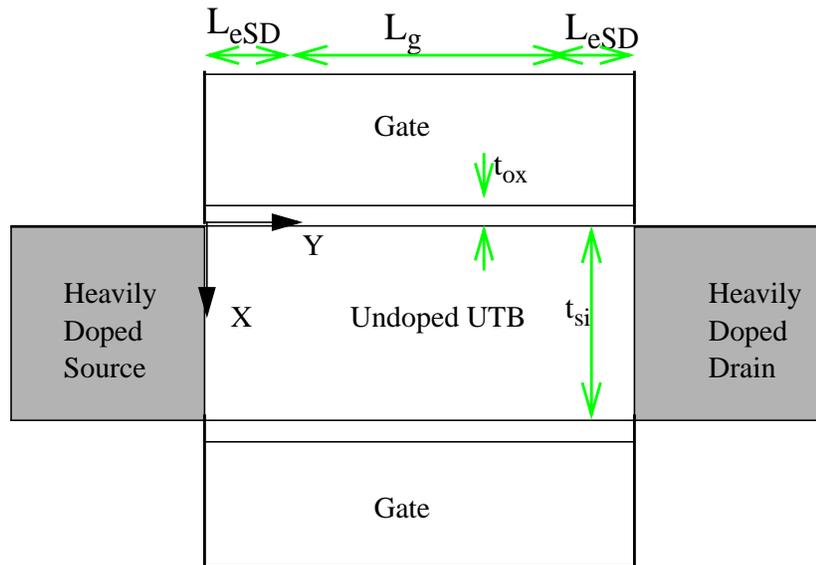
$k_{sp}$ (material)	$L_{eSD}$ (nm) Device I	$L_{eSD}$ (nm) Device II	$L_{eSD}$ (nm) Device III	$L_{eSD}$ (nm) Device IV
0	1.10	2.00	2.80	4.00
1 (Air)	1.20	2.20	3.10	4.50
3.9 (SiO <sub>2</sub> )	1.65	2.50	3.50	5.00
7.5 (Si <sub>3</sub> N <sub>4</sub> )	2.20	3.00	4.30	5.60

Table 4-3. UFDG/Spice3-predicted propagation delays of unloaded CMOS ROs comprising of Device III with different spacer materials

$k_{sp}$ (material)	$t_{pd}$ (ps)
1 (Air)	1.67
3.9 (SiO <sub>2</sub> )	2.22
7.5 (Si <sub>3</sub> N <sub>4</sub> )	2.78



A



B

Figure 4-1. The actual DG MOSFET device structure is portrayed in A. The spacer width defines the S/D extension length,  $L_{ext}$ . S/D dopant diffusion defines a gaussian S/D profile,  $N_{SD}(y)$ , in the extension region. B portrays the simplified structure with effective gate length,  $L_{eff(weak)} = L_g + 2L_{eSD}$  used to characterize the SCEs of the actual device. Note that different y coordinates are used for the actual and simplified structures.

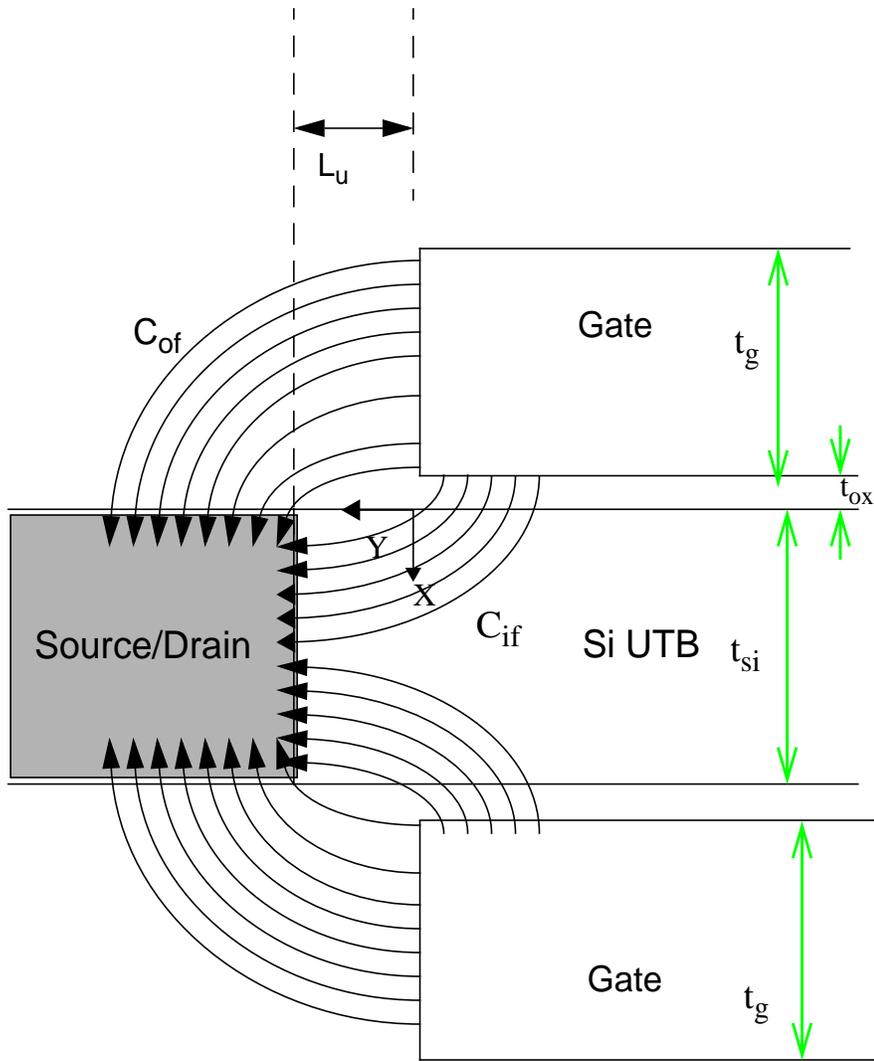


Figure 4-2. This figure portrays the fringing fields that underlie  $C_f$ . It is assumed that the inner fringing field is predominantly y-directional while the outer fringing field is x-directional. The depleted length in the y-direction,  $L_u$ , is defined by  $C_{if}$ .  $C_{of}$  is then defined by  $L_u$  and the device structure, as shown.

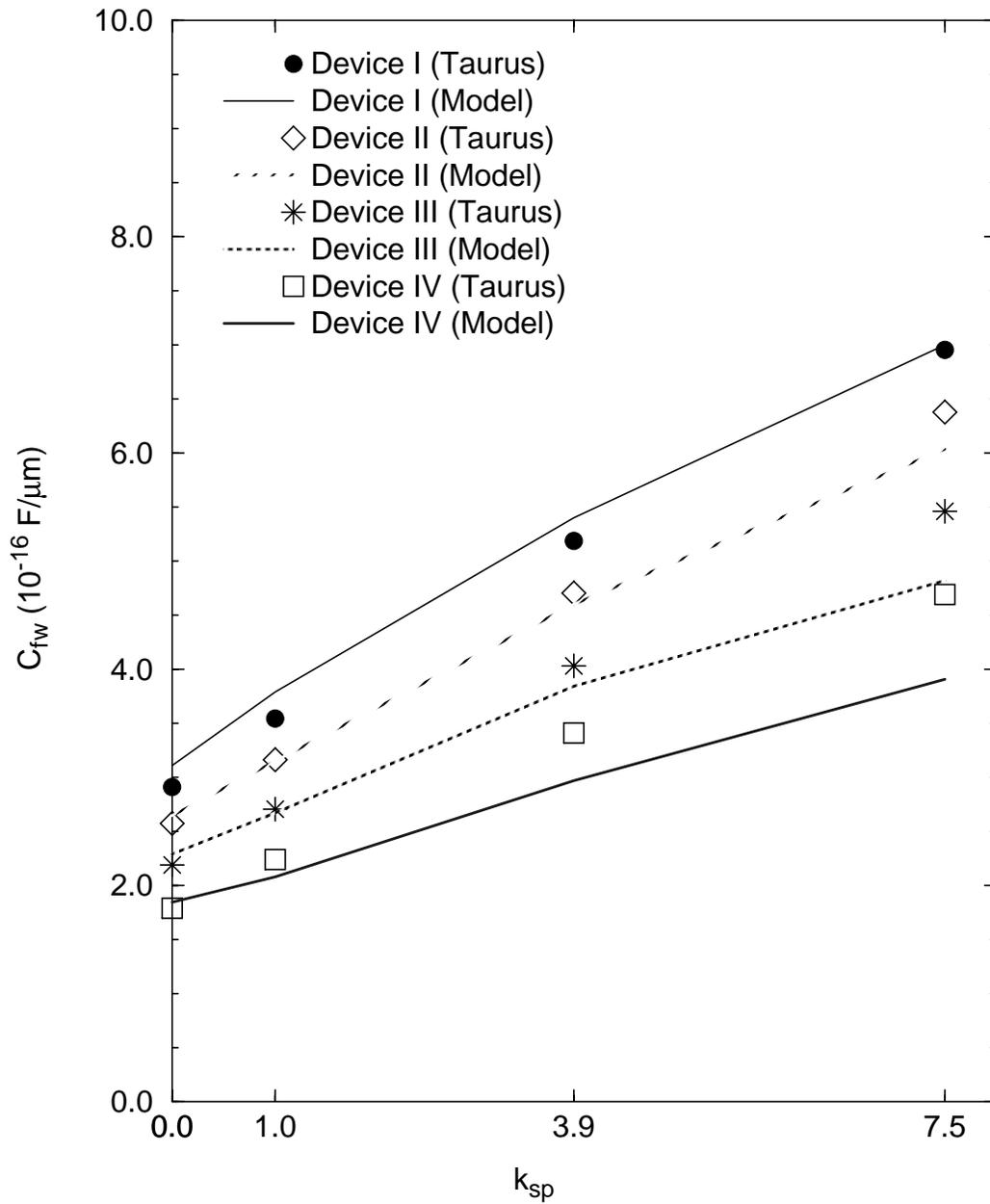


Figure 4-3. Taurus- and model-predicted weak-inversion fringe capacitance,  $C_{fw}$  v/s spacer permittivity,  $k_{sp}$  for Device I-IV.

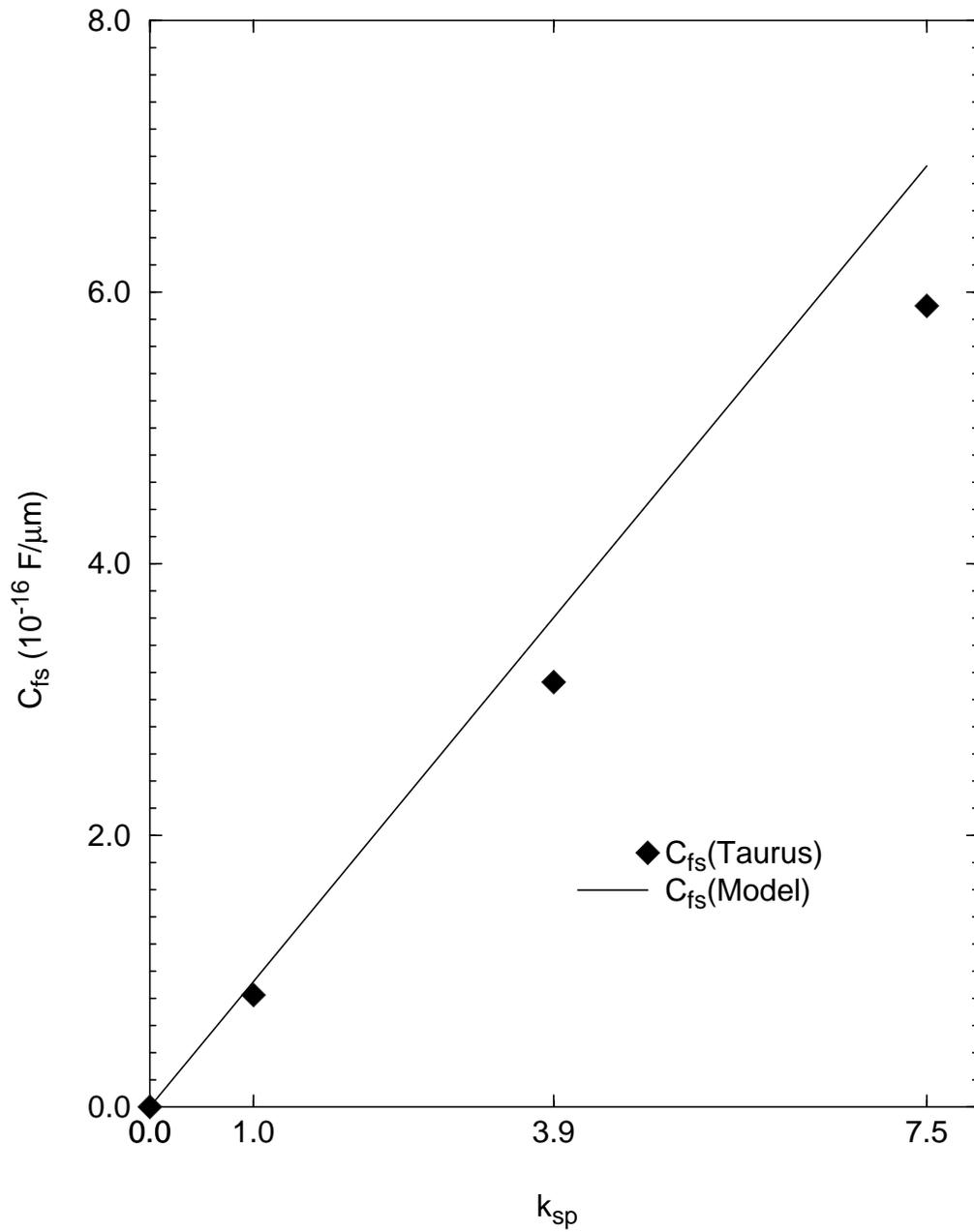


Figure 4-4. Taurus- and model-predicted strong-inversion fringe capacitance,  $C_{fs}$  v/s spacer permittivity,  $k_{sp}$ , for Device III.

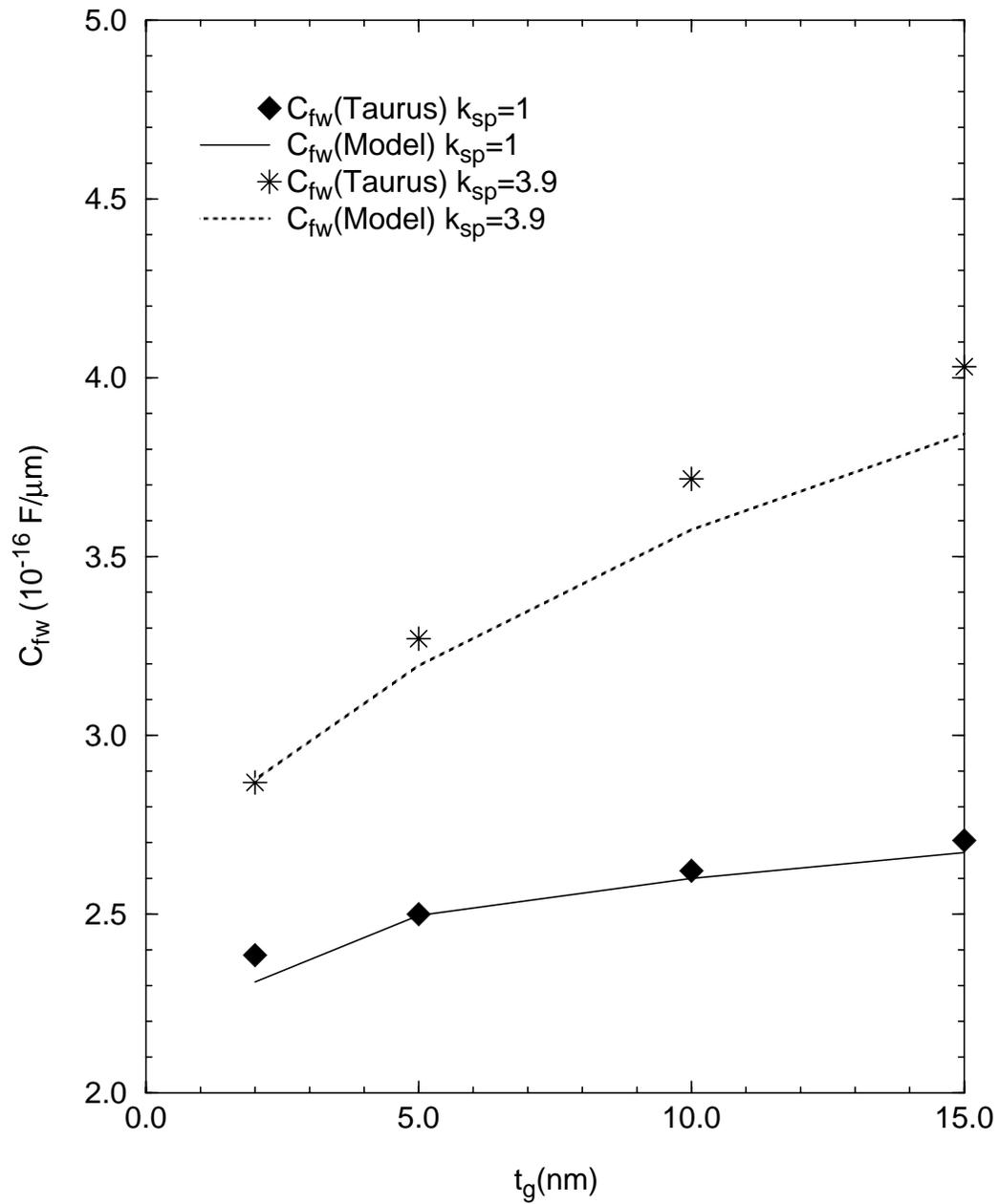


Figure 4-5. Taurus- and model-predicted weak-inversion fringe capacitance,  $C_{fw}$  v/s gate height,  $t_g$  for Device III with  $k_{sp} = 1, 3.9$ .

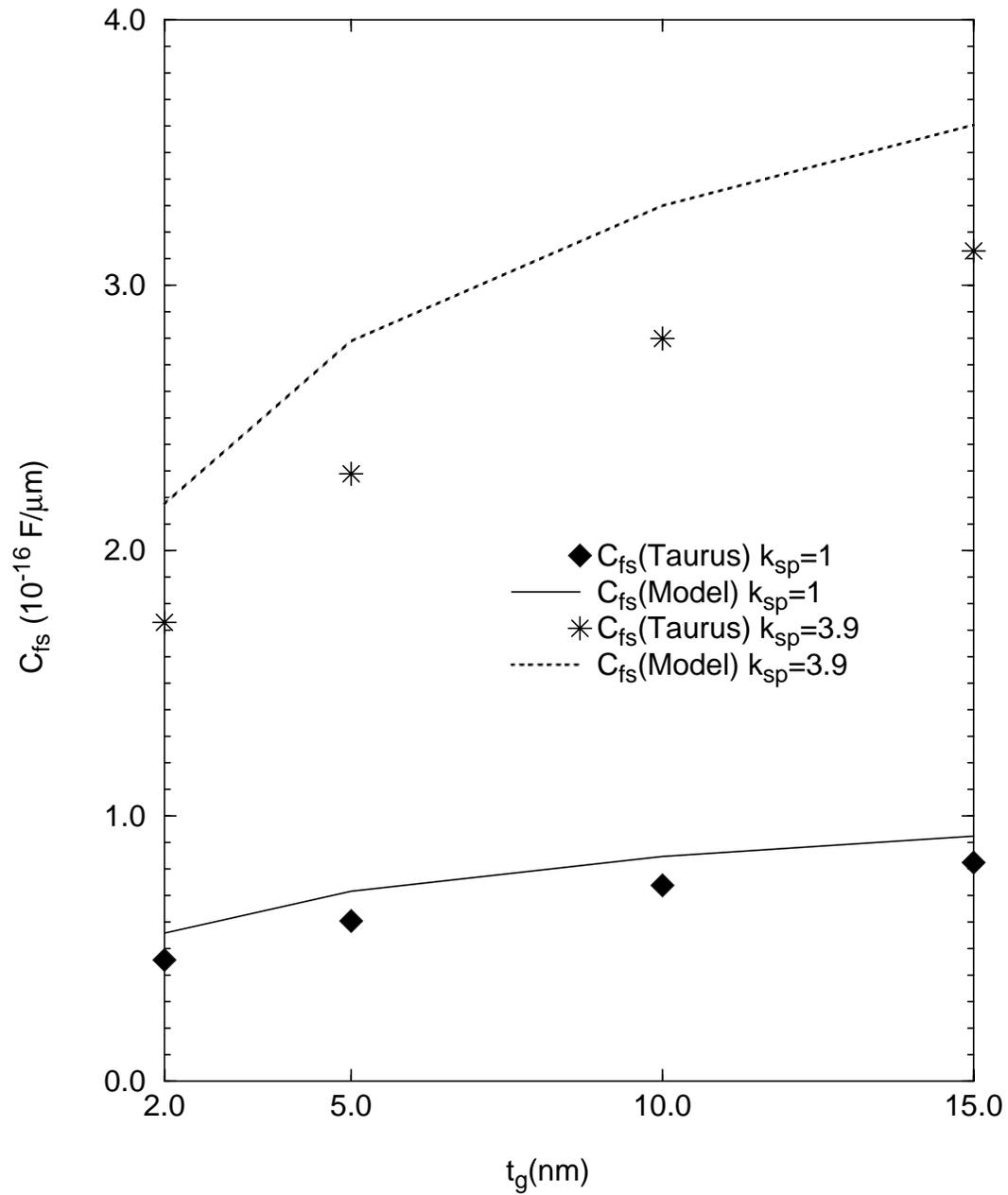


Figure 4-6. Taurus- and model-predicted strong-inversion fringe capacitance,  $C_{fs}$  v/s gate height,  $t_g$  for Device III with  $k_{sp} = 1, 3.9$ .

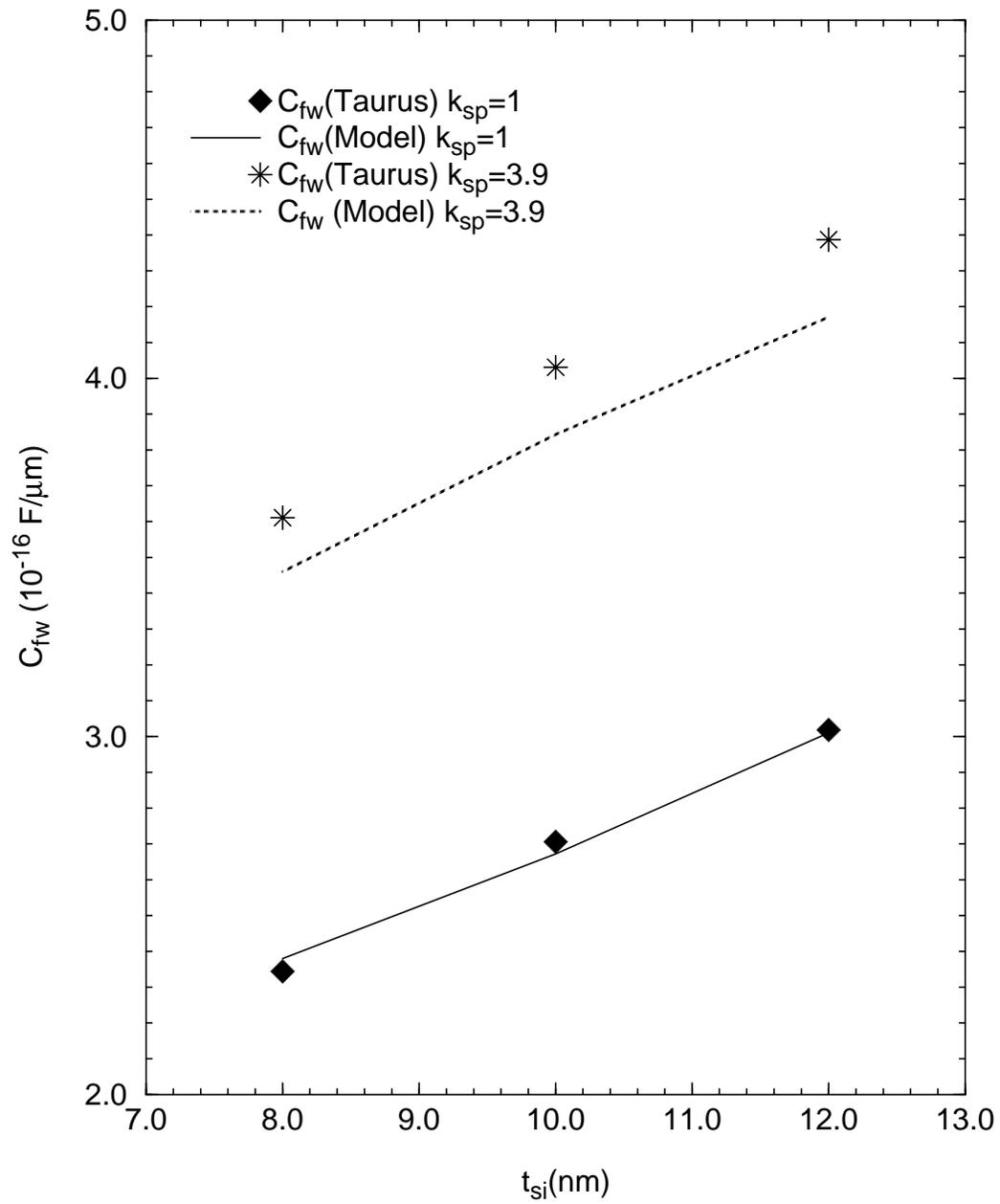


Figure 4-7. Taurus- and model-predicted weak-inversion fringe capacitance,  $C_{fw}$  v/s silicon width,  $t_{si}$  for Device III with  $k_{sp} = 1, 3.9$ .

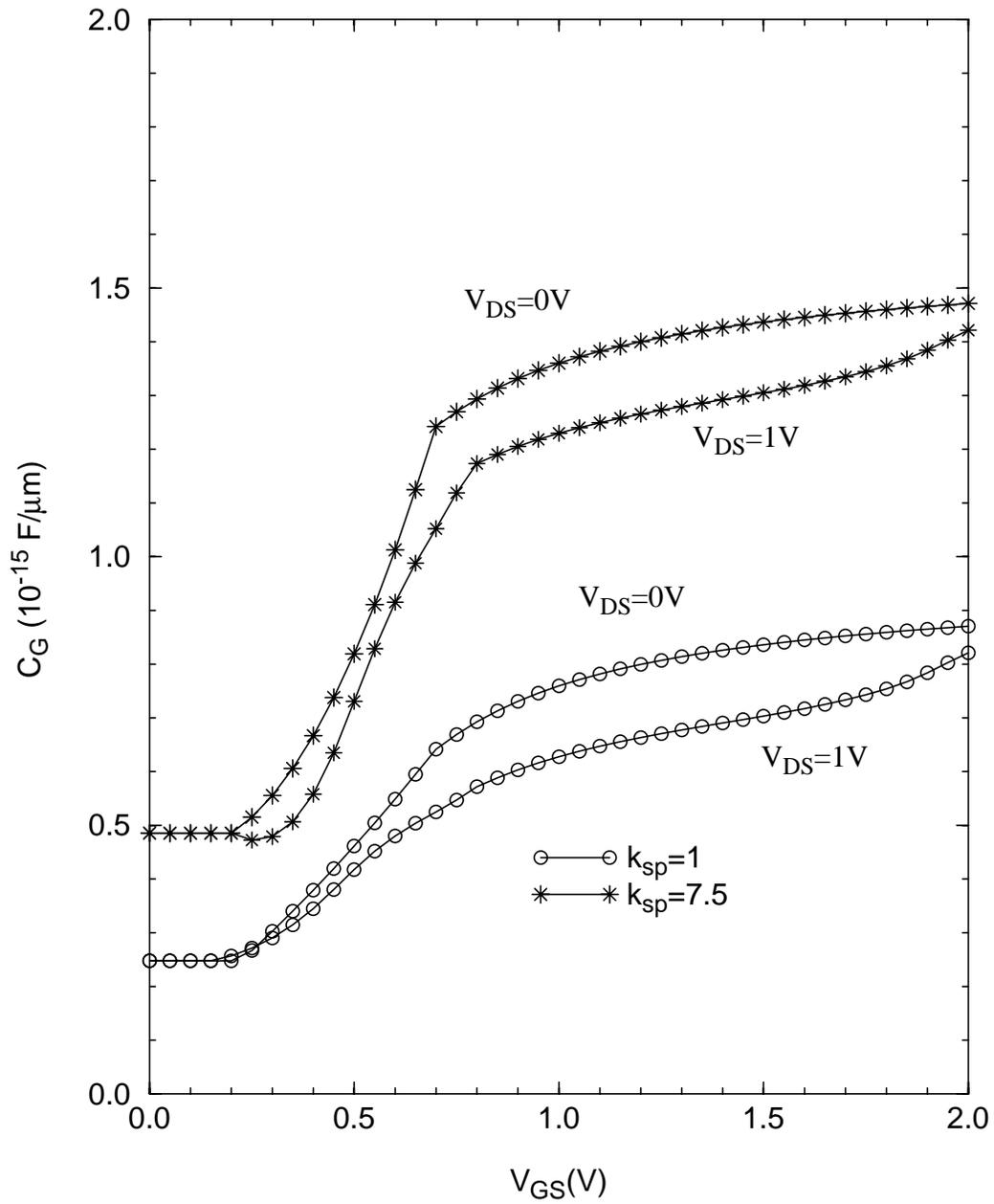


Figure 4-8. UFDG-predicted  $C_G$ - $V_{GS}$  characteristics for n-channel Device III with spacer permittivity,  $k_{sp} = 1, 7.5$  for  $V_{DS} = 0, 1\text{V}$ .

CHAPTER 5  
A REVERSE-ENGINEERING METHODOLOGY TO EXTRACT THE SOURCE/DRAIN  
DOPING PROFILE OF NANOSCALE FINFETS FOR OPTIMAL SOURCE/DRAIN PROCESS  
DESIGN

**5-1 Introduction**

Unlike conventional bulk-CMOS devices, UTB FinFETs are typically undoped; short-channel-effect (SCE) control is achieved by means of two gates on either side of a thin fin. The undoped nature of the body implies high carrier mobility [4] and also avoids random dopant effects [3]. To prevent source/drain (S/D) dopants from intruding into the channel region during the annealing process, the device needs to be designed with reasonably long S/D extensions. As discussed in Chapter 4, in devices with gate-source/drain (G-S/D) underlaps, the effective channel length in weak inversion,  $L_{\text{eff(weak)}}$ , is longer than the physical gate length,  $L_g$ , while the effective channel length in strong inversion,  $L_{\text{eff(strong)}} \cong L_g$ . Increase in  $L_{\text{eff(weak)}}$  is desirable because it improves SCE performance, implying lower  $I_{\text{off}}$ . It also implies reduced parasitic G-S/D fringe capacitances,  $C_{\text{if}}$  and  $C_{\text{of}}$ , as modeled in Chapter 4.  $L_{\text{eff(weak)}}$  is a function of the S/D extension length,  $L_{\text{ext}}$ , and the S/D doping straggle,  $\sigma_L$  [11]. Increasing  $L_{\text{ext}}$  and/or decreasing  $\sigma_L$  increases  $L_{\text{eff(weak)}}$ . However, it also increases the parasitic S/D resistance,  $R_{\text{S/D}}$ , implying lower on-state current ( $I_{\text{on}}$ ) [11]. Further, dopant atoms, or ions, in the channel region of the n-channel device can cause a reduction in the threshold voltage ( $V_t$ ) [12]. It has been shown [12] that it may be plausible to tailor  $V_t$  by allowing controlled densities of S/D dopants into the channel region. Clearly, the doping straggle needs to be carefully controlled to trade-off the SCEs with  $I_{\text{on}}$ .

To better understand the above design tradeoffs, we seek to relate the S/D process to the S/D doping profile in the extensions and channel region. We therefore propose a reverse-engineering methodology to extract the S/D lateral doping profile,  $N_{\text{SD}}(y)$ , from FinFET  $C_G$ - $V_{\text{GS}}$  and  $I_{\text{DS}}$ - $V_{\text{GS}}$  data. The extracted  $N_{\text{SD}}(y)$ , for specific processes, can be used to redesign the S/D

process to effect a better tradeoff between SCEs and  $I_{on}$ . In Sec. 5-2, we briefly describe our methodology, and in subsequent sections, we demonstrate it using FinFET data from Sematech.

## 5-2 Description of Reverse-Engineering Methodology

Our methodology consists of two steps. The first step consists of calibrating our physical/process-based compact model, UFDG [10], to FinFET  $C_G$ - $V_{GS}$  and  $I_{DS}$ - $V_{GS}$  data. UFDG is a quasi-predictive model which models FinFET characteristics using a small number of process-dependent parameters. UFDG has physical, QM-based models for carrier mobility ( $\mu_{eff}$ ) [32], quantum-mechanical  $V_t$  shift [19], inversion-layer capacitance [33], velocity overshoot [34], and ballistic-limit current [4]. We turn on all these physical models for this step (QMX = QMD = 1, VO = 1, BLIM = 1). UFDG calibration helps define the device structure and extract crucial parameters like  $\mu_{eff}$  and  $R_{S/D}$ . The simple algorithm to do so is defined as follows.

### Step 1: Calibration of UFDG to Device Data

1. We use the  $C_G$ - $V_{GS}$  characteristics of a long- $L_g$  device to tune the fin height,  $h_{Si}$ , effective gate oxide thickness, EOT, and gate work function,  $\Phi_M$ .
2. We match the weak-inversion  $I_{DS}$ - $V_{GS}$  characteristics by tuning the G-S/D underlap in weak inversion,  $L_{eSD}$  [11]. The effective channel length in weak inversion is  $L_{eff(weak)} = L_g + 2L_{eSD}$ .
3. We now tune the  $\mu_{eff}$  parameters, UO and  $\Theta$ . UO is the low-field, thick-Si mobility, and  $\Theta$  is a tuning parameter for surface-roughness scattering. To tune UO and  $\Theta$ , we match the  $g_m/I_{DS}^2$ - $V_{GS}$  curves of data with UFDG. This characteristic is virtually independent of the parasitic  $R_{S/D}$ .
4. We tune  $R_{S/D}$  by matching the strong-inversion low- $V_{DS}$  current and the linear  $I_{DS}$ - $V_{DS}$  characteristics.
5. We then match the strong-inversion, high- $V_{DS}$  current by calibrating the drain-induced charge enhancement (DICE) [44] and velocity overshoot (VO) [34] parameters.

## Step 2: Matching UFDG and Taurus to Get $N_{SD}(y)$

UFDG does not characterize devices in terms of  $N_{SD}(y)$ ; rather it uses an effective underlap,  $L_{eSD}$ , as discussed above. (Sec. 4-1 has a more detailed discussion.) Thus, in step 2 of our methodology, we use the Taurus device simulator [9] to simulate the refined device structure, and tune  $N_{SD}(y)$  to match the UFDG-predicted weak-inversion characteristics. We are careful to avoid errors associated with Taurus' inadequate modeling of quantization and transport phenomena for ultra-thin-body (UTB) devices. Thus, we turn off QM and transport models in UFDG (QMX=QMD=0, VO=BLIM=0) and use a constant,  $V_{GS}$ -independent  $\mu_{eff}$  model (tune UO to get certain  $\mu_{eff}$ ,  $\Theta = 0$ ). We also use a constant mobility model in Taurus and turn velocity-saturation off to avoid mobility model-related errors in weak-inversion (Appendix A). We assume that  $N_{SD}(y)$  is gaussian,  $N_{SD}(y) = N_{SD0}\exp(-y^2/\sigma_L^2)$ . The gaussian roll-off is assumed to start at the edge of the spacer, and  $N_{SD0}$  is assumed from the process information. We tune  $\sigma_L$  so as to match the UFDG-predicted weak-inversion  $I_{DS}-V_{GS}$  characteristics. Fig. 5-1 summarizes the methodology used for the  $N_{SD}(y)$  extraction. Starting in the next section, we demonstrate the above methodology using FinFET data from Sematech. The data consists of  $C_G-V_{GS}$  characteristics of a  $L_g = 10\mu\text{m}$  device and  $I_{DS}-V_{GS}$  characteristics for four different gate lengths ( $L_g = 10\mu\text{m}, 1\mu\text{m}, 75\text{nm}, 32\text{nm}$ ).

### 5-3 Calibration of UFDG to C-V Characteristics (Step 1)

In this section, we use the  $C_G-V_{GS}$  characteristics of a  $L_g = 10\mu\text{m}$  device, with number of fins,  $N_f = 200$ , to refine the device structural parameters:  $h_{Si}$ , EOT, and  $\Phi_M$ . The preliminary process information provided to us is given in Table 4-1. Thus, we calculate the inversion charge density,

$$N_{inv}(V_{GS}) = \frac{\int_0^{V_{GS}} C_G dV_{GS}}{qh_{Si}N_fL_g} \quad (5-1)$$

and match it to UFDG-predicted  $N_{inv}$ - $V_{GS}$  data. Figs. 5-2 and 5-3 show the match for nMOS and pMOS respectively. Interestingly, we find that using  $h_{Si} = 40\text{nm}$  (as given), leads to needed tuning of  $\Phi_M = 4.3\text{V}$  for nMOS and  $\Phi_M = 4.8\text{V}$  for pMOS to match the charge level. This is not physical because the gate metal (TiN) used for both devices is the same and is known to be near midgap. Also, it is not possible to match the slope of the  $N_{inv}$ - $V_{GS}$  data by tuning EOT  $\sim 1.2$ - $1.4\text{nm}$  (as given). Thus, we tune  $h_{Si} = 60\text{nm}$  and EOT =  $1.2\text{nm}$  so as to adequately match the slope of the graph (also in Figs. 5-2 and 5-3) as well as obtain a reasonable  $\Phi_M$  ( $4.5\text{V}$  for nMOS,  $4.55\text{V}$  for pMOS). We renormalize all the  $I_{DS}$ - $V_{GS}$  data, which was given to us per unit device width ( $=N_f \times h_{Si}$  as noted) to reflect the calibrated value of  $h_{Si}$ .

#### 5-4 Calibration of UFDG to I-V Characteristics (Step 1)

We now calibrate UFDG to the  $L_g = 1\mu\text{m}$   $I_{DS}$ - $V_{GS}$  data. We tune the weak-inversion characteristics by tuning  $\Phi_M$ . We then tune the UFDG [10] mobility parameters,  $U_0 = 900\text{cm}^2/\text{V}\cdot\text{s}$  and  $\Theta = 0.32$ , by matching the  $g_m/I_{DS}^2$ - $V_{GS}$  characteristics, shown in Fig. 5-4 for n-channel  $L_g = 1\mu\text{m}$ . Fig. 5-5 shows the measured and UFDG-predicted  $I_{DS}$ - $V_{GS}$  characteristics of the n-channel  $L_g=1\mu\text{m}$  device. The high- $V_{DS}$  current does not match because UFDG, which is primarily aimed at nanoscale devices, does not model the effect of  $\mu_{eff}$  increasing towards the drain end of the channel in long- $L_g$  devices due to decreasing transverse electric field.

The short- $L_g$  devices display a degradation in  $\mu_{eff}$  as compared to the long  $L_g = 1\mu\text{m}$  device. This is evident from the channel resistance,  $R_{ch}$ , versus  $L_g$  plot (Fig. 5-6). This effect has been observed in earlier work [13][14][15][16], and we will discuss it in Sec. 5-6, and in greater

detail in Chapter 6. For  $L_g = 75\text{nm}$ , we obtain  $UO = 350\text{cm}^2/\text{V-s}$  and  $\Theta = 0.2$ , by matching the  $g_m/I_{DS}^2$ - $V_{GS}$  graph as illustrated in Fig. 5-7. Having determined the  $\mu_{\text{eff}}$  parameters, we match (very well) the linear-region current by tuning  $R_{S/D} = 100\Omega\text{-}\mu\text{m}$ , as shown in Fig. 5-8. For high  $V_{DS}$ , there is significant discrepancy between the measured and predicted current in the high- $V_{GS}$  region. We attribute this to self-heating of the device, which reduces mobility and current. The  $L_g = 32\text{nm}$  device shows significant, but not excessive, short-channel effects (see Fig. 5-9). We therefore tune short-channel effects (SCEs) using effective underlap,  $L_{\text{eSD}} = 9.25\text{nm}$ , so that the effective channel length in weak inversion is  $L_{\text{eff(weak)}} = L_g + 2L_{\text{eSD}}$ ; in strong inversion,  $L_{\text{eff(strong)}} = L_g$ . The  $g_m/I_{DS}^2$  calibration, shown in Fig. 5-10, gives  $UO = 110\text{cm}^2/\text{V-s}$  and  $\Theta = 0.2$ , and the linear  $I_{DS}$ - $V_{GS}$  calibration, shown in Fig. 5-11, gives  $R_{S/D} = 120\Omega\text{-}\mu\text{m}$ . Again the measured and UFDG-predicted characteristics match well, except for the self-heating effect.

For the p-channel  $L_g = 1\mu\text{m}$  device, the  $\mu_{\text{eff}}$  parameters obtained from the  $g_m/I_{DS}^2$ - $V_{GS}$  calibration do not provide a good  $I_{DS}$ - $V_{GS}$  match. We think this is because  $R_{S/D}$  is high and is a function of  $V_{GS}$  in the p-channel devices, implying that  $g_m/I_{DS}^2$ - $V_{GS}$  is not independent of  $R_{S/D}$ . Thus, we instead used a split-CV measurement on the  $L_g = 10\mu\text{m}$  device to calibrate the  $\mu_{\text{eff}}$  parameters,  $UO = 205\text{cm}^2/\text{V-s}$ ,  $\Theta = 0.2$  (Fig. 5-12). The corresponding  $I_{DS}$ - $V_{GS}$  match is shown in Fig. 5-13. The  $R_{\text{ch}}$  vs.  $L_g$  plot (Fig. 5-14) suggests little or no  $\mu_{\text{eff}}$  degradation for  $L_g = 75\text{nm}$ , but huge degradation for  $L_g = 32\text{nm}$ . Thus, we use  $UO = 205\text{cm}^2/\text{V-s}$  and  $\Theta = 0.2$  for calibration to the  $L_g = 75\text{nm}$  device, and tune  $R_{S/D} = 250\Omega\text{-}\mu\text{m}$  to match the linear-region current, as shown in Fig. 5-15. We see discrepancy between the measured and predicted linear-region characteristics, which confirms our insight that  $R_{S/D}$  is varying with  $V_{GS}$ . For increasing  $V_{GS}$ , the gate induces free carriers in the extension via the outer fringing field through the spacer, thereby reducing the extension resistance and therefore,  $R_{S/D}$ . This effect can be significant in devices with long G-S/D underlap since the lightly doped extension implies that gate-induced free carriers

would contribute significantly to the extension resistance. We are unable to tune the  $\mu_{\text{eff}}$  parameters for  $L_g = 32\text{nm}$  since, as discussed above,  $g_m/I_{\text{DS}}^2 - V_{\text{GS}}$  is not independent of  $R_{\text{S/D}}$ . However, if we let  $R_{\text{S/D}} = 250\Omega\text{-}\mu\text{m}$  (ascertained from  $L_g = 75\text{nm}$ ), we can tune  $UO = 90\text{cm}^2/\text{V-s}$  to approximately match the strong-inversion characteristics, as shown in Fig. 5-16. We note a significant self-heating related discrepancy between the  $L_g = 32\text{nm}$  data and the model at high  $V_{\text{GS}}$ ,  $V_{\text{DS}} = 1\text{V}$ . The discrepancy is less pronounced than in n-channel devices due to the lower  $I_{\text{DS}}$  in p-channel devices. Also, the weak-inversion calibration, illustrated in Fig. 5-17, yields an underlap of  $L_{\text{eSD}} = 8.0\text{nm}$ .

### 5-5 Determination of SDE Doping Profile (Step 2)

We now seek to estimate  $N_{\text{SD}}(y)$  of the n-channel FinFETs. To estimate  $N_{\text{SD}}(y)$ , we match the short-channel effects of the  $L_g = 32\text{nm}$  device with a Taurus-simulated device of the refined structure (Table 5-2). We use the UFDG model card obtained from calibration of the  $L_g = 32\text{nm}$  device, turn the QM and transport models off and use a constant,  $V_{\text{GS}}$ -independent  $\mu_{\text{eff}} = 107\text{cm}^2/\text{V-s}$  (corresponding to  $UO = 110\text{cm}^2/\text{V-s}$ ). We use a constant mobility model with velocity saturation and QM models off in Taurus as well (Appendix A). We assume a gaussian doping profile and estimate that  $N_{\text{SD0}} \sim 1 \times 10^{20}\text{cm}^{-3}$ . We then tune  $\sigma_L = 20\text{nm}$  in Taurus to match (very well) the UFDG-predicted SCEs, as indicated in Fig. 5-18. Since we recognize that our estimate of  $N_{\text{SD0}}$  is subject to error, we repeat this procedure for several values of  $N_{\text{SD0}}$  near the nominal value. Table 5-3 indicates the different combinations of  $N_{\text{SD0}}$  and  $\sigma_L$  which would give the predicted SCEs. We find that for reasonable perturbations about  $N_{\text{SD0}} \sim 1 \times 10^{20}\text{cm}^{-3}$ ,  $\sigma_L$  varies by  $\sim \pm 10\%$ . We repeat the above procedure for the p-channel devices and obtain  $\sigma_L = 22\text{nm}$  corresponding to  $N_{\text{SD0}} = 1 \times 10^{20}\text{cm}^{-3}$ .

## 5-6 Carrier Mobility Degradation in Short-Channel FinFETs

Fig. 5-19 shows the UFDG-predicted  $\mu_{\text{eff}}N_{\text{inv}}$  for devices of all three gate lengths. We note that there is some degradation in  $\mu_{\text{eff}}$  for  $L_g = 75\text{nm}$  and huge degradation for  $L_g = 32\text{nm}$ . One possible explanation for this effect is that it may be caused by Coulomb scattering; either dopants in the channel, or long-range Coulomb scattering from the S/D. However, Fig. 5-19 shows that the degradation is almost independent of the inversion charge density,  $N_{\text{inv}}$ . This is inconsistent with the Coulomb scattering explanation because Coulomb scattering tends to get screened out at high  $N_{\text{inv}}$ . Also, the doping profile extracted in Sec. 5-5 implies a very low dopant density,  $\sim 5 \times 10^{17} \text{ cm}^{-3}$  at the edge of the channel. This dopant density would not explain the strong  $\mu_{\text{eff}}$  degradation via Coulomb scattering. Recent works [13][14] attribute this short- $L_g$   $\mu_{\text{eff}}$  degradation to “neutral defects” associated with the S/D, but there is no insight on what may be the nature of these defects; [13] also suggests that  $\mu_{\text{eff}}$  improves with an increase in annealing temperature. We will discuss this effect in more detail in Chapter 6, and offer suggestions to diagnose and possibly mitigate the  $\mu_{\text{eff}}$  degradation. For now, we note that  $\mu_{\text{eff}}$  is probably dependent on the S/D process, and a more comprehensive study should examine the effects of different S/D processes on  $\mu_{\text{eff}}$ .

## 5-7 Device Redesign for Better $I_{\text{on}}/I_{\text{off}}$ Tradeoff

Using our UFDG-data calibrations and observations, we attain good insights into FinFET design. We observed in Section 5-4 that  $L_{\text{eSD}} = 9.25\text{nm}/8.0\text{nm}$  for the n/p-channel devices. This implies excellent SCEs as demonstrated in Figs. 5-9 and 5-17. However, it also leads to large  $R_{\text{S/D}} = 120/250\Omega\text{-}\mu\text{m}$  for the n/p-channel devices. For high-performance (HP) design, the device can be redesigned to effect a better tradeoff between  $I_{\text{on}}$  and SCEs [11]. This can be achieved by using a shorter spacer ( $L_{\text{ext}} \cong 50\text{nm}$  for the actual FinFETs) and/or increasing the annealing time/temperature to allow more dopants in the S/D extension (SDE), and perhaps a controlled number

of dopants in the channel. A higher doping density in the SDE would reduce  $L_{\text{eff(weak)}}$ , thereby degrading SCEs in weak inversion. However, in strong inversion,  $L_{\text{eff(strong)}} \cong L_g$  and a higher doping density in the SDE would imply higher carrier density, thereby reducing  $R_{S/D}$ . Also, S/D dopants in the channel would imply a reduction in the strong-inversion  $V_t$ , causing more  $I_{\text{on}}$  enhancement [12]. To prevent reduction in the weak-inversion  $V_t$  and avoid random dopant fluctuations, we must ensure that the doping density in the center of the channel  $\leq 5 \times 10^{17} \text{ cm}^{-3}$  [12].

Based on our estimated  $N_{\text{SD}}(y)$ ,  $N_{\text{SD0}} = 1 \times 10^{20} \text{ cm}^{-3}$  and  $\sigma_L = 20 \text{ nm}$  for the nFinFET, we use Taurus to estimate the spacer width,  $L_{\text{ext}}$ , that would give reasonable SCEs (DIBL  $< 100 \text{ mV/V}$  and  $S < 80 \text{ mV}$ ) and  $I_{\text{off}} < \cong 100 \text{ nA}/\mu\text{m}$ . We project that reducing the spacer width to  $L_{\text{ext}} = 40 \text{ nm}$  would yield reasonable SCEs (DIBL =  $85 \text{ mV/V}$ ,  $S = 73 \text{ mV}$ ) and give  $I_{\text{off}} \cong 100 \text{ nA}/\mu\text{m}$ . The reason for the relatively high  $I_{\text{off}}$  is that the gate work function is  $\sim 100 \text{ mV}$  below midgap ( $\Phi_M = 4.5 \text{ V}$ ). The resulting higher doping density in the extension would significantly reduce  $R_{S/D}$ . Fig. 5-20 demonstrates Taurus-predicted weak-inversion  $I_{\text{DS}}-V_{\text{GS}}$  characteristics for the redesigned device ( $L_{\text{ext}} = 40 \text{ nm}$ ). We calibrate the weak-inversion characteristics of the redesigned device to UFDG to obtain  $L_{\text{eSD}} = 4.5 \text{ nm}$  (instead of  $L_{\text{eSD}} = 9.25 \text{ nm}$  for the actual device), which implies  $L_{\text{eff(weak)}} = L_g + 2L_{\text{eSD}} = 41 \text{ nm}$ . The value of  $L_{\text{eff(weak)}}$  is consistent with the thumb rule that for reasonable SCEs,  $t_{\text{Si}}/L_{\text{eff(weak)}} \cong 0.5$ . We note from Fig. 5-20 that  $I_{\text{off}}$  increases by almost two orders of magnitude (from  $\cong 2 \text{ nA}/\mu\text{m}$  to  $\cong 100 \text{ nA}/\mu\text{m}$ ) in the redesigned device as compared to the nominal device. Fig. 5-21 portrays our estimated  $N_{\text{SD}}(y)$  for the actual device and also the projected  $N_{\text{SD}}(y)$  for the redesigned device ( $L_{\text{ext}} = 40 \text{ nm}$ ). The projected doping density in the center of the channel for the redesigned device is  $\cong 8 \times 10^{16} \text{ cm}^{-3}$ ; this should avoid excessive random dopant fluctuations [12].

Since the redesigned device has more dopants in the extension region, the extension resistance ( $R_{\text{ext}}$ ) will be reduced resulting in lower  $R_{\text{S/D}}$ . It is difficult to estimate the reduction in  $R_{\text{S/D}}$  in the redesigned device since that would depend on the relative contributions of  $R_{\text{ext}}$  and the other  $R_{\text{S/D}}$  components, primarily the contact resistance,  $R_{\text{con}}$ . To quantify the  $I_{\text{on}}/I_{\text{off}}$  tradeoff, devices with different  $L_{\text{ext}}$  should be fabricated as part of a more comprehensive study. We outline such a study in Sec. 7-2 that would provide a clearer picture into the S/D design tradeoffs for optimal FinFET design. To get a sense of the typical  $I_{\text{on}}/I_{\text{off}}$  tradeoff at  $L_g = 32\text{nm}$ , we consider the  $L_g = 30\text{nm}$  bulk device in [45]. Fig. 3 in [45] shows that an increase in  $I_{\text{off}}$  from  $\sim 1\text{nA}/\mu\text{m}$  to  $100\text{nA}/\mu\text{m}$  leads to  $\sim 30\%$  increase in  $I_{\text{on}}$  (from  $\sim 1.2\text{mA}/\mu\text{m}$  to  $1.55\text{mA}/\mu\text{m}$ ). We recognize that the comparison between a bulk device and a FinFET is not exact, but the above illustration does imply a crude estimate of the improvement in  $I_{\text{on}}$  corresponding to the increased  $I_{\text{off}}$ .

S/D dopants in the channel can cause a reduction in the strong-inversion threshold voltage ( $V_{\text{ts}}$ ) and therefore enhanced  $I_{\text{on}}$ . The reduction in  $V_{\text{ts}}$  ( $\Delta V_{\text{ts}}$ ) can be estimated [12] using the average doping density in the channel

$$\Delta V_{\text{ts}} \cong - \frac{q\bar{N}_{SD}t_{\text{Si}}}{2C_{\text{ox}}} \quad (5-2)$$

where

$$\bar{N}_{SD} \cong \frac{1}{L_g} \int_0^{L_g} N_{SD}(y)dy. \quad (5-3)$$

For the redesigned device ( $L_{\text{ext}} = 40\text{nm}$ ), we use Eq. 5-2 and Eq. 5-3 to estimate a negligible  $\Delta V_{\text{ts}} \cong -30\text{mV}$ . The negligible  $\Delta V_{\text{ts}}$  can be attributed to the long  $L_{\text{eSD}} = 4.5\text{nm}$  needed to control SCEs; for HP devices, a shorter  $L_{\text{eSD}} \cong 1-2\text{nm}$  may be desirable. For a well-tempered device with shorter  $L_{\text{eSD}}$ ,  $t_{\text{Si}}$  should be scaled down for better SCE control. For example, if the device is designed with  $L_{\text{eSD}} = 2\text{nm}$ ,  $L_{\text{eff(weak)}} = L_g + 2L_{\text{eSD}} = 36\text{nm}$  which would imply  $t_{\text{Si}} \sim 0.5L_{\text{eff(weak)}} = 18\text{nm}$ . A

shorter  $L_{eSD}$  would allow shorter  $L_{ext}$  which will imply higher  $\Delta V_{ts}$  and greater reduction in  $R_{S/D}$  leading to more enhancement in  $I_{on}$ .

The  $\mu_{eff}$ -degradation effect has important consequences for the design of short- $L_g$  FinFETs. The severe degradation in  $\mu_{eff}$  undermines an important benefit of FinFET scaling and raises questions about the viability of nanoscale FinFET technology, especially for HP logic applications. Reducing  $L_{ext}$  in order to improve  $I_{on}$  as suggested above would increase the proximity of the channel to the S/D, potentially causing even worse  $\mu_{eff}$  degradation. This might then imply a tradeoff between  $\mu_{eff}$  and  $R_{S/D}$ . We analyze the  $\mu_{eff}$  degradation effect in more detail in Chapter 6 and offer suggestions in Sec. 6-5 to diagnose and possibly mitigate it.

### **5-8 Summary**

In this Chapter, we developed a reverse-engineering methodology to extract the source/drain doping profile from the I-V and C-V characteristics of nanoscale FinFETs. We demonstrated the methodology using FinFET data from Sematech. Based on the extracted doping profile, we provided insights on device redesign for better  $I_{on}/I_{off}$  tradeoff. The reverse-engineering methodology discussed in this chapter can be used in a more comprehensive study of optimal S/D process design for nanoscale FinFETs. We also noted the unexplained degradation of carrier mobility in short- $L_g$  FinFETs (Fig. 5-19) which is possibly related to S/D processing. We discuss the significance and possible causes of this effect in Chapter 6 and propose solutions to possibly mitigate it.

Table 5-1. Approximate device and process information available for Sematech n- and p-channel FinFETs

Device Physical Parameter Name	Device Physical Parameter Value
Physical gate length, $L_g$	$\sim 1\mu\text{m}$ , $\sim 75\text{nm}$ , $\sim 32\text{nm}$
Fin width, $w_{\text{Si}}$	$\sim 20\text{nm}$
Effective oxide thickness, EOT	$\sim 1.2\text{-}1.4\text{nm}$
High-k dielectric, K	$\sim 15$
Fin height, $h_{\text{Si}}$	$\sim 40\text{nm}$
Gate work function, $\Phi_M$	near midgap
Channel doping	Undoped
Surface orientation	{110}
Spacer width, $L_{\text{ext}}$	$\sim 50\text{nm}$
Spacer material	$\text{Si}_3\text{N}_4$
Peak S/D doping density, $N_{\text{SD0}}$	$\sim 1 \times 10^{20} \text{cm}^{-3}$
Source/drain implant dose (As/B)	$3 \times 10^{15} \text{cm}^{-2}$
Source/drain implant energy	35 keV/4 keV for As/B at $0^\circ$ tilt
Source/drain anneal	1070°C spike anneal (4 sec for nMOS, 1 sec for pMOS)

Table 5-2. Device structural and physical parameters obtained after calibration of  $C_G$ - $V_{GS}$  and  $I_{DS}$ - $V_{GS}$  data to UFDG

Device Physical Parameter Name	Device Physical Parameter Value		
Physical gate length, $L_g$	1 $\mu$ m	75nm	32nm
UFDG mobility parameters (n-channel)	UO = 900	UO = 350	UO = 110
	$\Theta = 0.32$	$\Theta = 0.2$	$\Theta = 0.2$
UFDG mobility parameters (p-channel)	UO = 205	UO = 205	UO ~ 90
	$\Theta = 0.2$	$\Theta = 0.2$	$\Theta \sim 0.2$
S/D resistance ( $R_{S/D}$ ) (n-channel) ( $\Omega$ - $\mu$ m)		100	120
S/D resistance ( $R_{S/D}$ ) (p-channel) ( $\Omega$ - $\mu$ m)		250	~250
Fin width, $w_{Si}$ (nm)	20	20	20
Effective oxide thickness, EOT (nm)	1.2	1.2	1.2
Fin height, $h_{Si}$ (nm)	60	60	60
Gate work function, $\Phi_M$	4.5	4.5	4.5
Channel doping	Undoped	Undoped	Undoped
Spacer width, $L_{ext}$ (nm)	~50	~50	~50
Spacer material	$Si_3N_4$	$Si_3N_4$	$Si_3N_4$
Peak S/D doping density, $N_{SD0}$ ( $cm^{-3}$ )	$\sim 1 \times 10^{20}$	$\sim 1 \times 10^{20}$	$\sim 1 \times 10^{20}$

Table 5-3. This table portrays the estimated doping profiles,  $N_{SD}(y)$ , which would match the SCEs of the n-channel  $L_g=32$ nm device.  $N_{SD0}$  is the peak doping density and  $\sigma_L$  is the doping straggle.

Peak Doping density, $N_{SD0}$	Doping straggle $\sigma_L$ (nm)
$8 \times 10^{19}$	21
$1 \times 10^{20}$	20
$2 \times 10^{20}$	18

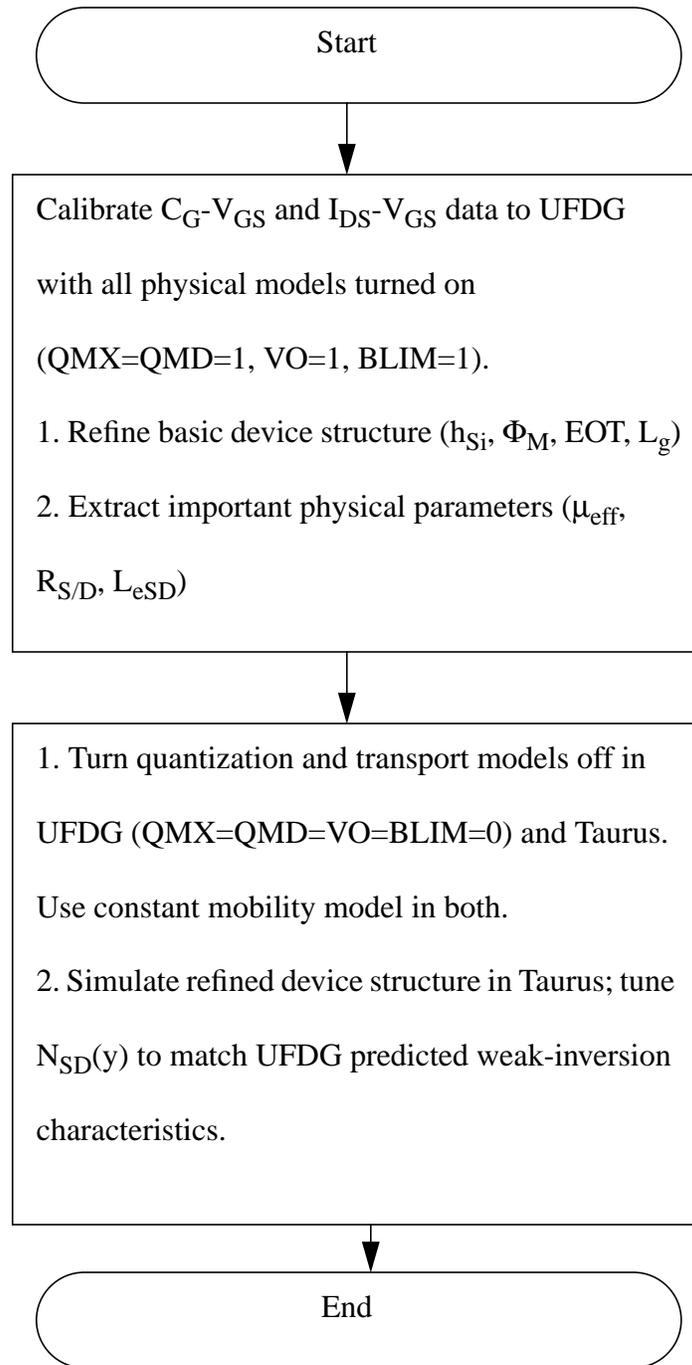


Figure 5-1. The above flowchart demonstrates the reverse-engineering methodology to extract the source/drain doping profile,  $N_{SD}(y)$ , from FinFET  $C_G$ - $V_{GS}$  and  $I_{DS}$ - $V_{GS}$  characteristics.

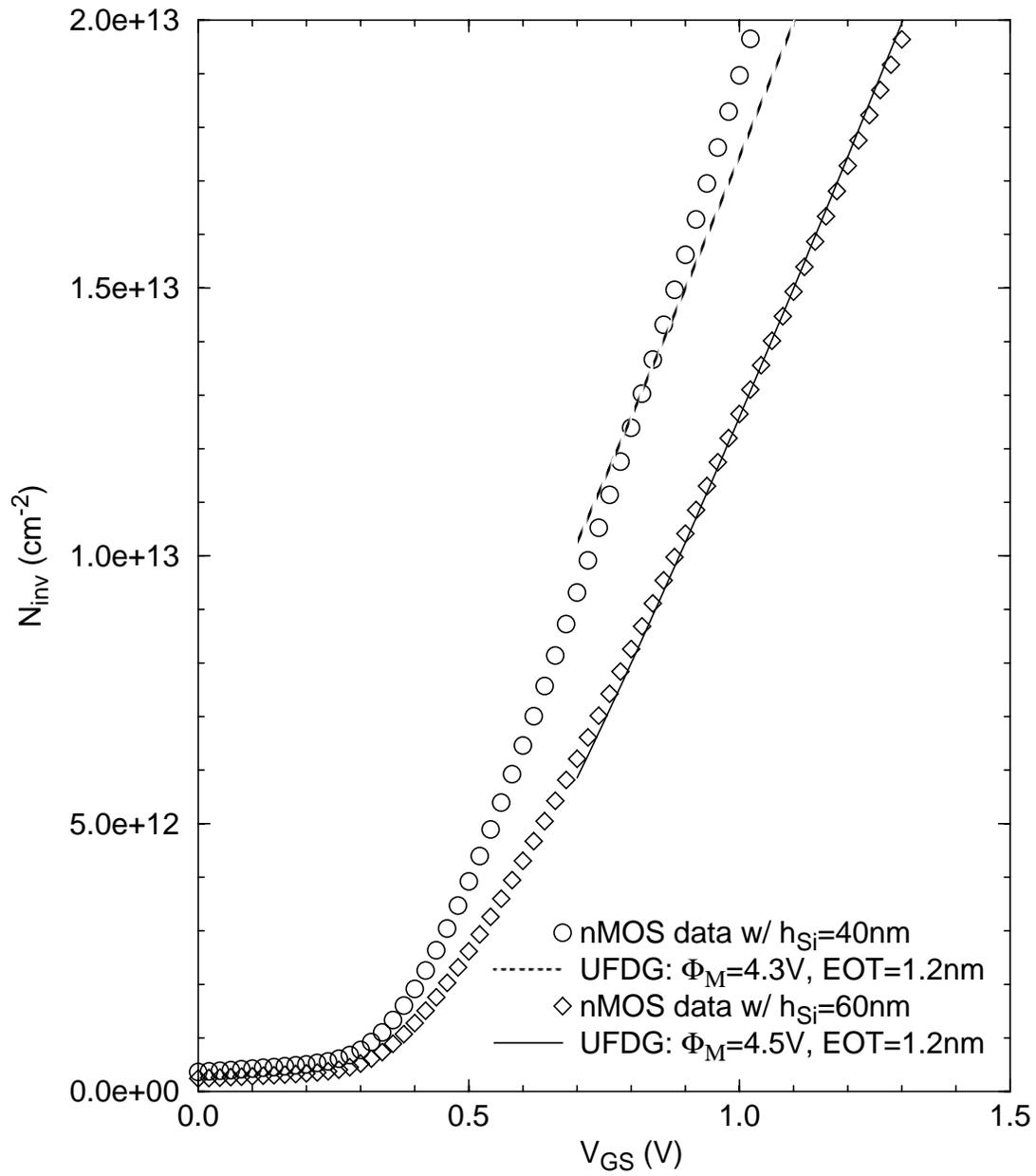


Figure 5-2. Calibration of UFDG to n-channel ( $L_g = 10\mu\text{m}$ )  $N_{inv}$ - $V_{GS}$  data for two different values of  $h_{Si}$ . Here,  $N_{inv}$  was calculated using  $C_G$ - $V_{GS}$  data and Eq. 5-1.

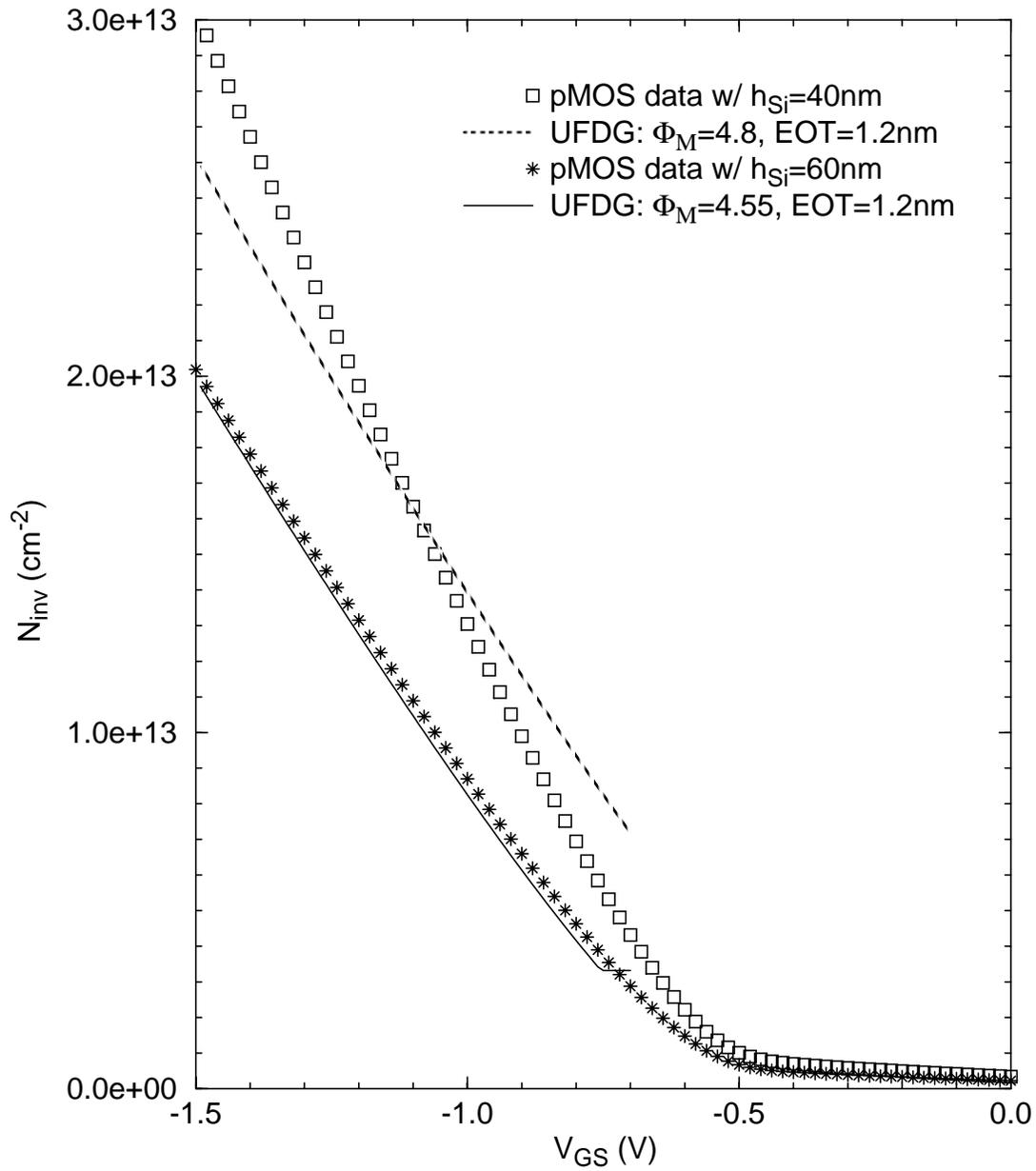


Figure 5-3. Calibration of UFDG to p-channel ( $L_g = 10\mu\text{m}$ )  $N_{inv}$ - $V_{GS}$  data for two different values of  $h_{Si}$ . Here,  $N_{inv}$  was calculated using  $C_G$ - $V_{GS}$  data and Eq. 5-1.

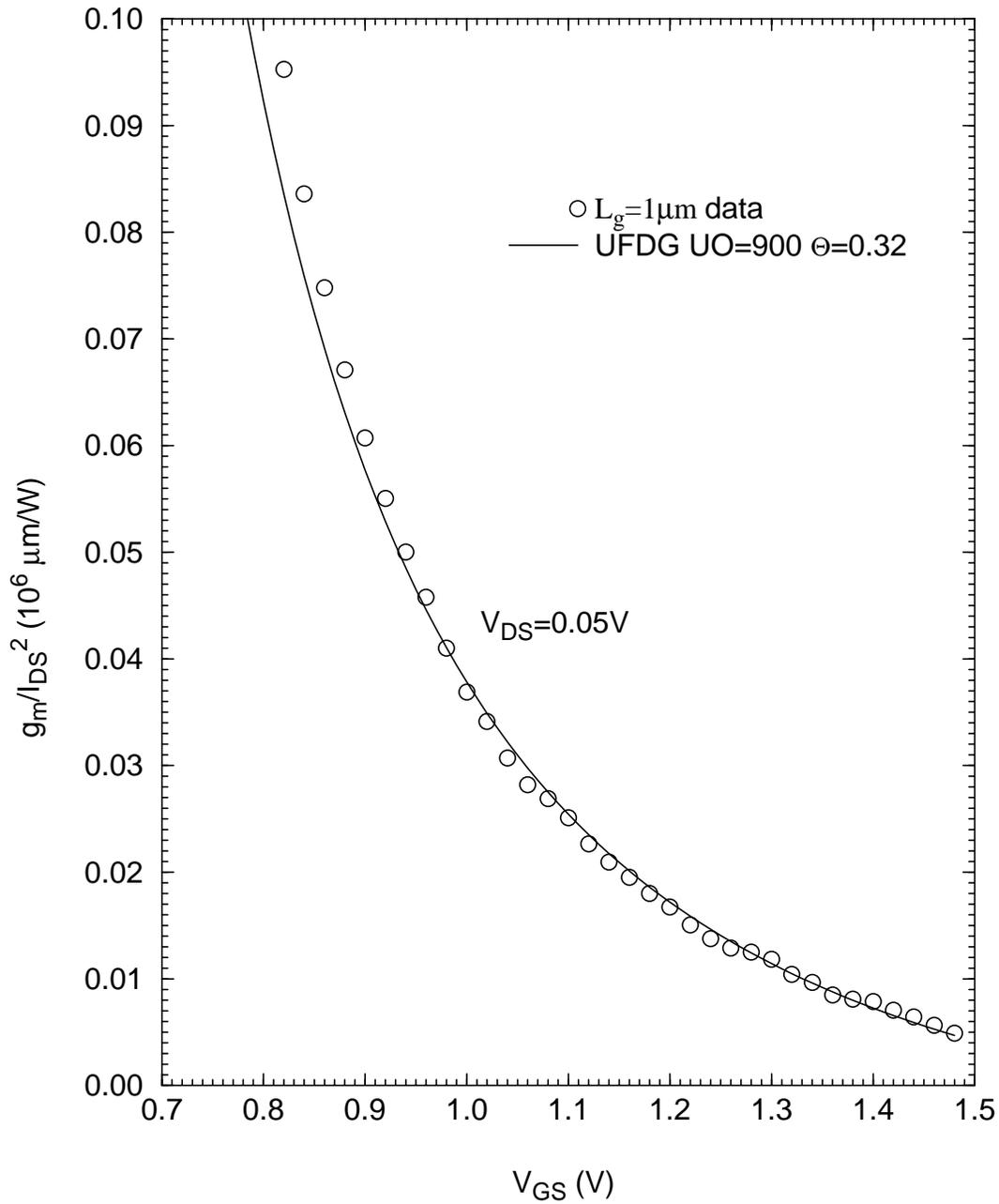


Figure 5-4. Measured and UFDG-predicted low  $V_{DS}$ ,  $g_m/I_{DS}^2$  vs.  $V_{GS}$  used to tune the UFDG mobility parameters:  $UO = 900\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.32$  for the  $L_g = 1\mu\text{m}$  nFinFET.

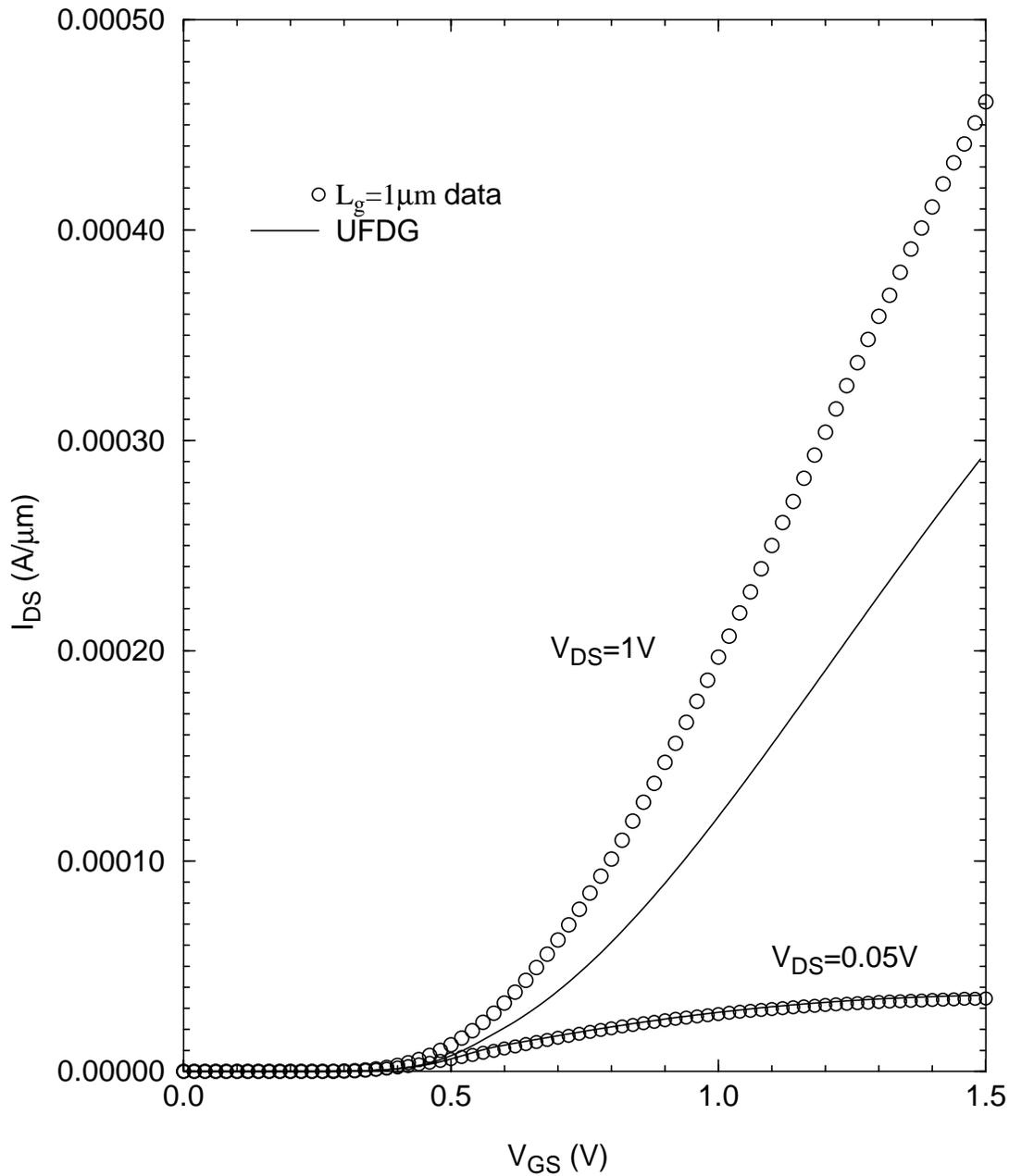


Figure 5-5.  $L_g = 1\mu\text{m}$  nFinFET  $I_{DS}$ - $V_{GS}$  data, with corresponding UFDG calibration results. The  $\mu_{\text{eff}}$  parameters extracted, via  $g_m/I_{DS}^2$ - $V_{GS}$  in Fig. 5-4 ( $UO = 900\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.32$ ) were used in the calibration.

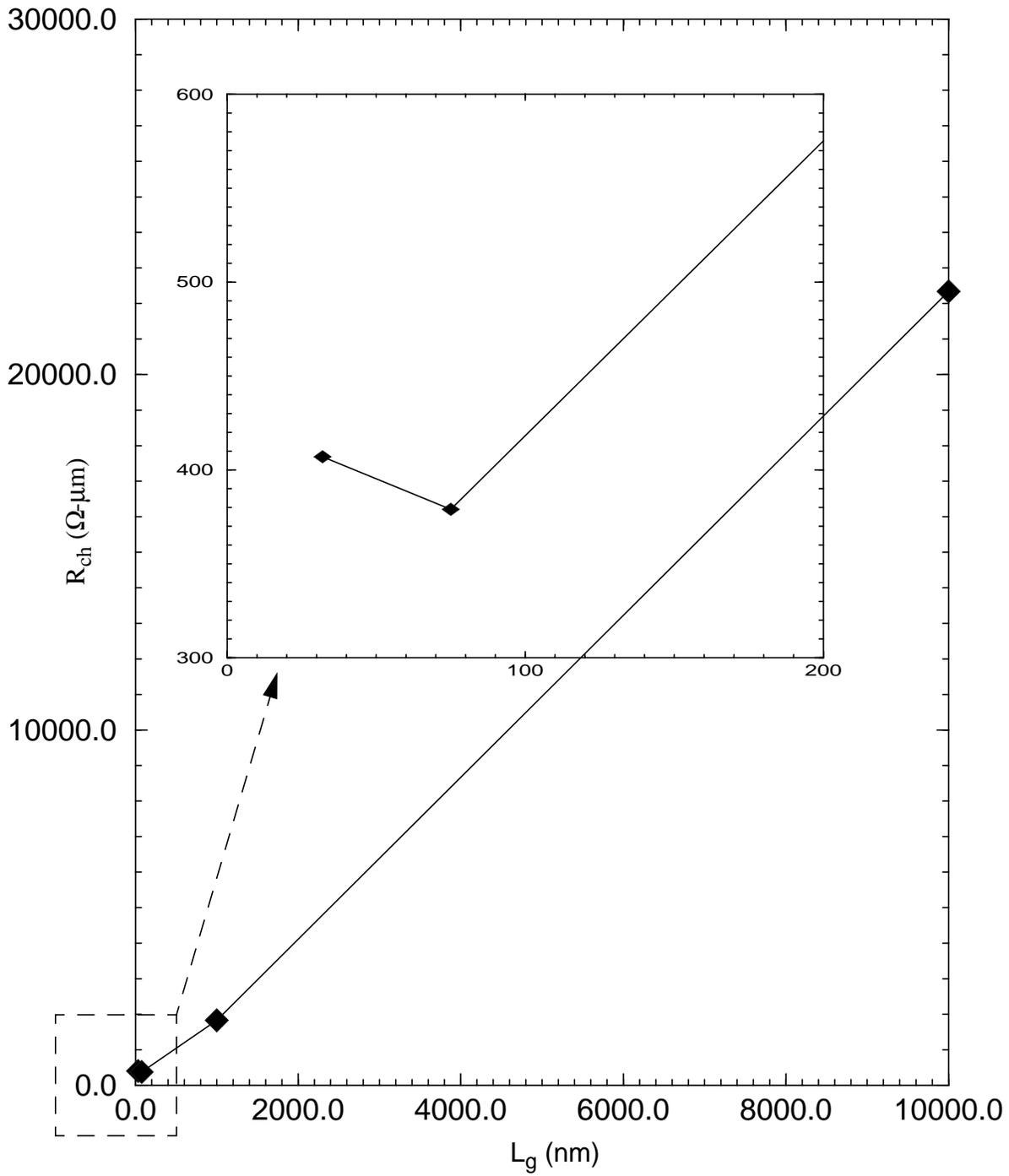


Figure 5-6. Channel resistance of n-devices,  $R_{ch}$  ( $V_{GS} = 1\text{V}$ ,  $V_{DS} = 0.05\text{V}$ ) plotted v/s  $L_g$ . The inset shows the increase in  $R_{ch}$  for  $L_g = 32\text{nm}$  compared to  $L_g = 75\text{nm}$ .

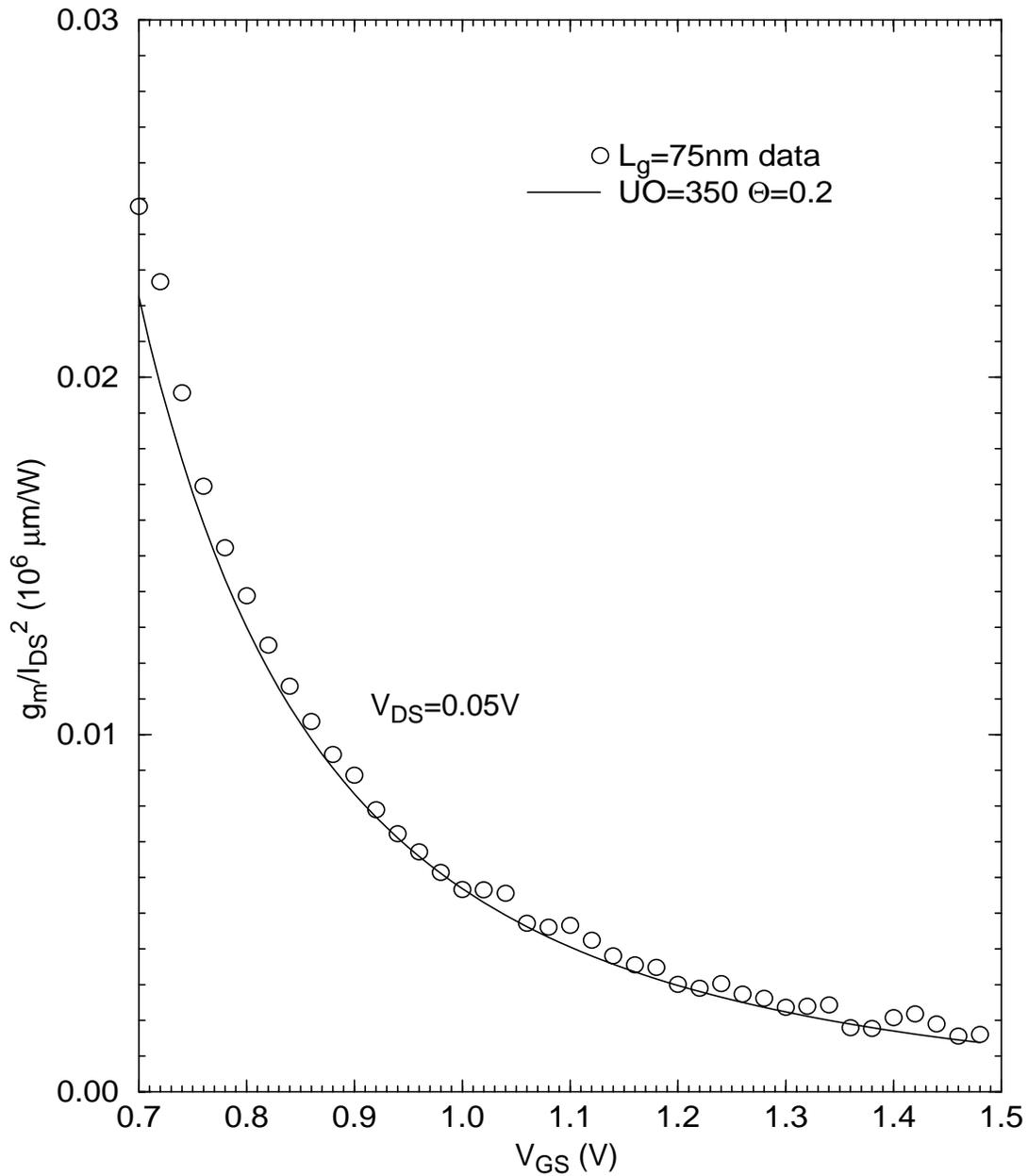


Figure 5-7. Measured and UFDG-predicted low- $V_{DS}$   $g_m/I_{DS}^2$  vs.  $V_{GS}$  used to tune the UFDG mobility parameters:  $UO = 350\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.2$  for the  $L_g = 75\text{nm}$  nFinFET.

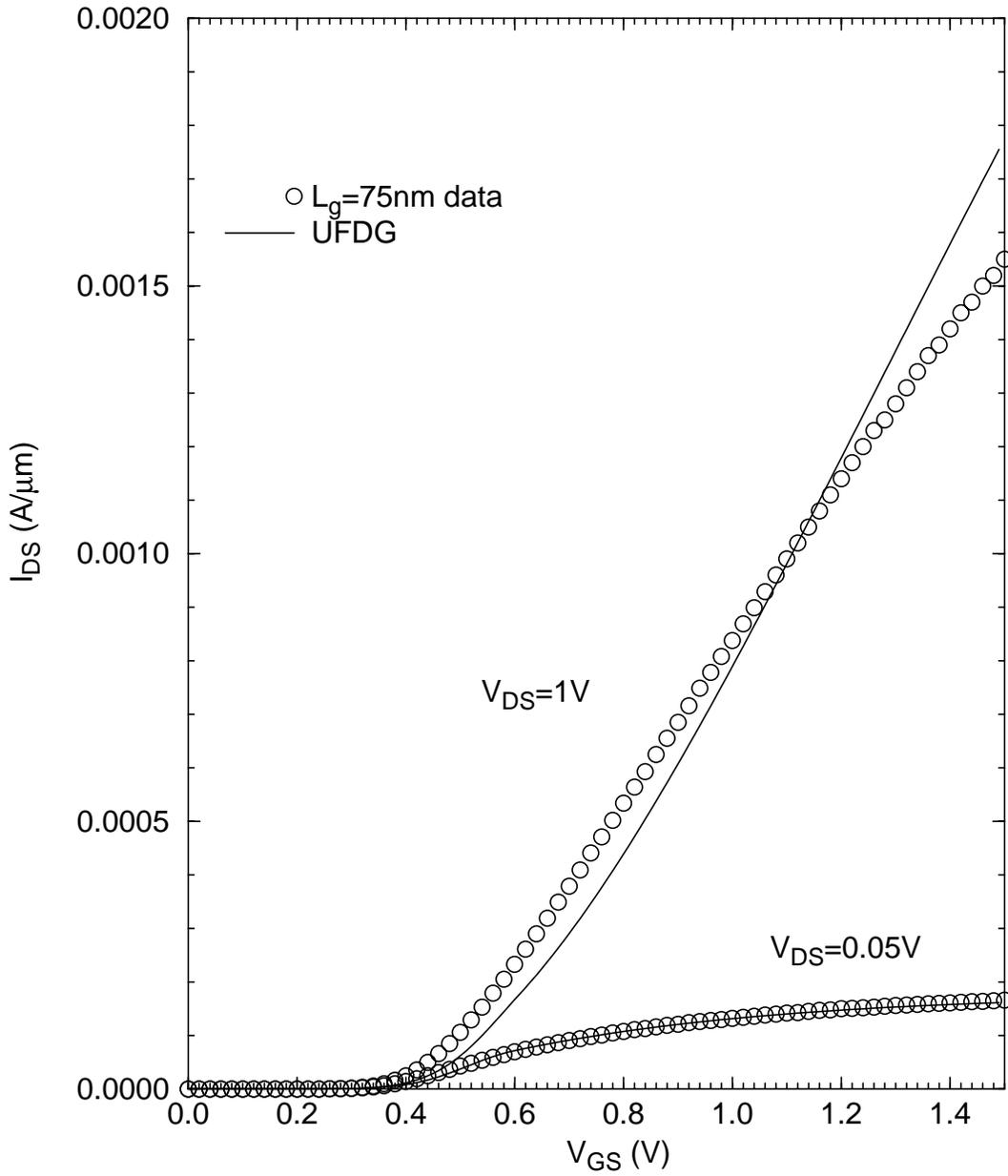


Figure 5-8.  $L_g = 75\text{nm}$  nFinFET  $I_{DS}$ - $V_{GS}$  data, with corresponding UFDG calibration results. The  $\mu_{\text{eff}}$  parameters extracted, via  $g_m/I_{DS}^2$ - $V_{GS}$  in Fig. 5-7 ( $UO = 350\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.2$ ) and  $R_{S/D} = 100\Omega\text{-}\mu\text{m}$  were used in the calibration.

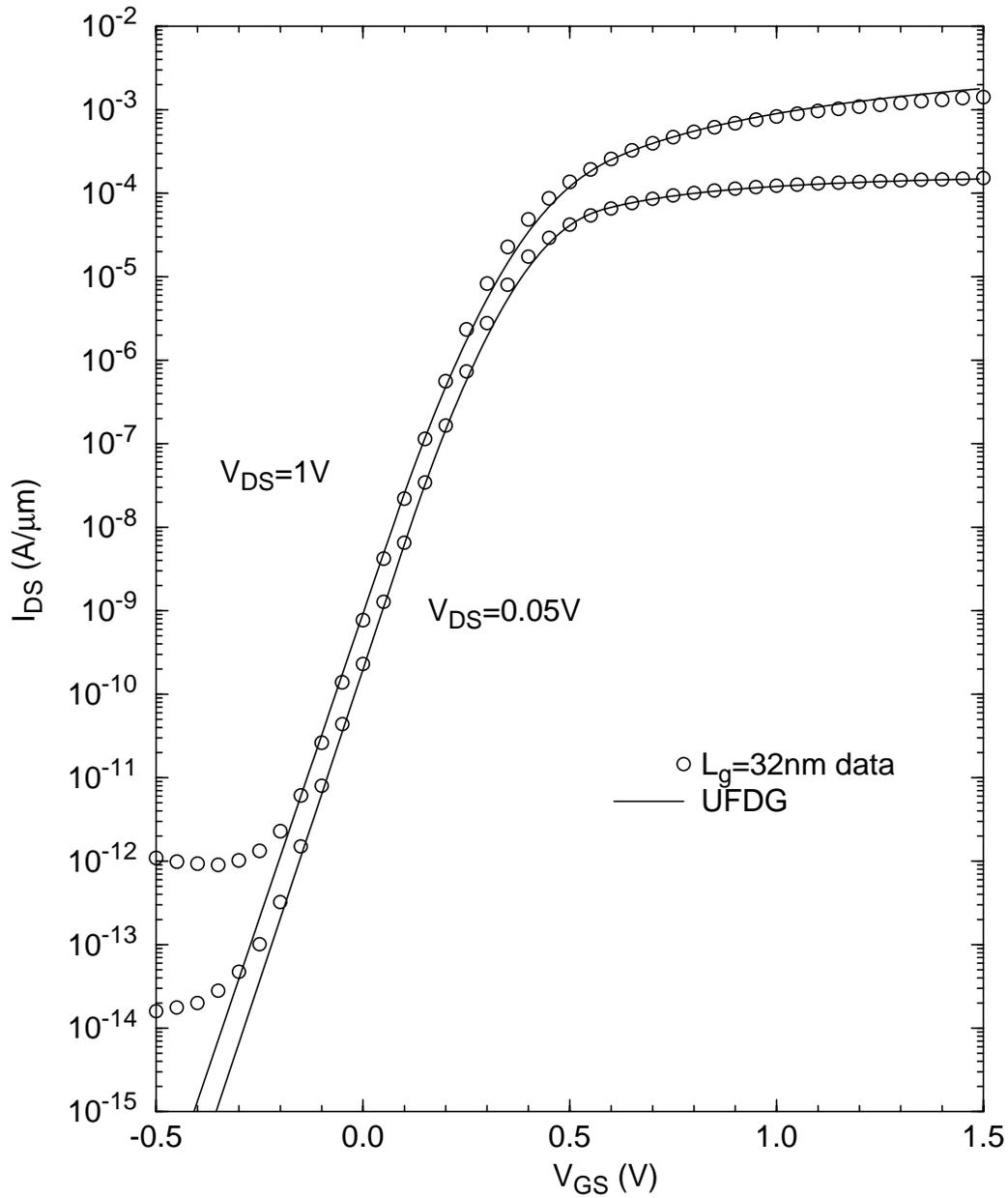


Figure 5-9.  $L_g = 32\text{nm}$  nFinFET  $I_{DS}$ - $V_{GS}$  data (on a log scale), with corresponding UFDG calibration results. The effective extension length,  $L_{eSD}$  is tuned to be  $9.25\text{nm}$ , implying  $L_{\text{eff(weak)}} = L_g + 2L_{eSD} = 50.5\text{nm}$ .

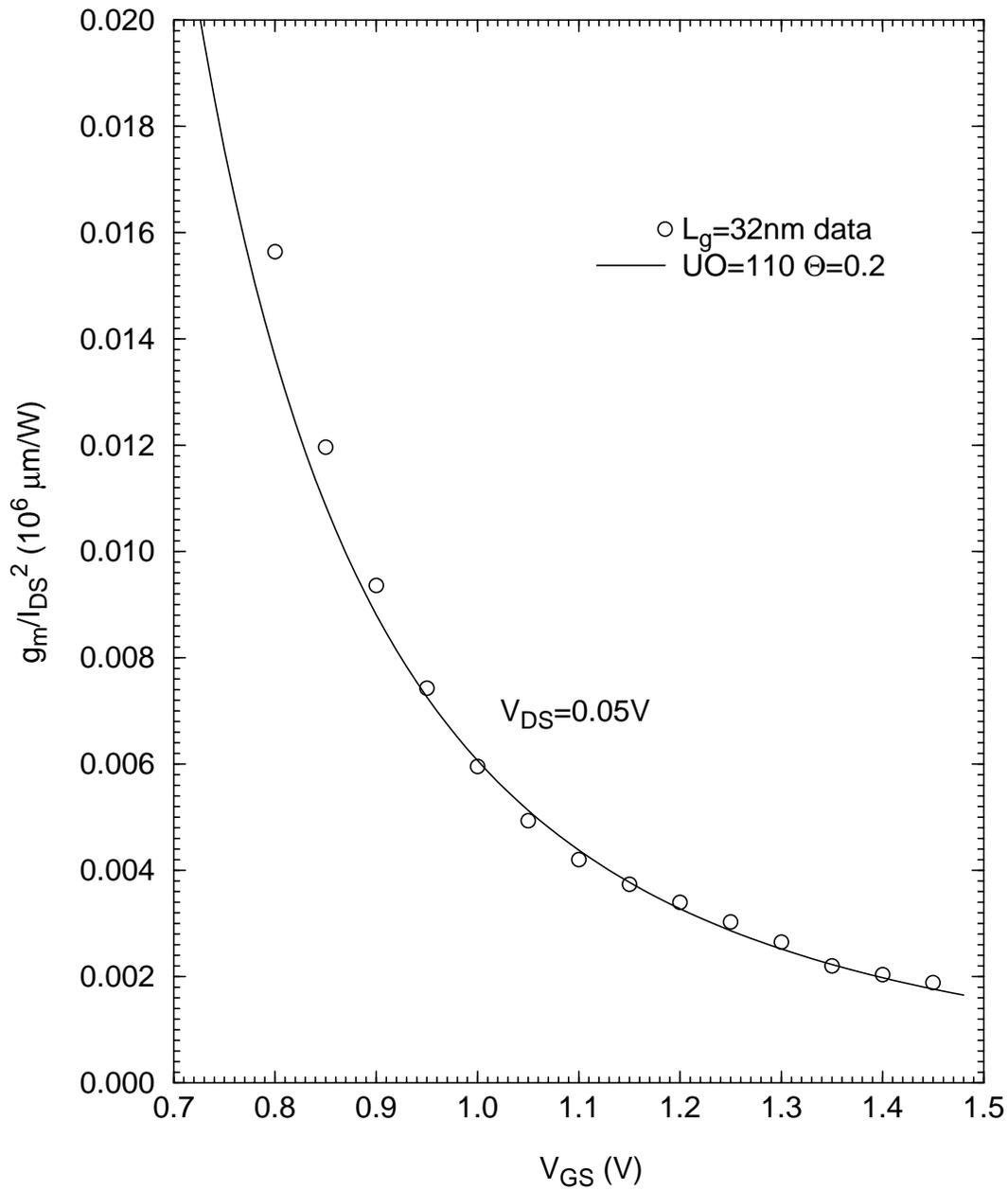


Figure 5-10. Measured and UFDG-predicted low- $V_{DS}$   $g_m/I_{DS}^2$  vs.  $V_{GS}$  used to tune the UFDG mobility parameters:  $UO = 110\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.2$  for the  $L_g = 32\text{nm}$  nFinFET.

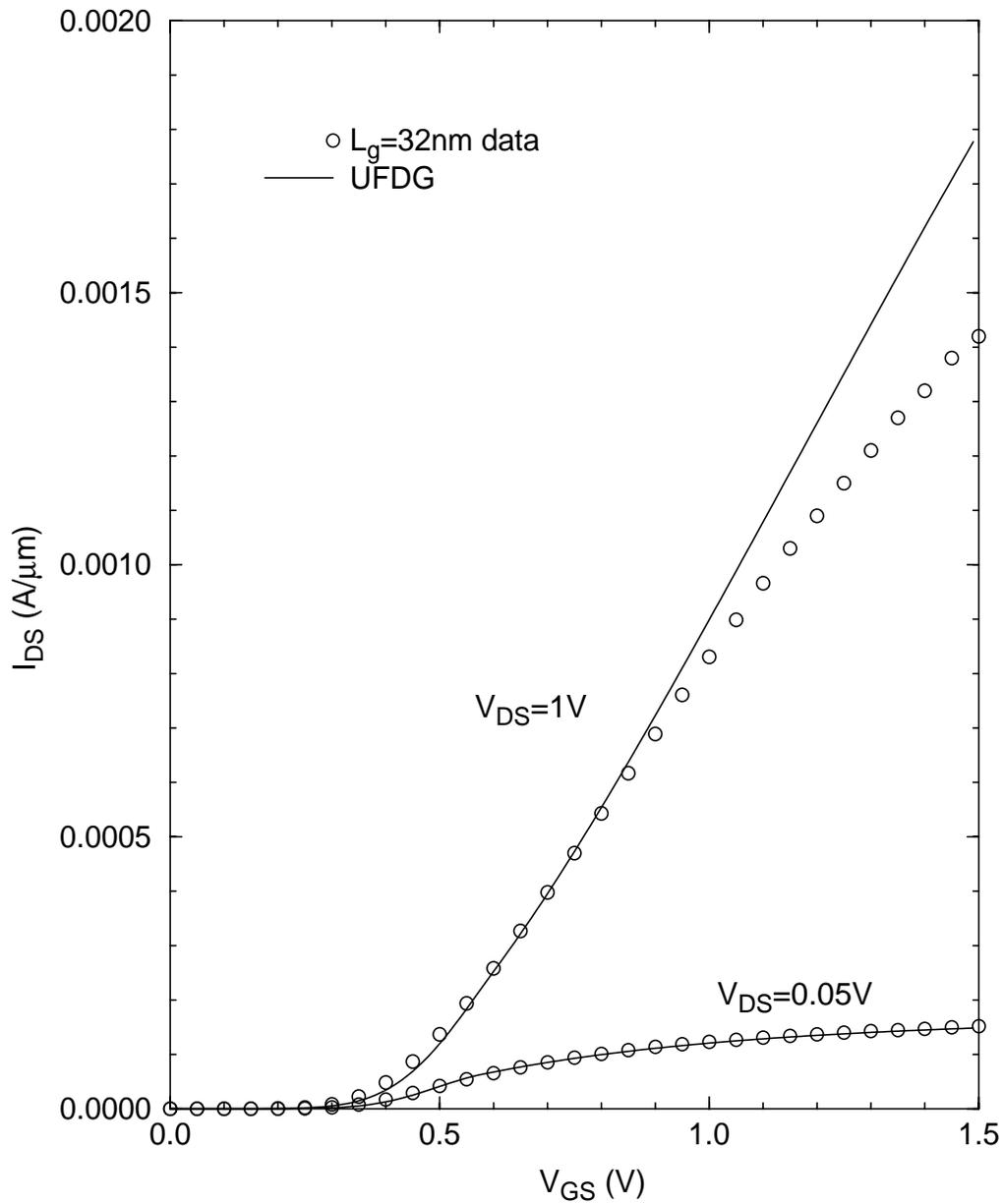


Figure 5-11.  $L_g = 32\text{nm}$  n-FinFET  $I_{DS}$ - $V_{GS}$  data with corresponding UFDG calibration results. The  $\mu_{\text{eff}}$  parameters extracted, via  $g_m/I_{DS}^2$ - $V_{GS}$  in Fig. 5-10 ( $UO = 110\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.2$ ) and  $R_{S/D} = 120\Omega\text{-}\mu\text{m}$  were used in the calibration.

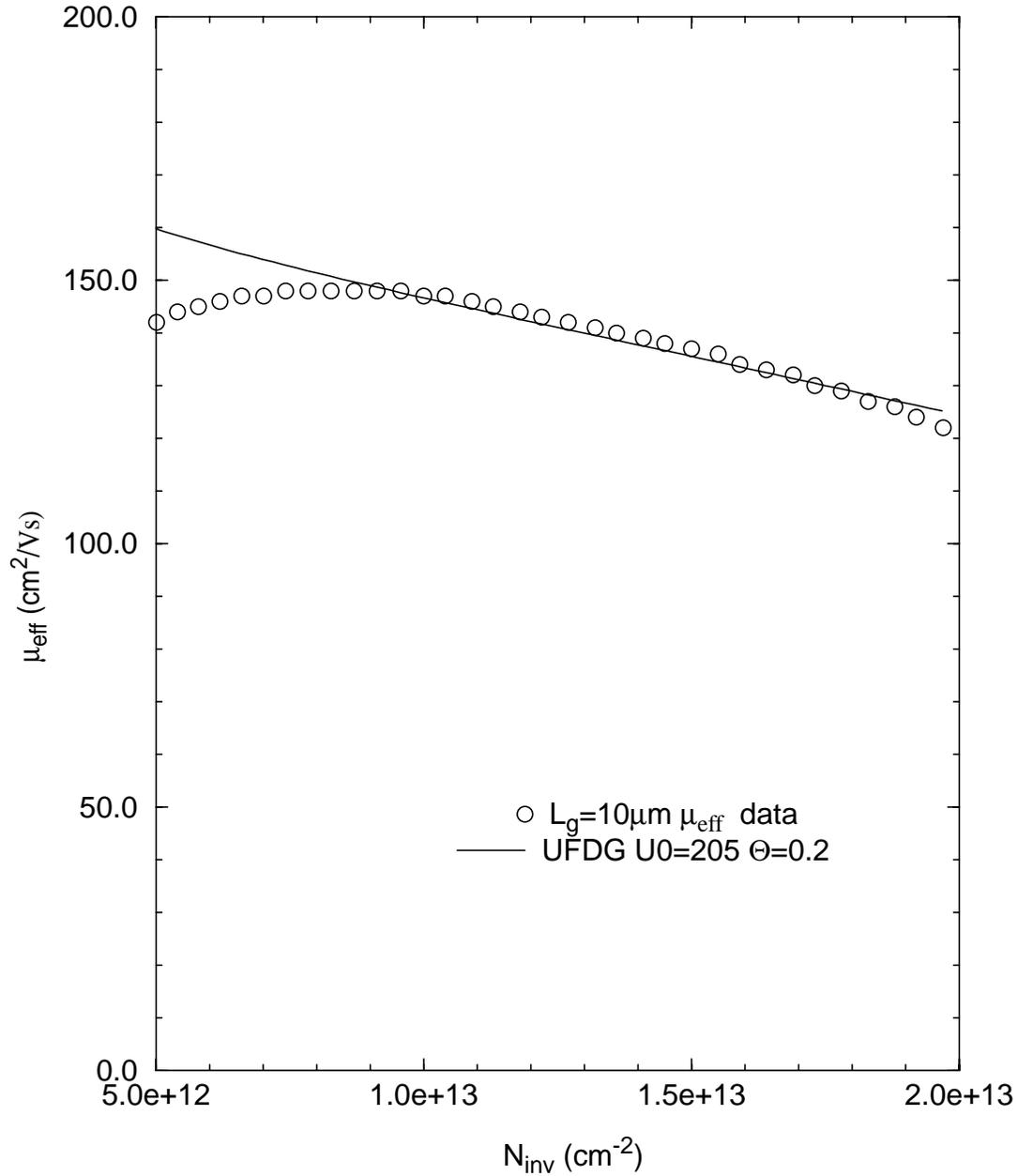


Figure 5-12.  $\mu_{\text{eff}} \sim N_{\text{inv}}$  for p-channel  $L_g = 10\mu\text{m}$  device extracted using split-CV method. UFDG  $\mu_{\text{eff}}$  parameters ( $U_0 = 205 \text{ cm}^2/\text{Vs}$ ,  $\Theta = 0.2$ ) are tuned to match data.

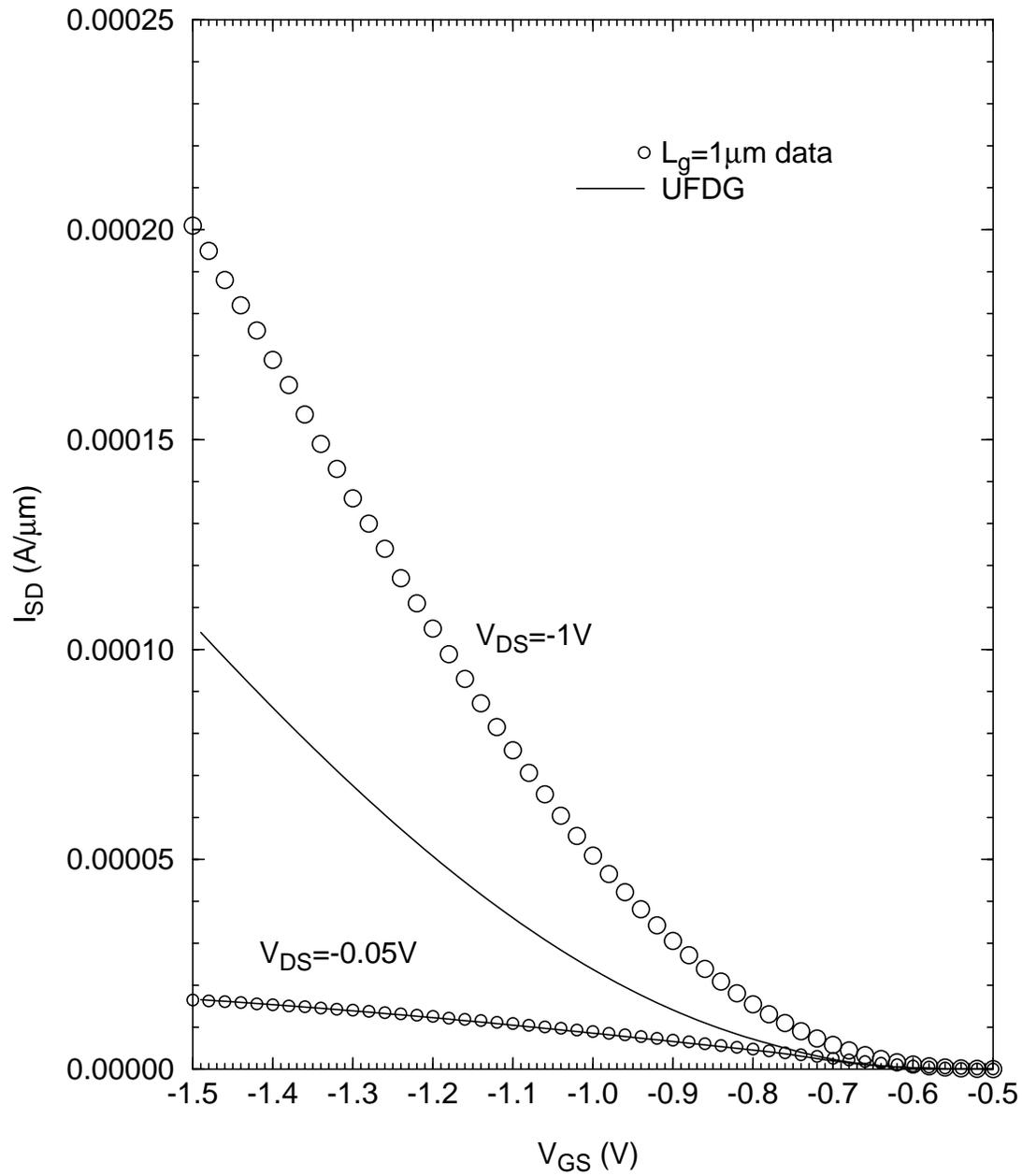


Figure 5-13.  $L_g = 1\mu\text{m}$  p-FinFET  $I_{SD}$ - $V_{GS}$  data, with corresponding UFDG calibration results. The  $\mu_{\text{eff}}$  parameters extracted via split-CV in Fig. 5-12 ( $UO = 205\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.2$ ) were used in the calibration.

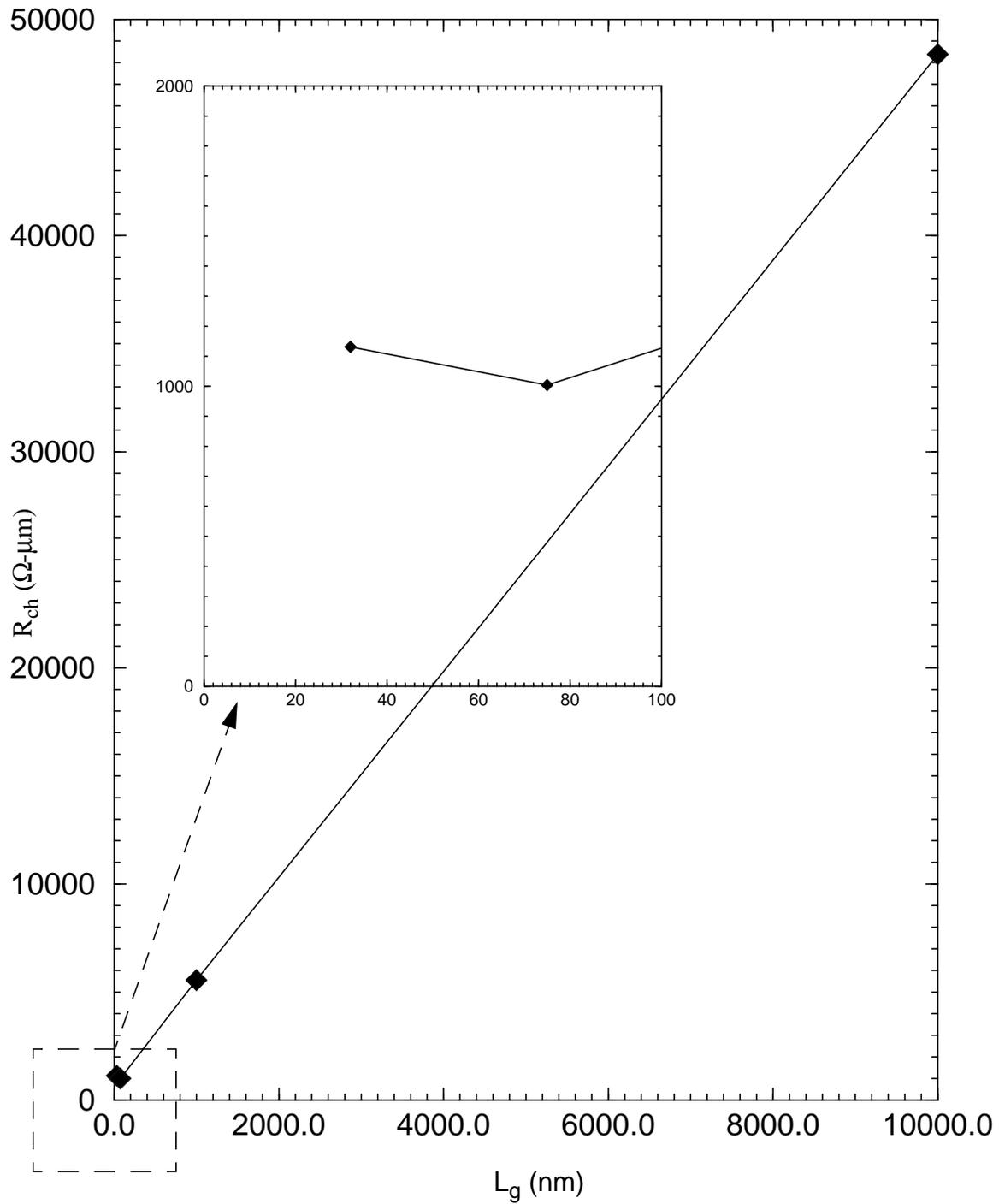


Figure 5-14. Channel resistance of p-channel devices,  $R_{ch}$  ( $V_{GS} = -1V$ ,  $V_{DS} = -0.05V$ ) plotted versus gate length,  $L_g$ . The inset shows the increase in  $R_{ch}$  for  $L_g = 32\text{nm}$  compared to  $L_g = 75\text{nm}$ .

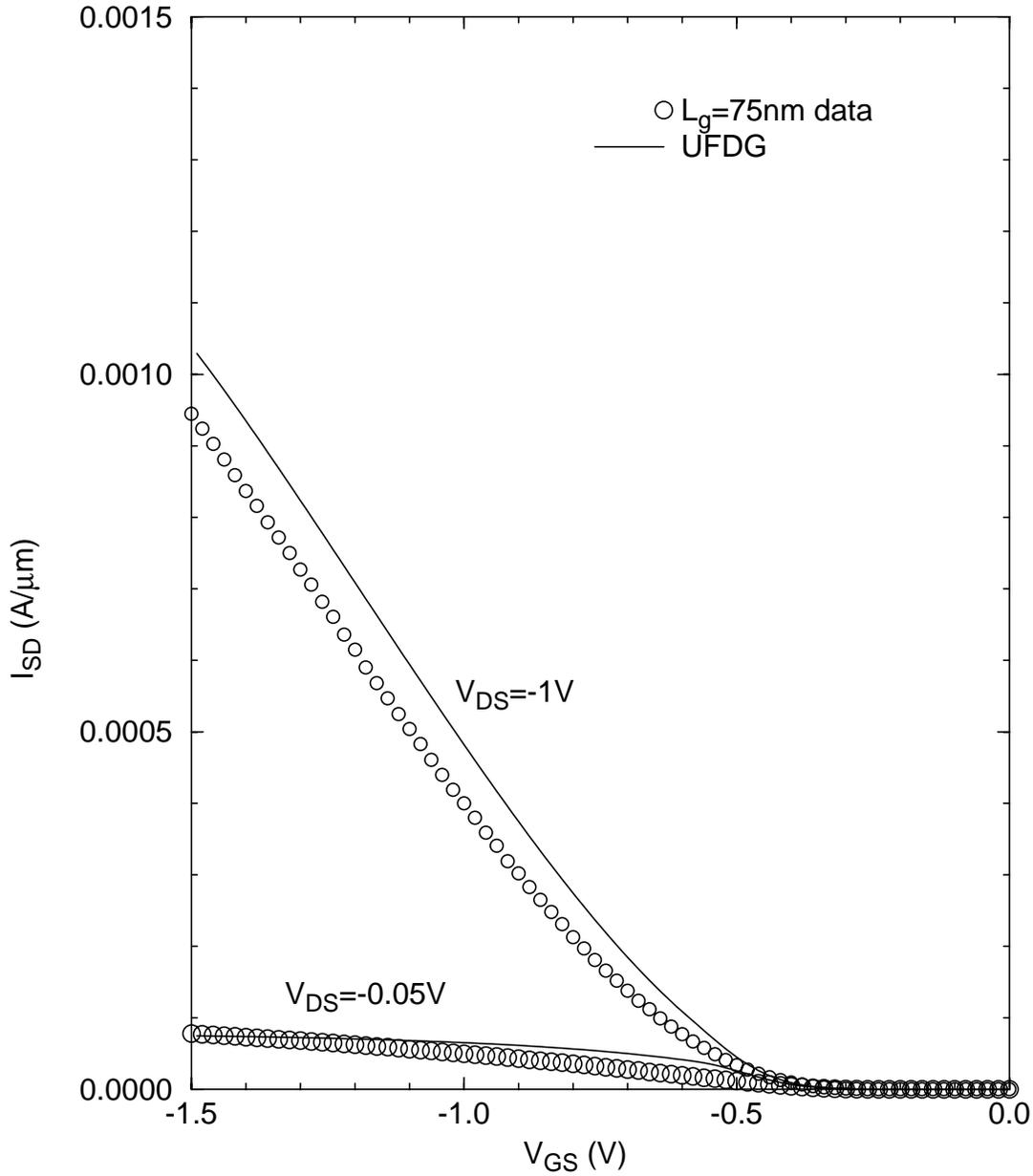


Figure 5-15.  $L_g = 75\text{nm}$  pFinFET  $I_{DS}$ - $V_{GS}$  data, with corresponding UFDG calibration results. The  $\mu_{\text{eff}}$  parameters extracted for  $L_g = 1\mu\text{m}$  ( $UO = 205\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.2$ ) and  $R_S = R_D = 250\Omega\text{-}\mu\text{m}$  were used in the calibration.

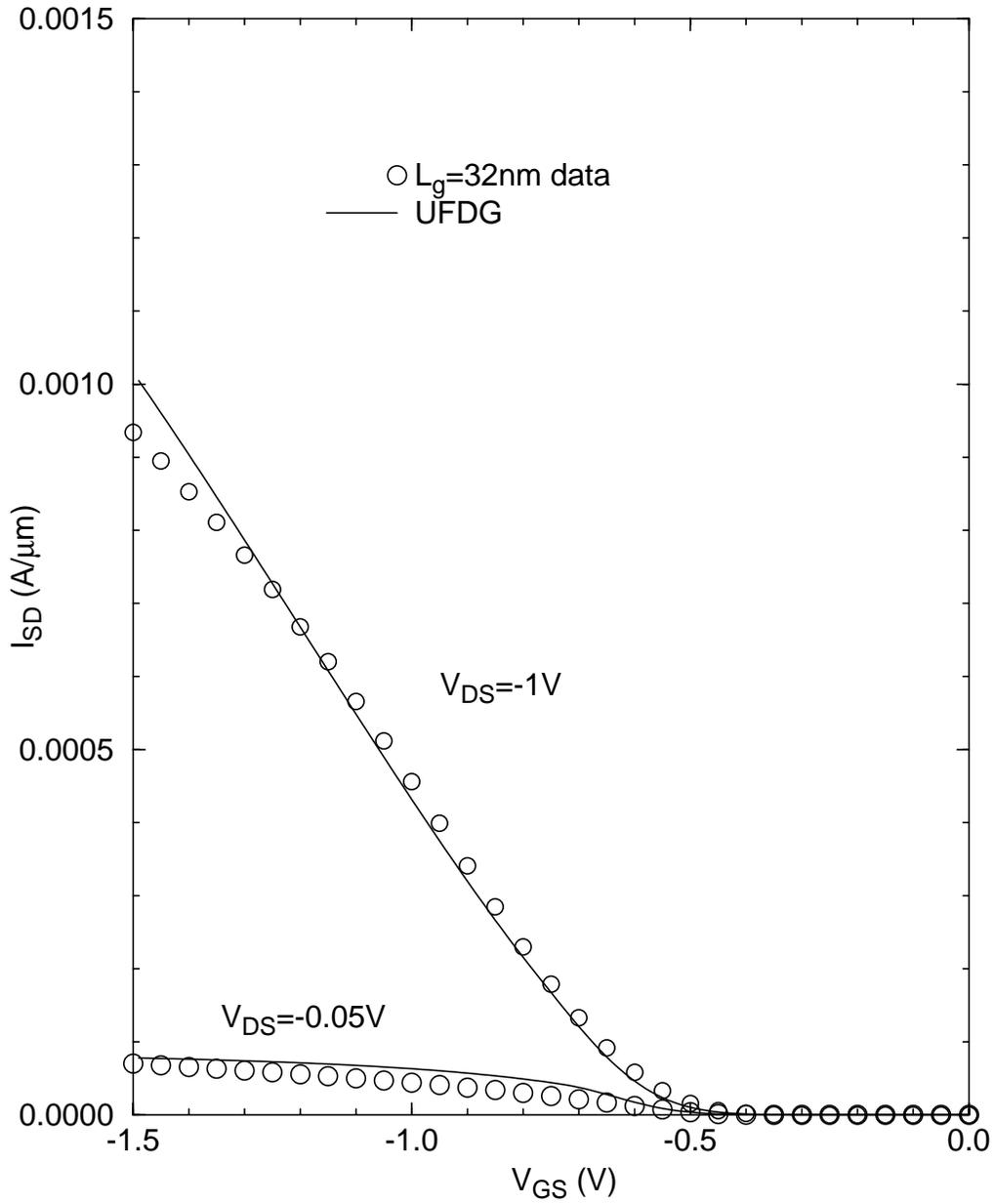


Figure 5-16. p-channel  $L_g = 32\text{nm}$  nFinFET  $I_{DS}$ - $V_{GS}$  data, with corresponding UFDG calibration results. The  $\mu_{\text{eff}}$  parameters ( $UO = 90\text{cm}^2/\text{V}\cdot\text{s}$ ,  $\Theta = 0.2$ ) and  $R_{S/D} = 250\Omega\text{-}\mu\text{m}$  were used in the calibration.

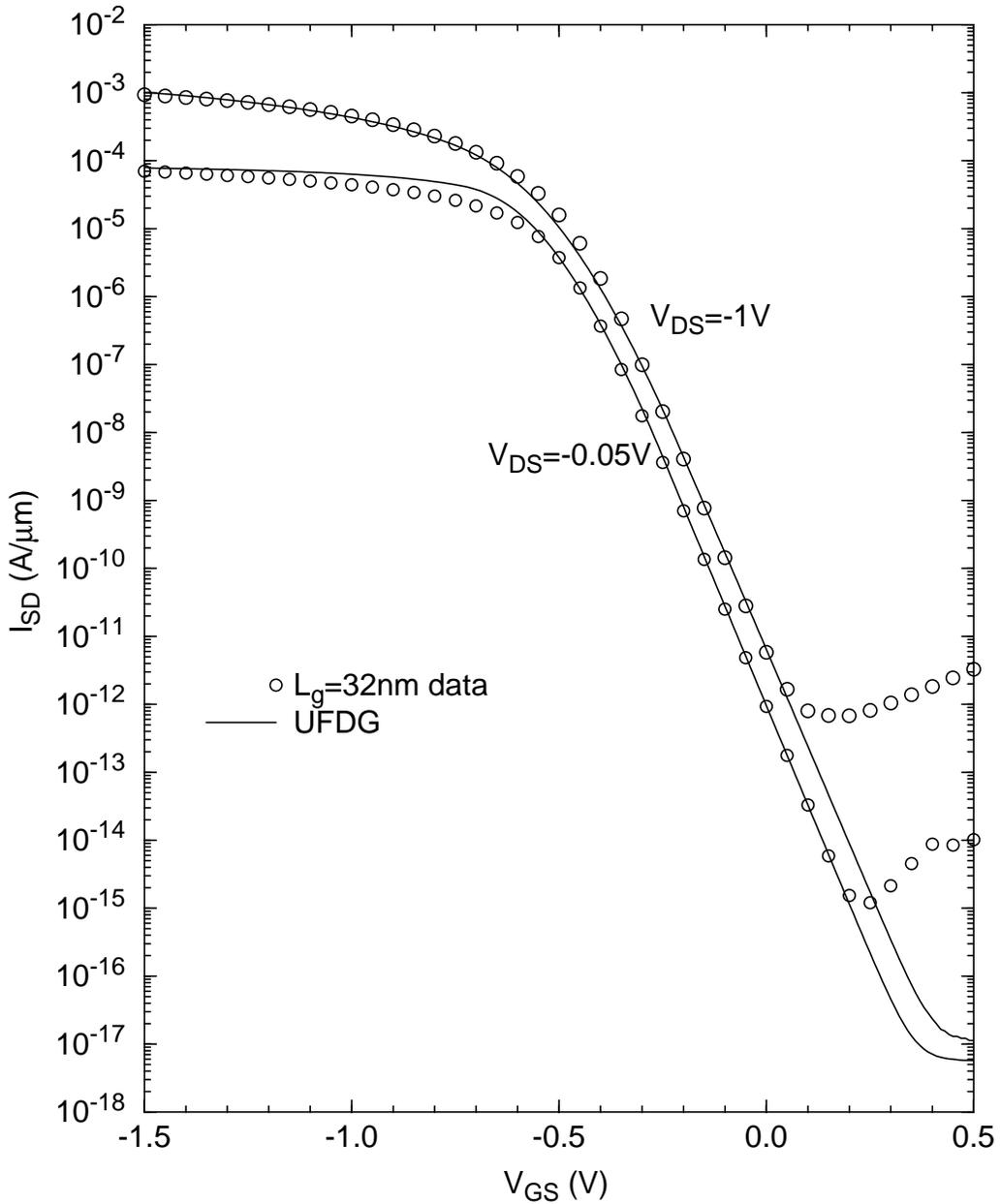


Figure 5-17.  $L_g = 32\text{nm}$  p-FinFET  $I_{DS}$ - $V_{GS}$  data (on a log scale), with corresponding UFDG calibration results. The effective extension length,  $L_{eSD}$  is tuned to be 8nm, implying  $L_{\text{eff(weak)}} = L_g + 2L_{eSD} = 48\text{nm}$ .

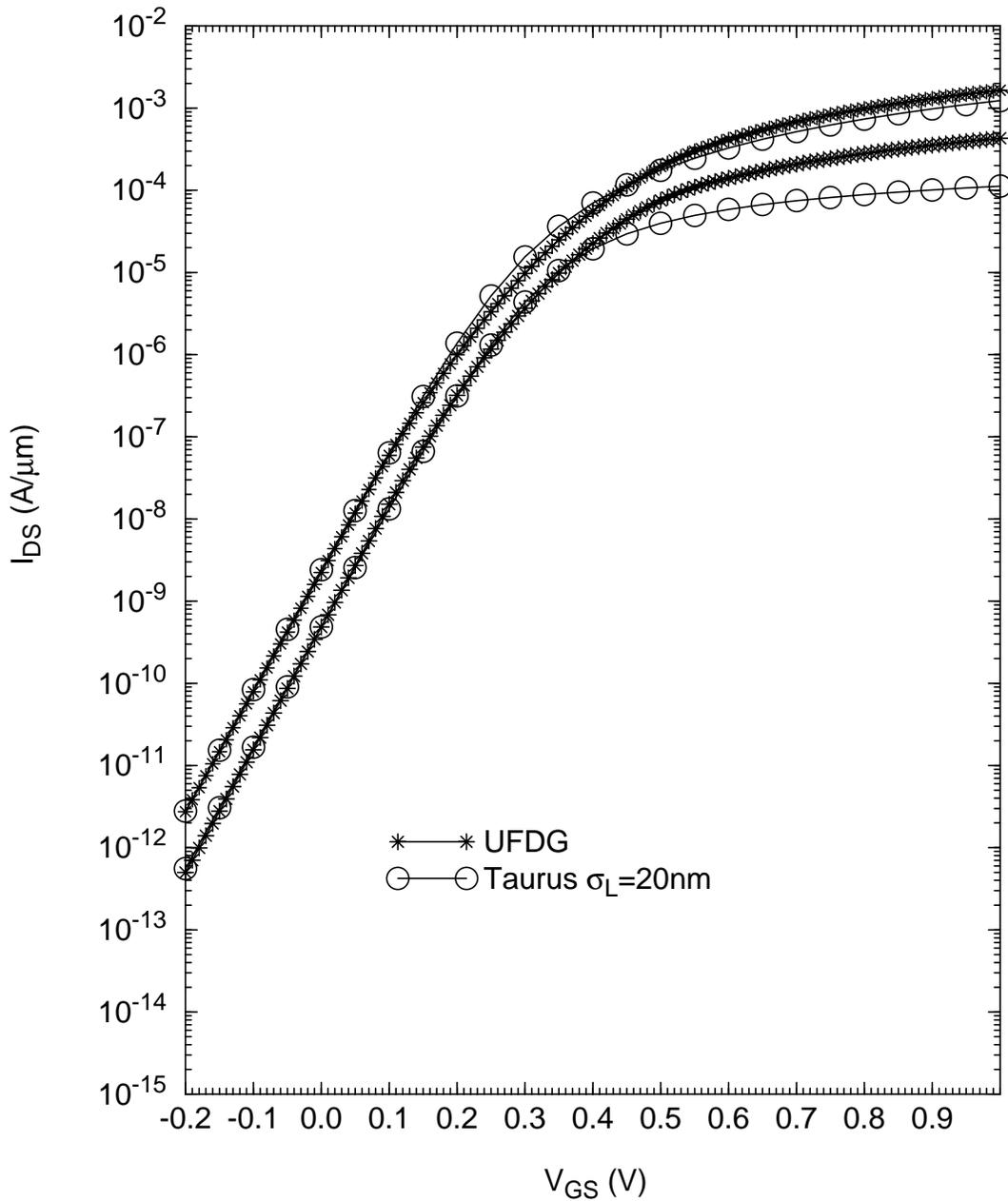


Figure 5-18. This figure portrays the weak-inversion match between UFDG and Taurus-predicted characteristics to estimate the S/D doping profile,  $N_{SD}(y)$ . For  $N_{SD0} = 1 \times 10^{20} \text{ cm}^{-3}$ ,  $\sigma_L = 20 \text{ nm}$ . Both Taurus and UFDG use a constant mobility model with  $\mu_{\text{eff}} = 107 \text{ cm}^2/\text{V}\cdot\text{s}$ .

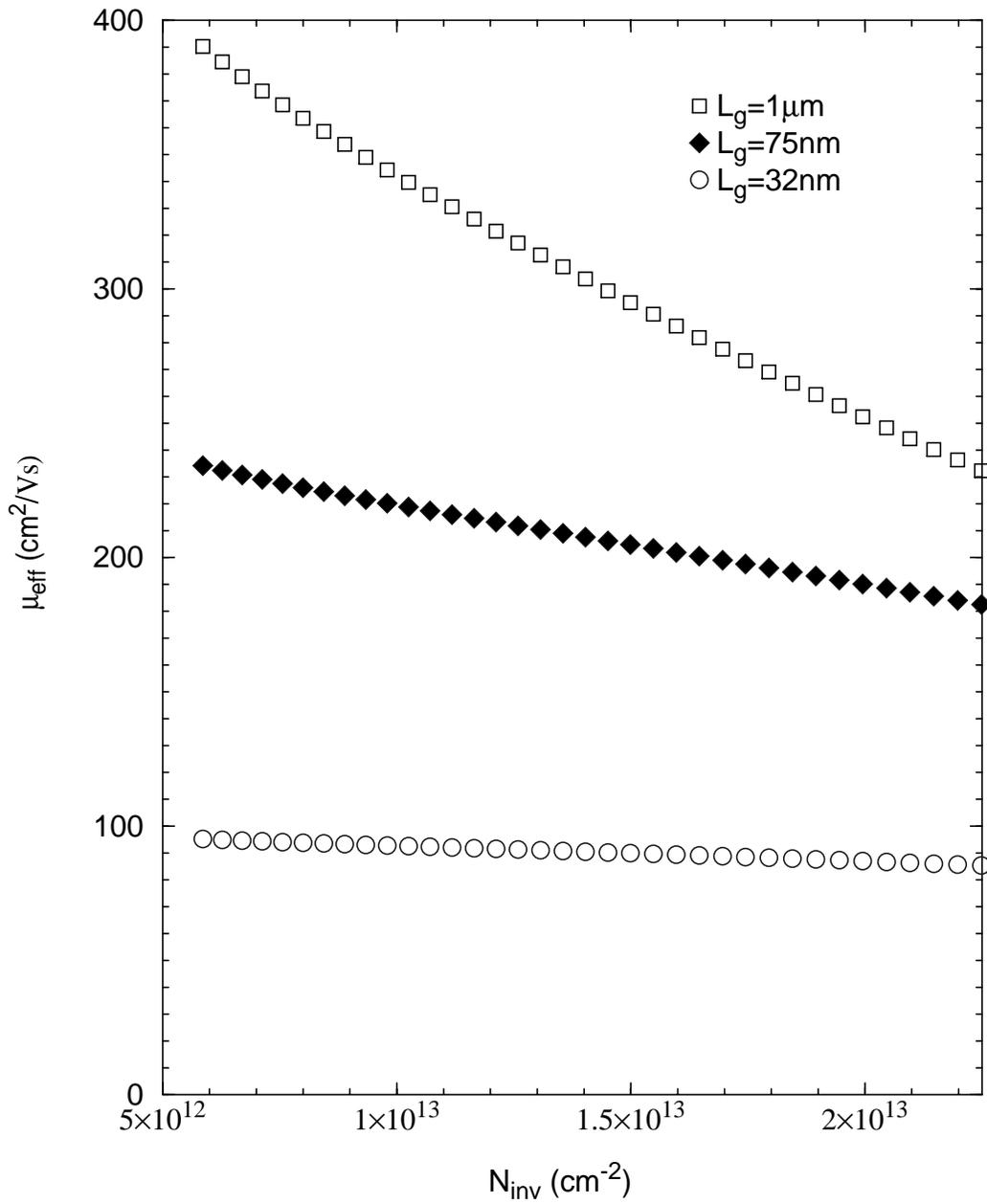


Figure 5-19. UFDG predicted  $\mu_{\text{eff}} \sim N_{\text{inv}}$  data for nFinFETs of various  $L_g$ . We observe a strong degradation of  $\mu_{\text{eff}}$  with reducing  $L_g$ . Further the degradation seems almost independent of  $N_{\text{inv}}$ .

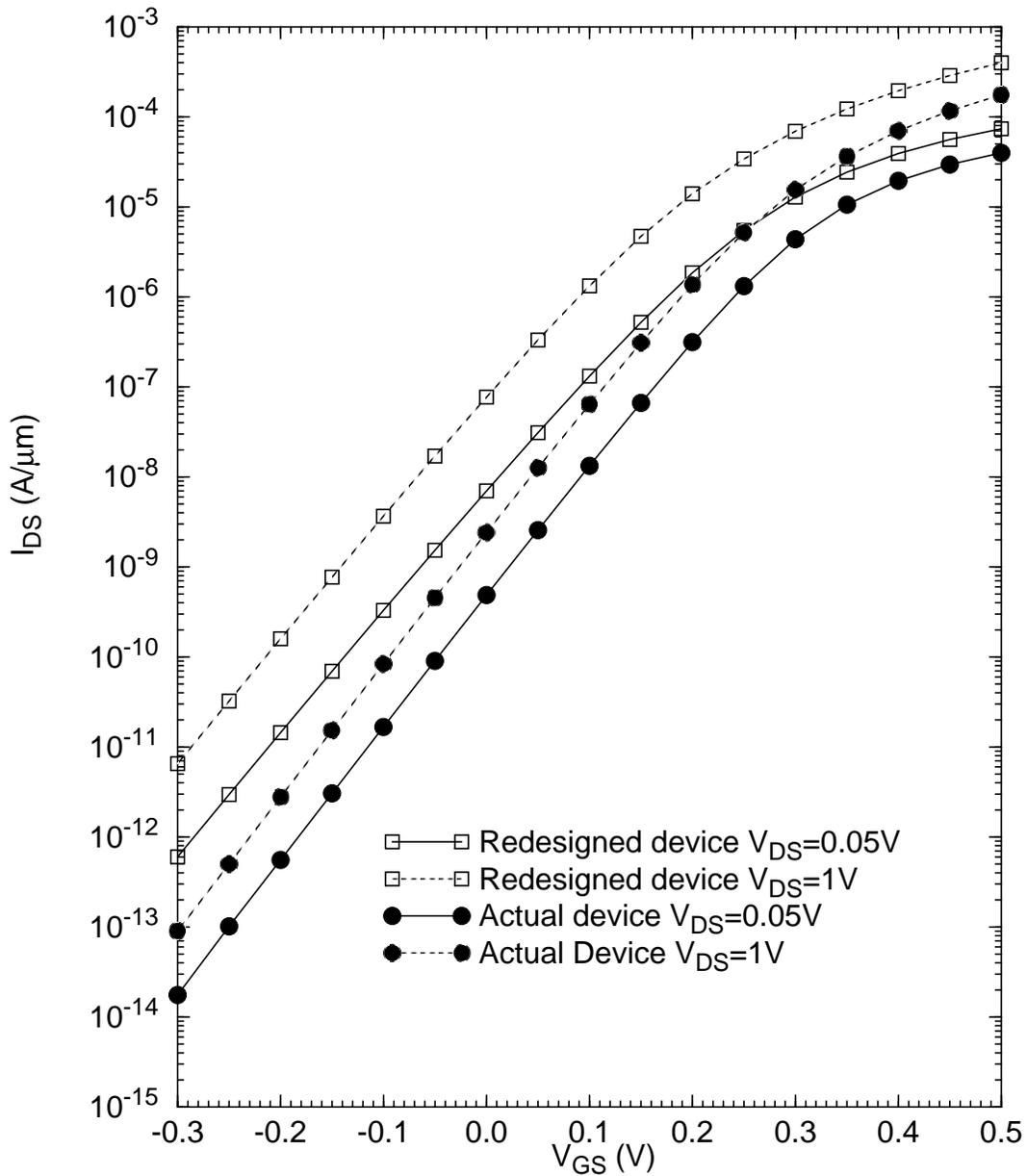


Figure 5-20. This figure portrays the Taurus-projected weak-inversion  $I_{DS}$ - $V_{GS}$  characteristics (on log scale) for the redesigned n-channel device. For reference, the  $I_{DS}$ - $V_{GS}$  characteristics of the actual n-channel device are also shown.

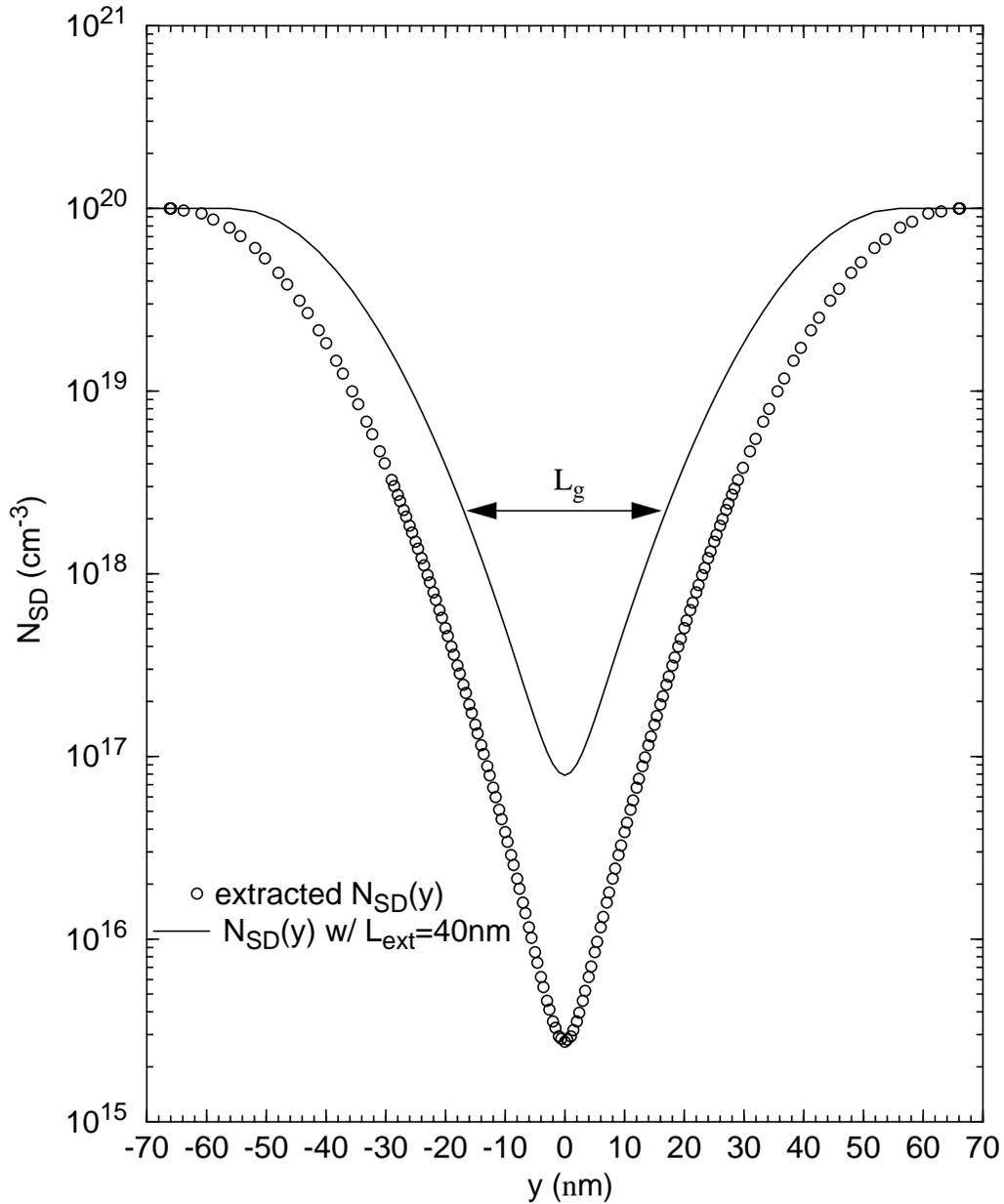


Figure 5-21. This figure portrays the extracted  $N_{SD}(y)$  v/s distance from the center of the channel (in channel direction). It also portrays the projected  $N_{SD}(y)$  for an improved device design with  $L_{ext} = 40\text{nm}$ .  $y = 0$  corresponds to the center of the channel, so  $N_{SD}(y)$  rolls off at  $y = L_g/2 + L_{ext}$ . For  $L_{ext} = 50\text{nm}$ , this corresponds to  $y = 66\text{nm}$  while for  $L_{ext} = 40\text{nm}$ , the corresponding point is  $y = 56\text{nm}$ .

CHAPTER 6  
PHYSICAL INSIGHTS ON CARRIER MOBILITIES IN SHORT-CHANNEL FINFETS

**6-1 Introduction**

Scaling the conventional bulk device implies the need to reduce depletion width with reducing gate length ( $L_g$ ) via increased channel doping [35]. The increased channel doping implies degraded effective carrier mobility,  $\mu_{\text{eff}}$ , due to higher impurity scattering as well as higher surface roughness scattering due to increased transverse electric field,  $E_x$  [35]. In the FinFET device structure, the body can be left undoped since short-channel-effect (SCE) control is obtained via two coupled gates on either side of a thin fin. Thus, SCE control for scaled  $L_g$  is maintained by scaling down the fin width,  $t_{\text{Si}}$ , not by increased channel doping. Since the strong-inversion  $\mu_{\text{eff}}$  is not a strong function of  $t_{\text{Si}}$  [32],  $\mu_{\text{eff}}$  in strong inversion should not reduce with scaling  $L_g$ . While high  $\mu_{\text{eff}}$ , about three-times that of conventional bulk devices, has been widely reported for long  $L_g$  ( $\sim 1\mu\text{m}$ ) devices [4][13][14][15], we have observed strong degradation in  $\mu_{\text{eff}}$  for  $L_g < \sim 100\text{nm}$  (Fig. 5-19); further, such degradation has also been noted in the literature [13][14][15]. In fact, for  $L_g < \sim 100\text{nm}$ , reported  $\mu_{\text{eff}}$  is comparable to that of bulk devices at the same  $L_g$  [13]. This trend is worrisome because it undermines an important benefit of FinFET technology and also raises questions about the potential benefits of FinFET scalability. Various explanations have been advanced to explain this effect. They include Coulomb scattering due to source/drain (S/D) dopants/defects [16] and also “neutral defects” related to S/D processing [13][14].

In this chapter, we study the significance, and possible causes of the mobility degradation for undoped short- $L_g$  FinFETs. In Sec. 6-2, we use the  $L_g = 32\text{nm}$  device data from Sematech (Table 5-1) and our physical/process-based model, UFDG [10], to discuss the significance of attaining high mobility in short- $L_g$  FinFETs. In Sec. 6-3, we perform a literature survey focused

on various papers that deal with the mobility-degradation issue. In Sec. 6-4, we discuss our insights based on UFDG calibration of long- and short- $L_g$  FinFET data at different temperatures. In Sec. 6-5, we use the insights gained to discuss possible causes of this effect and suggest ways of addressing it. Assuming that the mobility-degradation issue can be resolved, we then perform a comparison (Sec. 6-6) between a *pragmatic*  $L_g=32\text{nm}$  FinFET technology with state-of-the-art bulk-MOSFET technology. Based on the comparison, we discuss technological requirements for FinFET-CMOS to eventually replace bulk-CMOS.

### **6-2 Significance of High Mobility for FinFET CMOS Technology**

We use our UFDG model card obtained by calibration of the n-channel  $L_g = 32\text{nm}$  device (Table 5-2), and substitute the mobility parameters with those obtained for the  $L_g = 1\mu\text{m}$  device (Sec. 5-4). The UFDG projection using the (ideal) high mobility is shown in Fig. 6-1, along with the actual UFDG calibration. We observe that the degradation in  $\mu_{\text{eff}}$  accounts for  $\sim 40\%$  degradation in  $I_{\text{on}}$ , from  $I_{\text{on}} = 1.48\text{mA}/\mu\text{m}$  to  $0.9\text{mA}/\mu\text{m}$ , in the n-channel device and  $\sim 30\%$ , from  $I_{\text{on}} = 0.61\text{mA}/\mu\text{m}$  to  $I_{\text{on}} = 0.43\text{mA}/\mu\text{m}$ , in the p-channel device (Fig. 6-2). This severe degradation in  $I_{\text{on}}$  highlights the importance of achieving high mobility in short- $L_g$  FinFETs. Further, the trend (Fig. 5-6) implies a reduction in  $I_{\text{on}}$  with reducing  $L_g$ ; if unresolved, it would undermine an important benefit of scaling. For many applications, especially high-performance (HP) logic, the degradation in mobility threatens to be a “show-stopper” for FinFET CMOS technology.

### **6-3 Literature Survey**

The degradation of mobility for short- $L_g$  undoped DG MOSFETs was first observed in [13], where the low-field mobility,  $\mu_0$ , of long- and short- $L_g$  FinFETs was extracted from measured data using the Y-function method [46] at different temperatures. The authors found a substantial, temperature (T)-independent degradation in  $\mu_0$  which got worse with reducing  $L_g$ . This led them to conclude that the mobility at short- $L_g$  was being limited by “neutral defects”

associated with the source/drain (S/D), since Coulomb-limited mobility,  $\mu_{co}$ , typically has  $\sim T^{1-1.5}$  dependence [47]. The authors also reported some improvement in mobility with increasing anneal temperature which they attributed to “neutral defect” healing. In [14],  $\mu_{eff}$  measurements were done for both the front as well as the back channel of an undoped FD/SOI MOSFET.

Interestingly, the authors obtained similar results as [13] for the front channel but observed little or no degradation in  $\mu_{eff}$  for the back channel. Both papers [13][14] did not check the dependence of  $\mu_{eff}$  on  $N_{inv}$ ; that would have solidified (or not) their insights since  $\mu_{co}$  tends to increase with  $N_{inv}$ . In [16],  $\mu_{eff}$  measurements were done on undoped FD/SOI MOSFETs using the magnetoresistance method and the authors concluded that the degradation for short  $L_g$  was actually caused by Coulomb scattering due to S/D dopants/defects. They suggested that their methodology is more accurate than [13][14] because i) it is not limited by uncertainties in  $L_g$ , ii) is more accurate at low  $N_{inv}$ , and iii) T dependence of  $\mu_{co}$  is higher for magnetoresistance method [48]. Fig. 6 in [16] suggests that the  $\mu_{eff}$  degradation at room temperature is pronounced only for low  $N_{inv}$ ; for high  $N_{inv}$ , the  $L_g = 40\text{nm}$  MOSFET displays about the same  $\mu_{eff}$  as the  $L_g = 1\mu\text{m}$ . However,  $\mu_{co}$  is about 3x higher in magnetoresistance measurements as compared to split-CV measurement [48]. So, the  $\mu_{eff}$  degradation might well be significant even in strong inversion.

#### 6-4 UFDG Calibration of Long- and Short-Channel FinFETs

UFDG has a physical, QM-based model for  $\mu_{eff}$  [32] with just two parameters:  $UO$ , the thick- $t_{Si}$ , low- $E_x$  mobility and  $\Theta$ , which is the surface-roughness tuning parameter. Typical values are  $UO = 1100\text{cm}^2/\text{V-s}$  and  $\Theta = 0.8$  for n-FinFETs and  $UO = 190\text{cm}^2/\text{V-s}$  and  $\Theta = 1$  for p-FinFETs [32]. Table 6-1 shows a summary of the  $\mu_{eff}$  parameters derived from the calibration of both long- and short- $L_g$  FinFETs. Let us first look at the three devices ( $L_g = 1\mu\text{m}$ , 75nm, 32nm) from Sematech. All of these devices were manufactured using the same S/D process (Table 5-1). Fig. 5-19 shows the  $\mu_{eff}-N_{inv}$  graph for the three devices. For the  $L_g = 1\mu\text{m}$  FinFET, the  $\mu_{eff}-N_{inv}$

plot shows (as expected) high  $\mu_{\text{eff}} = 330\text{cm}^2/\text{V}\cdot\text{s}$  at high  $N_{\text{inv}} = 10^{13}\text{cm}^{-2}$ . The low value of  $\Theta$  suggests relatively clean surfaces implying low surface roughness scattering. It is also clear that for progressively shorter  $L_g$ , there is a scattering component caused by proximity of channel carriers to the S/D. The degradation is pronounced at  $L_g = 32\text{nm}$  and seems almost independent of  $N_{\text{inv}}$ . Let us now look at the other devices in Table 6-1 and 6-2 with  $L_g \sim 60\text{-}70\text{nm}$ . We notice that the degradation in mobility corresponds directly to the effective underlap,  $L_{\text{eSD}}$ . Longer  $L_{\text{eSD}}$  implies lower degradation in  $\mu_{\text{eff}}$ . One possible explanation for this fact could be (as suggested in [16]) Coulomb scattering from S/D dopants in the channel or remote Coulomb scattering from the S/D. Shorter underlap implies more S/D dopants in the channel (or in close proximity to the channel). However, Coulomb scattering would tend to be screened out at high  $N_{\text{inv}}$ , while the  $\mu_{\text{eff}}$  degradation is almost independent of  $N_{\text{inv}}$  (Fig. 5-19).

We perform linear region  $I_{\text{DS}}\text{-}V_{\text{GS}}$  measurements and UFDG calibrations at different T for the  $L_g = 1\mu\text{m}$ , 32nm FinFETs from Sematech (Table 5-2). The corresponding  $\mu_{\text{eff}}\text{-}N_{\text{inv}}$  graphs are shown for  $L_g = 1\mu\text{m}$ , 32nm in Figs. 6-3 and 6-4, respectively. For the long- $L_g$  devices, we find that  $\mu_{\text{eff}} \sim \mu_{\text{eff}0} (T/300)^{-0.8}$ , which is close to the typical T dependence of acoustic phonon scattering. For short- $L_g$  devices, there is a much weaker dependence of  $\mu_{\text{eff}}$  on T for all values of  $N_{\text{inv}}$ . If we assume that the additional scattering component due to S/D defects/dopants is  $\mu_N$ , then

$$\frac{1}{\mu_N(L_g = 32\text{nm})} = \frac{1}{\mu_{\text{eff}}(L_g = 32\text{nm})} - \frac{1}{\mu_{\text{eff}}(L_g = 1\mu\text{m})} \quad (6-1)$$

Table 6-3 shows the so-computed values of  $\mu_N$  at different T and  $N_{\text{inv}}$ . We note a very weak dependence on T and almost no dependence on  $N_{\text{inv}}$ . Thus, our observations seem to concur with [13][14], which attribute the  $\mu_{\text{eff}}$  degradation to “neutral defects”, rather than Coulomb scattering. We do not draw any conclusions, however, since there is no insight on what might be the nature of

these “neutral defects” and how they might scatter charge carriers. Another observation from Table 6-1 and 6-2 is that  $\mu_N$  seems approximately the same for n- and p-channel FinFETs. For Sematech p-channel devices,  $\mu_N = 160\text{cm}^2/\text{V-s}$  from Eq. 6-1 ( $\mu_0 = 205\text{cm}^2/\text{V-s}$  for  $L_g = 1\mu\text{m}$ ,  $\mu_0 = 90\text{cm}^2/\text{V-s}$  for  $L_g = 32\text{nm}$  (Table 6-2)). This corresponds closely to the values of  $\mu_N$  for the n-channel devices (Table 6-3). Thus, the  $\mu_{\text{eff}}$  degradation seems to be independent of carrier/dopant type. In Sec. 6-5, we discuss possible explanations of the  $\mu_{\text{eff}}$  degradation and suggest next steps into further diagnosing the problem.

### **6-5 Possible Explanations of the Mobility Degradation**

Let us first discuss the possibility of “neutral defect” scattering as the cause of the mobility degradation. Since all scattering is essentially electrical in nature, a “neutral defect” would have to introduce electrical perturbations in the silicon lattice to effectively scatter charge carriers. One might imagine that the presence of copious numbers of vacancies/divacancies might have this effect. Since interstitials are electrically neutral, they would not be effective in scattering charge carriers. [49] suggests that, since divacancies are immobile at room temperature, high-temperature processing can cause large (T-dependent) concentrations of divacancies to get “frozen” when the temperature is reduced too quickly to room temperature. The suggested solution is a long, relatively low-temperature ( $\sim 350\text{-}400^\circ\text{C}$ ) anneal which would restore the equilibrium detailed balance between vacancies and dopants. This seems to be a plausible explanation given that the S/D anneal used in the Sematech devices (Table 5-1) and also in [13][14][15] is a  $>1000^\circ\text{C}$  spike anneal. However, [49] also suggests that the divacancy concentration,  $N_d$ , is a strong function of donor doping,  $N_D$ , but a weak function of acceptor doping,  $N_A$ . This means that p-channel devices should be relatively immune from the mobility degradation if it is caused by divacancies, which is contrary to our observations in Sec. 6-4. Also, the S/D silicidation anneal is typically a long, relatively low-T anneal which should restore the equilibrium detailed balance between dopants

and vacancies. Nevertheless, we believe that this is a theory worth checking. We suggest that a future study of mobility in undoped short- $L_g$  MOSFETs should check the effect of a  $\sim 30$ min,  $350-400^\circ\text{C}$  anneal following the S/D processing on device characteristics.

Another possible explanation may be that ion-implantation damage, possibly amorphization at the edges of the channel, might account for the mobility degradation. This explanation could explain the observation of [14], where the front-channel carrier mobility of a FD/SOI MOSFET is degraded but the back-channel mobility is not. The problem with this explanation is that the spacer width of the Sematech devices is  $L_{\text{ext}} \sim 50\text{nm}$  (Table 5-1), and the S/D implant is carried out at a  $0^\circ$  vertical tilt. Thus, it seems unlikely that implantation damage could make it to the edges of the channel. To check this explanation, we suggest a reduction in the S/D implant dose. For the Sematech devices, the implant dose was  $3 \times 10^{15} \text{cm}^{-2}$  for a fin height,  $h_{\text{Si}} = 60\text{nm}$  (Table 5-1). This implies a S/D dopant concentration of  $\text{dose}/h_{\text{Si}} = 5 \times 10^{20} \text{cm}^{-3}$ , which is much larger than the active dopant concentration ( $\sim 1 \times 10^{20} \text{cm}^{-3}$ ). We suggest that a reduction of the total implant dose should also be accompanied by implantation at multiple energies to ensure that the S/D dopants are distributed evenly across the vertical height of the fin.

A third possible explanation might be that the mobility degradation is caused by the high- $k/\text{TiN}$  gate stack. It has been shown [50] that nitrogen diffusion to the Si-dielectric interface causes Si-N interface donor/acceptor-type defects, which can cause Coulomb scattering. It is possible that N-diffusion to the edges of the channel during the deposition the TiN gate might account for SiN interface defects, which account for the mobility degradation. This explanation is supported by the fact that, to the best of our knowledge, the papers reporting mobility degradation [14][15][16] as well as our studies here have used device with a TiN gate (gate material not mentioned in [13]). The obvious objection to this theory is that the  $T$ ,  $N_{\text{inv}}$  dependences of  $\mu_{\text{N}}$  make Coulomb scattering an unlikely explanation. Nevertheless, this explanation can be checked

out by fabricating devices with a SiO<sub>2</sub>/poly gate stack. If the TiN gate is determined to be the cause of the effect, it may be necessary either to switch to a different gate material (e.g., TaC) or to introduce a variation in the process which prevents N-diffusion. The high-k material might also potentially play a role in the mobility degradation. Based on our insights in Chapter 3, we think it is best to do away with high-k materials and use the conventional, relatively thick SiO<sub>2</sub>/SiON dielectric for nanoscale FinFET technology.

A slight improvement in short-L<sub>g</sub> mobility with increasing anneal temperature has been reported in [13]. A future study should also check the effect of different anneal temperatures to study any significant correlation with carrier mobility.

The above observations are summarized in Table 6-4. While we have not been able to establish the cause of the mobility degradation, we have laid out a plan for a future study of this effect. As noted in Sec. 6-2, the mobility-degradation issue needs to be resolved for FinFET technology to be viable in high-performance (HP) logic applications. Assuming that high carrier mobility can be realized in nanoscale FinFETs, we discuss technological requirements in Sec. 6-6 for FinFET-CMOS to eventually replace bulk-CMOS.

### **6-6 Comparison of a “Pragmatic” FinFET Technology to Bulk Technology**

We now seek to compare a hypothetical *pragmatic* FinFET technology [2], one with a common midgap metal gate, relatively thick SiO<sub>2</sub>/SiON gate dielectric, high mobility without any transport enhancement via strain, to a state-of-the-art HP bulk technology at L<sub>g</sub> ~ 32nm [45]. We assume that the mobility degradation issue for short-L<sub>g</sub> can be resolved based on our insights in Sec. 6-5 and future development work. We therefore use our UFDG model card obtained from calibration of L<sub>g</sub> = 32nm device data with ideal mobility parameters (obtained for L<sub>g</sub> = 1μm). The UFDG projected I<sub>DS</sub>-V<sub>GS</sub> characteristics of the n- and p-channel devices are shown in Fig. 6-1

and 6-2, respectively. Based on the comparison, we seek to understand the technological requirements for FinFET-CMOS to eventually replace planar bulk-CMOS.

Since FinFET device width is along the height of the fin, a fair comparison of the two technologies would entail comparison of  $I_{\text{on}}$  per unit lithographic pitch (P) [1]. This is discussed more in the next paragraph; we first look at  $I_{\text{on}}$  per unit device width. We use the (correct) device width definition,  $W = h_{\text{Si}}$ , for the FinFET [20] as compared to  $W = 2h_{\text{Si}}$  typically used in the literature. The HP n-channel  $L_g = 30\text{nm}$  bulk device displays  $I_{\text{on}} = 1.55\text{mA}/\mu\text{m}$  [45], while the pragmatic [2]  $L_g = 32\text{nm}$  Sematech FinFET with ideal mobility parameters would give  $I_{\text{on}} = 1.48\text{mA}/\mu\text{m}$  (Fig. 6-1). The HP p-channel  $L_g = 30\text{nm}$  bulk device displays  $I_{\text{on}} = 1.21\text{mA}/\mu\text{m}$  [45], while the pragmatic p-FinFET would give  $I_{\text{on}} = 0.61\text{mA}/\mu\text{m}$  (Fig. 6-2). The p-channel FinFET  $I_{\text{on}}$  is low, partly due to the high parasitic resistance,  $R_{\text{S/D}} = 250\Omega\text{-}\mu\text{m}$  (Table 5-2). Due to the long underlap, there is scope for improvement in the FinFET  $I_{\text{on}}$  for both p- and n-channel devices via the device redesign insights in Sec. 5-7. For example, in the p-channel device, a more reasonable  $R_{\text{S/D}} = 100\Omega\text{-}\mu\text{m}$  would give  $I_{\text{on}} = 0.82\text{mA}/\mu\text{m}$ .

Aside from higher  $R_{\text{S/D}}$  as a result of the long underlap, the reason that FinFET  $I_{\text{on}}$  per unit “gate width” ( $2h_{\text{Si}}$ ) is significantly lower than  $I_{\text{on}}$  per unit width in bulk devices is three-fold. One, the bulk-inversion effect [20] in undoped UTB MOSFETs implies that the charge centroid is significantly removed from the Si-dielectric interface, resulting in a lower gate capacitance than bulk devices. We believe that this effect leads to ~15% reduction in charge in undoped devices for the same gate overdrive ( $V_{\text{GS}} - V_t$ ) [51]. Also, the midgap gate typically used in FinFETs implies a relatively high threshold voltage in strong inversion,  $V_{\text{ts}} \sim 0.4\text{V}$ , causing lower  $I_{\text{on}}$ .  $V_t$  can potentially be tuned by allowing controlled densities of S/D dopants into the channel [12], or by using a dual-metal gate technology. Additionally, the absence of carrier transport enhancement via strain leads to p-channel  $I_{\text{on}}$  per unit device width being significantly lower in the FinFET as

compared to the bulk device. The reason for the significant difference in p-channel  $I_{on}$  rather than the n-channel  $I_{on}$  is because strain leads to mobility enhancement  $\sim 4x$  for holes as compared to only  $\sim 1.7x$  for electrons [52].

Based on the above insights, we can make the fair assumption that an optimized *pragmatic* n-channel FinFET would give somewhat better  $I_{on}$  per unit device width than a HP bulk device, provided the FinFET device width is defined as  $W = h_{Si}$ . A pragmatic p-channel FinFET would give significantly lower  $I_{on}$  per unit device width in comparison because of the absence of the transport enhancement via strain [52] as in a bulk device. It may be feasible to enhance p-FinFET  $I_{on}$  via strain-induced transport enhancement in the future [53].

The primary advantage of FinFET technology over a future bulk technology is the enhanced scalability due to excellent SCE control and reduced process-related variations from device to device. While bulk devices are not scalable beyond  $L_g < \sim 30nm$ , FinFETs can be scaled to  $L_g < 10nm$  [2]. Given this enhanced scalability, we reasonably assume that for HP FinFET-CMOS technology to be competitive with HP bulk technology would require the total device width per pitch to be about the same as the pitch. For a nominal technology with one fin per lithographic pitch, this translates to  $h_{Si} = P$  while for a spacer lithography technology [54] with two fins per pitch,  $2h_{Si} = P$  or  $h_{Si} = P/2$ . While the challenges for future bulk transistors are novel dielectric and channel materials and a dual metal gate technology, pragmatic FinFET devices could use the conventional  $SiO_2/SiON$  gate dielectric (Chapter 2) and a midgap metal gate. The important challenges for the commercialization of FinFET-CMOS technology include optimal S/D engineering (Chapter 5), resolution of the mobility degradation issue, and fabrication of thin, tall fins for adequate on-state current drive.

## 6-7 Summary

In this chapter, we obtained physical insights into the significance and possible causes of carrier mobility degradation in short-channel FinFETs. We offered processing suggestions which can be used to diagnose and possibly mitigate the effect. We then compared our  $L_g = 32\text{nm}$  FinFET device, with high mobility, with a state-of-the-art bulk device, noting technological requirements for FinFET-CMOS to eventually replace conventional bulk-CMOS.

Table 6-1. Summary of mobility parameters for n-channel FinFETs. The table also gives the source of the data along with relevant device characteristics.

Data Source	n-channel $L_g$ (nm)	Mobility parameters	Device characteristics
Sematech	1 $\mu$ m	UO = 900 $\Theta$ = 0.32	{110} surface  $L_{eSD} = 9.25\text{nm}$
Sematech	75nm	UO = 350 $\Theta$ = 0.2	{110} surface $L_{eSD} = 9.25\text{nm}$
Freescale	60nm	UO = 200 $\Theta$ = 0.2	underlapped
Freescale	70nm	UO = 125 $\Theta$ = 0.4	$L_{eSD} = -5$ to $-10\text{nm}$
Sematech	32nm	UO = 110 $\Theta$ = 0.2	{110} surface  $L_{eSD} = 9.25\text{nm}$
Sematech	32nm	UO = 80 $\Theta$ = 0.2	{100} surface  $L_{eSD} = 9.25\text{nm}$

Table 6-2. Summary of mobility parameters for p-channel FinFETs. The table also gives the source of the data along with relevant device characteristics.

Data Source	p-channel $L_g$ (nm)	Mobility parameters	Device characteristics
Sematech	1 $\mu$ m	UO = 205 $\Theta$ = 0.2	{110} surface  $L_{eSD} = 8\text{nm}$
Sematech	75nm	UO $\cong$ 205 $\Theta$ $\cong$ 0.2	{110} surface $L_{eSD} = 8\text{nm}$
Freescale	60nm	UO = 140 $\Theta$ = 0.2	$L_{eSD} = 3\text{nm}$
Sematech	32nm	UO $\cong$ 90 $\Theta$ = 0.2	{110} surface  $L_{eSD} = 8\text{nm}$
Sematech	32nm	UO $\cong$ 90 $\Theta$ = 0.2	{100} surface  $L_{eSD} = 8\text{nm}$

Table 6-3. The extracted mobility due to undetermined degradation mechanism ( $\mu_N$ ) for  $L_g = 32\text{nm}$  as a function of temperature (T) and inversion charge density,  $N_{\text{inv}}$ .

T (K)	$\mu_N$ ( $N_{\text{inv}}=6 \times 10^{12} \text{ cm}^{-2}$ )	$\mu_N$ ( $N_{\text{inv}}=1 \times 10^{13} \text{ cm}^{-2}$ )	$\mu_N$ ( $N_{\text{inv}}=1.5 \times 10^{13} \text{ cm}^{-2}$ )
125	161	160	164
175	178	182	184
200	194	192	206
225	194	197	208
250	200	203	208
275	154	157	165

Table 6-4. Possible causes and solutions to the problem of mobility degradation in nanoscale FinFETs

Possible causes of mobility degradation	Possible Solutions
High concentration of vacancies/divacancies during high temperature processing, which get “frozen” due to quick reduction to room temperature.	A long (~30min), relatively low temperature (350-400° C) anneal to restore equilibrium detailed balance between dopants and vacancies.
Ion implantation damage, e.g. amorphization, which affects the edges of the channel	Reduce implant dose; implant at multiple energies to distribute S/D dopants uniformly along fin height.
Nitrogen diffusion to channel edges during deposition of TiN metal gate, which results in interface states that may cause Coulomb scattering.	Fabricate devices with SiO <sub>2</sub> /poly gate stack to check explanation. If true, use a different gate material or change process to prevent N-diffusion.

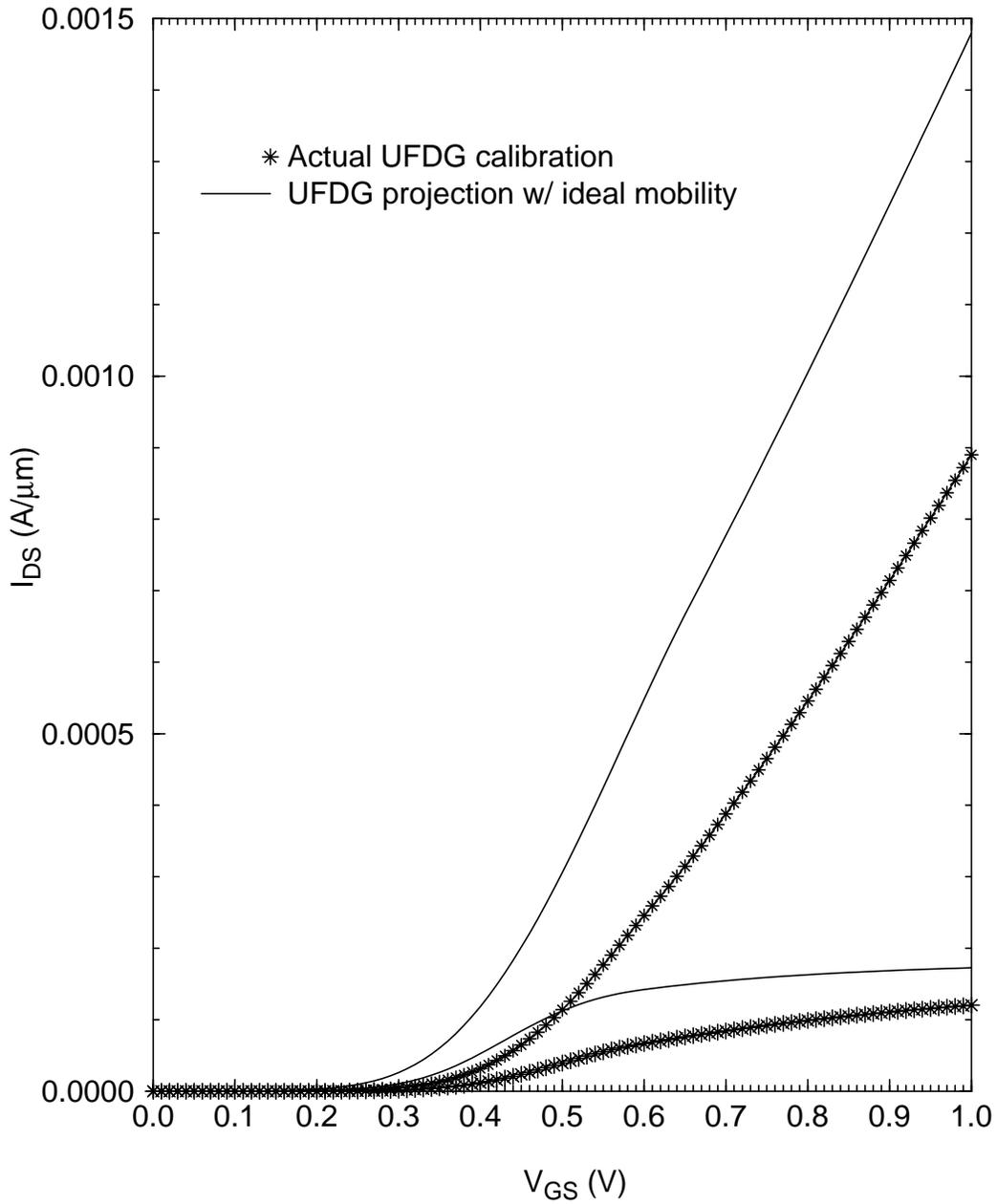


Figure 6-1. This figure portrays UFDG calibration of  $I_{DS}$ - $V_{GS}$  characteristics of the n-channel  $L_g = 32\text{nm}$  device in Table 5-1 along with UFDG projected  $I_{DS}$ - $V_{GS}$  characteristics with ideal mobility parameters (same as  $L_g = 1\mu\text{m}$ )

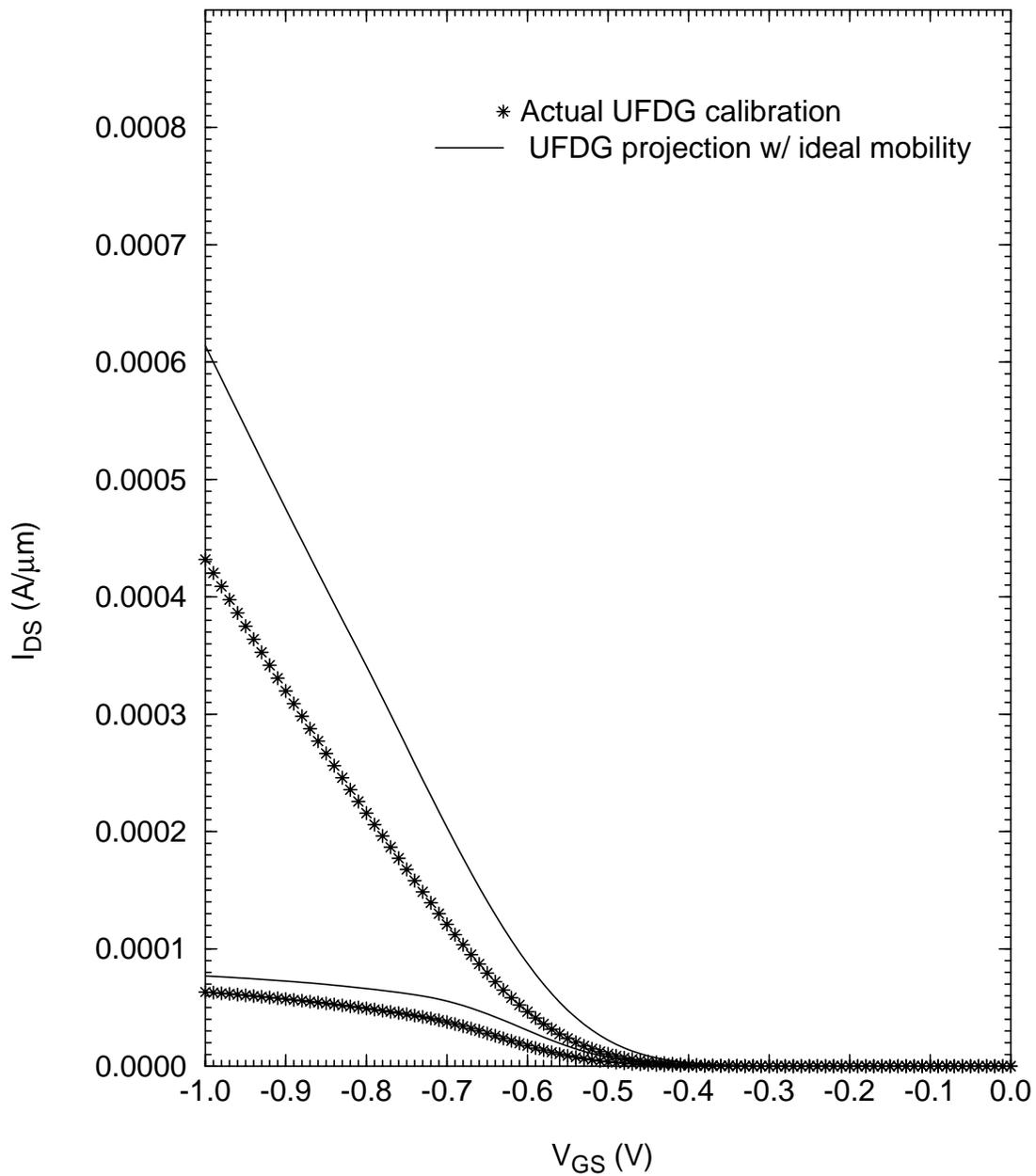


Figure 6-2. This figure portrays UFDG calibration of  $I_{DS}$ - $V_{GS}$  characteristics of the p-channel  $L_g = 32nm$  device in Table 5-1 along with UFDG projected  $I_{DS}$ - $V_{GS}$  characteristics with ideal mobility parameters (same as  $L_g = 1\mu m$ )

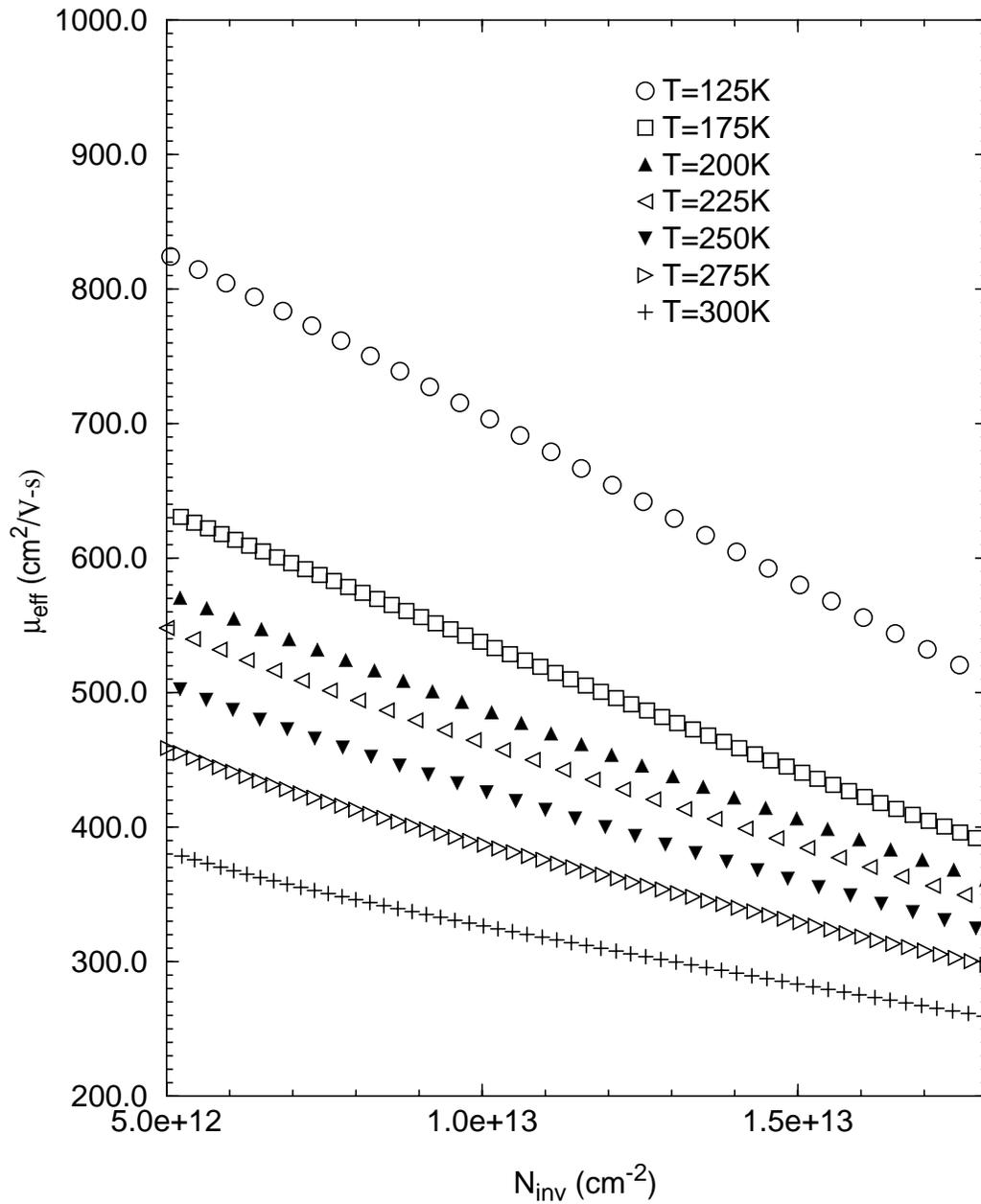


Figure 6-3. This figure portrays UFDG-predicted effective mobility,  $\mu_{\text{eff}}$  vs. inversion charge density,  $N_{\text{inv}}$  at different temperatures for n-channel  $L_g = 1\mu\text{m}$  FinFET.

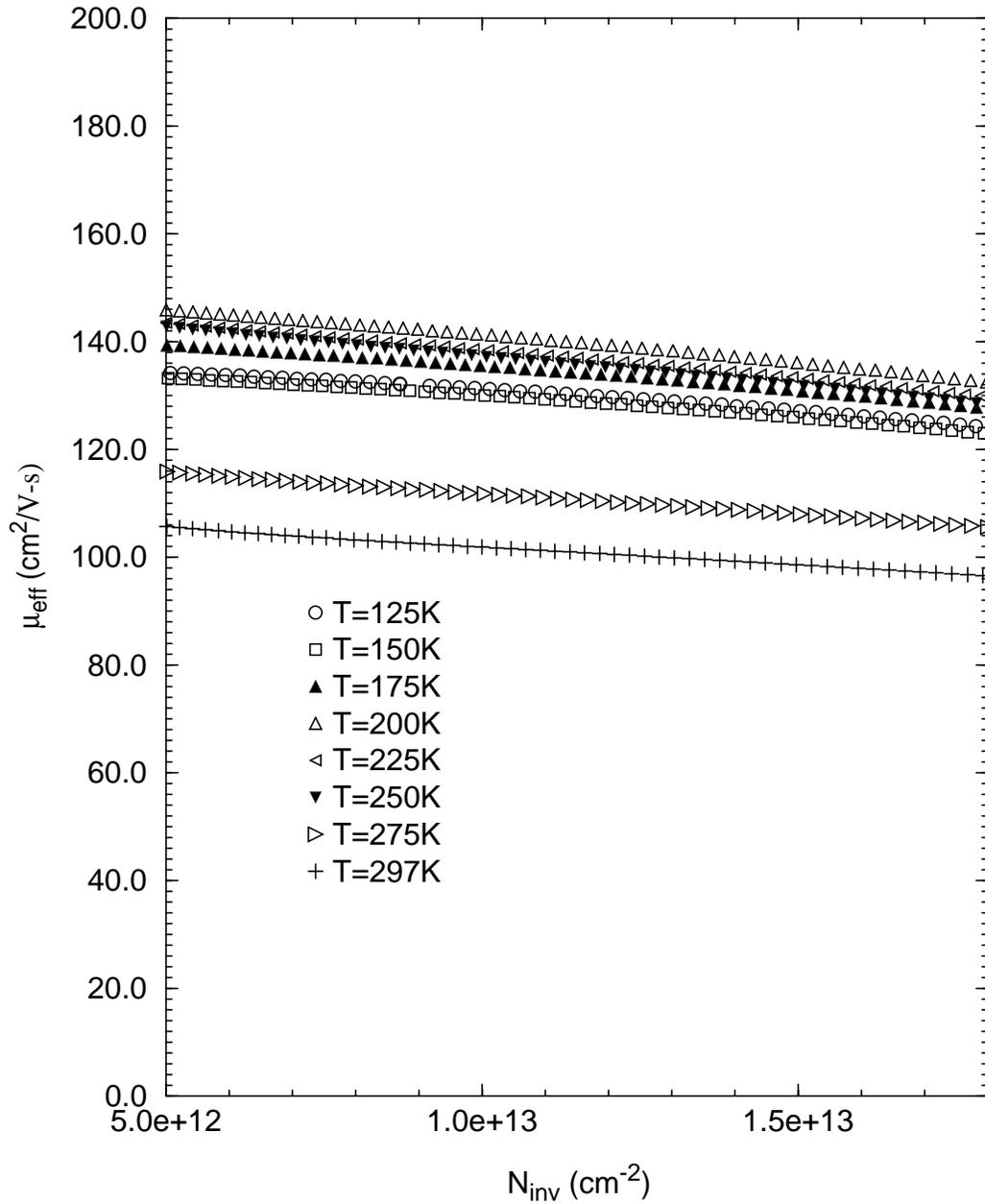


Figure 6-4. This figure portrays UFDG-predicted effective mobility,  $\mu_{\text{eff}}$  vs. inversion charge density,  $N_{\text{inv}}$  for n-channel  $L_g = 32\text{nm}$  FinFET.

## CHAPTER 7 SUMMARY AND FUTURE WORK

### 7-1 Summary

This dissertation focused on physical modeling and optimal design insights for nanoscale Double-Gate CMOS devices and technology. The major contributions of the research are summarized as follows.

In Chapter 2, we proposed and demonstrated via device simulations a novel ITFET design which is more scalable than a conventional ITFET. The design used a midgap metal gate on the FinFET in conjunction with a higher work function gate on the FDFET ( $p^+$  poly gate on the nITFET and vice versa). This design allows for a “poorly tempered” FDFET since the ITFET weak-inversion characteristics are not affected by the short-channel effects of the FDFET. The resulting device gives an improvement in on-state current of ~20-35% over the conventional FinFET. The proposed design is scalable to gate length  $L_g \sim 9\text{nm}$ , which is near the end of the SIA ITRS roadmap.

In Chapter 3, we studied the impact of a high-k gate dielectric on the performance and scalability of nanoscale DG-FinFET CMOS. We designed a high-k FinFET, and compared it with a *pragmatic* FinFET having thicker ( $> EOT$ ), conventional  $\text{SiO}_2$  gate dielectric at the HP-45nm node ( $L_g = 18\text{nm}$ ) of the 2005 SIA ITRS. We found that the high-k dielectric actually degrades CMOS-speed performance due to two heretofore unacknowledged compromising effects of the high-k. One is additional  $C_G$  reduction due to quantization because of higher transverse electric field resulting from thinner EOT; the  $I_{on}$  enhancement expected for the thinner EOT is thus compromised. The second is additional loss of effective gate bias due to source resistance,  $R_S$  because of the higher  $I_{on}$ ; the enhancement in  $I_{on}$  is thus compromised more. Further, in the pFinFET, hole mobility is degraded due to the higher transverse field, compromising  $I_{on}$  even

more. Our simulation-based study did imply a small improvement in the scalability of DG-FinFET CMOS due to the thinner EOT. However, given the reduced circuit performance and the technological challenges of integrating a high-k dielectric into the process flow, we concluded that it is not worthwhile. We believe, also based on our study, that the pragmatic approach to nanoscale FinFET CMOS should be taken, with good performance expected to the end of the ITRS.

In Chapter 4, we presented, for the first time, a model for parasitic fringe capacitance ( $C_f$ ) in DG MOSFETs with (realistic) non-abrupt source/drain junctions. Our model is physically based and therefore quasi-predictive. We evaluated our model for different device structures and found good agreement with numerical simulation results. We implemented the model for  $C_f$  in our physical/process-based model, UFDG. We used quasi-predictive ring oscillator simulations using UFDG/Spice3 to study the impact of a low-k spacer on FinFET CMOS speed performance. We found that using a  $\text{SiO}_2$  ( $k_{sp} = 3.9$ ) spacer instead of a  $\text{Si}_3\text{N}_4$  ( $k_{sp} = 7.5$ ) reduced RO delay by ~25%, while using an air ( $k_{sp} = 1$ ) spacer reduced RO delay by ~40%.  $C_f$  can also be reduced by increasing the effective underlap,  $L_{eSD}$ ; however, that also increases the S/D resistance ( $R_{S/D}$ ) implying a design tradeoff. Our model for  $C_f$  will be helpful in engineering this design trade-off for optimal speed performance.

In Chapter 5, we discussed the problem of S/D process design for nanoscale FinFET CMOS technology. FinFETs have undoped bodies; so, the S/D lateral doping profile,  $N_{SD}(y)$ , defines the tradeoff between  $L_{eSD}$  (which defines SCEs and  $C_f$ ) and  $R_{S/D}$  (which defines on-state current,  $I_{on}$ ). For optimal S/D process design, we developed and demonstrated a methodology to reverse-engineer  $N_{SD}(y)$  from FinFET  $C_G$ - $V_{GS}$  and  $I_{DS}$ - $V_{GS}$  characteristics. Using the extracted  $N_{SD}(y)$ , we provided insights on redesigning the device so as to obtain a better tradeoff between SCEs and  $I_{on}$ .

In Chapter 6, we studied the problem of mobility degradation for short- $L_g$  FinFETs. We observed that in the  $L_g = 32\text{nm}$  device data obtained for Sematech, the mobility degradation is responsible for ~40% and ~30% degradation in  $I_{\text{on}}$  for n- and p-devices, respectively. We attained insights into this phenomenon via a thorough literature survey as well as UFDG calibration to device data at different temperatures. Based on these insights, we discussed possible causes of the problem, which included high divacancy concentration related to high-temperature S/D processing, ion-implantation damage at the edges of the channel, and Coulomb scattering due to interface states caused by nitrogen diffusion. We also proposed solutions to help diagnose and possibly mitigate the effect. We then compared a pragmatic FinFET technology based on our  $L_g = 32\text{nm}$  device data (assuming ideal mobility) to state-of-the-art  $L_g = 30\text{nm}$  bulk technology. Based on the comparison, we discussed technological challenges for FinFET CMOS to eventually replace bulk CMOS.

## 7-2 Future Work

In Chapter 4, we developed a model for  $C_f$  in symmetrical DG MOSFETs. The model should be generalized to include asymmetrical DG and FD/SOI MOSFETs.

In Chapter 5, we outlined a methodology to extract  $N_{\text{SD}}(y) = N_{\text{SD0}}\exp(-y^2/\sigma_L^2)$  from device  $C_G$ - $V_{\text{GS}}$  and  $I_{\text{DS}}$ - $V_{\text{GS}}$  data. A comprehensive study on S/D processing is required to study the variation of device characteristics to different spacer widths ( $L_{\text{ext}}$ ) as well as different anneal conditions (anneal temperature, time and type). To study the variation of device characteristics with  $L_{\text{ext}}$ , devices should be fabricated with different  $L_{\text{ext}}$  but using the same S/D anneal conditions (so they have approximately the same  $\sigma_L$ ). Depending on  $L_{\text{ext}}$ , some of these devices will have underlap, others will have overlap. With the methodology in Chapter 5, the S/D doping straggle,  $\sigma_L$ , can be obtained for each device. The average value of  $\sigma_L$  can be used to plot, via Taurus simulations,  $L_{\text{eSD}}$  as a function of  $L_{\text{ext}}$ . In the case of overlapped devices,  $R_{\text{S/D}} \cong R_{\text{con}}$ ,

where  $R_{\text{con}}$  is the contact resistance. In the case of underlapped devices,  $R_{\text{S/D}} = R_{\text{con}} + R_{\text{ext}}(L_{\text{ext}})$ , where  $R_{\text{ext}}$  is the extension resistance as a function of  $L_{\text{ext}}$ .  $R_{\text{con}}$  can be obtained from overlapped devices, which can then be used to obtain  $R_{\text{ext}}$  as a function of  $L_{\text{ext}}$  for underlapped devices.  $R_{\text{con}}$ ,  $R_{\text{ext}}(L_{\text{ext}})$ , and  $L_{\text{eSD}}(L_{\text{ext}})$  can be used to calibrate UFDG and perform RO simulations to design  $L_{\text{ext}}$  for optimal circuit performance. Fig. 7-1 summarizes the methodology which would adequately characterize the S/D process. This methodology should be repeated for different S/D anneal cycles with variations in S/D anneal temperature/time to characterize the dependence of anneal conditions on  $R_{\text{S/D}}$  and  $L_{\text{eSD}}$ . It will be interesting also to see how the mobility,  $\mu_{\text{eff}}$ , varies with  $L_{\text{ext}}$  and as a function of S/D anneal conditions.

In Chapter 6, we made conjectures into the possible causes of mobility degradation in short- $L_g$  FinFETs. We also offered processing suggestions to diagnose and possibly mitigate the effect of mobility degradation in short- $L_g$  FinFETs. We believe that these processing suggestions should be tried in a comprehensive study of mobility in short- $L_g$  FinFETs. To test the theory that  $\mu_{\text{eff}}$  degradation is related to divacancies caused by high-temperature processing, we propose that the devices be subject to a long (~30 min), relatively low temperature (350-400° C) anneal after the S/D processing. To test the theory that the  $\mu_{\text{eff}}$  degradation is caused by ion-implantation damage, we propose that the S/D implant dose be reduced and the implant be carried out with different energies to ensure relatively uniform distribution of S/D dopants along the vertical height of the fin. To check the possibility of scattering by interface states caused by nitrogen diffusion, we propose that devices be fabricated with  $\text{SiO}_2$ /poly gate stack. If substantial improvement in  $\mu_{\text{eff}}$  is observed, the TiN gate should be replaced by a different material, or the process should be varied to prevent N diffusion.

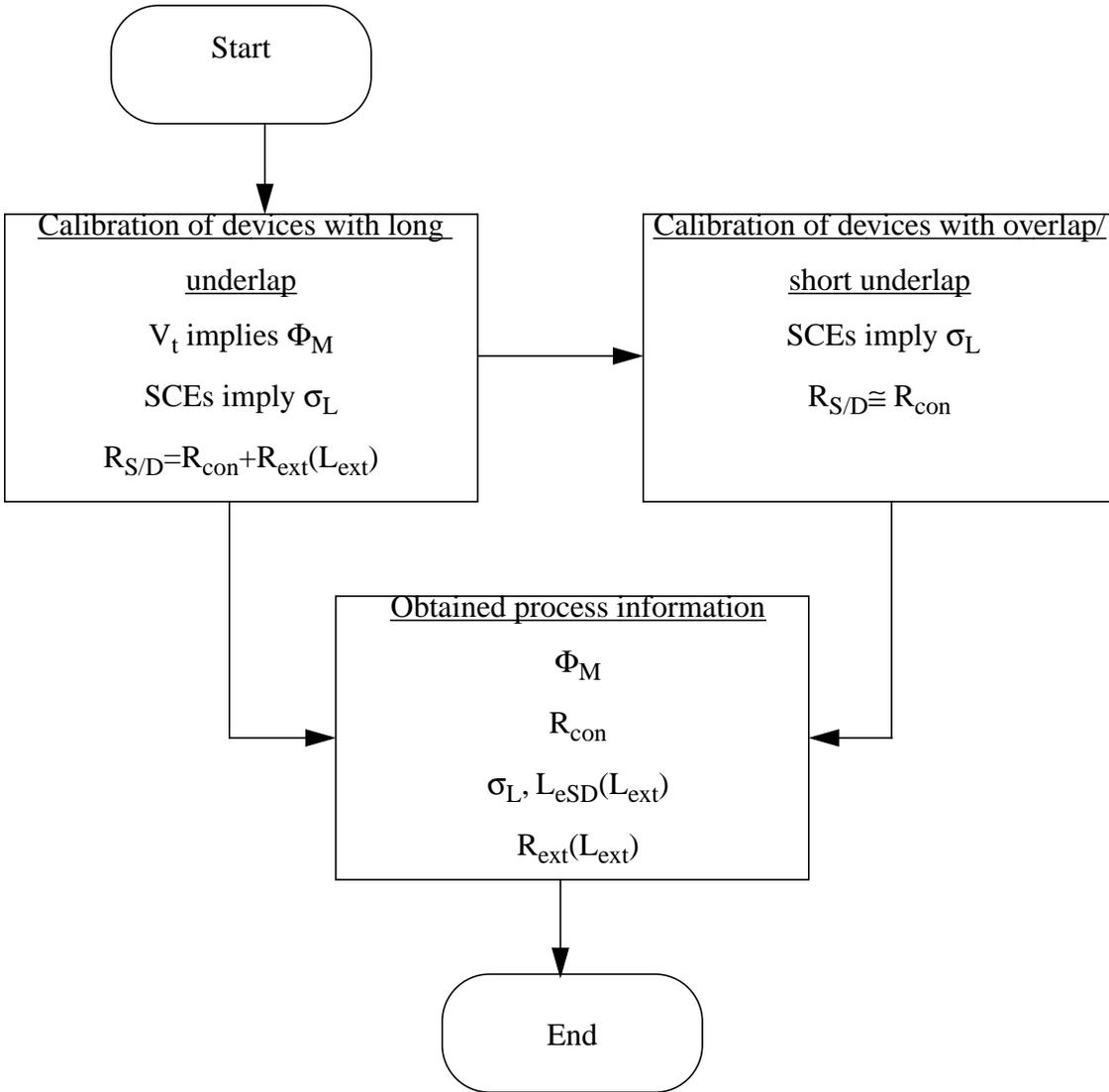


Figure 7-1. This flowchart describes a potential future study to optimize the S/D process using devices with different spacer width,  $L_{ext}$ .

## APPENDIX A USING TAURUS AND UFDG FOR UTB DEVICE SIMULATION

### A-1 Introduction

This appendix discusses how we use the Taurus device simulator [9] and our compact model, UFDG [10], in conjunction with each other for performance projections of ultra-thin-body (UTB) devices. Both UFDG and Taurus have unique strengths which are extremely useful in device modeling. Taurus has the advantage of the superior 2-D electrostatics modeling typical of device simulation programs. However, the quantization and transport models in Taurus are not well-suited to UTB devices. Also, the modeling of the carrier velocity-saturation effect in Taurus can introduce inadvertent errors in I-V projections as we will see later (This has been fixed in the newer version of Taurus ( Sentaurus [55]) as will be discussed). On the other hand, UFDG has physical, well-calibrated models for mobility [32], velocity overshoot [34], ballistic-limit current [4], quantum-mechanical threshold shift [19], and inversion layer capacitance [33].

To take advantage of the superior electrostatics of Taurus as well as the physical transport and QM models of UFDG, we often compare I-V characteristics predicted by them. When comparing UFDG- and Taurus-predicted characteristics, we turn off QM ( $QMX = QMD = 0$ ) and transport ( $VO = 0, BLIM = 0$ ) models in UFDG. We also use comparable mobility models in both UFDG and Taurus with the same electric field ( $E_x, E_y$ ) dependence. Here,  $E_x$  refers to the transverse field perpendicular to the current direction and  $E_y$  refers to the longitudinal field in the current direction. Section A-2 briefly describes the relevant Taurus and UFDG mobility models and how to relate them to each other. Section A-3 and A-4 describe how the models are used in weak and strong inversion, respectively.

## A-2 A Brief Description of Taurus and UFDG Mobility Models

Taurus offers several low-field mobility models which model mobility as a function of transverse field,  $E_x$ , doping,  $N_D$  and temperature  $T$  (i.e.,  $\mu_{\text{eff}}(E_x, N_D, T)$ ). (Note that the low-field here implies low longitudinal field,  $E_y$  not transverse field,  $E_x$ ). None of these are particularly well suited to UTB devices. Thus, we rely on the default low-field mobility model which assumes no  $E_x$ ,  $N_D$  or  $T$  dependence. The default low-field mobility can be specified by using  $MUN0$  ( $\mu_{N0}$ ) and  $MUP0$  ( $\mu_{P0}$ ) parameters for electron and hole mobility respectively. Let us assume that we specify  $MUN0 = 500\text{cm}^2/\text{Vs}$ . In addition, there are two options to specify longitudinal field ( $E_y$ ) dependence. The first one is the “constant mobility” model which assumes no  $\mu_{\text{eff}}(E_y)$  dependence. Thus, it does not model the velocity saturation effect for high  $V_{DS}$ . For the rest of this appendix, we refer to this model as the “constant mobility” model. This model is specified in the Taurus model card as follows:

```
Constant=True MUN0=500
```

The other high- $E_y$  model offered by Taurus is the CaugheyThomas model which can be tuned to model the  $E_y$  dependence the same way as UFDG. This model specifies  $\mu_{\text{eff}}$  as follows:  $\mu_{\text{eff}} = \mu_{N0} / [1 + (v_{\text{sat}}/\mu_{N0}E_y)^\beta]^{1/\beta}$ . To tune this model to UFDG, we specify  $\beta = 1$ .  $v_{\text{sat}}$  can also be set according to our convenience. The default value is  $v_{\text{sat}} = 1 \times 10^7 \text{cm/s}$ . The model code is as follows.

```
MUN0=500
HighFieldMobilityActive=True,
HighFieldMobility (
HighFieldModel=CaugheyThomasModel,
heatingField=EdgeField,
vsatModel=constantvsatmodel,
```

```

CaugheyThomasModel(betan=1),
constantvsatmodel(vsatn=7e6)
)

```

Through the rest of this appendix, this model is referred to as the “velocity saturation” model for convenience. As described in the following sections, we need to use both the “constant mobility” model and the “velocity saturation” model selectively to accurately tune Taurus-predicted characteristics to UFDG.

The mobility model in UFDG is QM-based [32]; when the QM model is turned off (QMX=QMD=0), the mobility model reverts to the simple model specified in [41]. Both these mobility models have two parameters:  $UO$ , which is  $\mu_{\text{eff}}$  in a thick body ( $t_{\text{Si}} \rightarrow \infty$  and  $E_x = 0$ ) and  $\Theta$ , which is the surface roughness tuning parameter. When tuning UFDG-predicted characteristics to Taurus, we always turn the QM model off as discussed in Sec. A-1. Specifying  $\Theta = 0$  removes the  $E_x$  dependence of  $\mu_{\text{eff}}$ .  $\mu_{\text{eff}}$  is then a function of  $UO$  and  $t_{\text{Si}}$ . We then choose  $UO$  so that it gives us  $\mu_{\text{eff}} = 500 \text{cm}^2/\text{Vs}$  (as specified in Taurus model card). We achieve this by printing out  $\mu_{\text{eff}}$  from the UFDG code for different values of  $UO$ .

### **A-3 Calibration of Weak-Inversion I-V Characteristics**

We first describe calibration of weak-inversion characteristics of an abrupt junction device predicted by Taurus and UFDG. We assume a nominal DG MOSFET structure with abrupt junction,  $L_g = 25 \text{nm}$ , UTB width  $t_{\text{Si}} = 10 \text{nm}$ , gate oxide thickness  $t_{\text{ox}} = 1.2 \text{nm}$  and midgap gates. Since weak-inversion current is diffusion limited,  $\mu_{\text{eff}}(E_y)$  should not have a significant effect on weak-inversion I-V characteristics. Hence, it is surprising that Taurus predicts very different I-V characteristics when using the “constant mobility” and “velocity saturation” models (Fig. A-1). On closer inspection, we find that when using the “velocity saturation” model, Taurus predicts a 4x reduction in carrier mobility because of the junction electric field between the source and

channel. Fig. A-2 shows the electric potential as a function of position with  $y = 0$  corresponding to the center of the channel. If we consider the diffusion length to be  $\sim 5\text{nm}$ ,  $E_y$  at the source end of the diffusion length is  $6 \times 10^4 \text{ V/cm}$ . Plugging  $E_y$  into the expression for  $\mu_{\text{eff}}(E_y)$  we calculate a 4x reduction in mobility. This is an invalid effect because the electric field in the source to channel junction is a built-in equilibrium field which does not heat carriers; thus it does not lead to  $\mu_{\text{eff}}$  reduction. We note, therefore, that calibration of UFDG with Taurus using the Taurus “velocity saturation” model will give erroneous results. Ideally, if the gradient of the electron quasi-Fermi potential,  $dF_n/dx$ , is used (instead of  $E_y$ ) to define  $\mu_{\text{eff}}$ , the above-mentioned effect will be eliminated. Since Taurus does not provide the user with this option, we use the “constant mobility” model in weak inversion to remove this dubious  $E_y$  dependence. A good match between UFDG and Taurus-predicted characteristics is obtained as observed in Fig. A-3. (The recent upgraded version of Taurus, Sentauros [55], provides the user with the option of using  $dF_n/dx$  to define  $\mu_{\text{eff}}$ . Thus, device simulations using Sentauros should use the “velocity saturation” model with  $\mu_{\text{eff}}$  defined by  $dF_n/dx$  instead of  $E_y$ ).

With the “constant mobility” model turned on, Taurus is very useful in predicting weak-inversion I-V characteristics for devices with a non-abrupt source/drain (S/D) doping profile. This is due to the superior electrostatics modeling which is typical of device simulation tools. UFDG does not model device characteristics in terms of their S/D doping profile; instead it uses a tunable effective underlap,  $L_{eS}/L_{eD}$ . To tune UFDG to Taurus-predicted I-V characteristics of devices with non-abrupt S/D junctions, we tune the mobility models as discussed above and additionally tune  $L_{eS}/L_{eD}$ .

#### **A-4 Calibration of Strong-Inversion I-V Characteristics**

As discussed above, since QM effects like inversion-layer capacitance [33] and transport effects like  $\mu_{\text{eff}}$  [32], velocity overshoot [34] and ballistic-limited current [4] are not well modeled

in Taurus, it is not possible to obtain realistic I-V projections in strong inversion. Typically, we use UFDG with all physical models turned on for quasi-predictive I-V projections in strong-inversion. However, we still use Taurus to predict *relative* currents for different device structures which cannot be reliably modeled using UFDG (Chapter 2). The “velocity saturation” mobility model is used inspite of some error introduced due to the invalid  $\mu_{\text{eff}}$  degradation effect discussed above. We note that this error is reduced in strong inversion for abrupt S/D junctions because the potential barrier between the S/D and the channel is greatly reduced (Fig. A-4). Figs. A-5 and A-6 show the match between Taurus and UFDG-predicted strong-inversion characteristics with  $R_{\text{S/D}}$  tuned to  $25\Omega\text{-}\mu\text{m}$  in UFDG. For devices with non-abrupt S/D junctions, there is still a significant potential barrier in the S/D extension (Fig. A-7) which amounts to an invalid  $R_{\text{S/D}}$  component. However, we believe that trends predicted by Taurus for different device structures with the same S/D doping profile,  $N_{\text{SD}}(y)$ , would still be meaningful since they would have the same (dubious)  $R_{\text{S/D}}$ . As mentioned in Sec. A-3, if Sentaurus [55] is used for device simulations, the invalid  $R_{\text{S/D}}$  component can be eliminated by using  $dF_n/dx$  (instead of  $E_y$ ) to define  $\mu_{\text{eff}}$ . While predicting strong-inversion currents, we use realistic values for  $MUN0$ ,  $MUP0$  based on UFDG’s mobility model predictions. For example, typical  $\mu_{\text{eff}}$  parameters for n-channel DG MOSFETs ( $UO = 1100\text{cm}^2/\text{V}\cdot\text{s}$  and  $\Theta = 0.8$ ) predicts  $\mu_{\text{eff}} \sim 300\text{cm}^2/\text{V}\cdot\text{s}$  for an undoped DG MOSFET with midgap gates at  $V_{\text{GS}} = 1\text{V}$ .

In summary, it is possible to predict trends in strong inversion between different device structures using Taurus as long as the S/D doping profile is the same for all given structures.  $MUN0$ ,  $MUP0$  should be chosen to be realistic and the “velocity saturation” model should be turned on. Table A-1 summarizes our findings.

### A-5 Additional Note

When the doping profile is assumed to be gaussian, Taurus assumes the lateral straggle  $\sigma$  as is given in the equation,  $N_{SD}(y) = N_{SD0}\exp(-y^2/2\sigma^2)$ . Medici [56], however, assumes the  $\sigma$  as is given in the equation,  $N_{SD}(y)=N_{SD0}\exp(-y^2/\sigma^2)$  (as is assumed in our work). Thus,  $\sigma$ (as given by Medici) = $\sqrt{2}$ \* $\sigma$  (as given by Taurus).

Table A-1. A summary of the mobility models to be used for Taurus predictions.

Type of Characteristic	Mobility Model
Weak inversion $I_{DS}-V_{GS}$ characteristics	Use “constant mobility” model, i.e. constant mobility with velocity saturation turned off
Strong inversion $I_{DS}-V_{GS}$ and $I_{DS}-V_{DS}$ characteristics	Use “velocity saturation” model i.e. constant mobility with velocity saturation turned on

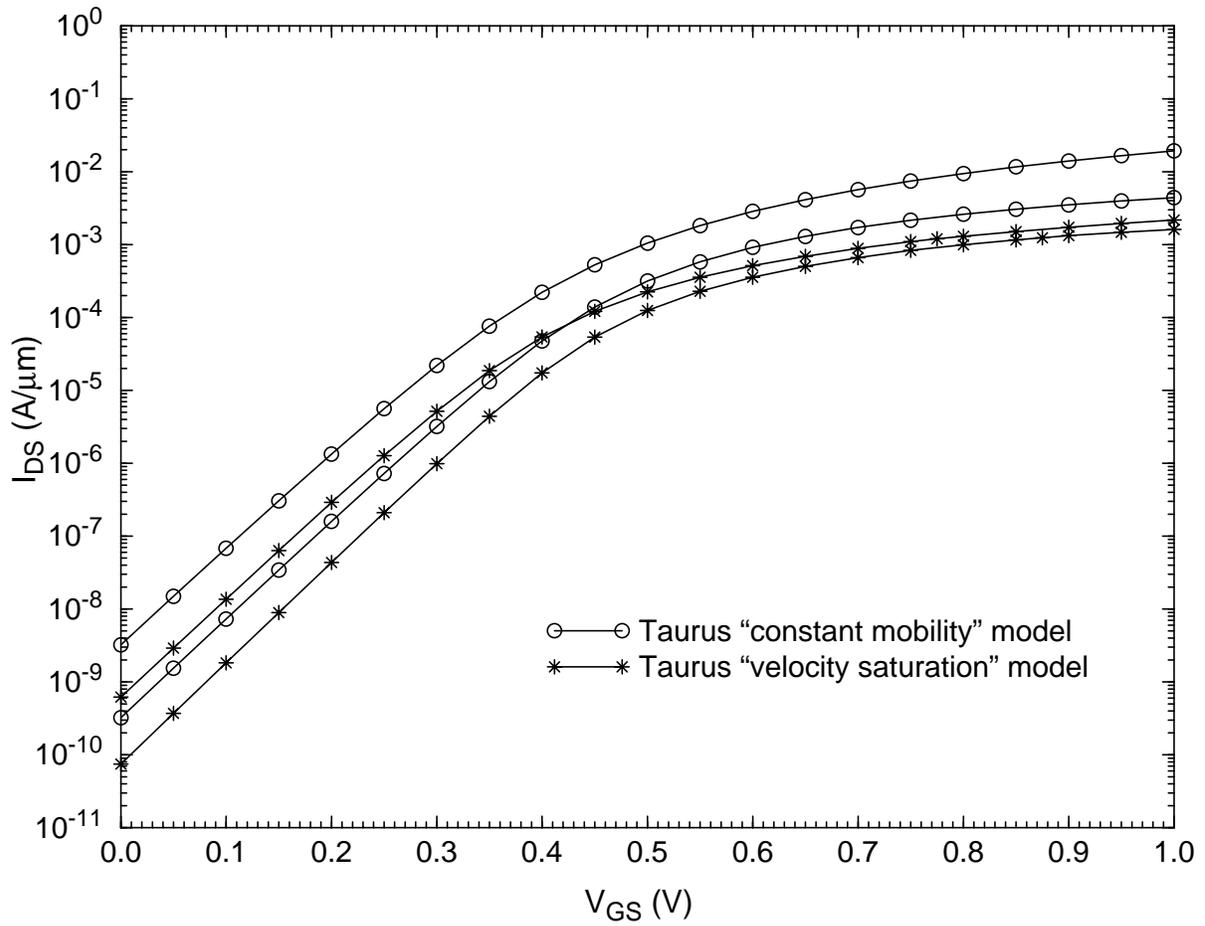


Figure A-1. A comparison of the Taurus-predicted  $I_{DS}$ - $V_{GS}$  characteristics using “constant mobility” and “velocity saturation” models for an abrupt DG MOSFET ( $L_g = 25\text{nm}$ ,  $w_{Si} = 10\text{nm}$ ,  $t_{ox} = 1.2\text{nm}$ , midgap gates).

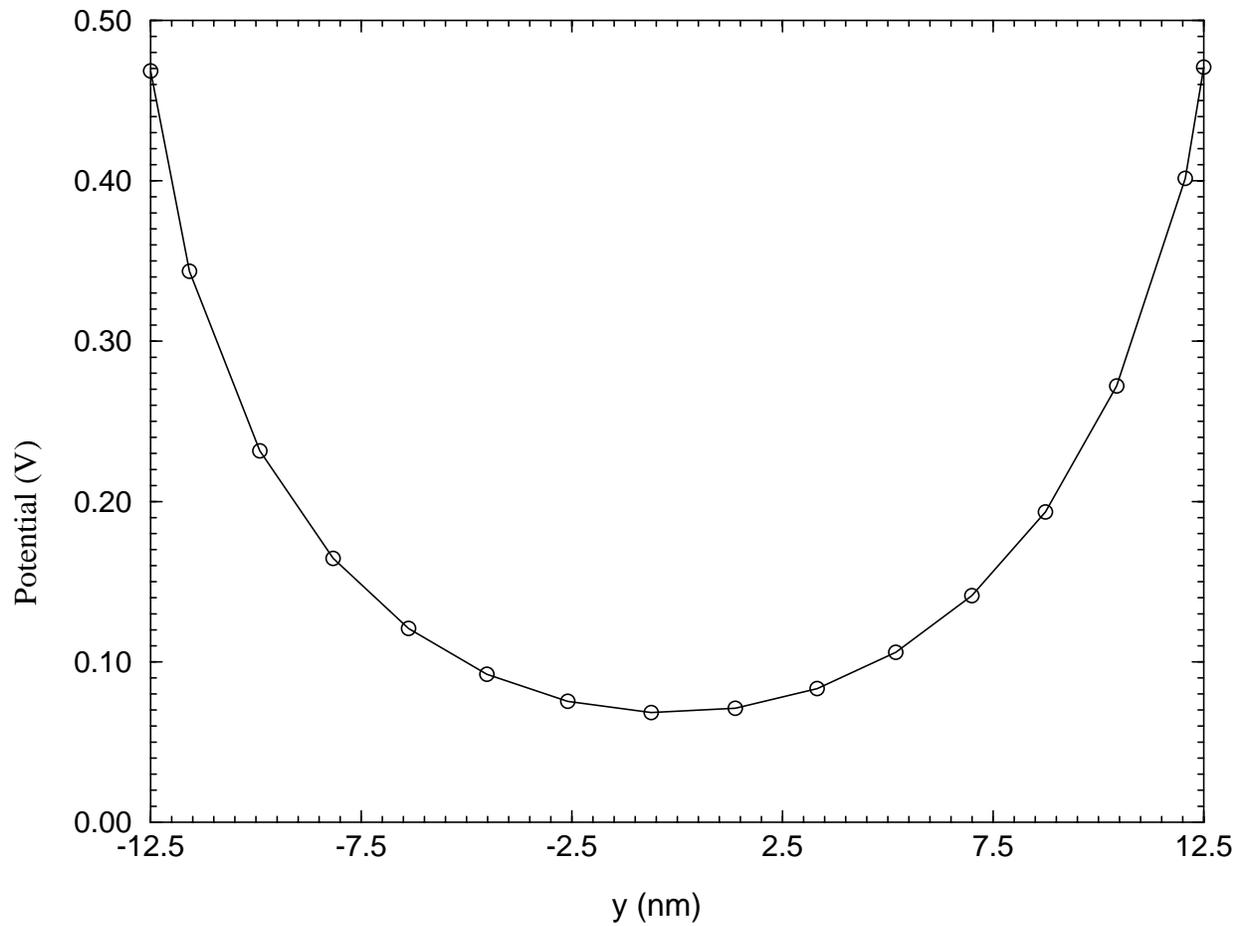


Figure A-2. The electric potential as a function of distance along the center of the channel of an abrupt junction DG MOSFET ( $L_g = 25\text{nm}$ ,  $w_{\text{Si}} = 10\text{nm}$ ,  $t_{\text{ox}} = 1.2\text{nm}$ , midgap gates). The bias conditions used are  $V_{\text{GS}} = 0\text{V}$ ,  $V_{\text{DS}} = 100\text{mV}$ .  $y = 0$  corresponds to the center of the channel.

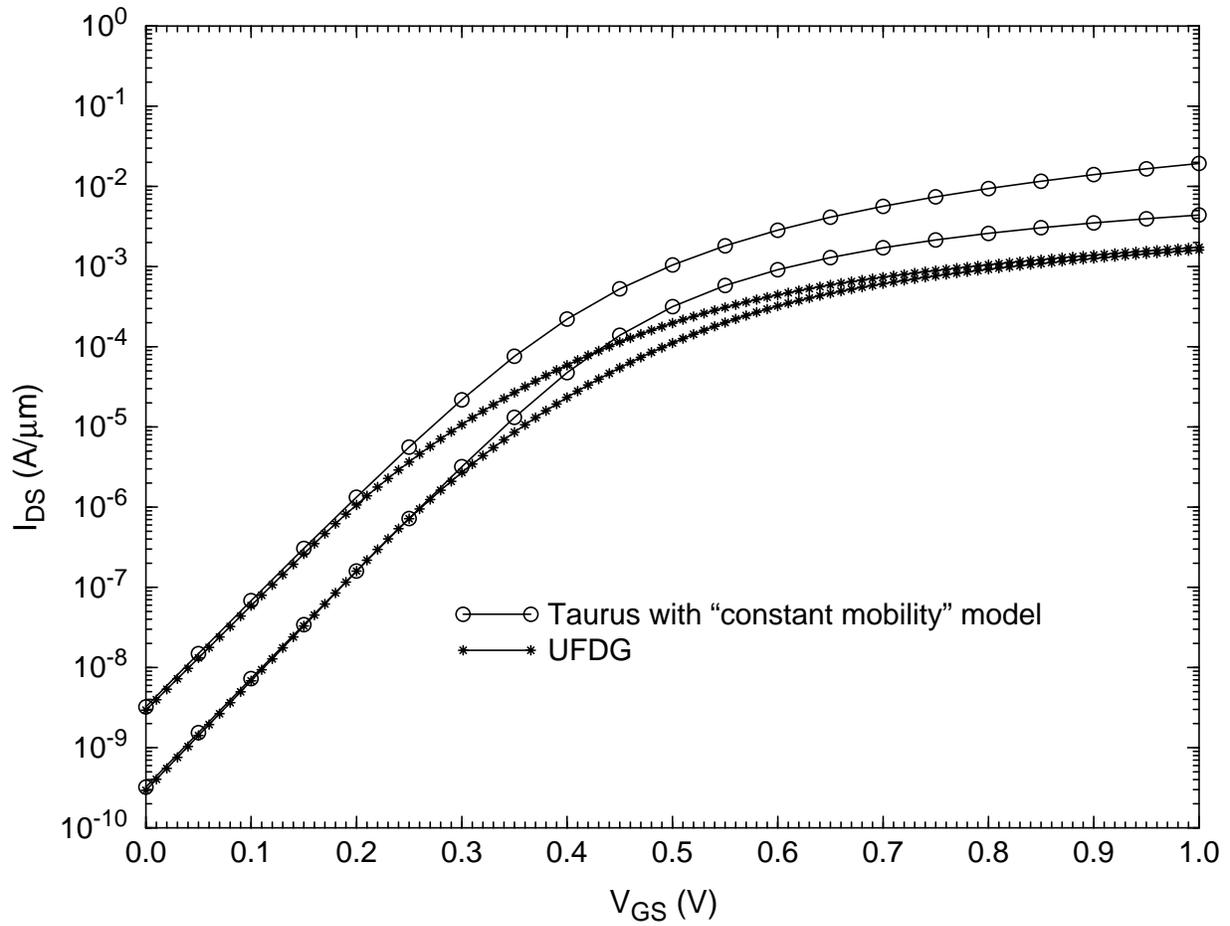


Figure A-3. Calibration between Taurus and UFDG-predicted weak-inversion  $I_{DS}$ - $V_{GS}$  characteristics of an abrupt junction DG MOSFET using the “constant mobility” model. The device used has  $L_g = 25\text{nm}$ ,  $w_{Si} = 10\text{nm}$ ,  $t_{ox} = 1.2\text{nm}$ , abrupt S/D junctions, and mid-gap gates.

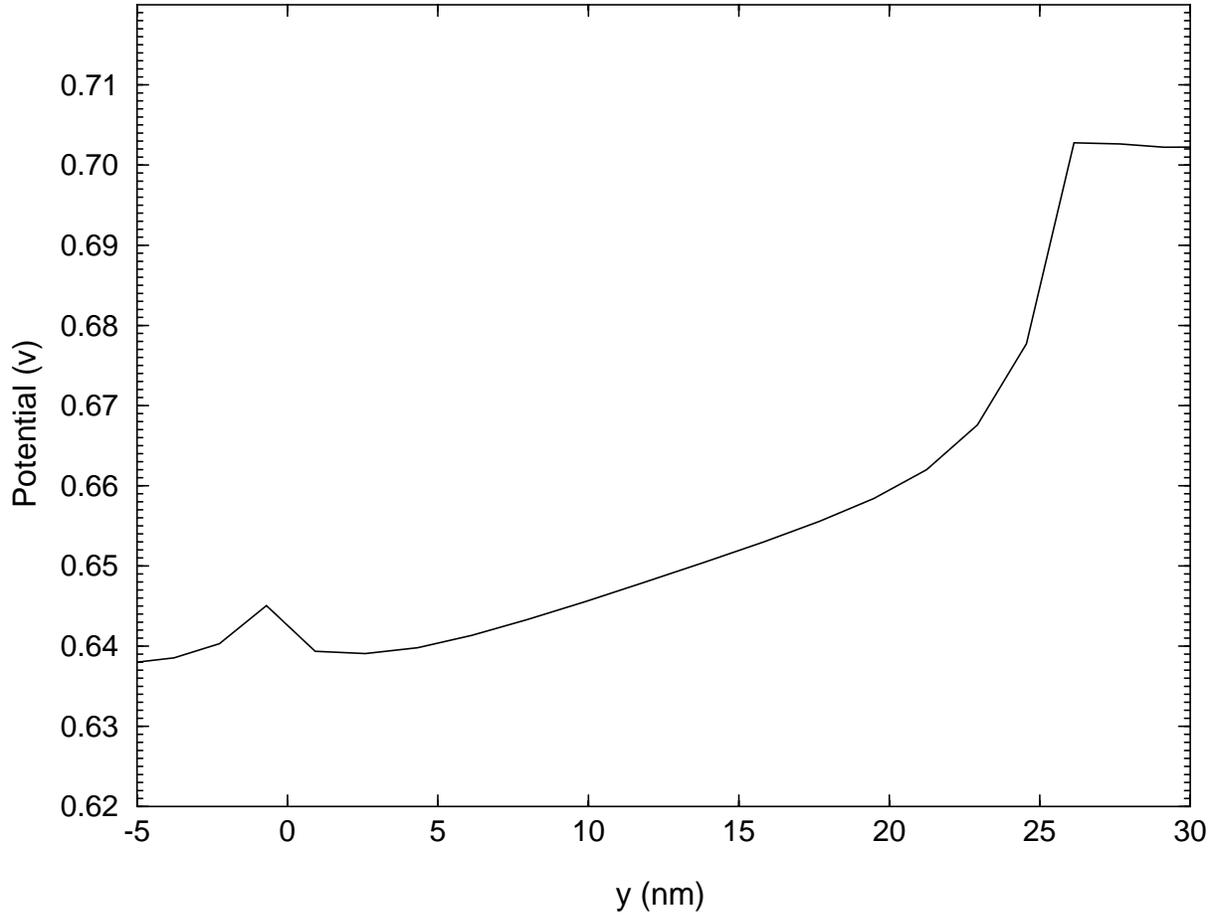


Figure A-4. The electric potential as a function of distance along the center of the channel of an abrupt junction DG MOSFET ( $L_g = 25\text{nm}$ ,  $w_{\text{Si}} = 10\text{nm}$ ,  $t_{\text{ox}} = 1.2\text{nm}$ , midgap gates). The bias conditions used are  $V_{\text{GS}} = 1\text{V}$ ,  $V_{\text{DS}} = 100\text{mV}$ .  $y = 0$  is at the edge of the gate

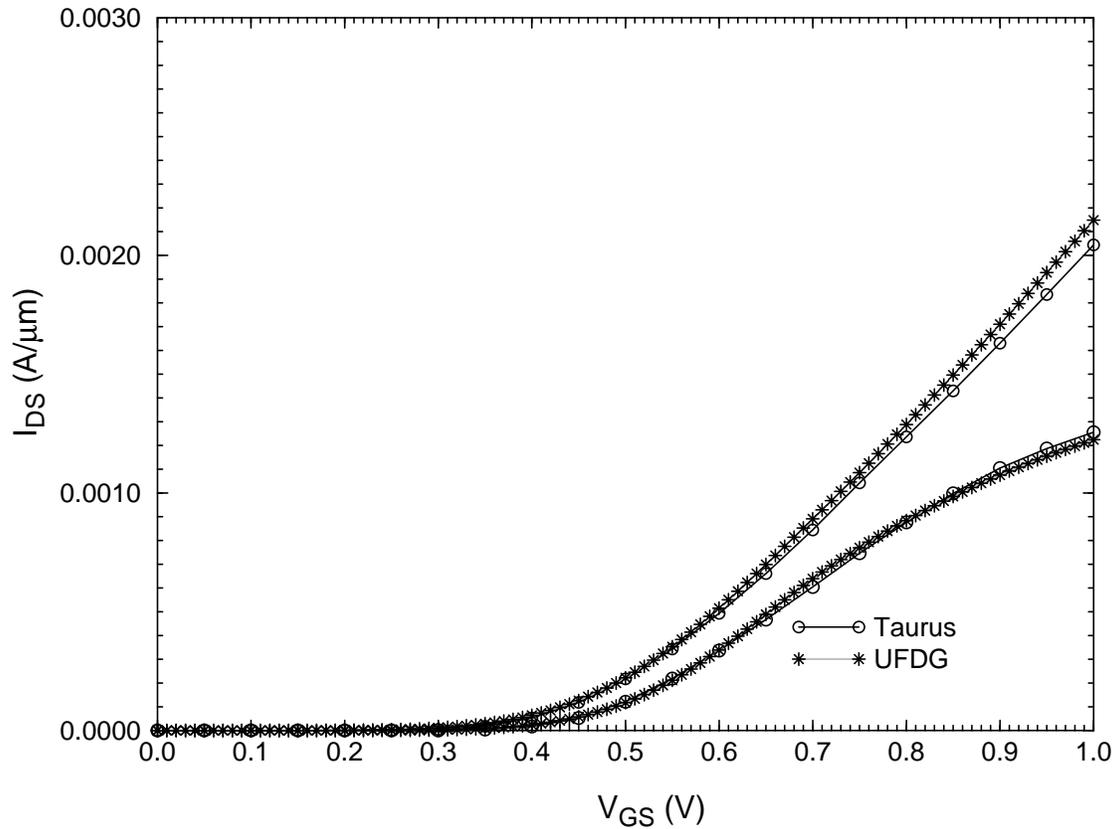


Figure A-5. Calibration between Taurus and UFDG-predicted  $I_{DS}$ - $V_{GS}$  strong-inversion characteristics of an abrupt junction DG MOSFET using the “velocity saturation” model. The S/D resistance is tuned to  $R_{S/D} = 25\Omega\text{-}\mu\text{m}$  in UFDG. The device used has  $L_g = 25\text{nm}$ ,  $w_{Si} = 10\text{nm}$ ,  $t_{ox} = 1.2\text{nm}$ , midgap gates.

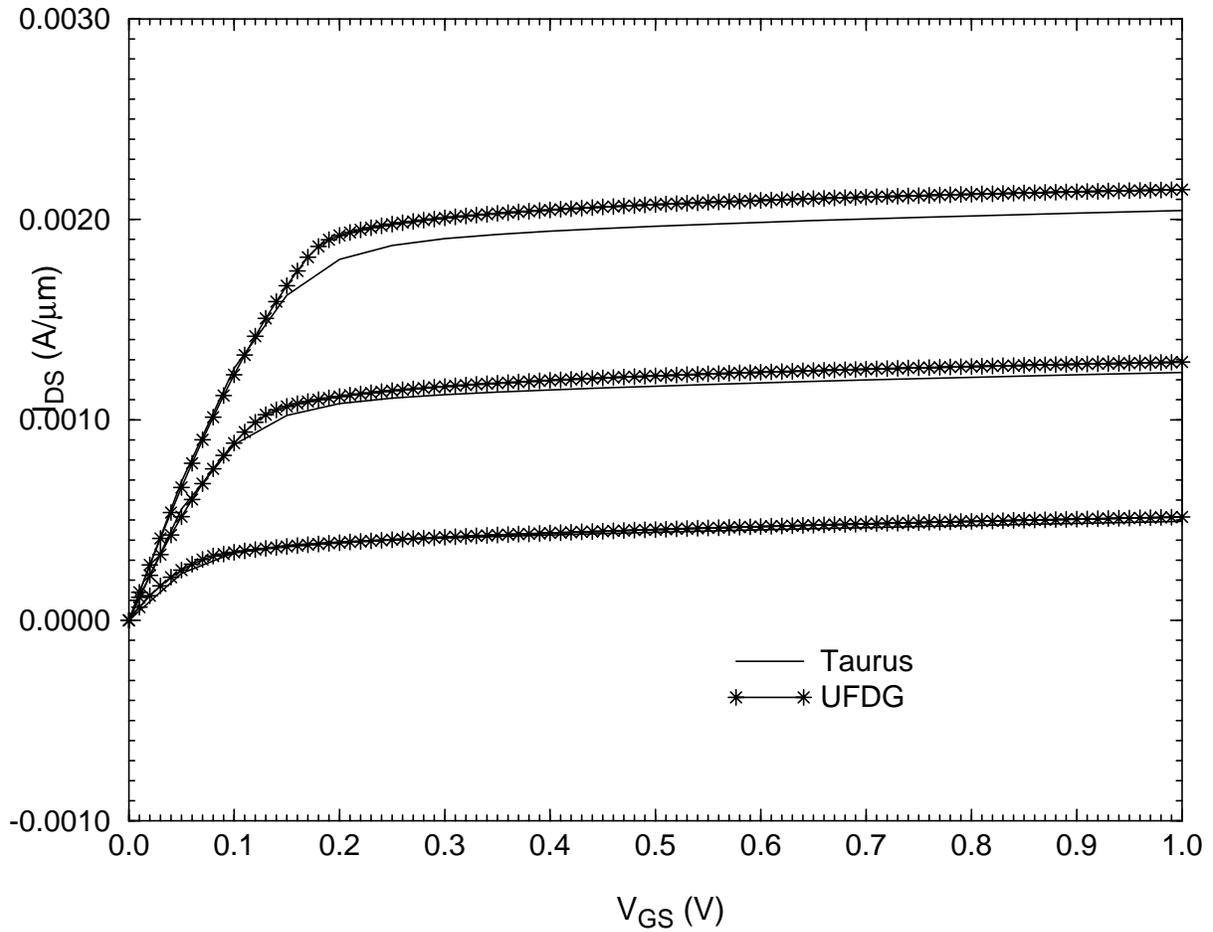


Figure A-6. Calibration between Taurus and UFDG-predicted  $I_{DS}$ - $V_{DS}$  strong-inversion characteristics of an abrupt junction device using the “velocity saturation” mobility model. The S/D resistance is tuned to  $R_{S/D} = 25\Omega\text{-}\mu\text{m}$  in UFDG. The device used has  $L_g = 25\text{nm}$ ,  $w_{Si} = 10\text{nm}$ ,  $t_{ox} = 1.2\text{nm}$ , abrupt S/D junctions, and midgap gates.

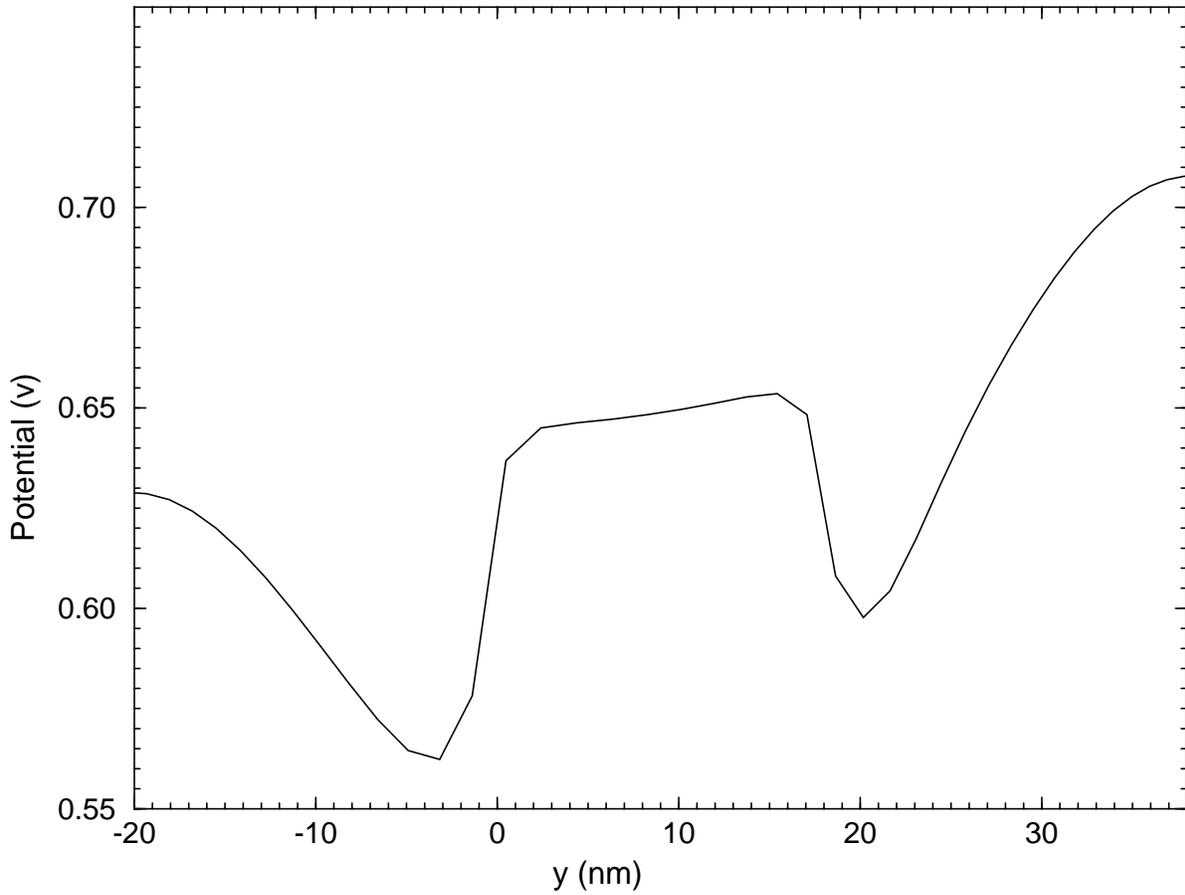


Figure A-7. The electric potential as a function of distance along the center of the channel of a non-abrupt junction DG MOSFET ( $L_g = 18\text{nm}$ ,  $w_{\text{Si}} = 10\text{nm}$ ,  $t_{\text{ox}} = 1.2\text{nm}$ ,  $L_{\text{ext}} = 20\text{nm}$ ,  $\sigma = 10.7\text{nm}$ , midgap gates). The point  $y = 0$  is at the edge of the gate.

## APPENDIX B UFDG CODE REVISIONS FOR HIGH-K DEVICE SIMULATION

### **B-1 Introduction**

In Chapter 3, we discussed the suitability of high-k gate dielectrics in nanoscale FinFET CMOS technology. To evaluate the performance of FinFETs with a high-k gate dielectric, we performed quasi-predictive simulations of these devices using UFDG [10]. However, UFDG implicitly assumes that the gate dielectric material is SiO<sub>2</sub>. This Appendix discusses the UFDG code changes that were made to account for the high-k dielectric material. We note that these code revisions are not permanent; they were done only for the purpose of this work. The revisions included defining a new constant to account for the gate dielectric material (Sec. B-2) and changes in the parasitic fringe capacitance model (Sec. B-3).

### **B-2 Code Revision to Account for Gate Dielectric Material**

As of version 3.7, UFDG assumes that both the gate dielectric and the spacer material are SiO<sub>2</sub>. In the UFDG header file a constant, OXIDE\_PERMITTIVITY, is defined and its value is set to  $3.9\epsilon_0$  (corresponding to SiO<sub>2</sub>), where  $\epsilon_0$  is the permittivity of vacuum. This constant is used both for the gate dielectric permittivity as well as the spacer permittivity. In our high-k device simulation, we assume that the gate dielectric has  $k=25$  (corresponding approximately to hafnium dioxide) and the spacer material is SiO<sub>2</sub> ( $k=3.9$ ). Thus, we set the value of OXIDE\_PERMITTIVITY to  $25\epsilon_0$ , corresponding to the high-k gate dielectric material. We also define a new constant for the spacer permittivity. We call it SPACER\_PERMITTIVITY and set the value to  $3.9\epsilon_0$ . The new variable, SPACER\_PERMITTIVITY, is used in the parasitic capacitance model of UFDG to define the outer fringe capacitance ( $C_{of}$ ).

### B-3 Code Revisions for Fringe Capacitance Model

The basic fringe-capacitance model [24] in UFDG stems from that defined by two separated conducting plates at an angle  $\theta$  as shown in Fig. B-1. The capacitance per unit width of this system is given by [57]

$$C = \frac{\epsilon}{\theta} \ln\left(\frac{r_2}{r_1}\right) \quad (\text{B-1})$$

where  $\epsilon$  is the permittivity and  $\theta$ ,  $r_1$  and  $r_2$  are the geometrical parameters defined in Fig. B-1. The underlap structure in weak inversion is approximated by an abrupt junction at distance  $L_{eSD}$  from the gate edge.  $L_{eSD}$  is defined by the lateral doping profile,  $N_{SD}(y)$ , and is a function of the source/drain extension length,  $L_{ext}$ , and doping straggle,  $\sigma_L$  (Chapter 4). The system is shown in Fig. B-2. The model for  $C_{if}$  reduces the plate-plate angle from  $\pi/2$  to  $\beta$  (Fig. B-3), to effectively account for the silicon permittivity ( $\epsilon_{Si}$ ) being about three times that of the gate dielectric ( $\epsilon_{di}$ )(SiO<sub>2</sub> in this case).

$$\beta = \frac{\pi \epsilon_{di}}{2 \epsilon_{Si}} \cong \frac{\pi}{6} . \quad (\text{B-2})$$

However, when we use the high-k gate dielectric ( $\epsilon_{di}=25\epsilon_0$ ) instead of SiO<sub>2</sub>,  $\beta>\pi$ , and the construction in Fig. B-3 is not physical. To account for  $C_{if}$ , we use another construction, shown in Fig. B-4, where the thickness of the gate dielectric is scaled by  $\alpha$ , where  $\alpha=\epsilon_{di}/\epsilon_{Si}$ . This is done to define an effective system consisting of only one dielectric material (silicon), while keeping the capacitance of the gate dielectric constant. From Fig. B-4, we see that,

If  $L_{eSD}>t_{di}/\alpha$ ,  $C_{if}$  is given by,

$$C_{if} = \frac{2\epsilon_{Si}}{\pi} \log\left(\frac{t_{Si} + t_{di}/\alpha}{f_{if} L_{eSD}}\right) . \quad (\text{B-3})$$

If  $L_{eSD}<t_{di}/\alpha$ ,

$$C_{if} = \frac{2\epsilon_{Si}}{\pi} \log\left(\frac{t_{Si} + t_{di}/\alpha}{t_{di}/\alpha}\right) \quad (B-4)$$

In the above expressions,  $t_{di}$  is the thickness of the gate dielectric and  $t_{Si}$  is the thickness of the silicon fin.  $f_{if}$  is a tuning parameter used to account for uncertainty in  $L_{eSD}$  and is determined by matching model predictions with numerical simulations. To tune  $f_{if}$  for the high-k device in Chapter 3, we simulated the high-k device structure (Table 3-2) using Taurus [9] and tuned  $f_{if}=1.3$  in UFDG to match the weak-inversion fringe capacitance.

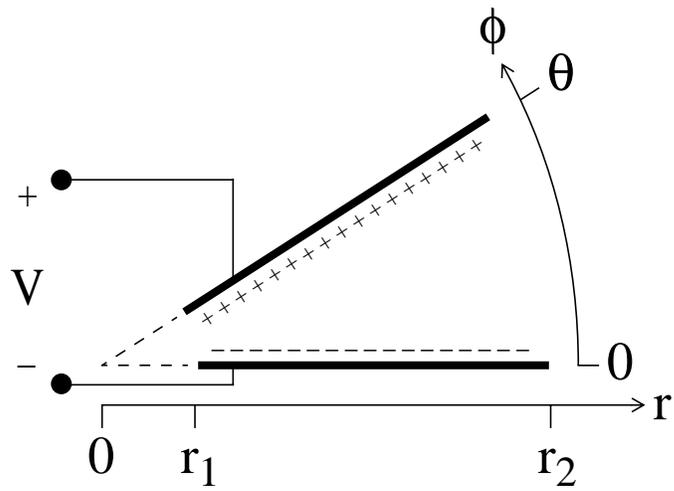


Figure B-1. Basic two-plate model for fringe capacitance (per unit width in  $z$ ), with the cylindrical coordinates ( $r$  and  $\phi$ ) used in the analysis shown.

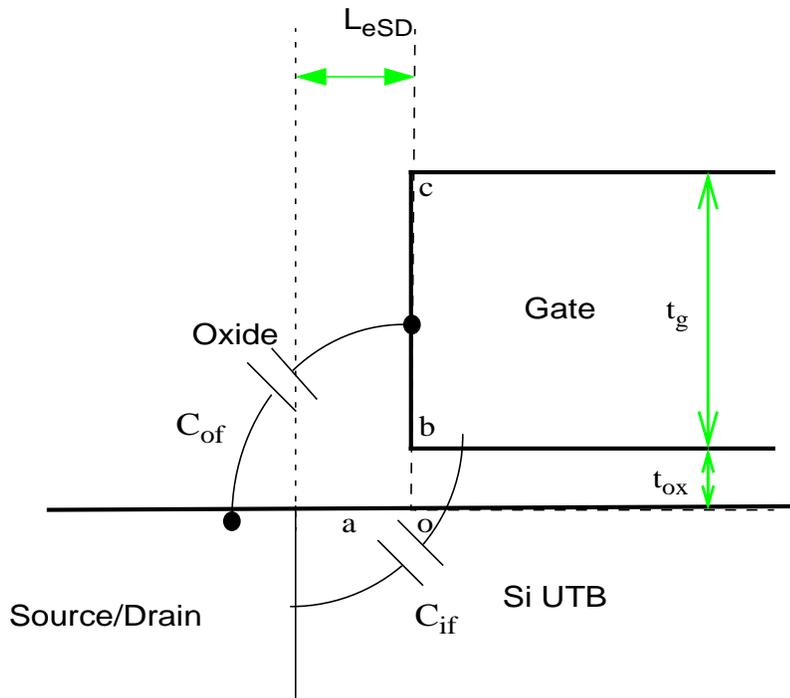


Figure B-2. A schematic diagram of the gate-source/drain structure of a DG MOSFET, indicating the G-S/D underlap (with effective length,  $L_{eSD}$ ) and the two components of the parasitic fringe capacitance.

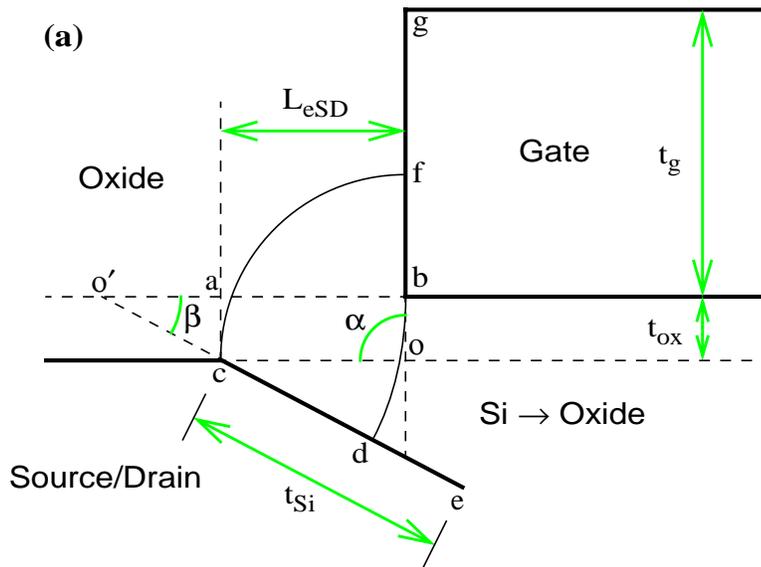


Figure B-3. Schematic of the G-S/D underlap structure used in derivation of  $C_{if}$  for device with  $SiO_2$  dielectric. The reduced angle  $\beta$  is defined after replacing the higher-permittivity silicon with oxide.

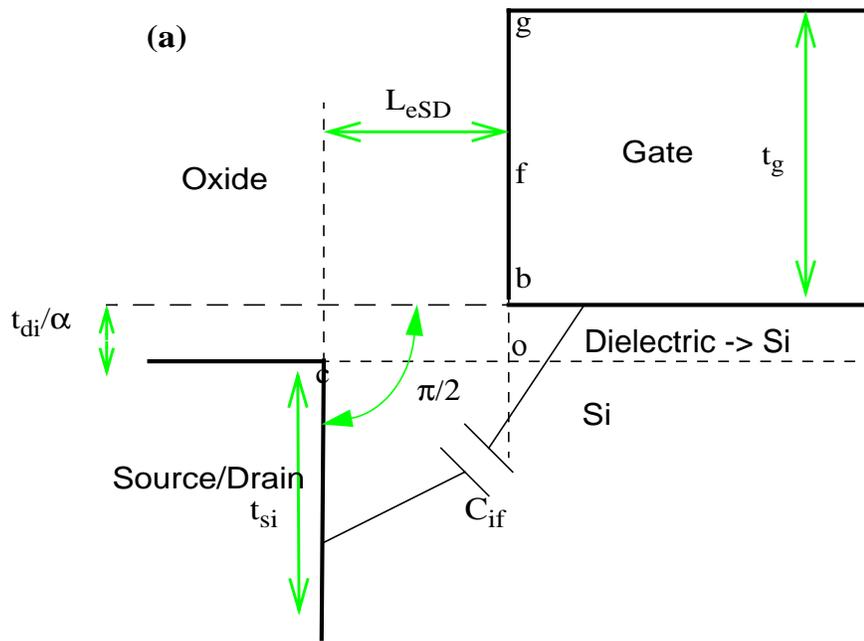


Figure B-4. The effective G-S/D underlap structure used in derivation of  $C_{if}$  for device with high- $k$  dielectric. The effective dielectric thickness is reduced to  $t_{di}/\alpha$  to account for permittivity of dielectric.

## LIST OF REFERENCES

- [1] *The International Technology Roadmap for Semiconductors*, 2005, Austin, TX: Semiconductor Industry Association.
- [2] J. G. Fossum, L. Q. Wang, J.-W. Yang, S.-H. Kim, and V. P. Trivedi, "Pragmatic design of nanoscale multi-gate CMOS," in *IEDM Tech. Dig.*, Dec. 2004, pp. 613-616.
- [3] V. P. Trivedi, J. G. Fossum, and W. Zhang, "Threshold voltage and bulk inversion effects in nonclassical CMOS devices with undoped ultra-thin bodies," *Solid State Electronics Journal*, vol. 26, no. 8, pp. 170-178, Jan. 2007.
- [4] M. M. Chowdhury, V. P. Trivedi, J. G. Fossum, and L. Mathew, "Carrier mobility/transport in undoped-UTB DG FinFETs," *IEEE Trans. Electron Devices*, vol. 54, no. 1, pp. 1125-1131, May 2007.
- [5] D. Hisamoto *et al.*, "A folded-channel MOSFET for deep-sub-tenth micron era," in *IEDM Tech. Dig.*, Dec. 1998, pp. 1032-1034.
- [6] X. Huang *et al.*, "Sub 50-nm FinFET: PMOS," in *IEDM Tech. Dig.*, Dec. 1999, pp. 67-70.
- [7] W. Zhang, J. G. Fossum, and L. Mathew, "The ITFET: A novel FinFET-based hybrid device," *IEEE Trans. Electron Devices*, vol. 53, no. 9, pp. 2335-2343, Sep. 2006.
- [8] L. Mathew *et al.*, "ITFET: Inverted T Channel FET, A novel device architecture and circuits based on the ITFET," in *IEEE Int. Conf. on Integrated Circuit Design and Technology*, May 2006, pp. 1-4.
- [9] *Taurus-Device User Guide*, Synopsys, Inc., Mountain View, CA, Sep. 2004.
- [10] *UFDG MOSFET MODEL User's Guide (Ver. 3.7)*, SOI Group, Univ. Florida, Gainesville, FL, Jul. 2007. [Online]. Available: <http://www.soi.tec.ufl.edu>.
- [11] V. P. Trivedi and J. G. Fossum, "Nanoscale FinFETs with gate-source/drain underlap," *IEEE Trans. Electron Devices*, vol. 52, no. 1, pp. 56-62, Jan. 2005.
- [12] S. Chouksey, J. G. Fossum, S. Agrawal, and L. Mathew, "LP-HP nanoscale FinFET-CMOS design via source/drain engineering," in *IEEE Int. SOI Conf.*, Oct. 2008, pp. 125-126.
- [13] A. Cros *et al.*, "Unexpected mobility degradation for very short devices: A new challenge for CMOS scaling," in *IEDM Tech. Dig.*, Dec. 2006, pp. 1-4.
- [14] Nguyen *et al.*, "Direct comparison of Si/high-k and Si/SiO<sub>2</sub> channels in advanced FD SOI MOSFETs," in *IEEE Int. SOI Conf.*, Oct. 2008, pp. 25-26.
- [15] Ramos *et al.*, "Effective mobility extraction based on a split RF C-V method for short-channel FinFETs," in *IEEE Solid-State Device Research Conf.*, Sep. 2006, pp. 363-366.

- [16] M. Casse *et al.*, “Mobility of strained and unstrained short channel FD-SOI MOS-FETs: New insight by magnetoresistance,” in *2008 Symp. on VLSI Technology*, Jun. 2008, pp. 170-171.
- [17] V. P. Trivedi and J. G. Fossum, “Scaling fully depleted SOI CMOS,” *IEEE Trans. Electron Devices*, vol. 50, no. 10, pp. 2095-2103, Oct. 2003.
- [18] L. Mathew *et al.*, “Undoped channel PMOS FinFET with deposited titanium nitride gate electrode over SiO<sub>2</sub> gate dielectric for low leakage applications,” in *Proc. IEEE Si Nanoelectronics Workshop*, Jun. 2006, pp. 7-8.
- [19] V. P. Trivedi and J.G. Fossum, “Quantum-Mechanical effects on the threshold voltage of undoped Double-Gate MOSFETs,” *IEEE Electron Device Letters*, vol. 26, no. 8, pp. 579-582, Aug. 2005.
- [20] S.-H. Kim, J. G. Fossum, and V. P. Trivedi, “Bulk inversion in FinFETs and implied insights on effective gate width,” *IEEE Trans. Electron Devices*, vol. 52, no. 9, pp. 1993-1997, Sep. 2005.
- [21] J. C. Lee *et al.*, “Hafnium-based high-k dielectrics,” in *VLSI Symp. Tech. Dig.*, Apr. 2005, pp. 122-125.
- [22] S.-C. Song *et al.*, “Highly manufacturable advanced gate-stack technology for sub-45-nm self-aligned gate-first CMOSFETs,” *IEEE Trans. Electron Devices*, vol. 53, no. 5, pp. 979-989, May 2006.
- [23] G. Vellianitis *et al.*, “Gatestacks for scalable high-performance FinFETs,” in *IEDM Tech. Dig.*, Dec. 2007, pp. 681-684.
- [24] S.-H. Kim, J. G. Fossum, and J. Yang, “Modeling and significance of fringe capacitance in nonclassical CMOS devices with gate-source/drain underlap,” *IEEE Trans. Electron Devices*, vol. 53, no. 9, pp. 2143-2150, Sep. 2006.
- [25] C. R. Manoj and V. R. Rao, “Impact of high-k gate dielectrics on the device and circuit performance of nanoscale FinFETs,” *IEEE Electron Device Letters*, vol. 28, no. 4, Apr. 2007.
- [26] N. R. Mohapatra, M. P. Desai and V. R. Rao, “Detailed analysis of FIBL in MOS transistors with high-k gate dielectrics,” in *Proc. 16th Annual Conf. VLSI Design*, Jan. 2003, pp. 99-104.
- [27] W. Zhu, J. P. Han, and T. P. Ma, “Mobility measurement and degradation mechanisms of MOSFETs made with ultrathin high-k dielectrics,” *IEEE Trans. Electron Devices*, vol. 51, no. 7, pp. 98-105, Jan. 2004.
- [28] M. V. Fischetti, D. A. Neumayer, and E. A. Cartier, “Effective electron mobility in Si inversion layers in metal-oxide-semiconductor systems with a high-k insulator: The role of remote phonon scattering,” *Journal of Applied Physics*, vol. 90, no. 9, pp. 4587-4608, Nov. 2001.

- [29] V. S. Chang *et al.*, "Optimization and scaling limit forecast of nitrided gate oxide using an equivalent nitride/oxide (N/O) stack model," in *Proc. IEEE Int. Conf. Integrated Circuit Design Technology*, Mar. 2004, pp. 363-366.
- [30] S. -H. Kim and J. G. Fossum, "Design optimization and performance projections of Double-Gate FinFETs With gate-source/drain underlap for SRAM application," *IEEE Trans. Electron Devices*, vol. 54, no. 8, pp. 1934-1942, Aug. 2007.
- [31] N. D. Arora, J. R. Hauser, and D. J. Roulston, "Electron and hole mobilities in silicon as a function of concentration and temperature," *IEEE Trans. Electron Devices*, vol. 29, no. 2, pp. 292-295, Feb. 1982.
- [32] V. P. Trivedi, J.G. Fossum, and F. Gamiz, "A compact QM-based mobility model for nanoscale ultra-thin-body CMOS devices," in *IEDM Tech. Dig.*, Dec. 2004, pp. 763-766.
- [33] L. Ge and J. G. Fossum, "Analytical modeling of quantization and volume inversion in thin Si-film DG MOSFETs," *IEEE Trans. Electron Devices*, vol. 49, no. 2, pp. 287-294, Feb. 2002.
- [34] L. Ge, J. G. Fossum and B. Liu, "Physical compact modeling and analysis of velocity overshoot in extremely scaled CMOS devices and circuits," *IEEE Trans. Electron Devices*, vol. 48, no. 9, pp. 2074-2080, Sep. 2001.
- [35] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*. Cambridge: Cambridge University Press, 1998, pp. 132-133.
- [36] G. Ribes *et al.*, "Review on high-K dielectrics reliability issues," *IEEE Trans. Device and Materials Reliability*, vol. 5, no. 1, pp. 5-19, Mar. 2005.
- [37] S.-H. Kim and J. G. Fossum, "Nanoscale CMOS: Potential nonclassical technologies versus a hypothetical bulk-silicon technology," *Solid-State Electronics Journal*, vol. 49, no. 4, pp. 595-605, Apr. 2005.
- [38] A. S. Roy, C. C. Enz, and J. M. Sallese, "Compact modeling of gate sidewall capacitance of DG-MOSFET," *IEEE Trans. Electron Devices*, vol. 53, no. 10, pp. 2655-2657, Oct. 2006.
- [39] A. Bansal, B. C. Paul, and K. Roy, "Modeling and optimization of fringe capacitance of nanoscale DGMOS devices," *IEEE Trans. Electron Devices*, vol. 52, no. 9, pp. 256-262, Feb. 2005.
- [40] P. C. Yeh and J. G. Fossum, "Physical subthreshold MOSFET modeling applied to viable design of deep-submicrometer fully depleted SOI low-voltage CMOS technology," *IEEE Trans. Electron Devices*, vol. 42, no. 9, pp. 1605-1613, Sep. 1995.
- [41] Meng-Hsueh Chiang, "Process-based compact modeling and analysis of silicon-on-insulator CMOS devices and circuits, including Double-Gate MOSFETs," Ph.D. Dissertation, Univ. of Florida, Gainesville, FL, 2001.

- [42] C. Yin, P. C. H. Chan, and M. Chan, "An air spacer technology for improving short-channel immunity of MOSFETs with raised source/drain and high-k gate dielectric," *IEEE Electron Device Lett.*, vol. 26, no. 5, pp. 323-325, May 2005.
- [43] N. R. Mahopatra, M. P. Desai, S. G. Narendra, and V. Ramgopal Rao, "Modeling of parasitic capacitances in deep submicrometer conventional and high-k dielectric MOS transistors," *IEEE Trans. Electron Devices*, vol. 50, no. 4, Apr. 2003.
- [44] S. Chouksey and J. G. Fossum, "DICE: A beneficial short-channel effect in nanoscale Double-Gate MOSFETs," *IEEE Trans. Electron Devices*, vol. 55, no. 3, pp. 796-802, Mar. 2008.
- [45] S. Natarajan *et al.*, "A 32nm logic technology featuring 2<sup>nd</sup>-generation high-k+metal-gate transistors, enhanced channel strain and 0.171 $\mu\text{m}^2$  SRAM cell size in a 291Mb array," in *IEDM Tech. Digest*, Dec. 2008, pp. 941-943.
- [46] G. Ghibaudo, "New method for the extraction of MOSFET parameters," *Electronics Letters*, vol. 24, no. 9, pp. 543-545, Apr. 1988.
- [47] M. Lundstrom, *Fundamentals of Carrier Transport*. Cambridge: Cambridge University Press, 2000, pp. 67-72.
- [48] Y. M. Meziani, J. Lusakowski, W. Knap, N. Dyakonova, and F. Teppe, "Magnetoresistance characterization of nanometer Si Metal-Oxide-Semiconductor Transistors," *J. Applied Physics*, vol. 96, no. 10, pp. 5761-5765, Nov. 2004.
- [49] J. G. Fossum and D. S. Lee, "A physical model for the dependence of carrier lifetime on doping density in nondegenerate silicon," *Solid State Electronics Journal*, vol. 25, no. 8, pp. 741-747, Aug. 1982.
- [50] X. Garros *et al.*, "Guidelines to improve mobility performances and BTI reliability of advanced high-k/metal gate stacks," in *2008 Symp. VLSI Technology*, Jun. 2008, pp. 68-69.
- [51] V. P. Trivedi, "Physics and design of nonclassical nanoscale CMOS devices with ultrathin bodies," Ph.D. Dissertation, Univ. of Florida, Gainesville, FL, 2005.
- [52] S. E. Thompson, S. Suthram, Y. Sun, G. Sun, S. Parthasarthy, M. Chu, and T. Nishida, "Future of strained Si/Semiconductors in nanoscale MOSFETs," in *IEDM Tech. Dig.*, Dec. 2006, pp. 1-4.
- [53] S. Suthram *et al.*, "Understanding strain effects on Double-Gate FinFET drive-current enhancement, hot-carrier reliability and ring-oscillator delay performance via uniaxial wafer bending experiments," in *2008 Intl. Symp. VLSI Technology, Systems, and Applications*, Apr. 2008, pp. 163-164.
- [54] Yang-Kyu Choi, Tsu-Jae King, and Chenming Hu, "Nanoscale CMOS spacer FinFET for the terabit era," *IEEE Electron Device Lett.*, vol. 23, no. 1, pp. 25-27, Jan. 2002.
- [55] *Sentaurus User's Guide*, Synopsys, Inc., Mountain View, CA, Sep. 2007.

- [56] *MEDICI-4.0 User's Manual*, Synopsys, Inc., Mountain View, CA, Sep. 2004.
- [57] M. Zahn, *Electromagnetic Field Theory: A Problem Solving Approach*. New York: Wiley, 1979, pp. 272-273.

## BIOGRAPHICAL SKETCH

Shishir Agrawal was born in Pune, India. He received his Bachelor of Technology degree from Banaras Hindu University, Varanasi, India in 2002, M.S. degree from Rutgers University in 2005, and Ph.D. degree from the University of Florida in 2009.

In the Summer of 2008, he was a summer intern in the “Compact Modeling and Characterization” group at “Advanced Micro Devices” (AMD) in Sunnyvale, CA. His research interests are in the areas of modeling and design of non-classical CMOS devices, circuits and technology.