INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700  800/521-0600
RICE UNIVERSITY

PROCESS DESIGN AND CIRCUIT MODEL DEVELOPMENT

by

JUAN ANTONIO RODRÍGUEZ

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
MASTER OF SCIENCE

APPROVED, THESIS COMMITTEE

[Signatures]

William L. Wilson, Jr., Professor, Director
Electrical and Computer Engineering

Frank Tittel, Professor, Chair
Electrical and Computer Engineering

Joseph Cavallaro, Assistant Professor,
Electrical and Computer Engineering

Michael C. Smayling, Lecturer,
Electrical and Computer Engineering

Houston, Texas

April, 1994
ABSTRACT

Process Design and Circuit Model Development

by

Juan Antonio Rodriguez

Process design for integrated circuit manufacturing has traditionally been implemented with little simulation prior to fabrication. As with circuit design of a decade ago, the available simulation tools were mainframe-based, often incompatible, and lacked accurate physical models. Recent developments in process and device simulation allow accurate process modeling which reflect actual fabrication plant capabilities. A highly structured simulation environment implemented for development of Texas Instruments' PRISM™ technology is described, together with results of a simulation approach to circuit model development for a new class of silicon power transistors. A new analytical model for field effect transistor modeling is also proposed. This new model preserves continuity of both the drain current and conductance over all bias conditions. It also accurately models the effects of substrate bias on device behavior.
ACKNOWLEDGMENTS

A sincere show of appreciation is extended to everyone who made this project possible and its completion a reality. In particular, Michael Smayling of Texas Instruments and Professor William Wilson of Rice University have been extremely supportive throughout the duration of this work; their guidance, encouragement and challenges allow me the opportunity to learn and to grow in my dual role as a student and an engineer.

I thank Don Redwine for sharing his technical interests and allowing me the opportunity to work with him on an interesting and meaningful project. It is my hope that our discussions will continue and that opportunities will exist for more collaborative work in the near future.

Professor Richard Tapia of Rice has been a mentor and a friend since my undergraduate years and I would like to thank him and his family for the many great times we have shared. I admire and learn from his commitment to excellence on all levels.

Everyone in the PRISM Process Development Lab has been helpful in one way or another, especially Donna Cavness, Jim Francis, Michael Dickens and Alister Young. I would never have been able to complete every foil, take all the data and organize it all without their help. Michael Smayling and William Wilson have also generously shared their time and efforts in generating foils and viewgraphs. This is very much appreciated.

I would like to acknowledge my parents, Donato y Gloria Rodriguez and the rest of my familia including my sister Sonia and my brother Miguel. You all have always been supportive and mom and dad thanks for caring.
TABLE OF CONTENTS

1. Title Page i
2. Abstract ii
3. Acknowledgments iii
4. Table of Contents iv
5. Illustrations vii
6. Tables x

CHAPTER

I. Introduction 1
   1.1 Brief History on the use of Computer Tools in IC Design 4
   1.2 The Increasing use of Computer Tools for Process Design 4
   1.3 Modern Process and Device Simulation Capabilities 5
   1.4 Use of Process and Device Simulation in Technology Development 7
      1.4.1 Design of Experiments and Sensitivity Analysis 7
      1.4.2 Circuit Model Development 11
   1.5 Computer Aided Modeling for MOS Field Effect Transistors 12
   1.6 Linking Process Simulation with Circuit Model Development 14
   1.7 Process Simulation in Texas Instruments' Technology Development 15

II. Computational Tools for Process Development 17
   2.1 Metal-Oxide-Semiconductor Field Effect Transistor Concepts 17
   2.2 Modern CMOS Process Engineering Concepts 23
      2.2.1 Oxidation 23
      2.2.2 Lithography 23
      2.2.3 Nitride Deposition 24
      2.2.4 Ion Implantation 24
      2.2.5 Anneal/Diffusion 24
2.2.6 Etching 24
2.2.7 Chemical Vapor Deposition 25
2.2.8 Physical Vapor Deposition 25

2.3 1-Dimensional Process Modeling 26
2.4 2-Dimensional Process Modeling 33
2.5 2-Dimensional Device Modeling 34
2.6 Device Measurement Tools and Model Fitting 36
2.7 Examples of Simulation Output 36

III. Circuit Model Development 43
3.1 Lateral DMOS Transistor Structure Description 43
3.2 Circuit Model Development Methodology 58
3.3 Statistics and Process Capability 62
   3.3.1 Variations in Fabrication Processes 62
   3.3.2 Process Capability 63
   3.3.3 Statistical Process Control 63
3.4 Nominal Device Modeling 72
   3.4.1 Statistical Database 72
   3.4.2 Selecting the Nominal Device 72
   3.4.2 Characteristics Measurements and SPICE Model Fitting 72
   3.4.3 Simulation Strategy 97
   3.4.4 Comparison of Simulated and Measured Data 101
3.6 Weak and Strong SPICE Models 103
   3.6.1 Comparison of Simulated and Measured Data 103
3.7 Summary and Concluding Remarks 107

IV. Hyperbolic Model for MOS Transistor Modeling 109
4.1 Objectives of MOSFET Modeling 109
4.2 The Four Terminal MOS Transistor
   4.2.1 Shichman-Hodges Model 110
   4.2.2 Models Incorporating Short Channel Effects 113
   4.2.3 Subthreshold Region of Operation 115
   4.2.4 Linear and Saturation Regions of Operation 117
   4.2.5 Improving the SPICE Models 121

4.3 The Hyperbolic Model 129
   4.3.1 Derivation of the Drain Conductance Equation 132
   4.3.2 Subthreshold and Substrate Bias Modeling 136
   4.3.3 Derivation of the Transconductance Equation 144
   4.3.4 Derivation of the Substrate Bias Conductance Equation 146
   4.3.5 Conclusions 149

V. Conclusions 150

VI. Appendix 155

VII. References 162
ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure No.</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Number of devices per chip over three decades.</td>
<td>3</td>
</tr>
<tr>
<td>2.</td>
<td>Process development flow chart.</td>
<td>10</td>
</tr>
<tr>
<td>3.</td>
<td>Linking process and circuit simulation.</td>
<td>13</td>
</tr>
<tr>
<td>4.</td>
<td>Basic structure of an NMOS Field Effect Transistor.</td>
<td>18</td>
</tr>
<tr>
<td>5.</td>
<td>Biasing Scheme for an NMOS FET.</td>
<td>19</td>
</tr>
<tr>
<td>6.</td>
<td>Punchthrough during device operation.</td>
<td>21</td>
</tr>
<tr>
<td>7.</td>
<td>Cross section of a CMOS twin well structure</td>
<td>22</td>
</tr>
<tr>
<td>8.</td>
<td>Simulation environment for PRISM.</td>
<td>32</td>
</tr>
<tr>
<td>9.</td>
<td>Cross section of a PMOS transistor.</td>
<td>38</td>
</tr>
<tr>
<td>10.</td>
<td>Plan view of a PMOS transistor.</td>
<td>39</td>
</tr>
<tr>
<td>11.</td>
<td>Channel region profile for a PMOS transistor.</td>
<td>40</td>
</tr>
<tr>
<td>12.</td>
<td>Source/Drain profile for a PMOS transistor.</td>
<td>41</td>
</tr>
<tr>
<td>13.</td>
<td>Field region profile for a PMOS transistor.</td>
<td>42</td>
</tr>
<tr>
<td>14.</td>
<td>Cross section of an NMOS transistor.</td>
<td>44</td>
</tr>
<tr>
<td>15.</td>
<td>PN junction with no external bias.</td>
<td>48</td>
</tr>
<tr>
<td>16.</td>
<td>PN junction with an applied reverse bias.</td>
<td>49</td>
</tr>
<tr>
<td>17.</td>
<td>Cross section of a PRISM LDMOS transistor.</td>
<td>51</td>
</tr>
<tr>
<td>18.</td>
<td>Cross section of a MOS capacitor.</td>
<td>52</td>
</tr>
<tr>
<td>19.</td>
<td>Electron current vectors in an LDMOS transistor.</td>
<td>54</td>
</tr>
<tr>
<td>20.</td>
<td>LDMOS potential contours at $V_{ds} = 42$ volts.</td>
<td>55</td>
</tr>
<tr>
<td>21.</td>
<td>LDMOS transistor in a &quot;High Side&quot; driver configuration.</td>
<td>57</td>
</tr>
<tr>
<td>22.</td>
<td>On resistance vs. gate bias for the LDMOS transistor.</td>
<td>59</td>
</tr>
<tr>
<td>23.</td>
<td>On resistance vs. temperature for the LDMOS transistor.</td>
<td>60</td>
</tr>
<tr>
<td>24.</td>
<td>Flow chart for LDMOS circuit model development.</td>
<td>61</td>
</tr>
</tbody>
</table>
48. Shichman-Hodges MOS equivalent circuit. 112
49. SPICE subthreshold characteristics as a function of substrate bias. 116
50. SPICE drain characteristics for zero substrate bias. 119
51. SPICE drain characteristics for non-zero substrate bias 120
52. Channel region doping profile for an NMOS device. 122
53. Measured MOSFET drain conductance characteristics. 123
54. Comparison of measured and simulated drain conductance. 124
55. SPICE definition of the saturation voltage. 125
56. New definition of the saturation voltage. 130
57. A generalized hyperbola. 133
58. Translation of the hyperbola to empirically match MOSFET drain conductance. 135
59. Drain conductance comparison using HYPMOD. 137
60. Drain characteristics for zero substrate bias using HYPMOD. 138
61. Subthreshold characteristics for HYPMOD with non-zero substrate bias. 142
62. Drain curves for non-zero substrate bias using HYPMOD. 143
<table>
<thead>
<tr>
<th>Table No.</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>Trade-Offs Between Simulation and Processing.</td>
<td>9</td>
</tr>
<tr>
<td>3.</td>
<td>PRISM Low Voltage NMOS Threshold Voltage Comparison between simulation and meaurement.</td>
<td>31</td>
</tr>
<tr>
<td>4.</td>
<td>MOS Transistor Parameter Dependence on Process Conditions.</td>
<td>45</td>
</tr>
<tr>
<td>5.</td>
<td>Process Parameter Conditions for Matching Weak and Strong LDMOS Transistors.</td>
<td>104</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

Integrated circuit processing is the manufacturing and fabrication of integrated electronic circuits on semiconductor material. Of all semiconductor materials available for building integrated circuits, silicon dominates the industry because its physical properties make it possible to build and package reliable circuits on a single “chip.” Although much of the early work in semiconductor technology centered on germanium, the unique ability to use silicon-dioxide as an insulator and diffusion mask meant that reliable, complex circuits could be built using silicon as the semiconductor material.$^1$

The initial idea for combining typical circuit elements on a single silicon wafer came from Jack Kilby of Texas Instruments in 1958. On July 24 of that year he described in his lab notebook how resistors, capacitors, distributed capacitors and transistors could be designed into a single slice of silicon$^2$. He successfully demonstrated circuits with different elements comprised of only silicon, thus showing that integration was possible. The problem of interconnection was later solved by the use of oxide as an insulator. An explosive growth in semiconductor technology occurred soon after these early demonstrations of integrated circuits. A short history of silicon IC technology applications over the past three decades is presented in Table 1.

The table shows how much technology has become a part of everyday life and the evolution of the present age of microelectronics, which is commonly referred to as the “information age.” These advancements in IC technology and its applications have been driven by the continual progress achieved in manufacturing and process technology$^3$. The driving force has been the enormous increase in the number of devices which can be placed on a single chip. Figure 1 shows how the number of devices that can be placed on a single chip has risen dramatically over the past three decades.
<table>
<thead>
<tr>
<th>Year</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>1958</td>
<td>First demonstration of an integrated device.</td>
</tr>
<tr>
<td>1960-69</td>
<td>Television, radio and minicomputers.</td>
</tr>
<tr>
<td>1970-79</td>
<td>Calculators, personal computers, automobile engine control and television games</td>
</tr>
<tr>
<td>1980-94</td>
<td>Defense, home electronics, worldwide information networks, electronic banking</td>
</tr>
</tbody>
</table>
Figure 1. Number of devices per chip over three decades.
Brief History on the use of Computer Tools in IC Design

Progress in IC design technology has been synergistically linked with the digital computing and data storage capability of each era. In the 1960’s, for example, advances were made in circuit simulation for functionality verification, however, circuit simulators did not become widely available until the following decade. The number of devices on a single chip were limited enough that patterns could be drawn by hand on rubylith and then subsequently photographically reduced in order to generate the IC masks. As device density and functional complexity of the circuits increased, computers began to play an important role in IC design. For example, mask layout and circuit design information was stored in computer databases by the early 1970’s; by the middle of the decade, tools became available for verifying that design rules were followed in the circuit layouts.

The use of these tools became important as the design stage and cycle time for production of silicon chips increased and the debugging effort became difficult due to the increasing numbers of devices present in the circuits. Efforts for design automation of integrated circuits continued to progress and by the mid 1980’s standard methods for synthesis and optimization of gate-level, register level and high level design were in place.

The Increasing use of Computer Tools for Process Design

Process design, however, has traditionally been implemented with little or no simulation prior to running silicon in a wafer fabrication plant. There are several reasons for this. The simulation tools that were available 10-15 years ago were mainframe based, often incompatible, and lacked accurate physical models. In addition, the physical constants for basic process steps such as oxidation and impurity re-distribution were not well characterized, inspiring little confidence in the results. Thus process design has traditionally been confined to experimentation through actual fabrication of test wafers. With lack of simulation capability process engineers had no choice but to implement
silicon test wafers under a variety of process conditions in order to optimize and finalize the fabrication method.

The need for process simulation tools became evident by the mid 1970's as functional requirements of integrated circuits became more complex and the feature sizes of the MOS and bipolar devices continued to shrink. In addition, structural requirements on the devices, such as minimum junction depth, made the process definition a critical part of the overall circuit design. Process definition and control became increasingly more important as IC fabrication moved into high volume production and worldwide competition in the microelectronics industry emerged.

More recent trends such as the need for high speed memories and the rise of ASIC (application specific integrated circuits) chips, especially mixed signal (analog and digital) chips, have also made the use of software tools, both for process and design automation, mandatory. These types of complex, mixed function circuits are examples of applications where modern CAD tools in general can facilitate the design process and encourage the interaction of circuit/device and process engineers leading to more efficient design cycles for IC's. For example, it has been demonstrated that for ASIC chips, high performance is tied to how well the digital architectures are optimized for the specific application as opposed to the scaling of the feature sizes.

**Modern Process and Device Simulation Tool Capabilities**

Recent developments in process and device modeling have led to considerable improvements in standard simulation tools. Physical models such as those used for simulating oxidation, etching, diffusion, ion implantation, epitaxial growth and deposition of nitride and polysilicon have improved significantly in recent years. These are processes that are essential for fabricating integrated circuits.
Physical and empirical models implemented in device simulators have also improved. A variety of reasonably accurate carrier mobility models which take into account concentration dependence as well as electric field dependence are available. There are also models for investigating parasitic effects such as hot carrier generation and band to band tunneling. In addition, standard tools are now UNIX based, offering design engineers workstation flexibility and multi-window environments leading to a reduction in cycle time by allowing different design tasks to be worked on concurrently. With modern processor capability, process and design engineers can run extensive experiments using the simulation tools and find target values for the parameters important in the process design in about a week’s time, given an efficient simulation format. This is about 1/6 the time it takes to run modern VLSI silicon wafers through a fabrication line.

Simulators also offer the opportunity to investigate details about device structures that are difficult to measure or which can only be inferred or assumed: cross sections of the impurity distributions, potential contours, electric field intensities, current flowlines, etc. inside of the integrated structures under any operating conditions can be obtained. These capabilities are extremely useful for optimizing the devices and even for finding out why a circuit does not work in a failure analysis investigation.

Another key improvement has been the implementation of a general interchange format for importing process simulation data such as doping profiles and geometry information to device simulators\(^{10}\). This link between process and device simulators offers important integration to simulation schemes. Furthermore, device simulators can now be linked to device parameter extractors making it possible to directly compare electrical characteristics gathered from the simulation to data measured from fabricated devices\(^{11}\).
Use of Process and Device Simulators in Technology Development

The advances in process and device simulation have made a "virtual factory" available to modern process development laboratories and computer experimentation is now an everyday part of technology development work. Process development engineers can make use of simulation tools in a variety of ways, including: "design of experiments," process based sensitivity analysis and circuit model development.

Design of Experiments and sensitivity analysis. IC fabrication processes consist of a series of steps which are formally defined as recipes. For example, recipes exist for gate oxidation, polysilicon deposition, ion implant steps, etc. As the name implies, each recipe outlines the specifics of how that process is carried out, including the temperature settings over time and at what pressure that process may take place. Defining the process and the recipes for each step is one of the most difficult tasks in developing new technologies. Design of experiments in the IC industry is a structured approach towards finding the optimal set of recipes for fabricating silicon chips, given the technology requirements. One of the basic premises of design of experiments is that the manufacturing process can be improved and made more efficient if statistically designed experiments are carried out to eliminate dependencies on inherent process variations early in the development cycle. For example, in order to find the appropriate conditions that will result with a specific channel profile in a MOS transistor, substrate doping, threshold voltage adjust implants and anneal conditions (time and temperature) are variables that need to be investigated. With design of experiments one is able to examine the impact of each of these variables on the process and find the optimum set of conditions that will result in the required channel profile.

Design of experiments, along with other modern tools of manufacturing such as Statistical Process Control and the use of Pareto Diagrams are tools commonly used in manufacturing industries to improve quality and productivity. This is where sensitivity
analysis also plays a role in the process definition. Wafer fabrication plants operate with a set of tolerances for the process parameters. Given those inherent process variations, devices should be able to operate within tolerance levels set by the specifications. With a sensitivity analysis, devices can be designed so that they will meet the specifications, taking into account inherent variations during the processing. Design of experiments analysis helps define process recipes which can yield devices within those specified performance levels. Any well understood and established process can then serve as a basis for future technologies.

The use of process simulation tools offers a parallel path in the process development cycle. Table 2 compares the tradeoffs between choosing simulation or actual processed wafer data for process design information. It is important to understand the advantages and limitations for both cases. While the processed silicon will be the final test of any experiment, the simulation tools can offer extensive experimental data at minimal cost. This data can then be used for sensitivity analysis and interpolated in design centering schemes. Once target values for different process variables are found, they can then be run in “split lots” in the fabrication line for final verification or further tuning.

There are a few standard methods for defining new technology processes. Figure 2 shows a flow chart of the necessary steps required in developing a new process including the use of modeling tools. Typically an older version of a similar technology is enhanced in order to accommodate the needs for the new chip designs. Standard requirements for designing a process include 1) outlining the customer requirements on the circuits to be manufactured, 2) isolating the process parameters which affect those specifications and 3) defining the process steps to ensure that all of the specifications are met. This type of analysis is called Quality Function Deployment and is carried out in conjunction with design of experiments. It helps assure that the key process steps are identified and optimized to meet customer requirements or “care-abouts”.
<table>
<thead>
<tr>
<th>Criteria</th>
<th>Simulation</th>
<th>Processed wafer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle time</td>
<td>~1 week per component to set up, &lt; 1 day per component for DOE</td>
<td>~1 week per component for test die design, ~40-60 days for fabrication; ~1 week per component for test verification, deconvolution</td>
</tr>
<tr>
<td>Insight</td>
<td>Internal node values</td>
<td>Need inference</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Good for interpolation</td>
<td>Good if process is known</td>
</tr>
<tr>
<td>Cost</td>
<td>~75K software cost</td>
<td>~1K/Development wafer</td>
</tr>
</tbody>
</table>
Figure 2. Process Development Flow Chart.
Circuit Model Development. Circuit simulation and verification was one of the earliest tools used by engineers in IC design. The work by Shichman and Hodges in developing an equivalent circuit model for MOS transistors became a standard which has been used in circuit simulation tools since they first published their work in 1968\textsuperscript{16}. Modern MOS transistor models use the Shichman-Hodges model along with parameters which take into account effects due to velocity saturation, mobility degradation and other effects present in modern short channel devices\textsuperscript{17}. By measuring the device characteristics from the processed silicon wafers, device engineers are able to extract the key parameters necessary for representing the devices accurately. Accurate models are necessary as this is the primary tool for simulating and verifying that circuit designs meet specifications before the mask set is generated for the IC's. Most circuit simulators used today are variations of the SPICE2 program developed at the University of California at Berkeley, published in 1975\textsuperscript{18}.

Circuit model development is an area where process and device simulation can be linked with circuit performance evaluation. Green and Fossum have demonstrated the ability to develop bipolar transistor circuit models using a program they developed for combining the 1-dimensional simulator SUPREM-3 with MMSPICE, a "mixed-mode" device and circuit simulator\textsuperscript{19}. Their program, SUMM, takes one dimensional profile data and extracts structure dependent physical parameters which can then be used by the mixed-mode simulator for obtaining characteristics such as gain bandwidth product for an individual device and delay characteristics for subcircuits. Their approach eliminates the need for circuit model fitting and parameter extraction; it also has the speed advantage over 2-dimensional numerical process and device simulators by using the 1-dimensional program SUPREM-3 and the analytical models implemented in MMSPICE.

At AT&T Bell Laboratories, integration of process, device and circuit simulation tools is a key part of their technology development program\textsuperscript{20}. Their simulation
environment includes the use of a numerical optimization system for process design centering and also a parameter extraction tool which can be used with arbitrary device models. Their general approach in process development is to define an initial process and allow the process and mixed-mode device/circuit simulators to iterate, given the technology requirements and objectives. Once the process has been optimized the next step is characterization which includes parameter extraction, circuit simulation and verification with measurements.

In general, computer aided process design and device modeling requires process simulation coupled with device and circuit simulators. Figure 3 shows how the modeling efforts from process to circuit verification are linked. Wan, et al., have demonstrated the possibility of linking standard tools for process and device simulation with parameter extraction tools via their program MOSGEN\textsuperscript{21}. By accessing device simulator data taken at different operating conditions they are able to extract SPICE parameters such as junction depth, channel length and threshold voltage as well as parasitic effects such as channel length modulation. Once all the parameters are identified they can then compare their model curves with measured data.

**CAD Modeling for MOSFET Devices**

MOS transistor modeling is another area in which considerable improvement has been made over the past decade. For example, the ability of SPICE simulators to accurately handle charge conservation in MOS circuits was a considerable step towards simulation and verification for DRAM circuits\textsuperscript{22}. Recent improvements have been driven largely by the needs of digital circuit designers as this has been one of the areas most affected by recent trends in the industry.
Figure 3. Linking process and circuit simulation.
The increasing use of bipolar devices for mixed signal applications and high speed circuits indicates that bipolar device and analog MOS models will have to be improved in order for circuit simulators to handle those types of circuits accurately. Modeling issues for analog applications which still exist include the poor capabilities of standard simulators such as SPICE for handling the transition from linear to saturation current in a MOSFET. G. T. Wright developed an analytical model in which non-uniform substrate doping can be accounted for in a continuous drain current equation by using the threshold voltage for finding the bulk charge\textsuperscript{23}. Wright's results indicate good agreement with measured drain current data as a function of gate and body bias.

Recently Tsividis and Suyama identified benchmarks that all device circuit models should try to meet\textsuperscript{24}. Specific issues identified include the lack of continuity of drain current derivatives with respect to all MOS transistor terminal potentials in current versions of SPICE model equations. These are calculations that are important for analog applications.

**Linking Process Simulation to Circuit Model Development**

The underlying theme in the work to be presented in this manuscript can be summarized in the following way: computer tools for process and device simulation have reached a sophistication such that when coupled with physical measurement they can offer insight that is of use to process and device engineers. As was mentioned earlier, one way in which process simulation tools can be used effectively is in circuit model development.

Specifically, the development of circuit models which reflect the manufacturing process is an important task as circuit designers require the latest information on the fabrication plant capability in order to produce designs which will yield functioning integrated circuits. A methodology for circuit model development that takes into account fabrication line statistical fluctuations will be described, and it will be shown that
sufficiently accurate I-V data ready for SPICE model fitting can be obtained directly from process and device simulation.

Additionally, analytical equations which can model MOS devices are extremely important for fast and accurate simulation of circuits which can have up to thousands of transistors. One problem with modern SPICE programs is their inability to model drain current with a single continuous function. This can lead to erroneous simulation, especially with analog designs. A model will be presented which maintains continuity of the drain current in going from the linear to saturation region, eliminating the need for circuit simulators to use two different equations for drain current in the above threshold regime. Also, the model has been developed to accurately model the effect of substrate bias on drain current output, both in subthreshold and above threshold bias conditions. The small signal conductances have also been derived and characterized with respect to measured data.

To summarize, this work is concerned with employing computer experiment as part of technology development. Comparison of simulated and measured data will show that the manufacturing process can be modeled accurately by readily available tools, and therefore can lead to meaningful information during the development of new technologies. Lastly, an analytical model for MOS transistor modeling has been developed which overcomes limitations imposed by currently available models.

**Process Simulation in Texas Instruments’ Technology Development**

A process and device simulation environment has been implemented in conjunction with the development of Texas Instruments’ PRISM™ technology. It includes the use of standard tools and efficiently supports the complex processing options available in PRISM. The technology is the first of its kind to support low voltage logic, nonvolatile memories, linear functions and power devices, all on the same chip. The simulation
environment implemented makes it possible to carry out design of experiments simulations in a very efficient manner for all of these applications.

An example of the simulation approach for circuit model development for the case of lateral diffused MOS transistors will be described. Integration of device simulation with standard semiconductor characterization and "fitting" tools offers the opportunity to calibrate device simulations to fabrication line capabilities. Very good agreement is found between simulated and measured data for typical fabricated transistors.

The device modeling work includes the use of a new analytical equation for MOS drain conductance which can model linear and saturation regions with one continuous function, eliminating the need for two different conductance models in circuit simulation tools. It has been shown that having to match a discontinuous function from the linear to saturation region can result in serious error for drain characteristics modeling\textsuperscript{26}. The model maintains continuity between these two regions and shows excellent promise as an analytical form of MOS drain conductance in all regions of operation.
CHAPTER 2

COMPUTATIONAL TOOLS FOR PROCESS DEVELOPMENT

The design of integrated circuit processes is becoming more sophisticated as improvements in semiconductor modeling have led to flexible and accurate structure and device simulation tools. An introduction to MOS devices will be given. Standard process steps required for CMOS fabrication are then described, followed by a description of simulation tools used in process development.

Metal-Oxide-Silicon Field Effect Transistor Concepts

The most widely used device in modern VLSI chips is the MOS transistor. Figure 4 shows the cross section of an idealized NMOS device. Several key features are shown. The transistor is constructed in p-type silicon, formed by introducing boron into the substrate. The source and drain regions are created by “n+” diffusions, consisting typically of high concentrations of arsenic or phosphorus ions. The depletion regions result from n type ions interacting with the p type substrate at the junction. A thin oxide layer is used to separate the metal gate contact from the silicon surface in the active device. The region between the source and drain up near the surface of the silicon is commonly referred to as the “channel.”

Figure 5 shows a typical biasing scheme for this transistor. The grounded source grounded it will serve as the potential reference. The transistor action is not too difficult to visualize: with the gate terminal held at ground, the source and drain are separated by the p region in between and there can be no conductance between them, even if the battery is connected to the drain. The two n-type regions and the substrate from “back to back” diodes which can not conduct in either direction.
Figure 4. Basic Structure of an N-Channel MOS Field Effect Transistor.
Figure 5. Biasing Scheme for an NMOS FET.
If the gate potential is raised, however, the electric field across the gate oxide will attract electrons to the surface and when the gate bias reaches “threshold” the device will be turned on and current will begin to flow. The minority carriers in the channel region form what is called an inversion layer; that is, the p-type substrate has now become n-type just near the surface. The name “field effect transistor” comes from the fact that the transistor action depends on the gate potential modulating the carriers in the channel region.

There are several important parameters that must be considered when designing MOS transistors. The threshold voltage, $V_t$, must be defined, and this value can be finely tuned by the use of boron implants for adjustment. The channel length is another important parameter and this is typically controlled by the minimum feature size available with the process. The substrate doping will affect the threshold voltage and thus will also determine how high of an implant dose is required for $V_t$ adjustment.

The substrate doping level will also affect the punchthrough breakdown of the transistor. Punchthrough refers to the lateral and vertical movement of the drain depletion edge at high potential. Figure 6 shows the depletion edge advancing towards the source and into the substrate region. In this case, the carriers in the channel create a depletion edge with the substrate. At high drain potential the depletion edge moves into the grounded bulk region and substrate currents are created. For 5 volt logic transistors breakdown voltage is usually about 10 volts or so. Other potential problems include hot carrier generation providing gate currents through the oxide. This type of condition can exist for high gate potentials providing an electric field strong enough to pull electrons from the silicon through the oxide.

Modern VLSI chips consist of both NMOS and PMOS transistors. PMOS transistors depend on positive charge flow for currents as opposed to NMOS depending on negative charge flow. This type of technology is called Complementary MOS and it places strict requirements on the process design. Figure 7 shows a cross section for
Figure 6. Punchthrough during device operation.
Figure 7. Cross Section of a Complementary MOS Twin-Well Structure.
typical CMOS structures. Particular features include the use of “wells” for creating the individual transistor substrates. Also, notice the thick field oxide separating the individual transistors. This is known as LOCOS field oxide for LOCal Oxidation of Silicon. LOCOS oxide helps isolate active devices and prevents the creation of parasitic transistors. Although CMOS processes are more difficult to implement than NMOS or PMOS alone, the rewards include lower power dissipation in the circuits, higher reliability and cost effective designs\textsuperscript{27}.

**Modern CMOS Process Engineering Concepts**

There are many process steps that must be implemented in order to get from single crystalline silicon to a functioning integrated circuit. Several key process steps will be described. For more detailed and specific information the reader is asked to consult the literature\textsuperscript{28}.

**Oxidation.** As was mentioned earlier, oxidation of silicon was one of the important developments in silicon based integrated circuit technology. Oxides play several key roles. As the figures presented show, gate oxides serve as the insulator between the silicon and the gate for MOS transistors, and thick oxides are used to isolate individual devices and prevent the creation of parasitic devices. Pad oxides are used to reduce damage to the silicon during a high energy ion implant step, and they also serve as passivation layers to protect the final circuits and prepare them for packaging.

**Lithography.** Lithography refers to the process of creating patterns which will serve to introduce dopants into the silicon and indicate the interconnections of devices and circuits. Photoresist is first “spin coated” over the entire wafer. Selected areas are then exposed to radiation and the patterns will be created once these exposed regions are developed. Any region which was not exposed will remain and serve to protect the silicon underneath. After the processing is completed, such as an ion implant step, oxide etch,
etc., the rest of the photoresist is dissolved away and the wafer gets cleaned in preparation for the further processing.

**Nitride Deposition.** Nitride layers play a very important role in VLSI processing. As shown in Figure 8, the LOCOS oxide regions defined areas of the wafer which did not contain active devices. The key to defining active regions is to deposit nitride over a pad oxide and subsequently pattern and etch over all of the inactive regions. The field oxide is then grown, and the nitride protects the areas which will contain every active device. After this step, the remaining nitride is stripped and the wafer is ready for the gate level metal.

**Ion Implantation.** In modern processes dopants are typically introduced into the silicon substrate through ion implantation. During this step the wafer literally serves as a target as the required ions are accelerated and aimed directly at it. Photoresist protects the regions which do not require the implant, and pad oxide helps protect the exposed silicon surface. Although some damage to the wafer does occur, this can be subsequently annealed. Ion implants offer a very controlled manner of introducing dopants into the substrate and the resulting profiles in the silicon can be fine tuned as necessary, i.e., the junction depth and concentration levels of the source/drain diffusions and the wells can be controlled by the energy of the implant and the subsequent anneal.

**Anneal/Diffusion.** Silicon wafers require annealing and diffusion cycles during the processing. These are high temperature cycles (~900 to 1000°C) used to: 1) drive in dopants to required junction depths, 2) anneal out damage to the silicon after an implant step and 3) to ensure the implanted impurities become electrically active in the silicon.

**Etching.** Etching refers to the removal of material from the surface of the wafer. Often masked layers are used to protect certain parts of the wafer while the exposed areas are etched. For example, to create the polysilicon gates, the poly is first deposited over the entire wafer and masks are used to etch away all of it except for the gate areas. Etch processes must have high selectivity against material underneath that being etched; i.e.,
during a poly gate etch, the gate oxide should be left intact. There are both wet and dry etch techniques. Wet etching is a chemical process by which etchants react with the material to be removed and a subsequent diffusion carries the by-products from the surface. Wet etch processes are typically isotropic, meaning vertical and horizontal etching occur at the same rate. For feature sizes of about 3 microns or less it becomes necessary to use anisotropic etches and these are accomplished by dry etch techniques. In a dry etch, plasmas are used to create reactive species which then must reach the wafer and be adsorbed by the surface. A chemical reaction follows and the by-products are then desorbed from the surface.

**Chemical Vapor Deposition.** CVD is used for growing single crystal silicon films and also for depositing polysilicon as the gate material. CVD works by introducing reactant gases into the process chamber and allowing them to move towards the substrate. These atoms are then adsorbed by the target wafer and undergo chemical reactions which lead to solid thin films. Low pressure CVD offers many advantages including uniform films and parallel processing. After the substrate structures are in place, CVD is used to deposit oxide to insulate the lower level poly. Contact holes are then etched in order to make connections to the poly and also for the source and drain contacts.

**Physical Vapor Deposition.** PVD by sputtering is the most common way of depositing aluminum and aluminum alloy layers for interconnects. In this case energetic ions are used to dislodge atoms from the material to be deposited. The ejected atoms travel to the wafer where they will adsorb and form a film. One of the advantages of sputtering is that it can be carried out at room temperature and the deposited films exhibit characteristics required of thin films in VLSI including uniform step coverage, thickness uniformity, low resistivity and good adhesion to the underlying and overlying surfaces.

The design of a technology process for IC fabrication then must take into account all issues related to operation and reliability of the transistors and other electronic devices.
to be supported by the process. As has been shown, the process definition will indicate the doping levels in the devices, the thickness of the gate oxides and the materials to be used for interconnect between the transistor gates and the metal lines which will eventually connect to package output pins. These issues related to process design and implementation can be investigated with modern tools as part of the development process. The following sections will describe how these tools work and ways they are used in an process development setting.

1-Dimensional Process Modeling

SUPREM-3 is the oldest program; the initial version of this 1-D simulator, SUPREM I, was the first commercially available process simulator and it was introduced in 1977 by Antoniadis and Dutton from Stanford University\textsuperscript{29}. Key process steps supported in the modern version include ion implantation, oxidation, diffusion, low and high temperature deposition, epitaxial growth of silicon, selective etching and masking\textsuperscript{30}. SUPREM-3 simulates these process steps along the direction perpendicular to the surface of the silicon. This is an important limitation as any doping profile which varies along the lateral dimension will not be predicted correctly.

The continuous physical processes are approximated using finite difference numerical methods. The simulation structure can consist of up to a maximum of 10 different layers. Each layer in the structure will consist of a series of cells which are centered about a single solution node. The maximum allowed number of nodes per simulation structure is 1000. The user defines the grid spacing between the nodes in any of the layers. This flexibility allows one to place more nodes to increase the accuracy in areas of interest, such as near metallurgical junctions to determine junction depth or near the surface to calculate surface concentrations. The impurity concentrations and physical constants are assumed to be constant within each cell.
For impurity re-distribution, SUPREM-3 solves the one dimensional continuity equation,

\[ \frac{\partial C}{\partial t} = -\frac{\partial J}{\partial x} \]  

(1)

where \( C \) is the chemical impurity concentration and the impurity flux \( J \) has both diffusion and drift components, i.e.,

\[ J = -D \frac{\partial C_m}{\partial x} + \frac{qD}{kT} E C_a, \]  

for donor impurities, and

\[ J = -D \frac{\partial C_m}{\partial x} - \frac{qD}{kT} E C_a \]  

(3)

for acceptor impurities.

Here \( E \) is the electric field, \( C_a \) is the active concentration, \( C_m \) is the mobile impurity concentration and \( D \) is the diffusion coefficient for each impurity. The dependence of the diffusion coefficients and electric field on impurity and electron concentrations makes these equations nonlinear. The solution proceeds by dividing the total diffusion time into a series of smaller time steps \( \Delta t \) which have default calculated values but can also be specified by the user. After the time increment is calculated, an integration is performed to find the new value for the concentration \( C \), and this will be the starting point for the next step. The value of the next time step will depend on the maximum value of \( \Delta C \) for all impurities.

As said above, impurities are typically introduced into the silicon via ion implantation\(^{31}\). SUPREM-3 has a variety of analytical models available to simulate implantation processes including the use of Gaussian or Pearson distribution functions. As an alternative, Monte Carlo simulation of the implant step can be specified, but this can be
a computationally expensive simulation. One of the advantages of using the Monte Carlo method for simulating implant steps is that it can take into account damage to the lattice and channeling of impurities\textsuperscript{32}. While this helps to understand important physical phenomena, the use of well characterized analytical distribution functions in most cases is quite adequate for engineering applications.

Thermal oxidation growth is handled by simulating a diffusion step in an oxidizing ambient. The only materials which will oxidize in SUPREM-3 are nitride layers, polysilicon and silicon. Ambients the user can specify include those for wet or dry oxidation in addition to the flow of the ambient gases such as oxygen, hydrogen, chlorine or nitrogen. The thickness of the oxide layer grown depends on the partial pressure of the ambient gas and the oxidation rates for the ambients. Standard oxide thickness formulas are implemented as can be found in the SUPREM-3 User's Manual or in the literature\textsuperscript{33}. The time step used to calculate the impurity re-distribution is limited to prevent any material interface from moving past one node during a single time step.

SUPREM-3 also supports simple electrostatic calculations to allow estimating device parameters like MOS transistor threshold voltage. The form of Poisson's equation solved in this case is

\begin{equation}
\frac{\partial}{\partial x} \left( \varepsilon \frac{\partial \psi(x)}{\partial x} \right) = -q(p(\psi) - n(\psi) + N_a(x) - N_a(x)) \quad \text{for semiconductors, and}
\end{equation}

\begin{equation}
\frac{\partial}{\partial x} \left( \varepsilon \frac{\partial \psi(x)}{\partial x} \right) = 0 \quad \text{for insulators.}
\end{equation}

Here $\varepsilon$ is the relative dielectric constant of the material multiplied by $\varepsilon_0$, the dielectric constant in vacuum. The potential in a conductor will be constant throughout the entire
layer. Either Fermi-Dirac or Boltzmann statistics can be used to calculate the values of electron and hole concentrations \( n \) and \( p \) in the semiconductor. In addition, incomplete ionization of the impurities \( N_d \) and \( N_a \) concentrations can be calculated.

Newton’s method is used iteratively by linearizing the equation and using straight Gaussian elimination to solve the resulting system of equations at each iteration. MOS threshold voltage can be approximated by calculating the sheet conductance of the semiconductor for a number of gate electrode biases. Given the electron concentration for example, the sheet conductance will be given by

\[
G_n = \int q\mu n dx.
\]

The ability to simulate accurate substrate profiles makes it possible to predict process conditions for threshold voltage adjust implants, which are standard in modern VLSI processing. Table 3 shows the results of simulated and measured threshold voltage data for low voltage NMOS transistors. The implant doses were generated directly from SUPREM-3 runs.

The recent addition of masking statements and the ability to assign variables for switching process steps “on” or “off” has made it possible to support the complex process flow options available in PRISM technology with a single simulation deck. This has had a significant impact on the ability to make design of experiments runs with minimal input deck editing.

The file hierarchy used for the PRISM simulation environment is shown in Figure 8. Each of the components available with the PRISM process has its own simulation directory, located in the upper level directories “s3” or “s4” for process simulations and “md” for electrical simulations. Each component directory contains mask layout and processing information required for simulation. In addition, simulated structures are also
kept in each components' own subdirectory. All component mask files access the same process simulation files ("globals" for process parameters and "main" for the PRISM flow), with process switches defined in the mask file dictating the appropriate process flow steps required.

In the electrical simulation directory, individual component files access the appropriate structures, either 1-D or 2-D, from the process directories. Again, all the components have access to the upper level files such as "mos.cv" or "drain.iv," with internal switches setting the appropriate experiment conditions. UNIX compatibility ensures that simulation files in each component subdirectory can access any other file in the simulation environment. For example, to carry out a low voltage CMOS electrical simulation, the simulated structures located in the "s3" or "s4" directories can be accessed directly by the calling file located inside the LVCMOS directory under "md."

The files "globals" and "main" in the process simulation directories describe the specific process parameters and the process flow, respectively. These files are parameterized such that each component's mask file sets the process flow variables so that only those steps required for that device are turned on. For example, for a low voltage CMOS transistor, the process flow will indicate a gate oxidation step for a low voltage device as opposed to the oxides grown for high voltage or memory devices.

Simulated device structures and profile data is held in each device's subdirectory. These profiles can be accessed directly from the SUPREM-4 device directory when 2-D effects such as lateral diffusion of impurities or isolation oxide characteristics are of interest. This type of hierarchy eliminates the movement of profile data and processed structures; any profile and any structure can be accessed directly. This is one of the advantages of having UNIX compatibility.
**TABLE 3**

PRISM Low Voltage NMOS Transistor Threshold Voltage Comparison.

<table>
<thead>
<tr>
<th>Total Dose (ions/area)</th>
<th>Measured ( V_{\text{target}} / V_{\text{measured}} )</th>
<th>Simulated ( V_{\text{target}} / V_{\text{simulated}} )</th>
<th>% Difference between simulation and meas.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>+10%</td>
<td>0.95</td>
<td>0.95</td>
<td>0</td>
</tr>
<tr>
<td>-10%</td>
<td>1.06</td>
<td>1.04</td>
<td>2</td>
</tr>
</tbody>
</table>
Figure 8. Simulation environment for PRISM. The files accessible by all components are in lower case. Each component subdirectory is in uppercase. Profile, structure, and electrical results data are held in each device's subdirectory.
2-Dimensional Process Modeling

The 2-D process simulator SUPREM-4 solves the same basic continuity equations listed in (1) through (3), with additional models for point defect diffusion. The equations are similar, with recombination and generation terms included for the case of interstitial and vacancy diffusion,

\[
\frac{\partial C}{\partial t} = -\nabla \cdot J - R + I.
\]

Here \( R \) is the total rate of recombination including that due to interstitials with vacancies in bulk silicon, at interfaces, and the absorption of interstitials at stationary trapping sites\(^{34}\). \( I \) is the total rate of impurity injection at material interfaces. The motivation behind the development of SUPREM-4 was to solve coupled oxidation and diffusion equations on the same grid\(^{35}\). In particular, with the widespread use and acceptance of CMOS technology in industry during the past decade, it has become necessary to understand effects related to isolation such as LOCOS isolation between active device regions\(^{36}\).

Ion implantation can be modeled with analytical distribution functions or Monte Carlo techniques similar to the 1-D simulator. A key difference between the two simulators is in how they approximate oxidation. In SUPREM-4 oxidation is modeled based on the theory developed by Deal and Grove\(^ {37} \). From this theory, analytical equations can be derived for fast simulation of simple structures. Sophisticated numerical models can also be derived in order to accommodate arbitrary structures. With the "compress" numerical model, viscous flow of the oxide is simulated and movement of the oxide/silicon interface is two dimensional. The model can accurately simulate oxidation of non-planar structures and also structures containing polysilicon.
SUPREM-4 models the physical processes by using finite difference techniques for diffusion and finite element techniques for oxide flow. The mesh is allowed to change and adapt as the various process steps affect the device structure. For example, during oxidation, nodes at the oxide/silicon interface continuously move in towards the silicon while nodes in the oxide move with the direction of the oxide flow. Thus at any given solution step nodes can be added or removed from the mesh structure.

Besides the inclusion of physically based models for oxidation and diffusion, the use of a triangular mesh in SUPREM-4 is one major improvement over earlier 2-D process simulators such as SUPRA\textsuperscript{38}. This allows selective node elimination without leaving grid points disconnected. In addition, this type of mesh also allows better matching to non-planar surfaces such as LOCOS field oxide since it can conform to curves and shape changes.

2-Dimensional Device Modeling

One of the most significant improvements has been with the device simulator MEDICI, previously known as PISCES\textsuperscript{39}. The program can now accept structure inputs from SUPREM-4 by way of the standard PIF format, and it can produce outputs for curve fitting with the device characterization tool IC-CAP. Internal models have improved to account for field dependent mobility, impact ionization and lattice heating. Extensions to the basic program allow modeling EPROM and EEPROM structures as well as simple subcircuits. Given the concentrations of electron and hole impurities as well as any surface charge density, MEDICI solves the 2-D Poisson's equation,

\begin{equation}
\varepsilon \nabla^2 \Psi = -q(p - n + N_D - N_A) - \rho_S
\end{equation}

for the electrostatic potential distribution in the structure.
To account for large carrier concentrations under normal operating conditions, MEDICI also solves the continuity equations for electrons and holes, using, for example,

\[
\frac{\partial n}{\partial t} = \frac{1}{q} \nabla \cdot \mathbf{J}_n - \mathbf{R}_n
\]

for the case of electrons. The recombination models available include Shockley-Read-Hall and Auger terms and include concentration dependent lifetimes, surface recombination and tunneling effects. The program solves the coupled equations for self consistent solutions for electrons and holes.

A variety of numerical techniques are available for specifying solution methods including Newton iteration, straight Gaussian elimination and the continuation method. The continuation method is a powerful technique which can be used for tracing difficult current-voltage characteristics such as those exhibited under latchup or breakdown conditions. When using the continuation method, the solution is calculated by specifying both current and voltage boundary conditions. This allows it to automatically set the bias step in order to find accurate solutions. In cases of flat I-V characteristics the steps will be large, while smaller bias steps are taken in regions of rapid change in the current.

MEDICI also uses a triangular simulation grid which can handle arbitrary geometries, including planar and non-planar surface structures. The user can specify automatic refinement of the grid during the execution of the program, which allows additional nodes and mesh elements to be added when quantities of interest such as potential contours, doping profiles, band-to-band tunneling rates, etc. vary by more than a specified tolerance in the original mesh.

Finally, AC analysis can be performed on individual devices or simple circuits. This is useful for example, in generating frequency dependent parameters such as capacitances or conductances.
Device Measurement Tools and Model Fitting

IC-CAP is a standard tool for semiconductor measurement and analysis available from Hewlett-Packard. It contains drivers for standard measurement tools and automates the extraction of MOS or bipolar device I-V data, for example. In addition, it can be used to implement arbitrary device circuit models and optimize the model fitting task. Numerical techniques such as the Levenburg-Marquardt method for least squares optimization can be used to fit device models to measured data\textsuperscript{40}. Furthermore, MEDICI data can be directly imported into IC-CAP which allows calibration of electrical simulation files.

Once calibration with measured devices is done, the simulation programs are useful for generating the Weak/Nominal/Strong I-V curves needed for SPICE models. Using the known capability of the wafer fab, along with target conditions for key parameters like gate oxide thickness and channel length, models can be developed reflecting 6 sigma overall performance. These models can then be used for design and verification, and cross checked against fabrication statistical fluctuations.

Examples Of Simulation Output

For each component or device built in a process, it is useful to have 1-D cross sections of key regions. Figure 9 shows the cross section of a PMOS transistor with the regions of interest indicated by the arrows. 1-D cross sections looking at this device structure from the top are shown in Figures 10 through 13, which are a typical set of views coming out of SUPREM-3. Figure 10 shows the plan view showing the layout of the component and indicating where the cross sections are made. Figures 11 through 13 show doping profiles and boundaries for the channel, source/drain and field regions of the transistor, respectively. Junction depths, peak doping profiles and film thickness can be estimated from these plots. The “tick” marks along the bottom of the profiles indicate
what grid points were used in the simulations. All of these plots were generated with one
SUPREM-3 run, with multiple passes through the main deck. The masking file determines
which process flows are set, and establishes the lateral position of the simulation line.

These simulation tools offer many advantages during the development of a new
process. The computer experiment time, for example, is on the order of a couple of hours,
in contrast to the actual fabrication time for silicon wafers can take about 6 weeks or more
depending on the process. As will be shown in Chapter 3, simulated device characteristics
compare very well to the measured data, even when variations in the processing are taken
into account. Accuracy of the tools coupled with fast experiment time allow process and
device engineers the ability to test many conditions for processing and layout before
designs are implemented in silicon. This leads to a reduction in cycle time and process
development costs.
Figure 9. Cross section of a PMOS transistor, including a contact for the nwell.
Figure 10. Plan view of a PMOS transistor.
Figure 11. Channel region profile for a PMOS transistor.
Figure 12. Source/Drain region profile for a PMOS transistor.
Figure 13. Field region for a PMOS transistor.
CHAPTER 3
CIRCUIT MODEL DEVELOPMENT

A simulation approach to SPICE circuit model development has been implemented for lateral "double-diffused" MOS transistors as part of the development strategy for Texas Instruments' PRISM technology. Lateral DMOS transistors are important components for power IC's whose applications include automotive, personal computing and related electronics\(^{41}\). A brief description of the transistor structure and the process flow steps required for manufacture will be given. The methodology implemented for developing the circuit models will be described, together with the results obtained for simulated and measured I-V data.

**Lateral DMOS Structure and Process Description**

LDMOS transistors form part of the power components available in PRISM technology. As such they are fabricated using standard high voltage CMOS process steps compatible with 0.8 micron silicon technology\(^ {42}\). There are several key differences between high voltage, or power transistors, and low voltage transistors. Figure 14 and Table 4 will help illustrate this point.

The figure shows a cross section of a typical NMOS transistor. The doping level in the substrate is indicated by \( N_a \). The thickness of the gate oxide is indicated by \( t_{ox} \). The designed channel length is the length of the poly gate, but the effective channel length, \( L_{eff} \), is less than \( L \) due to the diffusion of the source and drain implants during the anneal and passivation steps which take place after the implant.

Transistor structures similar to this device dominate low voltage applications for several reasons. To begin with, low voltage transistors are used in core logic gating
Figure 14. Cross section of an NMOS transistor.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>gate oxide thickness</th>
<th>substrate doping</th>
<th>source/drain anneal time and temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k'$ threshold voltage</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>punchthrough voltage</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>
functions and signal processing applications which require one major characteristic: speed. These type of transistors are optimized with a short list of requirements, chief among them that they be able to handle low voltage signals very rapidly.

Table 4 indicates the process and parameters which affect device transfer (drain current output, etc.) characteristics. The ranking (Medium, High) indicates how much of an impact the process parameter makes on the output characteristic. Since the poly gates are typically driven by 5 volt signals, the gate oxide can be made thin without having to worry about breakdown. This helps increase the transistor gain parameter \( k \) (this will be defined later in this chapter) since it is inversely proportional to \( t_{ox} \). The substrate doping is optimized to avoid punchthrough only up to the maximum drain biases that are used. Typically this might be \(~2x\) the maximum gate bias. Punchthrough susceptibility is reduced for higher doping levels, however, higher doping also reduces the carrier mobility plus it increases the threshold voltage. This is a classic example of parameter trade-offs that must be examined carefully during the design of the process. Finally, another way to get maximum speed response is to reduce the channel length. However, this will also be conducive to punchthrough effects. Short channels also lead to higher leakage currents when the device is turned off which can cause reliability problems, especially in memory applications.

Integrated circuits fabricated for high voltage and/or high current applications have very different requirements. For example, the substrate doping levels are much higher than found in low voltage devices in order to reduce punchthrough breakdown at high drain bias. Typical applications in which power devices are needed often have to be able to handle drain voltages of 30, 40 or even 60 volts. Also, because of the higher substrate doping, the threshold voltage tends to be higher and so the response is not as fast as that of a low voltage device. To compensate for that, charge pumps are often used to provide
high gate signals in order to "drive" the transistors very quickly. This means that in order to avoid oxide breakdown there are limits imposed on the minimum gate oxide thickness.

As the drain voltage approaches the gate voltage \( V_{ds} \sim V_{gs} - V_t \), MOS transistors exhibit a phenomena called "pinch-off." This is a high drain electric field condition which leads to current saturation and also hot carrier effects. The electric fields can get so high that energetic electrons can be "pulled" into the gate oxide and result in gate current flow. If the gate current reaches pico-ampere \( (10^{-12} \text{ amperes}) \) levels this can cause permanent charge trapping in the oxide which will change the device characteristics permanently. These are some of the problems that must be overcome in power integrated circuits.

Two simple ideas will help introduce the Lateral DMOS transistor. In a circuit, one way to cause a voltage drop between two points is to place a resistor in between them. Also, p-n junction theory predicts that depletion regions will spread more "rapidly" towards the lighter doped side of the junction. Figures 15 and 16 attempt to pictorially describe this effect. In Figure 15, a p/n\(^-\) junction is shown under zero applied bias. The depletion region edges which occur due to the internal potential are also shown. In Figure 16, a reverse bias is applied, so the diode will not conduct. The effect is to cause the depletion edges to spread, but since the n region is much lighter doped than the p region, the depletion edge has spread much more towards the n\(^-\) area. The reason this occurs is that equal charges on both sides of the junction must be uncovered as the depletion region expands. If one side of the junction is several orders of magnitude higher doped, then the only way an equal amount of charge on the opposite side will be uncovered is for the depletion region to cover a wider distance. This type of behavior is very important for understanding power devices.

The Lateral DMOS transistor is one way to implement an integrated circuit device that will overcome MOS deficiencies at high drain voltage and current conditions.
The P region is heavier doped than the N region so the net effect is a wider depletion depth into the N- side such that all field lines terminate on a charge.

Electric Field due to the space charge in the depletion region at the junction.

Figure 15. A P/N- junction with no external bias.
At reverse bias, the depletion width will spread much more into the lighter doped region.

Figure 16. A reverse bias increases the depletion width towards the lighter doped regions.
Figure 17 shows a simulated cross section of the LDMOS transistor. This device has several unique features which serve to overcome MOSFET problems at high bias as have been described above. The region labeled “p backgate” serves the same purpose as the substrate in the MOS transistor of Figure 14; the notable difference is that in this case it reaches the surface level next to the N+ source, and in fact is always electrically tied to the source. These two regions remain grounded during operation.

Another interesting feature is the n- well region which occupies the rest of the area of the active device. This region serves two very important purposes: first, it helps reduce the high drain electric fields during pinch off by increasing the lateral space between the drain and the channel/gate oxide region; second, the n- well is doped much lighter than the p-type backgate, which means that at high drain bias, the depletion edge of the p/n- junction will spread towards the drain more than it will spread towards the n+ source contact. In addition, the large area (or “carrier drift region”) between the backgate and the n+ drain contact means that the nwell essentially plays the role of a big resistor meant to absorb the high drain potentials.

The thick oxide between the backgate area and the drain contact is LOCOS field oxidation. The extension of the poly gate over the field oxide creates a “field-plated” surface junction. This means that it resembles a capacitor with the poly and the semiconductor serving as the conducting plates.

Figure 18 shows a cross section of such a capacitor, with the field lines extending across the oxide dielectric and the bias across the plates such that the semiconductor is in depletion mode. The thickness of the dielectric and depletion regions are $d$ and $x_d$, respectively. The existence of the depletion region means that the p-type substrate has been depleted of holes at the surface; the potential on the top plate has attracted electrons to the surface and so the acceptor dopant ions located there have their vacancies filled.
Figure 17. Cross section of a PRISM LDMOS Transistor.
Figure 18. Cross section of a MOS capacitor with a positive potential on the top plate.
been depleted of holes at the surface; the potential on the top plate has attracted electrons to the surface and so the acceptor dopant ions located there have their vacancies filled.

For this type of structure, it has been demonstrated that 1) the normalized maximum electric field at the semiconductor surface is a function of the dielectric (silicon-dioxide) thickness and the width of the depletion region in the semiconductor\(^{43}\left(\frac{d}{x_0}\right)\), and 2) this field is minimized when the top plate structure is implemented using multiple electrodes in small steps\(^{44}\), and when the oxide dielectric increases outward away from the MOS structure\(^{45}\).

The LOCOS field oxidation thus serves a very important role: the gradually increasing oxide thickness will serve to optimize the field profile, and the poly gate extension will resemble a series of electrode steps over this region. This helps minimize the electric field build up at the surface of the semiconductor and give the transistor a higher, more reliable breakdown voltage.

Breakdown will occur when the depletion edge from the n+/n- drain contact in the n-well extends through the boron and into the n+(source)/backgate depletion edge. The following figures will help to demonstrate the LDMOS transistor action. Figure 19 shows vectors of the electron current in the transistor under low bias conditions: the gate potential is at 2 volts and the surface of the silicon is now inverted, and the drain is also at 2 volts. The plot clearly indicates the direction of the current flow inside the LDMOS transistor structure.

Figure 20 and shows simulation results of the internal potential contours for a drain bias of 42 volts, with the gate bias at zero volts. In this case there is no current flow but the depletion edges can be seen to move away from the p backgate/n-well junction. The key is that depletion edges for reverse biased junctions “spread” more rapidly through lighter (n+) doped regions as opposed to the higher doped (p) backgate region; in this case
Figure 19. Electron current vectors for $V_{gs} = 2$ volts, $V_{ds} = 2$ volts.
Figure 20. LDMOS potential contours for $V_{ds} = 42$ volts.
the large nwell area offers plenty of space for depletion region movement and therefore it can "absorb" the high drain voltages.

The field effect due to the poly gate extension helps keep the potential edges from moving too 'rapidly' across the drift region. This makes the device less vulnerable to premature breakdown which can result if the potential contours are not aligned to the top plate and thus allowed to curve in towards the drain diffusion edge. BV_{ds} values greater than 60 volts\(^{46}\) are obtained for high side drive devices and up to 90 volts\(^{47}\) breakdown for devices in low side drive configurations. Figure 21 shows the lateral DMOS transistor in a high side drive configuration\(^{48}\).

An important feature of power devices is the on-state resistance. For a lateral DMOS transistor this consists of two major components: the n\(^{-}\) drift region below the field oxide and the resistance due to the channel region\(^{49}\). The channel resistance can be expressed as

\[
R_{ch} = \frac{1}{W \frac{L_{eff}}{\mu_{eff} C_{ox}(V_{gs} - V_{t})}}
\]

where \(\mu_{eff}\) is the effective mobility of the carriers in the channel. \(W\) is the width of the transistor and \(L_{eff}\) is the effective channel length. This inversion layer mobility is a function of the dopant level and also the vertical gate electric field. It will be demonstrated later in this chapter that the gate bias has a dominant effect on this parameter. The resistance due to the drift region of the transistor can be approximated by

\[
R_{drift} = \rho_{s} \frac{L}{W} \quad (\rho_{s} = \frac{1}{q\mu N_{d}})
\]
Figure 21. LDMOS transistor in a “High Side” output driver configuration.
where $\rho_s$ is the bulk resistivity of the silicon. $N_d$ is the dopant level in the nwell drift region and $L$ in this case is the total length of the drift region.

Figure 22 shows the dependence of $R_{ds(on)}$ on the gate bias. It is interesting to note that this resistance will increase for higher temperatures, even though the general shape of the curve will remain the same. Part of this increase is due to the decrease of the carrier mobility in silicon at higher temperatures. Figure 23 shows how the resistance increases linearly with increasing operating temperature.

**Circuit Model Development Methodology**

A robust methodology was implemented for developing the LDMOS SPICE models. The main points are the following: 1) creation of a statistical database consisting of LDMOS transistor I-V data; 2) identifying via a histogram and normal distribution function the drive currents for "nominal" devices and the wafers on which these were measured; 3) for the nominal wafer selected, measured standard data required for device modeling such as gate oxide thickness, threshold voltage and C-V characteristics; and 4) I-V data measurement and model fitting using the parameter extraction tool IC-CAP$^{50}$. Figure 24 shows a flow chart detail including all relevant parameters necessary for a complete model description of the device. Process and device simulation was carried out in parallel with physical measurements as part of the modeling effort.

A discussion on statistical methods as they are employed in quality and productivity improvement in industry will now be given, with a particular emphasis on process variation and normal distribution processes. These topics are essential for introducing the device modeling strategy. This will be followed by the details of the modeling effort and comparisons between simulated and measured data for the LDMOS transistor.
Figure 22. On Resistance vs. Gate bias for the LDMOS transistor.
Figure 23. Variation of LDMOS on resistance with increasing temperature.
Figure 24. Flow Chart Detail for LDMOS SPICE Model Development.
A discussion on statistical methods as they are employed in quality and productivity improvement in industry will now be given, with a particular emphasis on process variation and normal distribution processes. These topics are essential for introducing the device modeling strategy. This will be followed by the details of the modeling effort and comparisons between simulated and measured data for the LDMOS transistor.

Statistics and Process Capability

The emergence of worldwide competition in manufacturing industries during the late 1970's brought dramatic changes to the way U.S. manufacturers approached efforts for quality and productivity improvement. In particular, the ideas of Deming51, Juran52 and Taguchi53 became very popular in the 1980's as U.S. companies tried to regain their share of world markets and build a strong quality base54. Several key ideas serve as the foundation of modern efforts in quality improvement including, 1) the use of statistically designed experiments for optimizing fabrication and manufacturing processes and 2) the use of statistical methods for characterizing process variation, process capability and process control. Statistically designed experiments were already discussed in Chapter 1 under the heading “Design of Experiments.” Statistics for process characterization and their relation to device circuit models will be described next.

Variations in fabrication processes. IC fabrication processes are defined by nominal parameters which are expected to yield devices with appropriate operating characteristics as required by the customer. It is a fact of life, however, that not all devices will behave in exactly the same manner; similar devices fabricated under the same process conditions will exhibit variations in their characteristics. There are many sources of noise, or variations, during the fabrication process. Differences in the way operators handle equipment is one source, and external factors such as equipment response to change in temperature or humidity are examples of other sources of noise introduced during
fabrication. Another source of variability is due to imperfections in the equipment itself. These type of variations can be taken into account during the IC design and production phase in order to reduce their impact on the final product.

Processing variations will always exist, but the important thing is that the final products meet the required specification limits. In other words, target specifications for a particular parameter are decided upon, with appropriate upper and lower tolerance limits. For example, a nominal gate oxide thickness of 500 Angstroms may have tolerance limits of +/- 10%. This means that any gate oxide measured between 450 and 550 Angstroms will be within tolerance and therefore will meet the process specifications. These type of design specifications are set in anticipation of equipment imperfections and other variations in the fabrication process.

In this example of gate oxide thickness, it is expected that many measurements of this parameter will yield a normal distribution centered about the target value of 500 Angstroms. The normal distribution, or “bell curve,” is one of the tools used for making judgments on the manufacturing process. The mathematical definition of the normal distribution is

\[
f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2}
\]

(13)

where the mean \( \mu \) is defined as

\[
\mu = \frac{\sum_{i=1}^{N} x_i}{N}
\]

(14)

and the standard deviation
\[
\sigma_x = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N-1}}.
\]

The standard deviation and the mean characterize the distribution; the mean gives the center of the distribution while the standard deviation gives an indication of how the process is spread around the center. Figure 25 gives the detail for this distribution, with the definitions for standard deviation and the mean.

**Process capability.** Process capability refers to the amount of variability which is found in measurements of a specific parameter. Mathematically it is defined as

\[
C_p = \frac{USL - LSL}{6\sigma}
\]

Where *USL* refers to the upper specification limit and *LSL* refers to the lower specification limit for the process. Figure 26 shows how the process capability, *Cp*, is calculated for a given distribution. For a symmetrical distribution, the area between the 3\(\sigma\) points on each end represent 99.7% of the measured quantity. If the process is in control and *Cp* is greater than 2.0, then the manufacturing process is said to be capable of meeting the specifications.

**Statistical process control.** This refers to the use of statistical methods for ensuring products are manufactured within specifications. In particular, the process should yield products which are centered about the nominal characteristics which are used to gauge quality, and these should exhibit 6\(\sigma\) performance; i.e., they should all be distributed within the specified tolerance limits.

Finally, process control aims to minimize losses in manufacturing by using a systematic approach including the use of Pareto charts and other tools for identifying root
Figure 25. Normal distribution and definitions for the mean and standard deviation.

\[ \mu = \frac{\sum_{i=1}^{N} X_i}{N} \]

\[ \sigma_x = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N - 1}} \]
$C_p = \frac{USL - LSL}{6\sigma}$

Figure 26. Definition of process capability.
causes in failures during the development stage. In this setting, Pareto diagrams measure
the number of occurrences, or frequency, of all possible failures. This is an organized way
of understanding how frequently a specific type of failure occurs relative to all others.
When using Pareto histogram charts in a diagnostic way, they help isolate the most
frequently occurring problems and therefore show what the priorities should be and where
energies should be expended in order to correct any problems.

The idea is that while failures of a particular device may be caused by several
defects, only a few of those defects may be causing the majority of the failures. Pareto
charts help identify these type of trends. Figure 27 is an example of a Pareto histogram
used to identify the failure mechanisms in a part. Concentrating on the first two defects
will clearly make the most difference in the yield for the parameters in question.

When process control tools are used in product development the goal is to
eliminate scrap material and continuously improve the process and its capability over time.
Figure 28 shows how employing statistical process control techniques and diagnostic tools
such as Pareto charts can lead to better product performance over time. The goal is to
minimize the amount of spread of measured parameter values and be centered as tightly as
possible about the mean.

In comparing the data on the transistors to be modeled, the two most important
characteristics for the LDMOS transistor were the drive current capability and the gate
oxide thickness. The data allowed a comparison of drive currents under different
conditions of gate and drain bias. This was important in order to ensure that the nominal
device selected exhibited nominal characteristics in both the linear and saturation regions
of operation. Typical drive current distribution is shown in Figure 29. As can be seen, not
all transistors exhibit the same drive current characteristics, even though they were
fabricated using the same process. The data instead shows that the distribution of currents
Figure 27. An example of a Pareto Chart identifying failure causes in a part.
Figure 28. Employing diagnostic tools aids in improving the process capability over time.
Figure 29. Typical distribution of the drive current for the LDMOS transistor.
approximates a normal distribution, and a similar plot can be made of the gate oxide thickness measurements.

The key is that the mean value, the point represented by the most common value of measured current, is the centered value of the drain current for this device. Of all the currents which were measured, that is the value which appeared with the highest frequency. The existence of transistors with "stronger" and "weaker" currents under similar bias conditions results from the tolerance limits on the fabrication process. In other words, the process is defined for a nominal set of parameters but tolerance specifications set by equipment limitations and other noise factors will lead to a distribution of these parameters.

Since the devices fabricated represent the process capabilities it is important to have a SPICE model which will account for the device performance spread. These are known as the Weak/Nominal/Strong circuit models and they represent the upper, lower and centered performance regions of the fabricated LDMOS transistors. Having these three models is very important for the circuit designers whom must have access to accurate data on the devices. The circuit simulators are capable of examining such details as time delays between logic gates and the number of devices which can be driven by a single output gate. These calculations, however, are only as good as the models they have available.

Circuit designers deal with two sources of perturbation of component parameters: environmental and process. Environmental factors of primary importance are supply voltage and junction temperature. Processing variations impact component parameters like threshold voltage and gain. When designers need to know the "worst-case" conditions under which a circuit must still operate as expected, it is important that they have available the weak and strong models of the devices they are using. If the concern is that high drive current may damage a particular gate or circuit being driven, then the strong model will
indicate what level of current can be expected if the driver happens to be a strong device. If a device has to drive more than one gate, then the weak model will indicate what the maximum fanout is under those conditions and the designs can be tuned accordingly.

Details of the modeling efforts for the LDMOS transistor follow. Processed wafers were tested for functionality and yield. Using modern data collection software such as ICMS, a statistical database was collected with critical data on all devices such as threshold voltage, drive current characteristics and gate oxide thickness. In this study the database consisted of 360 transistors taken from 45 wafers from 19 lots.

Nominal LDMOS Modeling

Statistical database. The data gathered for the transistors were analyzed in order to find the nominal device characteristics. The database included die, wafer, and lot number identification so that tracing it back for verification was feasible. A "lot number" refers to a specific batch of wafers that were fabricated as a group. Each wafer fabricated was labeled with a laser ablated number. Besides verification of the drive current data, measurements of gate oxide thickness and net interface charge were made using standard capacitance-voltage techniques.

Selection of nominal device from distribution. The nominal device was selected from the drain current data available for operating regions of interest to both digital and analog designers. A cross reference with the oxide thickness served to narrow the selection process. The device selected exhibited mean values for the measured currents and for the gate oxide thickness which matched the target values.

Characteristics measurement and SPICE model fitting. SPICE (Simulation Program with Integrated Circuit Emphasis) is perhaps the most widely distributed computer program for circuit design and simulation in the world\(^5\). The program can model different kinds of electronic devices including diodes, bipolar transistors, capacitors
and field effect transistors. It can simulate DC, transient, and AC characteristics for individual devices or simulate performance characteristics for circuits containing up to thousands of devices. In addition, it can simulate temperature effects as they impact diode and bipolar transistor currents and carrier mobility in MOS transistors.

The keys to gaining meaningful information from the program are the parameterized physical and empirical analytical models it contains for predicting the device and circuit performance characteristics. For example, in order to accurately model the DC characteristics of a MOS transistor, it is necessary to specify the parameters which will affect the threshold voltage and drain current calculations. In addition to device geometry and terminal bias information, process technology dependent parameters such as the mobility degradation factor need to be specified.

A review of the equations necessary for DC modeling of MOS transistors will be given next. Important physical parameters necessary for LDMOS model development will be identified, and the experimental techniques used in parameter extraction will be described. The methods of empirical parameter fitting to the SPICE equations will also be given.

The threshold voltage for MOS transistors is modeled by the equation,

\[
V_t = V_{to} + BE(\sqrt{V_{sb}} + 2\phi_f - \sqrt{2\phi_f}) - DE\sqrt{V_ds},
\]

(17)

where \(BE\) is the body effect parameter which helps model the impact of a substrate bias \(V_{sb}\) on MOS transistor threshold voltage, and \(V_{to}\) is the threshold voltage for \(V_{sb}=0\). For the case of LDMOS, however, the source is at the same potential as the backgate contact, so the impact of the body effect parameter in this case is zero. \(DE\) is the drain effect parameter and it accounts for the modulation of the threshold voltage by the drain depletion edge extending into the channel region at high drain bias. For the case of an
LDMOS transistor, this parameter is effectively zero since the drain drift region is lightly doped compared to the backgate region.

$V_{to}$ is given by the following equation,

$$V_{to} = V_{FB} + 2|\psi_0| + \frac{Q_{dep}}{C_{ox}} + \frac{Q_{ii}}{C_{ox}}$$

(18)

where the flatband voltage is given by

$$V_{FB} = \phi_m - \phi_t - \frac{Q_{net}}{C_{ox}}$$

(19)

the bulk, or substrate potential is

$$\psi_b = -\frac{KT}{q} \ln\left(\frac{N_a}{n_i}\right)$$

(20)

the depletion region charge is

$$Q_{dep} = \sqrt{2\varepsilon_0 q N_a 2|\psi_0|}, \text{ (units in charge/unit area)}$$

(21)

and $Q_{ii}$ is the $V_t$ adjust ion implant charge,

$$Q_{ii} = q N_u, \text{ (units in charge*ions/unit area)}$$

(22)

The difference between the gate work function and semiconductor bulk potential is shown as $\phi_m - \phi_t$ and the remaining term in (19) refers to a net interface charge located at the semiconductor surface. This net charge in the oxide and the silicon surface can consist
of fixed oxide charge, interface trapped charge, mobile oxide charge and trapped oxide charge\textsuperscript{56}. This net interface charge is different for different processes, but variation per process is minimal, usually on the order of 20 millivolts or less. The reader is asked to consult the reference for a detailed discussion on the origins of these various terms. It is important to note that the dominant terms will be the sum of the fixed oxide charge and the trapped oxide-semiconductor interface charge\textsuperscript{57}. $N_a$ refers to the net substrate doping concentration in units of ions per unit volume, $n_i$ is the intrinsic silicon carrier concentration level in ions per unit volume, and $\varepsilon_s$ is the dielectric constant of silicon (units of Farads/per unit length). The units of capacitance in the equation above are Farads/unit area.

SPICE uses two different model equations to predict drain current, one for the linear region and another for the saturation region. In the linear region, the drain current takes the form

\begin{equation}
I_{ds} = \left( \frac{\mu \varepsilon_{ox}}{t_{ox}} \right) \left( \frac{W_{eff}}{L_{eff}} \right) (V_{gs} - V_t - \frac{V_{ds}}{2}) V_{ds} (1 + \lambda V_{ds})
\end{equation}

where linear region is defined by ($V_{ds} < V_{gs} - V_t$), and ($V_{gs} > V_t$). $V_{ds}$ is the applied voltage between the drain and the source terminals. $V_t$ is the threshold voltage as defined above. $W_{eff}$ and $L_{eff}$ refer to the effective channel width and length of the transistor, $\mu$ refers to the carrier mobility in the channel region, and $\varepsilon_{ox}$ and $t_{ox}$ refer to the dielectric constant and thickness of the gate oxide layer, respectively. The empirical parameter $\lambda$ is an attempt to model the channel length modulation due to the drain depletion edge at high voltages.

In the saturation region, where ($V_{ds} > V_{gs} - V_t$), the drain current takes the form

\begin{equation}
I_{ds} = \left( \frac{\mu \varepsilon_{ox}}{t_{ox}} \right) \left( \frac{W_{eff}}{L_{eff}} \right) (V_{gs} - V_t)^2 (1 + \lambda V_{ds})
\end{equation}
The carrier mobility parameter $\mu$ has units of cm$^2$/volt*second, and is a complex parameter which depends on the substrate doping, temperature, and net electric fields. Models for bulk mobility in silicon as a function of temperature and doping concentration have been defined and are well known; in the channel region, however, the impact of the gate electric field must be accounted for as well as the fact that the inversion layer contains a high concentration of minority carriers. An empirical relation which tries to model these effects gives the mobility as

$$\mu = \frac{\mu_0}{(1 + \theta(V_{gs} - V_t))}$$

(25)

where $\mu_0$ is the bulk mobility. $L_{eff}$ is used for the channel length parameter since lateral diffusion of the source and drain often shorten the actual length of the channel from that given by the length of the gate. This can be modeled approximately as

$$L_{eff} = L - 2X_{jd}.$$  

(26)

Where $X_{jd}$ is the spread of the source/drain diffusions under the gate. This lateral diffusion effect is shown in Figure 14. Since the LDMOS transistor channel is formed by inverting the backgate diffused region under the poly, the channel length is fixed at about 1 micron so $L_{eff}$ in this case is the same as $L$. Additionally, the actual width of a MOS device may not be as indicated by the layout design. Imperfections in photoresist etch, LOCOS oxidation encroachment into the active area, or lateral diffusion of the channel stop implant may lead to an effective width which is smaller than the design indicates. Figure 30 gives an example of LOCOS and boron encroachment into a device active area along its width.
Figure 30. An example an $W_{eff} < W_{design}$ due to boron and LOCOS encroachment.
In the case of PRISM LDMOS transistors, effective width characterization is particularly important due to the layout of the device. Figure 31 shows the way these transistors are layed out using a "race-track" type of design. Notice the polysilicon gate and its curvature at the ends. This type of device layout makes the effective width calculation necessary as it is not immediately clear what effect the curved gate structure has on net current flow through the device.

One method of extracting the effective width is to measure the maximum gain of the transistor, $k'$, as a function of different design widths. This gain parameter is defined as

$$k = \left( \frac{\mu_{\text{ox}}}{t_{\text{ox}}} \right) \left( \frac{W}{L} \right),$$

and a plot of its maximum value, $k'$, as a function of design width will extrapolate linearly to the effective width correction factor $\Delta W$. The effective width can then be calculated as

$$W_{\text{eff}} = W_{\text{design}} - \Delta W.$$

Figure 32 shows an example plot of $k'$ vs. $W_{\text{design}}$ which results in an effective width smaller than the design width. The maximum value of $k$ can be extracted directly from the transconductance data for a MOS transistor, which is defined as the derivative of the drain current with respect to the gate bias,

$$g_m = \frac{\partial I_d}{\partial V_{\text{gs}}}.$$

Methods of measuring transconductance will be described shortly. Finding the effective width of the devices is important for scalable models. The device circuit models should be able to predict performance for transistors fabricated at different widths. During
Figure 31. Plan View of an LDMOS transistor. $W_{\text{design}}$ is $\sim 2W_o$. 
Figure 32. A plot of $k'$ vs. $W_{\text{design}}$ to find $W_{\text{eff}}$. 

$W_{\text{eff}} = W_{\text{design}} - \Delta W$

$k'$ (amps/volt)

$\Delta W$

$W_{\text{design}}$ (microns)
the model development it was found that models fitted the design width were not able to accurately predict characteristics for devices fabricated with different widths. Using the method outlined above, the model performance for devices with different widths improved significantly.

The form of the model equations as described above have not changed very much since they were first used in circuit simulators in the late 1960's. The only thing that has changed over the past 15 years or so has been the addition of "short channel" parameters (not necessarily indicated in the above equations) $\alpha$, $\lambda$, $\theta$, $\gamma$, and $DE$ in order to account for effects due to the shrinking of device feature sizes. The impact of these short channel effects on device characteristics will be described in the following chapter.

The model equations which are shown in (17), (23), and (24) are for the so-called Level 1 SPICE model set. SPICE provides different levels of model definition, each with different forms of equations, usually with parameters intended to increase the accuracy of the simulation. For example, Level 2 and Level 3 equations include terms for substrate bias effect on the drain current, etc. Other variations of similar models exist, each with their relative merits. The point is that 1) standard SPICE simulators use parameterized equations for modeling MOS transistors, and 2) these models are used to fit to the measured data in order to define the circuit models for those devices. Model definition implies finding the appropriate values for the different parameters in the equations. Integrated circuits fabricated using different processes must be modeled separately and the circuit designers must use the models developed for their specific technology during design.

Given the parameterized model equations, SPICE uses "equivalent circuits" to model the transistor action. This type of analysis was first defined by Shiichman and Hodges in 1968\textsuperscript{16}. The equivalent DC circuit model for the LDMOS transistor is shown in
Figure 33. Important parameters include the diode reverse saturation currents and the capacitance between the gate and source/drain diffusions.

The parameters which affect these internal structures must be measured in order to have an accurate SPICE model. For example, the diode saturation currents and capacitances are measured using special test structures built on the wafer. Test structures are a common way of finding parameters which are dependent on the process such as the resistivity of the polysilicon gate and the capacitances and currents associated with p-n junctions.

In the case of the p-n diode currents, the test is set up by examining the magnitude of the current under reverse bias. Diode current is defined as

\[ I = J_s (e^{qV_d / nkT} - 1) A, \]

(30)

where \( k \) is Boltzmann's constant, \( T \) is the temperature in Kelvins, \( q \) is the electron charge, \( V_d \) is the applied voltage, \( A \) is the diode area, and \( n \) is the "ideality" factor which at low and reverse currents has a value of about 1.5. \( J_s \) represents the saturation current density for a p-n junction and is usually in the range of \( \sim 10^{-16} \) amperes/cm\(^2\). Figure 34 shows how the reverse bias saturation current is measured experimentally.

Figure 35 shows how a MOS capacitor can be set up to make C-V measurements. Many important parameters can be measured from this type of test including the gate oxide thickness, \( t_{ox} \), the substrate doping \( N_A \) or \( N_d \), the flatband voltage, and net surface charge concentration. Figure 36 shows a typical C-V curve for a capacitor with a p type substrate. As the voltage is swept from negative to positive, the surface of the silicon goes from accumulation (holes attracted to the surface by the negative charges on the top plate) to depletion as the voltage starts becoming positive. Finally, at a high enough positive
Figure 33. Equivalent circuit model for the LDMOS transistor.
Figure 34. Measuring the reverse saturation current in a PN diode.
Figure 35. Capacitance-Voltage experiment with a MOS capacitor.
Figure 36. Capacitance-Voltage plot for a P type MOS capacitor.
voltage the surface is inverted and the total capacitance is the oxide capacitance in series with the depletion region capacitance.

It is important to note that this type of characteristic is only true for high frequency measurements. At low frequencies, recombination-generation kinetics of the electrons in the inversion region will vary with respect to the AC signal and the measurement will reflect these variations rather than the depletion region capacitance. In this case the resulting C-V curve will rise back up to its maximum value at high positive bias.

At negative bias, the measured capacitance is the maximum value and this is due mainly to the oxide serving as a dielectric between the poly and the surface which is in accumulation. This condition resembles a parallel plate capacitor. The gate oxide thickness can be calculated from the relation,

\[
\tau_{ox} = \frac{\varepsilon_{ox} A}{C_{max}}
\]

where \( C_{max} \) is equal to the maximum measured capacitance. The Area \( A \) is the junction area of the device.

A very important parameter to extract from the C-V curve is the net surface charge \( Q_{ss} \). Shown as \( Q_{net} \) in equation (19), the value of this charge will have a significant effect on the threshold voltage since it shifts the flatband voltage from the nominal value. If the workfunction difference \( \phi_{m} = \phi_{m} - \phi_{s} \) between the gate material and the substrate is known, then accurate measurements of this charge can be made using the same C-V plot data. This value can be extracted by calculating the shift in flatband voltage from the ideal case to the measured value for \( V_{FB} \).

\[
Q_{ss} = (\phi_{m} - V_{FB}) C_{max}.
\]
The measured value of flatband voltage is the point which corresponds to the flatband capacitance as shown in Figure 36. The capacitance in the semiconductor at this value of applied voltage is due to the series capacitance of the oxide and capacitance due to excess carriers generated at the surface\textsuperscript{58}. This carrier concentration will decay to zero after a characteristic length, known as the debye length\textsuperscript{59}, so that the total flatband capacitance will be given by

\begin{equation}
C_{FB} = \frac{EoxEsi}{EoxEsi + EoxLD}
\end{equation}

where the Debye length is defined as\textsuperscript{60}

\begin{equation}
LD = \frac{EoxkT}{q^2Na}.
\end{equation}

During a C-V measurement, the appropriate values for \(N_D\), the substrate doping level, as well as the bulk potential \(\phi_B\) can be found iteratively to obtain self consistent solutions for all of the measured and calculated quantities.

Two other parameters which are extracted from experiment are the threshold voltage and the maximum gain, \(\kappa'\). The threshold voltage is extracted from the gate characteristics, which is a measure of the drain current as a function of the gate bias. The drain terminal in this case is held constant and at very low bias, such as 100 millivolts. Figure 37 shows a schematic set up for measuring the gate characteristics. The threshold voltage is extracted using the linear extrapolation of the maximum slope of this curve. The intersect of this line with the voltage axis represents the threshold voltage. Figure 38 gives
Figure 37. Gate Characteristics Experiment for the Lateral DMOS FET.
Figure 38. Simulated gate characteristics for the LDMOS transistor.
an example of a gate characteristic curve with the linear extrapolation of its slope used to find \( V_T \). The transconductance is the derivative of the drain current with respect to the applied gate bias as shown previously in equation (28). Using equation (23), the derivative will yield

\[
g_m = \frac{\partial I_d}{\partial V_{gs}} = \left( \frac{\mu\varepsilon_{ox}}{t_{ox}} \right) \frac{W_{eff}}{L_{eff}} f(V_{ds})
\]

where \( f(V_{ds}) \) is a constant since \( V_{ds} \) is held constant during the experiment. The units for the transconductance are Amperes/volt. An example of this curve for the LDMOS transistor is shown in Figure 39. The peak value of the transconductance is defined as the maximum transistor gain, \( k' \), and it will increase for larger transistor widths. This parameter can be used to find the effective width for different transistor geometries as described earlier.

Since Lateral DMOS transistors are often subjected to high temperature conditions it is necessary to characterize the dependence of important parameters such as the maximum gain, \( k \), and threshold voltage, \( V_T \), on temperature. The experiments described earlier for measuring threshold voltage and transistor gain are then repeated at elevated temperatures in order to complete the characterization. Figure 40 shows the LDMOS threshold voltage dependence on temperature. The temperature dependence of \( k \) is given by

\[
k(T_2) = k(T_1)\left(\frac{T_1}{T_2}\right)^{1.5}.
\]

The threshold voltage varies with temperature via the bulk potential, \( \phi \), and this can be modeled approximately as
Figure 39. Transconductance characteristics for the LDMOS transistor. Dotted line indicates the measured data; straight line is the simulation.
Figure 40. LDMOS normalized threshold voltage as a function of temperature.
\[ \phi(T_2) = \frac{T_2}{T_1} \phi(T_1) - \frac{2kT_2}{q} \ln\left(\frac{T_2}{T_1}\right)^{1.5} . \]

With the process dependent parameters identified as described above, IC-CAP was then used to measure the DC characteristics directly from the wafer and the models fitted using the remaining parameters as the fitting coefficients. The mobility degradation factor proved to be a very important fitting parameter as most of the other parameters are intended to model short channel effects and were not critical in this case. The data which were used to find the remaining (mobility degradation, etc.) model parameters were the drain characteristics. In this kind of experiment a family of curves is obtained as the drain bias is swept from the linear into saturation regions for several gate biases. Figure 41 shows how the experiment is set up and Figure 42 shows typical LDMOS drain characteristics curves.

Parameter model fitting is typically done using least squares optimization, a standard numerical approach for semiconductor parameter extraction. For a given arbitrary model \( f \) of the form

\[ y = f(x, p) \]

with input vector parameters \( x \) and model parameters \( p \), the goal is to find the appropriate values of \( p \) which will match the measured values \( y \). The idea is to minimize the error function defined by

\[ E = \| y - Y \| . \]

The input vectors in this case are the applied voltages and the outputs are the predicted
Figure 41. Drain Characteristics Experiment for the Lateral DMOS FET.
Figure 42. Simulated drain characteristics for the LDMOS transistor. The plot shows drain current as a function of drain bias for different gate voltages.
and measured currents. In optimizing the fitting parameters, it was found that SPICE had difficulty in approximating and fitting to the saturation point, right around the pinch off. This is not surprising, since this is also the point where the discontinuity in the drain current model occurs. More will be said about this in Chapter 4.

Simulation strategy. The overall goal of the simulation approach was to match the current-voltage data measured for typically fabricated LDMOS transistors. The process flow, including thermal cycles and ion implant steps was simulated at nominal conditions to match both the substrate dopant level and the gate oxide thickness as measured from the nominal device.

The process flow steps required for LDMOS transistor fabrication are shown in Figure 43. All of these processes were simulated using the standard tools described in Chapter 2. Since the goal was to match the device transfer characteristics, issues related to isolation were not relevant, and it was only necessary to model the operating area of the device structure. This means that the only lateral diffusion and oxidation effects were those related to the boron backgate diffusion and the LOCOS oxidation. The process steps up to and including the HV Nwell implant and anneal which do not have any lateral dependence (for this case) were run using the one dimensional program SUPREM-3. This is straightforward to implement since the simulation environment has UNIX compatibility so that all structure files can be imported directly by using the “source” command in SUPREM-4.

The two dimensional simulator then completed the remaining process flow steps which included the lateral boron diffusion for the backgate and also the LOCOS oxidation step. The last step simulated was the passivation which takes place before first level metalization (not shown in the figure but comes after the source/drain contact etch). This is appropriate as the remaining steps are at low temperatures so the impact on impurity
ALIGN
N+ BURIED LAYER
EPI
High Voltage NTANKS
DEEP N+
TANK DRIVE 1
Low Voltage NTANKS
High Voltage PTANKS
Low Voltage PTANKS
DWELL
TANK DRIVE 2
DEEP P+
MOAT
CHANNEL STOP
FIELD OXIDE
FAMOS Floating Gate
HV Vt adjusts, HV gate ox
LV Vt adjusts, LV gate ox
Tunnel Diode
Poly Gates
EPROM STACK
SOURCE/DRAIN

Figure 43. PRISM Process Flow.
re-distribution in the substrate is minimal. The completed 2-D structure was the starting point for electrical simulations using MEDICI.

The next step was to match the nominal device electrical characteristics. The gate characteristics were an appropriate starting point in order to match the threshold voltage and the transconductance characteristics. This type of simulation only requires about 31 points to complete, as opposed to drain characteristics which could take up to 100 or more, and several simulations were required to evaluate the various mobility models available in MEDICI.

Following the experimental approach, gate characteristics simulations were carried out and compared to the measured data. After comparison of different mobility models, the models selected for the I-V simulations were the ARORA, PRPMOB and FLDMOB models available in MEDICI. The analytical models were found to have very similar transfer characteristics, and the field dependent models appeared to result with fairly accurate representations of the measured drain currents. The ARORA model is an analytical model which was derived from modified Brooks-Herring theory for mobility and empirical fitting to experimental data over different temperatures\(^ {63}\). Electron mobility using the ARORA model is defined as

\[
(40) \quad 88\left(\frac{T}{300}\right)^{-0.57} + \frac{1252 T_n^{-2.33}}{1 + \left(\frac{N(x, y)}{1.26 \times 10^{17} T_n^{2.4}}\right)^{887^n - 6144}}.
\]

This model was defined by minimizing the error between experimental data and the mobility expression by using the constant coefficients as fitting parameters. The model was found to be valid over a wide range of temperatures (250 - 500 K) and for doping levels in silicon as high as \(10^{20}\) cm\(^{-3}\). A similar expression is available for modeling hole mobility in PMOS devices.
The PRPMOB model takes into account the effect of the gate electric field on the carriers in the channel region. The expression for mobility takes the form,

$$\mu = \frac{\mu_0}{\sqrt{1 + \frac{E_1}{ECN.MU}}} \tag{41}$$

where $ECN.MU$ is the "critical" perpendicular electric field with default value 6.49E4 V/cm. When two mobility models are selected, MEDICI will select the "low field" model as the base mobility to calculate the more complicated field dependent mobility. In this case, for example, the ARORA predicted mobility will be $\mu_0$ in equation (41).

The FLDMOB model takes into account the lateral (drain to source) electric field effect on the mobility and it takes the form,

$$\mu = \frac{\mu_0}{[1 + \left(\frac{\mu_0 E_{II}}{v_{sat}}\right)^\beta]^{\frac{1}{\beta}}} \tag{42}$$

where $E_{II}$ is the parallel electric field through the structure, $v_{sat}$ is the carrier saturation velocity and $\mu_0$ is the low field mobility. The default value for $\beta$ is 2, and the saturation velocity is defined as $10^7$ cm/sec, which is the saturation velocity for bulk silicon. A similar expression also exists for hole mobility. This model is appropriate to use since the electric fields in the drain/drift region of the transistor can be very large under high drain bias.

The process and device simulations were run almost entirely using the default coefficients for diffusion, segregation, etc., as set by the vendor, TMA, including the coefficients in the mobility models. The electrical simulations allow user-defined values for various process-dependent quantities such as the surface charge density $Q_{ss}$ and the
contact resistance. Following the idea of using experimentally derived parameters to match measured results, the value for the net interface charge density extracted from the C-V measurements was used, and excellent agreement was found between the measured and simulated gate characteristics. In addition, distributed contact resistances were used to account for the series source/drain terminal resistance of the actual metal contacts. Appropriate values for these resistances were found by comparing the magnitude of the currents measured from the fabricated devices. The contact resistance values as were found from the gate characteristics were used throughout the rest of the computer experiments.

Using the measured surface charge density and the simulation observed contact resistance, the drain characteristics experiments were then repeated as on the actual device. Excellent agreement was found between the nominally simulated LDMOS transistor and the measured nominal device characteristics. A slightly smaller value for the saturation velocity in the FLDMOB model was used in order to get better agreement with the measured characteristics. Published values in the literature were used as a guide for finding a value which provided a better match with the measured data. In this case the literature indicates that electron saturation velocity in the channel inversion layer can be anywhere from $5 \times 10^6 \text{ cm/sec}^{66}$ to $6.5 \times 10^6 \text{ cm/sec}^{67}$. A possible explanation for having to adjust this parameter from the default value is that the carrier saturation velocity in an LDMOS might be lowered due to the high currents and electric fields causing lattice heating in the structure during high bias conditions$^{68}$.

Comparison of Simulated and Measured Data. Figure 44 shows the results of the drain characteristics simulation using the matched nominal structure. The gate and transconductance comparisons are in Figures 38 and 39, respectively. Good agreement can be found between the measured and simulated data for all cases. The only parameter which was adjusted was the carrier saturation velocity from the FLDMOB mobility model.
Figure 44. Comparison between measured and simulated drain curves for the nominal device. Dotted line indicates the simulation results.
Weak and Strong SPICE Models

From the statistical database, the data was cross checked, as with selecting the nominal device, between the measured currents and the gate oxide thicknesses in order to identify the weak and nominal devices. Similar measurements as with the nominal device were carried out for SPICE model characterization. Simulations were also done in order to match the fabrication process tolerance for the gate oxide thickness as measured on the devices. In this case the backgate boron dose was also set to upper and lower tolerance limits in order to match possible combinations of process variation according to the specifications. Table 5 shows the process conditions which were simulated for the weak/strong devices.

Comparison of Simulated and Measured Data. Figures 45 and 46 compare the results of the simulated and measured drain characteristics on the weak and strong devices. Very close agreement with the measured data can be seen in all cases. The processed structures gave I-V results comparable to that measured on typical fabricated devices. It is important to stress that these electrical simulations were run using the contact resistance calibration data found earlier for the nominal device. In addition, the surface charge density is a process dependent parameter and this value also remained the same. This means that no additional work was necessary to match the weak and strong selected structures.
**TABLE 5**

**PROCESS PARAMETER CONDITIONS FOR MATCHING WEAK AND STRONG LDMOS TRANSISTORS.**

(All values are relative to the nominal conditions).

<table>
<thead>
<tr>
<th>Gate Oxide Thickness</th>
<th>Backgate Boron Dose</th>
<th>Type of Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>+10%</td>
<td>+10%</td>
<td>Weak</td>
</tr>
<tr>
<td>-10%</td>
<td>+10%</td>
<td>Strong</td>
</tr>
<tr>
<td>+10%</td>
<td>-10%</td>
<td>Weak</td>
</tr>
<tr>
<td>-10%</td>
<td>-10%</td>
<td>Strong</td>
</tr>
</tbody>
</table>
Figure 45. Measured and simulated drain characteristics for the weak LDMOS device.
Figure 46. Measured and simulated characteristics for the strong LDMOS device.
Summary and Concluding Remarks

A robust methodology for SPICE model development was implemented for the case of LDMOS transistors. Extensive data was gathered from a database available from earlier measurements. A simulation approach was implemented in which fabrication characteristics were matched by the simulated structures. Typical I-V data as used in SPICE model fitting was compared and good agreement was found between simulation and measurement. The data gathered from these simulated structures can be used for sensitivity analysis and future device/circuit model development work for PRISM technology.

It is important to note that the simulation experiments were carried out over several days' time in comparison to the fabricated wafers which take about 6 weeks. This included the time for evaluation of various mobility models available in MEDICI. For LDMOS characterization the use of several models resulted in very accurate I-V results.

These type of simulations are important for several reasons. Process development work can be enhanced with the availability of powerful workstations and simulation tools which can accurately predict device sensitivities to changes in process conditions. In addition, computer tools offer internal structure information that is often not readily available from physical measurement, such as the exact location of current flow inside a device and the magnitudes of electric fields under different bias conditions. This type of information is extremely valuable during process development when layout rules and process guidelines are still being developed. To summarize, computer experiments coupled with physical measurement offer process and device engineers insights which otherwise would be very difficult to obtain.

In particular, when compared with the cost and time involved in processing experimental wafers and subsequent testing, measurement, and deconvolution of results, simulation experiments can provide meaningful information at a fraction of the cost which
can then be used by process, device and circuit design engineers, all in just a matter of
days. An advantage of computer experiments is that they are controlled experiments: the
user specifies exactly what process flow is implemented and how. Actual wafer fabrication
is subject to processing variation, and this uncertainty makes the extraction of relevant
information more difficult and time consuming.
CHAPTER 4
HYPERBOLIC MODEL FOR MOS TRANSISTOR MODELING

A new analytical model for MOSFET modeling has been developed and characterized. Unlike analytical drain conductance models available in circuit simulators such as SPICE, the "Hyperbolic Model" exhibits continuity of the drain conductance \( (g_{ds}) \) in transition from the linear to saturation region of operation, and integration of this analytical equation leads to a continuous form of the drain current equation. In addition, the new model accounts for non-uniform substrate doping leading to a more accurate simulation of substrate bias effects in all regions of device operation (subthreshold, linear and saturation) than currently available models. The small signal conductances with respect to the gate and substrate as predicted by the Hyperbolic Model have also been derived and characterized.

A review of standard MOS equations as currently available in circuit simulators such as SPICE will be presented along with the derivation of the proposed model and examples of its use in predicting MOS transistor characteristics.

Objectives of MOSFET Modeling

MOS transistor modeling is an integral part of VLSI design. As has been demonstrated, current-voltage data for MOS devices can be obtained from computer simulation experiments or measured directly from the fabricated wafers. This data yields useful process information such as threshold voltage and transistor gain which are some of the parameters in device circuit models. The device models are parameterized analytical equations which are used to represent the actual device characteristics in analytical form.

Besides process parameters such as threshold voltage and transistor gain, empirical parameters are also used by the models to account for parasitic effects such as mobility
degradation and carrier velocity saturation. These parameters are used to fit the device circuit models to the experimental data using standard least-squares algorithms. Device circuit models are very important as they are the only source of transistor level information available for verifying and simulating VLSI circuits\textsuperscript{69}.

**The Four Terminal MOS Transistor**

Figure 47 shows the cross section of a typical NMOS transistor. Included are all the biasing terminals including the gate, source, drain and substrate. A circuit model should be able to predict the output current characteristics as a function of any of these terminal voltages. A brief review of MOS transistor modeling methods follows. This will include a discussion of short channel effects in MOS transistors and currently available SPICE models for MOSFET operation.

**The Shichman-Hodges Equivalent Circuit Model.** Shichman and Hodges of Bell Laboratories developed a representation of MOS transistors which has become a standard tool for device and circuit simulation\textsuperscript{70}. Their “equivalent circuit” model divides a MOS transistor into a combination of discrete components including capacitances, diodes, current sources and resistors. Figure 48 shows a typical Shichman-Hodges model of a MOS transistor used for DC analysis. Since each of the discrete components are well understood, a linear system of equations can be solved to find the electrical characteristics of the transistor. LDMOS drain curves as in Figures 45 and 46 are typical of “long-channel” drain characteristics of MOS transistors. These type of characteristics were common for devices fabricated through the late 1970’s. Dramatic changes in minimum feature sizes and processing conditions for CMOS technologies led to secondary non-ideal effects which will be summarized in the following section.
Figure 47. Cross Section of an NMOS transistor including substrate contact.
Figure 48. Shichman-Hodges Equivalent Circuit for MOS transistors.
Models Incorporating Short Channel Effects. Deviations from the long channel characteristics are due to several side effects resulting from modern CMOS processing. Figures demonstrating short channel characteristics will be given shortly. Particular parameters which are important to model in short channel devices include\textsuperscript{71} 1) mobility degradation due to surface scattering and vertical electric field; 2) the drain effect (DE in the figures), or lowering of the threshold voltage due to the drain depletion region extending deep into the channel region at high drain bias; 3) the velocity saturation effects (\(\alpha'\) and \(\gamma\) in the figures) for carriers in the channel at high field conditions; 4) the channel length modulation (\(\lambda\)) due to the drain depletion spreading towards the source at high bias, and 5) the body effect, which accounts for the increasing of the threshold voltage when a substrate bias is applied.

Mobility degradation effects have been shown in the previous chapter for the case of LDMOS transistors. This type of effect on the MOS transistor is not too difficult to picture. Intuitively, it makes sense that as the channel region of a transistor is inverted, the high concentration of minority carriers will lead to a higher probability of collisions and scattering effects. Mobility degradation at high gate fields and also at high temperatures reduces the maximum gain, \(k'\).

At high drain bias, it has already been shown that the depletion region edge extends into the channel region of the device. As the channel length decreases, this leads to a lower threshold voltage. The lower effective threshold voltage leads to an increase in the drain current, even at high drain bias, since the drain current is proportional to \(V_{gs}-V_t\).

At high electric fields the carrier velocity in the channel region saturates as the lateral electric field reaches a "critical" value. The drift velocity can be written as
\[ v_d = \mu E_{\text{ill}} \]

where \( \mu \) is the mobility and the electric field \( E \) is along the lateral direction (parallel to current flow). The critical electric field magnitude can then be approximated given the mobility of the carriers and the fact that saturation velocity for electrons will be on the order of \( 10^7 \text{ cm/sec} \).

The channel modulation factor also affects the slope of the \( I_{dS} - V_{dS} \) curve in saturation. This effect is due to the drain depletion edge moving in towards the source depletion when the device is turned off. This leads to a smaller channel length which results in increased transistor gain and therefore an increase in the drain current when the device is turned on.

The body effect is not really a "short channel" effect, but rather just a consequence of the MOS transistor structure allowing for a substrate (or well) bias. This is important, for example, in floating gate memory applications, where voltage pumps are used to charge tunnel diode diffusions, forcing charge onto the storage cell via tunneling mechanisms. The effect of a reverse bias applied between the substrate (or a well) and the source (which is normally left at ground potential), is to increase the depletion width in the channel region, and this leads to an increase in the gate voltage necessary to cause strong inversion, so the result is an increase in \( V_T \).

Most of these effects are difficult to characterize and model using simple analytical equations, so device engineers have had to relate their functionality empirically from observations during experiment. In that sense some of the changes which have been made (and will be demonstrated shortly) to the simple drain current equations as they were presented in Chapter 3 are not supported directly from theoretical results but rather simply because they can reasonably model the actual characteristics. Device and circuit designers require predictive analytical models to get accurate results in a reasonable amount of time.
Subthreshold region of operation. Subthreshold operation refers to the condition such that the gate potential has not reached a level high enough to cause inversion in the channel. This condition can exist for either high or low drain bias, with or without substrate bias. One form of the drain current equation in this region is given by\(^7\)

\[
I_{ds,th} = kps \frac{W_{eff}}{L_{eff}} \left[ e^{-\frac{\frac{W_{eff}}{L_{eff}} (V_{bs} - \psi_s + 2\phi_f)}{\left[1 - e^{-\frac{W_{eff}}{L_{eff}} V_{ds}}\right]}} \right] (-NgKp \frac{W_{eff}}{L_{eff}} V_{ds})
\]

for the condition \((-V_{sb} + V_{FB}) \leq V_{gs} < V_t\). For \(V_{gs} > V_t\), this component becomes

\[
I_{ds,th} = kps \frac{W_{eff}}{L_{eff}} (-NgKp \frac{W_{eff}}{L_{eff}} V_{ds}) [1 - e^{-\frac{W_{eff}}{L_{eff}} V_{ds}}]
\]

Equation 45 is added to the total drain current at above threshold conditions in order to keep the transition from subthreshold to above threshold “smooth.” In the above equations, \(N_g\) is the subthreshold emission coefficient and \(K_{ps}\) is a subthreshold conduction factor. These fitting parameters were introduced by Yang and Chatterjee for subthreshold current modeling\(^7\).

Figure 49 shows the subthreshold characteristics as a function of gate and substrate bias. It is clear that the SPICE model has difficulty in predicting the overall effect of substrate bias on MOS operation in the “weak” inversion region.

In particular, the substrate bias causes the threshold voltage to increase, and this results in the “shift” of the curves as the substrate bias increases. In the SPICE equations,
Figure 49. SPICE (dotted line) subthreshold gate characteristics as a function of substrate bias compared to measured data.
the body effect is the parameter which modulates the substrate bias effects, and it is part of
the threshold voltage calculation. This parameter, however, is considered a constant which
depends on a uniform substrate doping. It will be shown later that an empirical
relationship can be derived for the body effect and its dependence on non-uniform
substrate doping which will lead to a more accurate simulation of the current
characteristics as a function of substrate bias.

Linear and saturation regions of operation. Drain current for above threshold
conditions include equation (45) added to

\begin{equation}
I_{ds} = \alpha' \beta \left[ 2(V_{gs} - V_t) V_{ds} - \alpha' V_{ds}^2 \right] / \left[ 1 + \theta(V_{gs} - V_t) \right]
\end{equation}

in the linear region where \( V_{ds} \leq \frac{V_{gs} - V_t}{\alpha'} \), and

\begin{equation}
I_{ds} = \beta \left( V_{gs} - V_t \right)^2 \left[ 1 + \lambda \left( \frac{\alpha}{\alpha'} \right)^2 \left( V_{ds} - V_{dsat} \right) \right] / \left[ 1 + \theta(V_{gs} - V_t) \right]
\end{equation}

in saturation, where \( V_{ds} > \frac{V_{gs} - V_t}{\alpha'} \).

The short channel parameters are defined in the following way,

\begin{equation}
\alpha' = \alpha + \gamma (V_{gs} - V_t),
\end{equation}

where \( \alpha \) is the mobility ratio for the carriers in going from the linear to the saturation
regions. This accounts for the lower effective mobility at high drain bias. The parameter \( \gamma \)
is the velocity saturation factor which accounts for the high lateral electric fields achieved
during saturation. For short channel devices, \( \alpha' \) is greater than 1 and it has the effect of
lowering the saturation voltage to below \( V_{gs} - V_t \):
\[ V_{dsat} = \frac{V_{gs} - V_t}{\alpha'}. \]

The parameter shown as \( \beta \) is just another way of expressing \( k \),

\[ \beta = \frac{\mu_{e} \varepsilon_{ox} W_{eff}}{2 t_{ox} L_{eff}} \]

As was previously shown, the threshold voltage for MOS transistors is modeled as

\[ V_t = V_{to} + BE(\sqrt{V_{bs} + 2\phi_f} - \sqrt{2\phi_f}) - DE \sqrt{V_{ds}}. \]

where the substrate bias is \( V_{bs} \) and the body effect parameter \( BE \) is intended to model the extent to which the substrate bias affects the channel inversion region. The body effect is normally defined as

\[ BE = \frac{\sqrt{2q\varepsilon_{ox}N_A}}{C_{ox}} \]

where \( C_{ox} \) is in units of Farads/unit area and the rest of the parameters take the usual meanings. The substrate doping \( N_A \), and hence, the body effect, is assumed to be constant in this case. \( DE \) refers to the drain effect factor which was described earlier.

Figures 50 and 51 show the drain current curves predicted by currently available models. The drain characteristics are shown for the case of zero substrate bias in Figure 50. The effect of a substrate bias on the same device are shown in Figure 51. These are submicron transistor characteristics and hence they demonstrate short channel effects in comparison to the LDMOS transistor curves. Since the threshold voltage, \( V_t \), increases for
Figure 50. SPICE model fit of the drain characteristics for zero substrate bias.
Figure 51. SPICE model fit of the drain characteristics for non-zero substrate bias.
increasing substrate bias, it makes sense that the drain current will decrease since $I_{ds}$ is proportional to $V_{gs}-V_t$. The SPICE equations do a reasonable job for predicting the drain characteristics with zero substrate bias as indicated by Figure 50. It is clear, however, that the model has difficulty in predicting the net effect of a substrate bias on the drain current, just as was true in the subthreshold region of operation.

Part of this difficulty is due to the fact that the channel region of modern MOSFETs includes ion implants for tuning the threshold voltage. In equation (52) it is predicted that the body effect term will depend on a constant substrate doping level. This has been demonstrated to cause errors in MOS modeling\textsuperscript{74}. Figure 52 shows the doping profile for a typical small channel NMOS transistor. The profile is seen to be non-uniform due to the threshold voltage adjust implant and because ion implants lead to gaussian type of impurity distributions in the substrate. Assuming a constant substrate doping is clearly invalid, especially when one considers that Figure 52 is a log plot of the concentration; the difference in the order of magnitude between surface and substrate doping is not trivial. An empirical formula for calculating the drain effect, together with a new form of the body effect will lead to better simulation results for MOS transistor characteristics as a function of both substrate and drain bias in all regions of operation.

Additionally, the small signal drain conductance, which is defined as the derivative of the drain current with respect to the drain bias is discontinuous in current SPICE models due to the way it models drain current. Figure 53 shows the drain conductance data as measured from an actual transistor, for different gate potentials. Figure 54 shows how the SPICE equations for drain conductance compare to the data in Figure 53. This type of comparison is the starting point towards improving the SPICE models.

**Improving the SPICE models.** One way to examine the error encountered in MOS drain current modeling is by examining the drain conductance as is predicted by the equations given in (46) and (47). This method is due to Redwine, published in 1985\textsuperscript{75}. The
Figures 52. Doping profile (logarithm scale) through the channel region of an NMOS transistor.
Figure 53. Measured drain conductance characteristics $\left( \frac{\partial I_{ds}}{\partial V_{ds}} \right)$. 
Figure 54. Fitted SPICE model for the drain conductance compared to measurement.
Figure 55. SPICE definition of the saturation voltage.
point can be made clear by examining Figure 55. In the figure, the saturation voltage as determined by SPICE is compared to the actual saturation in a real transistor. It can be seen that the error will be due to the discrepancy between the calculated area under the conductance curve and the actual area from measured data. The following definitions will make the writing easier,

\[
\alpha' = \alpha + \gamma (V_{gs} - V_t) \Rightarrow \alpha_g \\
\beta' = \frac{1}{1 + \theta (V_{gs} - V_t)} \Rightarrow \beta_g
\]

and additionally, \( V_{gs} - V_t \Rightarrow V_{gst} \).

In the linear region, the drain conductance will be given as

\[
G_{dl} = \frac{\partial I_{ds(\text{linear})}}{\partial V_{ds}} \big|_{V_t = \text{near}}.
\]

and applying this to equation (46) will yield

\[
G_{dl} = \frac{\partial}{\partial V_{ds}} (\beta_g (2 \alpha_g V_{gst} V_{ds} - \alpha_g^2 V_{ds}^2))
\]

and after rearranging the derivative calculation in the form of a line \( y = mx + b \) where the independent variable in this case is \( V_{ds} \), the following form of the drain conductance is obtained,

\[
G_{dl} = -(2 \beta_g \alpha_g^2) V_{ds} + (2 \beta_g \alpha_g V_{gst}) \quad \text{or} \\
G_{dl} = M_{gl} V_{ds} + G_{dlo}
\]
valid over the drain bias range $V_{ds} \leq V_{dsat}$. The key points here are the y-axis intercept given by

$$G_{d10} = 2\beta_8 \alpha_8 V_{gst},$$

and the x-axis intercept will be

$$V_{dsat} = \frac{V_{gst}}{\alpha_8}.$$ 

The drain conductance in the saturation region may be derived using equation (47). The derivation will yield,

$$G_{ds} = \frac{\partial I_{ds(saturation)}}{\partial V_{ds}}|_{(V_{gst} = \text{const.})} = \beta_6 \lambda V_{gst}^2 \frac{\alpha_8}{\alpha_g}.$$ 

The discontinuity in the drain conductance predicted by SPICE arises precisely at the drain voltage saturation point, $V_{dsat}$, as shown previously in Figure 55.

A way to remove the “jump,” and hence reduce the error due to the discontinuity of the drain conductance, is to define a new $V_{dsat}$, at the point where $G_{ds} = G_{d1}$, as shown in Figure 55. The new x-axis intercept will be given by

$$V_{dx} = \frac{V_{gst}}{\alpha_8}.$$ 

Using equations (55) and (58), and setting $G_{ds} = G_{d1}$ the appropriate value of $V_{ds}$ can be found (call it $V_{dsat}$).
\[ V_{ds_{\text{ain}}} = \left( \frac{V_{gs}}{\alpha_g} \right) \left[ 1 - \frac{\lambda V_{gs}}{2 \alpha_g} \right]^2. \]

The drain current will now be given by the integration of the new drain conductance in both the linear and saturation regions. This will be calculated from

\[ I_{ds} = \int_{0}^{V_{gs}} G_{dl} dV_{ds} + \int_{V_{gs}}^{V_{gs}} G_{ds} dV_{ds} \]

and referring back to equation (55),

\[ I_{ds} = \int_{0}^{V_{gs}} (M_{g} V_{ds} + G_{dlo}) dV_{ds} + \int_{V_{gs}}^{V_{gs}} G_{ds} dV_{ds}. \]

After evaluating the integrals we arrive at the new form of \( I_{ds} \),

\[ I_{ds} = M_{g} l \left( \frac{V_{ds_{\text{ain}}}^2}{2} \right) + G_{dlo} V_{ds_{\text{ain}}} + G_{ds} (V_{ds} - V_{ds_{\text{ain}}}) \]

From the slope as shown in Figure 55, it can be shown that

\[ G_{dlo} - G_{ds} = -M_{g} l V_{ds_{\text{ain}}} \]

so rearranging equation (63) and making the substitution in (64), the final form of the current becomes

\[ I_{ds} = \beta_{g} \lambda V_{gs}^2 \left( \frac{\alpha}{\alpha_g} \right)^2 V_{ds} + \beta_{g} \alpha_g^2 V_{ds_{\text{ain}}}^2 \quad \text{for} \quad V_{ds} \geq V_{ds_{\text{ain}}} \]

and
\begin{equation}
I_{ds} = \beta g (2\alpha g V_{gs} V_{ds} - \alpha g^2 V_{ds}^2) \quad \text{for } V_{ds} \leq V_{dsat}.
\end{equation}

A Plot of the drain conductance model as shown in Figure 55 using the new definition for the saturation current is shown in Figure 56. The fitting error is much smaller now and it would be expected that this method would lead to a better fit between the model and the measured data. The discontinuity problem still remains, however, and even if the fit is better, there is still error due to the model not approximating the curvature of a real device. The form of the drain conductance as derived above will subsequently be referred to as the "NSH" model of the drain conductance.

As device and circuit simulation become more important in custom circuit design and new technology development these type of modeling errors can lead to expensive delays in yielding functioning circuits. Redwine demonstrated examples of these errors in 1985. Tsividis and Suyama recently presented a comprehensive review of modern issues in MOSFET modeling, and these issues have received a considerable amount of attention in recent months. Ideally, all MOS models for the small signal conductances should exhibit continuity with respect to any of the four terminal voltages in a MOS device and in all regions of operation, including the subthreshold region and above threshold (linear) regions. A model which overcomes some of these limitations is proposed and will be described next.

**The Hyperbolic Model**

Redwine approached the problem of MOS modeling errors by attempting to gain insight from the form of the MOS drain conductance. This is an attempt to understand the origin of the errors that result in typical SPICE model equations. The drain conductance
Figure 56. A new definition of the saturation voltage for use with the Hyperbolic Model.
describes the change in drain current with respect to the drain potential. This is expressed as,

\[ g_{ds} = \frac{\partial I_{ds}}{\partial V_{ds}}. \] (68)

The drain current can then be calculated by integrating this expression,

\[ I_{ds} = \int g_{ds} dV_{ds}. \] (69)

The motivation behind this approach is that if one can find an accurate representation of equation (68), representing the drain conductance for MOS transistors, then the integration performed in (69) will lead to an accurate model for drain current.

Redwine's approach led him to the use of a hyperbolic empirical relationship for modeling MOSFET drain conductance. He showed that by integrating this form of the drain conductance an accurate model for drain characteristics resulted which gave excellent results in above threshold characteristics for the case of zero substrate bias.

In what follows, Redwine's method for deriving a hyperbolic empirical relationship for drain conductance is reviewed. Following this review, an updated version of the model is presented. The new version of the Hyperbolic Model includes accurate modeling capability of substrate bias effects in all regions of MOSFET operation: subthreshold, linear and saturation. The small signal conductance equations for the gate \( g_{m} \) and substrate \( g_{mb} \) have been derived and characterized as well. The details of these derivations will also be presented. Thus a new, complete analytical model for MOSFET simulation and modeling is presented which overcomes deficiencies in current MOS models.
Derivation of the drain conductance \((g_{ds})\) equation. The proposed model for MOSFET modeling begins with the drain conductance. The functional form of the drain conductance, \(g_{ds}\), for a MOS transistor was shown earlier in Figure 53. An empirical relationship can be made relating this type of characteristic to a hyperbola. By comparing standard short channel MOS model equations for the drain conductance and making the appropriate change of variables, a continuous analytical model of the drain conductance can be derived. The details of this follow.

Consider a generalized hyperbola in the x-y plane defined by the following equation,

\[
\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1,
\]

and shown in Figure 57. Making appropriate substitutions to rotate the axes about an angle \(\phi\), where \(\phi > 0\),

\[
x = x' \cos \phi - y' \sin \phi \\
y = y' \cos \phi + x' \sin \phi
\]

where \(\sin \phi = \frac{a}{r}\) and \(\cos \phi = \frac{b}{r}\) and \(r^2 = a^2 + b^2\),

the following equation can be derived for the rotated hyperbola,

\[
ktm \cdot y'^2 + 2 \cdot kcs^{-1} \cdot x' \cdot y' - r^2 = 0
\]

where

\[
ktm = \left( \frac{1}{\tan^2 \phi} - \tan^2 \phi \right)
\]

\[
kcs = \cos \phi \sin \phi \]

\[
\tan \phi = \frac{a}{b}
\]
Figure 57. A general hyperbola.
It will be demonstrated shortly why it is desirable to translate this function in both the x and y directions; letting \( y' = y - y_o \) and \( x' = x - x_o \), equation (70) becomes

\[
(74) \quad ktn(y - y_o)^2 + 2kcs^{-1} (x - x_o)(y - y_o) - r^2 = 0.
\]

The solution for \( (y - y_o) \) may be solved for using the quadratic formula,

\[
(75) \quad (y - y_o) = \left( \frac{1}{ktn \cdot kcs} \right) \left( \sqrt{(x - x_o)^2 + ktn \cdot kcs^2 \cdot r^2} - (x - x_o) \right).
\]

Using the following identities,

\[
(76) \quad \frac{ktn \cdot kcs^2}{ktn \cdot kcs} = \cos w
\]

\[
\frac{1}{ktn \cdot kcs} = -\frac{1}{2} \tan w
\]

\[
w = -2\phi,
\]

equation (75) then becomes

\[
(77) \quad y = y_o - \frac{1}{2} \tan w \sqrt{(x - x_o)^2 + r^2 \cos w - (x - x_o)}.
\]

If we transform the variables as shown in Figure 58 where the translated hyperbolic function is overlapped with the new drain conductance formulation, it can be seen that a match will exist if we let \( x_o = V_{dsatn} \) and \( y_o = G_{ds} \). \( G_{dlo} \) and the angle \( W \) are also marked on the figure in accordance with the drain conductance as derived from the “NSH” equations earlier (“Improving the SPICE models”). The final parameters and relationships which form part of the model are presented and defined in Appendix A.
Figure 58. Translated hyperbola to develop an empirical relationship with the MOSFET drain conductance.
Equation (77) written in terms of the "NSH" drain conductance model presented earlier becomes

\[ G_d = G_{ds} - \frac{1}{2} M_{gl} \sqrt{(V_{ds} - V_{dsath})^2 + Vr^2 \cos \theta - (V_{ds} - V_{dsath})} \]

and following the methodology stated in equations (68) and (69), integrating equation (78) will give the drain current for the above threshold conditions:

\[ I_{ds} = G_{ds} \cdot V_{ds} - \frac{1}{2} M_{gl} \cdot V_{dsath} \cdot V_{ds} - \frac{1}{4} M_{gl} \cdot V_{ds}^2 - \frac{1}{4} M_{gl} \cdot F(V_{ds}) + \frac{1}{4} M_{gl} \cdot F(0) \]

where each of the parameters is defined in the Appendix.

Figure 59 shows drain conductance data and the model prediction for 1 micron channel length transistors. The drain current predicted for this transistor is shown in Figure 60. Excellent agreement can be seen between the measured and simulated data.

**Subthreshold and substrate bias modeling.** The proposed model includes expressions for the subthreshold current and empirical relationships for the body and drain effect which lead to more accurate prediction of MOSFET characteristics. As was shown in Figure 52, the substrate doping in modern MOS transistors is not uniform. This is due to: 1) the threshold voltage tuning implants which lead to a non-uniform concentration of dopant just at the surface relative to the rest of the substrate, and 2) the fact that ion implants are used to create the well regions in modern processing technology leads to "gaussian" doping profiles as was described in Chapter 2. The proposed form of the body effect also takes into account the drain effect, that is, the potential in the channel region is
Figure 59. Hyperbolic Model fit of the drain conductance for 1 micron transistors.
Figure 60. Hyperbolic model fit of the drain characteristics at zero substrate bias.
also affected by the drain depletion region making the substrate conditions more complicated. The equation for the body effect takes the form,

\begin{equation}
B_{ef} = (BE + (BES - BE)e^{-(BEV*Vbs)^2})e^{-(BDS*\sqrt{Vds})}
\end{equation}

where \(BES\) is the maximum value of the body effect and is due to the surface concentration, and \(BE\) is the minimum body effect value due to the substrate doping level. The drain effect expression just follows the simple relationship that the drain effect varies with the square root of the drain bias as it currently exists in SPICE. \(BEV\) is a scaling factor for the substrate bias. The equation has been used to fit modern submicron MOS transistors and it will be shown that it leads to accurate prediction of drain current with respect to substrate, gate and drain bias in all regions of operation.

The equation was derived by setting limits for the value of the \(B_{ef}\) between the surface (the \(B_{ef}\) takes its upper limit) and the substrate (the lowest value of \(B_{ef}\)). The form of the body effect in between these two limits is assumed to be a gaussian distribution, similar to the shape of the implanted substrate profile. The square root dependence of the drain effect is well known and has been used by SPICE many years for predicting drain bias effects on threshold voltage and drain current. The addition of this term with a scaling factor \(BDS\) to the body effect leads to an accurate representation of the subthreshold and above threshold MOS characteristics.

The form of the subthreshold equation is similar to models that are currently available. It takes the form,

\begin{equation}
I_{subth} = kps \frac{W_{eff}}{L_{eff}} (1 - e^{-\frac{NG*V_{ds}}{Kdq*Tk}}) e^{\frac{NG(U_s - V_{bs} - 2\phi_f)}{Kdq*Tk}}
\end{equation}
where $Us$ is given by

$$Us = Vn + \left(\frac{B_{ef}^2}{2}\right)(1 - \sqrt{1 + 4\frac{Vn}{B_{ef}}}^2)$$

and $T_k$ refers to the temperature in Kelvins. $Vn$ in equation (81) is the total surface potential in the channel due to the net sum of all voltages,

$$Vn = V_{gs} + V_{bs} - V_{fb} + DE\sqrt{V_{ds}}.$$

In equation (81) the $K_{ps}$ term refers to the temperature dependence on the transistor gain and is given as

$$K_{ps} = k^* TCKP\left(\frac{T_k}{296.78}\right)$$

where $TCKP$ is the temperature fitting coefficient extracted from experiment as described in Chapter 3. The flatband voltage, $V_{fb}$, in this case is given by

$$V_{fb} = VTO - 2\phi_f - B_{ef}\sqrt{2}\phi_f,$$

where $VTO$ is the zero substrate bias threshold voltage.

Let equation (81) be referred to as $I_{sub1}$. During program execution, the total drain current will depend on the gate biasing conditions as established by the following outline:

if $V_{gs} - Vt < 0$, 


(86) \[ I_{ds} = I_{subth} \quad \text{(where } I_{subth} \text{ is } I_{sub1}) \]

\[ \text{else} \]

(87) \[ I_{subth} = K_{ps} \frac{W_{eff}}{L_{eff}} (1 - e^{-\left(\frac{N_{G} \cdot V_{ds}}{K_{q} \cdot q \cdot T_{k}}\right)}) \]

(88) \[ I_{dtotal} = I_{subth} + I_{ds} \]

(where \( I_{ds} \) is given by equation (79)).

The subthreshold component for the total drain current given in equation (88) (from now on referred to as \( I_{sub2} \)) is intended to provide a smooth transition between the subthreshold and above threshold conditions. Figures 61 and 62 show comparisons between measured and fitted data using the Hyperbolic Model, including the new forms of the body effect and subthreshold equations.

As can be seen, very close agreement can be achieved using this new model. In particular, the substrate bias effect simulation is greatly improved using the gaussian form of the body effect, both in the below threshold (Figure 61) and the above threshold characteristics (Figure 62).

Circuit simulators require models of the DC characteristics (as has been derived thus far for the new model) as well as the small signal equations (for AC modeling) in order to be able to completely predict MOS transistor characteristics. The small signal terms are the derivative of the drain current with respect to the drain, gate, and substrate bias. The model was derived from the drain conductance, so the remaining terms are the gate and substrate conductances. These equation as predicted by the Hyperbolic Model will be derived next.
Figure 61. Hyperbolic model fit of the subthreshold characteristics for non-zero substrate bias.
Figure 62. Hyperbolic model fit of the drain characteristics at non-zero substrate bias.
Derivation of the transconductance \( g_m \) equation. The transconductance is defined as the derivative of the drain current with respect to the gate bias. Mathematically this is shown as,

\[
g_m = \frac{\partial I_d}{\partial V_{gs}}.\tag{89}\]

For \( V_{gs} < V_t \), the drain current is \( I_{sub1} \) as defined in equation (81). Therefore it is shown as

\[
g_m = \frac{\partial I_{sub1}}{\partial V_{gs}} \quad \text{(for } V_{gs} \text{ below the threshold voltage).}\tag{90}\]

This derivative will evaluate as

\[
\frac{\partial I_{sub1}}{\partial V_{gs}} = I_{sub1}(\frac{NG}{K_{dq} * T_k}) \frac{\partial U_s}{\partial V_{gs}}\tag{91}\]

from equations (82) and (83) it can be seen that

\[
\frac{\partial U_s}{\partial V_{gs}} = 1,\tag{92}\]

therefore

\[
g_m = \frac{\partial I_{sub1}}{\partial V_{gs}} = I_{sub1}(\frac{NG}{K_{dq} * T_k}) \quad \text{for } V_{gs} < V_t.\tag{93}\]
The parameter \( V_{gs} - V_t \) appears frequently in equation (79) for the drain current in the above threshold conditions, so this will be shown simply as \( V_{gst} \). The equation for the transconductance will then become

\[
g_m = \frac{\partial I_{ds}}{\partial V_{gst}},
\]

and it will keep the same functionality as equation (86) since

\[
\frac{\partial I_{ds}}{\partial V_{gst}} = \frac{\partial I_{ds}}{\partial V_{gs}} \frac{\partial V_{gs}}{\partial V_{gst}},
\]

and

\[
\frac{\partial V_{gs}}{\partial V_{gst}} = 1.
\]

The units for the transconductance are Amps/volt.

The analytical form of the transconductance equation for \( V_{gs} > V_t \) will then consist of the derivative of each term in (79) which is a function of the gate voltage, \( V_{gs} \). This derivative will take the form given by

\[
\frac{\partial I_{ds}}{\partial V_{gst}} = V_{ds} \frac{\partial G_{ds}}{\partial V_{gst}} - \frac{V_{ds}}{2} \left( M_{gl} \frac{\partial V_{dsath}}{\partial V_{gst}} + V_{dsath} \frac{\partial M_{gl}}{\partial V_{gst}} \right) - \frac{V_{ds}^2}{4} \frac{\partial M_{gl}}{\partial V_{gst}}
\]

\[
+ \frac{1}{4} \left( M_{gl} \frac{\partial F(V_{ds})}{\partial V_{gst}} + F(V_{ds}) \frac{\partial M_{gl}}{\partial V_{gst}} \right) + \frac{1}{4} \left( M_{gl} \frac{\partial F(0)}{\partial V_{gst}} + F(0) \frac{\partial M_{gl}}{\partial V_{gst}} \right)
\]
in the above threshold conditions (since \( \frac{\partial I_{\text{sub}2}}{\partial V_{gs}} = 0 \)). The full details of the derivation is given in the Appendix.

**Derivation of the \( g_{mb} \) Equation.** This is the conductance of the drain current with respect to the substrate potential. Mathematically this is defined as,

\[
(98) \quad g_{mb} = \frac{\partial I_d}{\partial V_{bs}}.
\]

This derivative will depend heavily on both the subthreshold and above threshold equations due to the body effect dependence. Below threshold the equation becomes

\[
(99) \quad \frac{\partial I_{\text{sub}1}}{\partial V_{bs}} = I_{\text{sub}1} \left( \frac{NG}{k_{dq} T_k} \right) \left[ \frac{\partial U_s}{\partial V_{bs}} - \frac{\partial V_{bs}}{\partial V_{bs}} \right].
\]

From equations (82) and (83) the first derivative in (99) becomes

\[
(100) \quad \frac{\partial U_s}{\partial V_{bs}} = 1 - \frac{\partial V_{fb}}{\partial V_{bs}} + \frac{\partial}{\partial V_{bs}} \left[ \frac{B_{df}^2}{2} \left( 1 - \sqrt{1 + 4 \frac{V_{in}}{B_{df}^2}} \right) \right]
\]

and substituting this in to (99) the expression becomes

\[
(101) \quad \frac{\partial I_{\text{sub}1}}{\partial V_{bs}} = I_{\text{sub}1} \left( \frac{NG}{k_{dq} T_k} \right) \left[ -\frac{\partial V_{fb}}{\partial V_{bs}} + \frac{\partial}{\partial V_{bs}} \left[ \frac{B_{df}^2}{2} \left( 1 - \sqrt{1 + 4 \frac{V_{in}}{B_{df}^2}} \right) \right] \right].
\]

Each derivative term will be evaluated next. The derivative of the flatband voltage with respect to the substrate bias is given by (the flatband voltage is defined in (85))

\[
(102) \quad \frac{\partial V_{fb}}{\partial V_{bs}} = -\sqrt{2\Phi_r} \frac{\partial B_{df}}{\partial V_{bs}},
\]
where

\begin{equation}
\frac{dB_{ef}}{dV_{bs}} = (B_{ES} - BE)e^{-\frac{BEV_{bs}}{B_{es}}}e^{-\frac{B_{DS}V_{ds}}{B_{ef}}}\left(-2BEV^{2}V_{bs}\right).
\end{equation}

For simplicity, let \( F_1 \) be equal to part of the remaining derivative expression in (101):

\begin{equation}
F_1 = \left(1 - \sqrt{1 + \frac{4V_{bs}}{B_{ef}^{2}}}ight).
\end{equation}

That derivative expression will then evaluate to

\begin{equation}
\frac{\partial}{\partial V_{bs}}\left(\frac{B_{ef}^{2}}{2}F_1\right) = F_1\cdot B_{ef}\frac{\partial B_{ef}}{\partial V_{bs}} + \frac{B_{ef}^{2}}{2}\frac{\partial F_1}{\partial V_{bs}}
\end{equation}

where

\begin{equation}
\frac{\partial F_1}{\partial V_{bs}} = -\frac{1}{2}\left(1 + \frac{4V_{bs}}{B_{ef}^{2}}\right)^{\frac{1}{2}}\left[\frac{4}{B_{ef}^{2}}\left(1 - \frac{\partial V_{fs}}{\partial V_{bs}}\right) + 4V_{n}\left(-2B_{ef}^{-3}\frac{\partial B_{ef}}{\partial V_{bs}}\right)\right].
\end{equation}

This completes the terms for the subthreshold \( \xi_{mb} \) equation. At above threshold conditions equation (79) will have to be evaluated. It is of interest to note that in this case the substrate potential only influences the threshold voltage terms. The bulk conductance equation for above threshold conditions can therefore be derived using this fact and the following relationship,

\begin{equation}
\frac{\partial I_{ds}}{\partial V_{bs}} = \frac{\partial I_{ds}}{\partial V_{gst}}\frac{\partial V_{gst}}{\partial V_{t}}\frac{\partial V_{t}}{\partial V_{bs}}.
\end{equation}
The individual terms will evaluate as

\[(108)\]
\[
\frac{\partial V_{gst}}{\partial V_t} = -1,
\]

\[(109)\]
\[
\frac{\partial I_{ds}}{\partial V_{gst}} = g_m,
\]

therefore equation (107) will reduce to

\[(110)\]
\[
\frac{\partial I_{ds}}{\partial V_{bs}} = -g_m \frac{\partial V_t}{\partial V_{bs}},
\]

where \(g_m\) was already derived earlier. Equation (51) gives the complete expression for \(V_t\), and the derivative with respect to \(V_{bs}\) will yield

\[(111)\]
\[
\frac{\partial V_t}{\partial V_{bs}} = \frac{1}{C_{ox} \frac{\partial Q_{dep}}{\partial V_{bs}}} + \frac{1}{2 \sqrt{V_{bs} + 2 \phi_f} \frac{\partial B_{ef}}{\partial V_{bs}}}.
\]

The derivative of the depletion charge will evaluate as

\[(112)\]
\[
\frac{\partial Q_{dep}}{\partial V_{bs}} = \frac{qN_a \epsilon_s}{\sqrt{2qN_a \epsilon_s (2 \phi_f + V_{bs})}}.
\]

and so the resulting expression for \(g_{mb}\) in above threshold conditions will be
\[
\frac{\partial I_{ds}}{\partial V_{bs}} = -g_m \frac{q N_0 e_s}{C_{ox} \sqrt{2q N_0 e_s |2 \phi_f + V_{bs}|}} + \frac{1}{2 \sqrt{|V_{bs} + 2 \phi_f|}} \frac{\partial B_{ef}}{\partial V_{bs}},
\]

where the remaining derivative was defined earlier (103).

**Conclusions**

A new model for MOSFET simulation has been derived. The work builds on an earlier version of MOS drain conductance which led to accurate drain current modeling for zero substrate bias conditions. The updated Hyperbolic Model can now accurately represent drain currents as a function of gate and substrate potential in MOS transistors in all regions of operation: subthreshold, linear and saturation. Although currently available models can fit well to drain current data in above threshold conditions, errors exist due to the discontinuous from of the drain conductance.

In addition, currently available models for substrate bias effects on drain current are less than adequate. The proposed model accurately represents the effects of non-uniform substrate doping concentration and this leads to accurate representations of substrate potential effects on drain current. In addition, small signal equations for the gate and substrate terminals have been derived and were presented as well. The final form of the Hyperbolic model is a complete model for MOSFET characterization suitable for implementation in circuit simulation tools.
CHAPTER 5
CONCLUSIONS

The work presented in this thesis has demonstrated that the use of computational tools for process, device, and circuit simulation can make an important contribution to modern technology development. Simulation of silicon integrated circuit processing is one way in which mathematical tools can be combined with physical and empirical theory to obtain information about a phenomena which in the past could only be assumed or inferred.

Central to these capabilities are powerful workstations allowing engineers the ability to improve on traditional learning cycles, reinforcing the methods of statistical quality control and the use of designed experiments in the decision making process. This work demonstrated how variations affect manufacturing and fabrication processes, but more importantly it demonstrated that simulation tools are capable of accurately predicting the device characteristics as a result of those inherent variations.

The main point was not that the computer experiments reflected the measured results, although that was an important and necessary verification; the advantage, rather, lies in that process development staff now have access to information that in the past could only be inferred or assumed. In order for the simulations to predict impurity re-distributions and current flow through the structures, the tools had to know what the physical conditions were at each node in the structure. Now the design engineer has access to this same information: where exactly are the electric field magnitudes rising rapidly, leading to premature breakdown? What is causing a structure to latch up, and what can be done to prevent it?

This leads to another important point. Not only can the tools predict physical characteristics very accurately, the simulation experiments themselves can be carried out in
a relatively short amount of time. A device engineer is able to carry out about 4 separate computer experiments in a single day. Put another way: in terms of a learning cycle, engineers using modern computer tools have about 4 learning cycles a day.

Now consider the standard approach, without the use of simulation tools: when a device engineer designs an experiment, he or she has to wait about 6 weeks before the silicon gets processed. After the wafers are available, the test and measurement has to be set up, then the results have to be evaluated, for one condition. If split-lots were run during the processing, then perhaps 2 or 3 conditions could be evaluated. However, part of the evaluation entails learning where in the processing conditions may not have been run correctly. What if an anneal step is run at 1000°C when the process recipe called for an anneal at 1050°C? When process parameters are not on target, the analysis requires step by step verification of the processing for that particular lot of wafers. Additionally, the data for all wafers fabricated will have to be compared during the analysis for verification. Taking into account the time spent during fabrication, measurement and test set up, and deconvolution of results, the total time required can be on the order of 60 days. This gives about 6 learning cycles per year.

Consider a cost comparison: an engineer is paid to carry out experiments using the simulation tools, so the total cost is an engineer’s salary, the cost of the software plus maintenance, the cost of the workstation plus the cost of an upgrade and maintenance. Taking into account depreciation over several years, the cost to run computer experiments for process development is about 50 dollars per learning cycle.

The cost to run the silicon through the processing is about $1K per experimental wafer. Adding in the cost of labor, the cost of the hardware and software, the time for measurement and test set up, plus the time for analysis and evaluation, the cost will be about $10,000 per cycle. Again, this is the cost to get one condition. This type of analysis
makes quite clear the impact that simulation tools can have on the time and cost during the
development of a new process.

An efficient format for process and device simulation was presented which is
currently being used for the development of Texas Instruments’ PRISM technology. A
robust approach for circuit model development was described for the specific case of
Lateral DMOS transistors, which are part of the power components available with the
PRISM process.

The link between process simulation and circuit model development was also
described. Specifically, integration of device simulators with semiconductor
characterization and modeling tools have allowed the direct comparison of simulated and
measured data. A circuit model development approach was described in which physical
measurement combined with computer experiment can lead to accurate simulated I-V
characteristics, taking into account the specification limits and inherent variations during
the processing.

Semiconductor modeling and characterization is an important part of technology
development as circuit designers require accurate analytical models for simulation and
verification of new circuit designs. Analytical models, as opposed to numerical methods,
are important due to the extremely large sizes of circuits that need to be simulated.
Designs which include tens of thousands transistors are not uncommon, and in order to
obtain accurate results in a reasonable amount of time fast and accurate analytical
models are necessary.

SPICE is the most widely used tool for circuit simulation and verification.
Standard analytical models for MOS transistor modeling as they are available in SPICE
were described, including some of the reasons for errors which result from using these
models. Two of the main reasons for errors are the discontinuity in the drain current, and
the assumption that doping levels in the substrate are constant.
It was shown that the discontinuous form of the drain current leads to a discontinuity in the drain conductance, and this was the starting point for providing a possible solution. This methodology was first proposed by Redwine in 1985. The drain conductance as measured from actual devices follows a hyperbolic form, and an empirical relationship was made relating drain conductance models and parameters to a hyperbolic equation. With the addition of a fitting parameter, the new drain conductance equation maintained continuity and fitted to measured data very well. Integration of this function led to an accurate, continuous form the drain current equation in above threshold conditions and zero substrate bias for MOS transistors. This concluded the review of work done by Redwine in deriving the Hyperbolic Model as a way to eliminate common errors in model fitting.

A new and improved version of the Hyperbolic Model has been developed, including the addition of analytical relationships for the drain and body effect which leads to accurate simulation of substrate and drain bias effects in all regions of operation. It was demonstrated that the assumption of uniform substrate doping as current SPICE models maintain, is invalid, especially due to the dependence of modern CMOS technology processes on threshold adjust implants. This was the starting point for considering a more correct form of the body effect parameter commonly used by circuit simulators.

A Gaussian form for the range of values the body effect parameter normally takes was considered and found to be quite adequate in simulating substrate bias effects on the output characteristics for standard MOS transistors. In addition, incorporation of an expression for the drain effect coupled with the body effect equation led to an accurate representation of MOS device characteristics in all regions of operation: subthreshold, linear and saturation. The small signal conductances with respect to the gate and the substrate were also derived, making the model suitable for implementation in circuit simulators.
In summary, the Hyperbolic Model is a new and improved version of an analytical model developed earlier for accurate drain conductance modeling. The model maintains continuity of the drain conductance in the linear to saturation region, and it can now accurately model MOS electrical characteristics in both the subthreshold and above threshold (linear and saturation) conditions for any value of the substrate potential.
APPENDIX

Section 1. Definition of the Parameters for the Hyperbolic Model. The parameters used by the Hyperbolic Model are common to SPICE; however, new definitions are required since new parameters were introduced into the model. In some cases these will be repeated definitions and in others this will be the first time some parameters are introduced.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{ds}$</td>
<td>drain to source current</td>
</tr>
<tr>
<td>$V_{ds}$</td>
<td>drain to source potential</td>
</tr>
<tr>
<td>$V_{gs}$</td>
<td>gate to source potential</td>
</tr>
<tr>
<td>$V_{sb}$</td>
<td>bulk to source voltage</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>linear to saturation region mobility ratio</td>
</tr>
<tr>
<td>$\alpha_e$</td>
<td>gate voltage dependent $\alpha$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>transistor gain factor</td>
</tr>
<tr>
<td>$\beta_R$</td>
<td>gate voltage dependent from of $\beta$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>carrier velocity saturation factor</td>
</tr>
<tr>
<td>$\theta$</td>
<td>mobility degradation factor</td>
</tr>
<tr>
<td>$\phi_f$</td>
<td>Fermi potential in the substrate</td>
</tr>
<tr>
<td>$\varepsilon_{ox}$</td>
<td>dielectric constant of oxide</td>
</tr>
<tr>
<td>$\varepsilon_{si}$</td>
<td>dielectric constant of silicon</td>
</tr>
<tr>
<td>$BE$</td>
<td>body effect parameter</td>
</tr>
<tr>
<td>$DE$</td>
<td>drain effect parameter</td>
</tr>
<tr>
<td>$k_n$</td>
<td>non-geometric component of the gain</td>
</tr>
<tr>
<td>$t_{ox}$</td>
<td>gate oxide thickness</td>
</tr>
<tr>
<td>$R_n$</td>
<td>curvature control parameter for $g_{ds}$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>saturation region slope factor</td>
</tr>
</tbody>
</table>
Section 2. Geometrical relationships of the Hyperbolic Model.

From Figure 72 the following relationships can be derived,

\[ \tan W = M_{gl} = -(2\beta_2 \alpha_4^2). \]  

(114)

From this one can derive

\[ \cos W = \frac{1}{\sqrt{M_{gl}^2 + 1}}. \]  

(115)

In order to achieve a match at the boundaries between the Hyperbolic Model and the "NSH" theory, a new form of the saturation voltage is required. In this case, it will be called \( V_{dsath} \) and it will be defined as the NSH saturation voltage minus a correction factor due to the geometry,

\[ V_{dsath} = V_{dsat} - \delta V. \]  

(116)

The correction factor can be found by setting \( G_d = G_{dl0} \) and \( V_{ds} = \delta V \) and solving the equation for \( \delta V \):

\[ G_{dl0} = G_{ds} - \frac{1}{2} M_{gl} \left( \sqrt{(\delta V - V_{dsat})^2 + V_{r}^2 \cos W} - (\delta V - V_{dsat}) \right), \]  

(117)

\[ 2 \left( \frac{G_{dl0} - G_{ds}}{-M_{gl}} \right) = \left( \sqrt{(\delta V - V_{dsat})^2 + V_{r}^2 \cos W} - (\delta V - V_{dsat}) \right), \]  

(118)

and using the definition of the slope,
\[
\frac{G_{dl0} - G_{ds}}{-\text{Mgl}} = V_{dsat} \quad \text{then}
\]

\[
V_{dsat} + \delta V = \sqrt{(\delta V - V_{dsat})^2 + Vr^2 \cos W}
\]

\[
(V_{dsat} + \delta V)^2 - (\delta V - V_{dsat})^2 = Vr^2 \cos W
\]

and after some reduction this will lead to the expression for \(\delta V\),

\[
\delta V = \frac{Vr^2 \cos W}{4V_{dsat}}
\]

This leads to the final expression for the drain conductance as given in equation (78). \(Vr\) can be variable and affects the curvature of the drain conductance curve. An empirical relationship was found leading to a good fit with data,

\[
Vr = RN \ast (V_{gst} - Vr),
\]

where \(RN\) can be used during the fitting process as a control of the curvature around the new saturation point, \(V_{dsath}\). \(G_{ds}\) is defined as

\[
G_{ds} = \beta g_{m} V_{gst}^2 \left(\frac{\alpha}{\alpha_g}\right)^2.
\]

In equation (79) in Chapter 4 there were two terms that were used but have not been defined. These are

\[
F(V_{ds}) = \left((V_{ds} - V_{dsath})\sqrt{(V_{ds} - V_{dsath})^2 + (RN \ast V_{gst})^2 \cos W}ight.
\]

\[
+RN^2V_{gst}^2 \cos W \log((V_{ds} - V_{dsath}) + \sqrt{(V_{ds} - V_{dsath})^2 + RN \ast V_{gst} \cos W} )
\]
and $F(0)$ is the same expression evaluated at zero drain to source potential. The Hyperbolic saturation voltage is given in terms of the "NSH" saturation voltage, $V_{dsatn}$, which was defined in Chapter 4:

\begin{equation}
V_{dsath} = V_{dsatn} - \frac{(RN * V_{gsl})^2}{4V_{dsatn}\sqrt{(2\alpha_s^2\beta_s)^2 + 1}}
\end{equation}

Section 3. Derivation of the transconductance equation.

The individual derivative terms used in equation (89) are given by the following:

\begin{equation}
\frac{\partial \beta_s}{\partial V_{gsl}} = \frac{-\beta \Theta}{[1 + \Theta V_{gsl}]^2}
\end{equation}

\begin{equation}
\frac{\partial \alpha_s}{\partial V_{gsl}} = \gamma
\end{equation}

\begin{equation}
\frac{\partial M_{gs}}{\partial V_{gsl}} = -4\beta_s \gamma - 2\alpha_s^2 \frac{\partial \beta_s}{\partial V_{gsl}}
\end{equation}

\begin{equation}
\frac{\partial G_{ds}}{\partial V_{gsl}} = \lambda \left( \frac{\alpha}{\alpha_s} \right)^2 \frac{\partial \beta_s}{\partial V_{gsl}} + 2\lambda V_{gsl} \left( \frac{\alpha}{\alpha_s} \right)^2 \beta_s - 2\beta_s \lambda V_{gsl} \frac{\alpha^2}{\alpha_s^3} \gamma
\end{equation}

\begin{equation}
\frac{\partial V_{dsath}}{\partial V_{gsl}} = \frac{\alpha^2}{\alpha_s^3} - 3\frac{V_{gsl} \alpha^2 \gamma}{\alpha_s^4} - \frac{\lambda \alpha^2 V_{gsl}^2 \gamma}{\alpha_s^4} + \frac{2\lambda \alpha^2 V_{gsl}^2 \gamma}{\alpha_s^5}
\end{equation}

Define $f$ to be

\begin{equation}
f = (RN * V_{gsl})^2 \cos W = \frac{(RN * V_{gsl})^2}{\sqrt{(2\alpha_s^2\beta_s)^2 + 1}}
\end{equation}
this will be used to find the derivative of the saturation voltage, $V_{dsath}$.

\[
\frac{\partial V_{dsath}}{\partial V_{gst}} = \frac{\partial V_{dsath}}{\partial V_{gst}} \left[ \frac{1}{4V_{dsath}} \frac{\partial f}{\partial V_{gst}} - \frac{f}{4V_{dsath}^2} \frac{\partial f}{\partial V_{gst}} \right]
\]

where the derivative of $f$ with respect to $V_{gst}$ is given by

\[
\frac{\partial f}{\partial V_{gst}} = \frac{2RN*V_{gst}}{\sqrt{(2\alpha\varepsilon^2\beta_g)^2 + 1}} - \frac{f}{2((2\alpha\varepsilon^2\beta_g)^2 + 1)} \left[ 2(2\alpha\varepsilon^2\beta_g) \right. \\
\left. \times (4\alpha\varepsilon|\beta_g| + 2\alpha\varepsilon^2) \frac{\partial \beta_g}{\partial V_{gst}} \right].
\]

The derivative of $F(Vds)$ as given in (125) will be

\[
\frac{\partial F}{\partial V_{gst}} = Vds \times Fi - V_{dsath} \times Fi - \frac{\partial V_{dsath}}{\partial V_{gst}} \times \sqrt{(Vds - V_{dsath})^2 + f}
\]
\[
+ \frac{\partial f}{\partial V_{gst}} \times \log((Vds - V_{dsath}) + \sqrt{(Vds - V_{dsath})^2 + f})
\]
\[
- \frac{\partial V_{dsath}}{\partial V_{gst}} + Fi
\]
\[
+ f^* \left( \frac{\partial V_{dsath} + V_{dsath} + \sqrt{(Vds - V_{dsath})^2 + f}}{Vds - V_{dsath} + \sqrt{(Vds - V_{dsath})^2 + f}} \right)
\]

where

\[
Fi = \frac{2(Vds - V_{dsath}) \times \left( - \frac{\partial V_{dsath}}{\partial V_{gst}} + \frac{\partial f}{\partial V_{gst}} \right)}{2\sqrt{(Vds - V_{dsath})^2 + f}},
\]

(136)
\[(137) \quad f = Vr^2 \ast Csn \quad (Csn \equiv \cos W), \text{ and} \]

\[(138) \quad \frac{\partial f}{\partial V_{gst}} = (2 \ast Vr \ast Csn - \frac{1}{2} (Vr^2 \ast Csn^3)(2(2\alpha s^2 \beta s) \ast (4\alpha s \beta s + \frac{\partial \beta s}{\partial V_{gst}} 2\alpha s^2))). \]

\[F(0) \text{ is given by} \]

\[(139) \quad F(0) = -V_{death}\sqrt{V_{death}^2 + f} + f \log(-V_{death} + \sqrt{V_{death}^2 + f}) \]

and

\[(140) \quad \frac{\partial F(0)}{\partial V_{gst}} = -\frac{\partial V_{death}}{\partial V_{gst}} \sqrt{V_{death}^2 + f} - V_{death} \ast Foi \]

\[+ \frac{\partial f}{\partial V_{gst}} \log(-V_{death} + \sqrt{V_{death}^2 + f}) + f \frac{-\partial V_{death}}{-V_{death} + \sqrt{V_{death}^2 + f}} + Foi \]

where \(Foi\) is given by

\[(141) \quad Foi = \frac{1}{2} \frac{2V_{death} \frac{\partial V_{death}}{\partial V_{gst}} + \frac{\partial f}{\partial V_{gst}}}{\sqrt{V_{death}^2 + f}}. \]

The transconductance calculation is given by the subthreshold and above threshold equations derived previously (Chapter 4) with the parameters as defined above, as well as with the following special condition: (the following are special cases which had to be evaluated individually using L'Hospital's Rule)
For $V_{gs} - V_i = 0$,

(142) $V_{dsath} = 0$

(143) \[
\frac{\partial V_{dsath}}{\partial V_{gs}} = \frac{1}{\alpha} - 2\frac{\alpha^* R_N}{\sqrt{(2\alpha^2\beta)^2 + 1}} + \frac{\alpha^* R_N}{4\sqrt{(2\alpha^2\beta)^2 + 1}}
\]

(144) $\frac{\partial F(0)}{\partial V_{gs}} = 0$
REFERENCES


29. See reference 6, p. 654.


35 See Reference 26, Chapters 6 and 9.


38 SUPRA is now obsolete; it was very popular during the 1980’s due to the fast, empirical results it could generate based on its analytic and phenomenological models for oxidation and diffusion, etc.


42 See Reference 25 for more detail on other devices supported by PRISM.


47 See reference 25.


50 IC-CAP is a Hewlett-Packard Based software package for semiconductor modeling and characterization. It allows implementation of user defined models for modeling characteristics such as I-V data.


61See Reference 53, Chapter 4.


68Margaret E. Clarke, "Equivalent Circuit Models for Silicon Devices," Chapter 9 in *Semiconductor Device Modelling*. (Great Britain: Springer Verlag, 1989), Christopher M. Snowden, Editor.


70See Reference 15.


74See Reference 22.


76See Reference 23.