Gate-Oxide-Short Defect Analysis and Fault Modeling Based on FinFET's 3D Structure

by

Roya Dibaj

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Electronics

Carleton University
Ottawa, Ontario

To:

My parents, Ziba and Ali, for their endless love and sacrifice,

My husband, Bardia, for his vital love and continuous support,

My sons, Amin and Darius, for the happiness they brought to our lives.

Abstract

FinFET is one of the most promising candidates in replacing planar MOSFET beyond the 22nm technology node due to further improvements in the transistor performance. However, the complexity of FinFET manufacturing process due to its three-dimensional structure and reduced critical dimensions have caused new challenges in achieving reliable device testing. With the emergence of new types of defects and dominance of others, accurate modeling of the defects and generation of reliable fault models are essential to create realistic set of test vectors to detect the defects.

Automatic test pattern generation (ATPG) algorithms use traditional fault models that primarily capture the behavior of the circuit-under-test by introducing defects at the primary inputs and outputs. It has been found that many defects escape testing when they occur within the circuit structure. Recently, Cell-Aware Test (CAT) has been proposed to detect cell-internal defects by performing extensive analog simulations on post-layout standard circuit structures to generate the fault models, which are utilized by CAT-ATPG algorithm. Although CAT methodology has significantly improved the defect coverage of the generated test patterns in MOSFET-based circuits, the defect models utilized are obtained based on the defects injected at the layout level and primarily represented by fixed lumped passive components that cannot reflect the true defect nature in the complex 3D structure of FinFET. Gate-Oxide-Short (GOS) is one of the dominant defects, which has significant impact on circuit reliability. It is the most complex to analyze and most difficult to accurately model true behavior of the defective device.

This thesis presents a novel methodology for GOS defect injection and fault modeling in FinFETs by introducing the defect to a 3D structure of the device for a specific process

technology. The behavior of the defective device is captured through simulations in Sentaurus TCAD environment that lead to the generation of more accurate defect models. These defect models are used in circuit-level simulations to generate appropriate fault models for the circuit structures. These cell-aware models could be integrated in CAT environment to generate more realistic test patterns. This research will not only be used in test pattern generation, but it will aid in cell-aware diagnosis and yield analysis by improving the success rate and reducing the testing time and cost.

Acknowledgements

I would like to express my sincere gratitude to my supervisors: Professor Maitham Shams for his patience and guidance throughout this work, Professor Dhamin Al-Khalili for his invaluable advice, immense knowledge and continuous support in every step of this research, and Dr. Saman Adham for his insightful comments.

I am deeply appreciative to Professor Majid Ahmadi, Professor Voicu Groza, Professor Halim Yanikomeroglu, Professor Niall Tait, and Professor Ralph Mason for their constructive comments that helped me improve this thesis.

My appreciation goes out to CMC Microsystems and Synopsys Inc. for their valuable support throughout this research, as well as Mr. Scott Bruce, the System Administrator of Department of Electronics.

Last but not least, I am thankful to my family and friends for their endless love, patience, and encouragement.

Table of Contents

ABSTRACT 3

ACKNOWLEDGEMENTS 5

TABLE OF CONTENTS 6

LIST OF TABLES 10

LIST OF FIGURES 13

LIST OF FIGURES IN APPENDICES 18

LIST OF ACRONYMS 19

Chapter 1: Introduction 21

1.1 Motivation 23

1.2 Thesis Objective 24

1.3 Thesis Organization 25

Chapter 2: Background 26

2.1 FinFET Device and Manufacturing Process 26

2.1.1 FinFET Structure 26

2.1.2 FinFET DC Behavior 29

2.1.3 FinFET Manufacturing Steps 30

2.2 Manufacturing Defects 36
2.2.1 Defects' Classification 37
2.2.2 FinFET Device Defects 38
2.3 Fault Modeling 40
2.4 Design for Testability and ATPG Techniques 44
2.4.1 Design for Testability 44
2.4.2 Overview of ATPG Techniques 46
2.4.3 Cell-Aware Test (CAT) 47
2.5 Summary 49

Chapter 3: Literature Review 50
3.1 FinFET Defects 50
3.2 GOS Modeling in MOSFET 52
3.3 Previous Work on GOS Defect in FinFET and Shortcomings 55
3.4 FinFET Fault Models in CAT Methodology 57
3.5 Summary 59

Chapter 4: FinFET Defect Injection and Characterization Methodology 60
4.1 FinFET Device Models and Tools 60
4.1.1 Sentaurus Process Models 61
4.1.2 Sentaurus Device Models 63
4.1.3 HSPICE Model Parameters 64
4.2 GOS Defect Injection Strategy 66
4.2.1 GOS in Rectangular and Trapezoidal Fins 66
4.2.2 GOS in Triangular Fins 67
4.3 Defect Characterization Setup 69

4.4 Summary 72

Chapter 5: Device Level Simulation and Defect Modeling 74

5.1 Defect-Free FinFET 74

5.1.1 Single-fin Defect-Free FinFET 75

5.1.2 Double-fin Defect-Free FinFET 79

5.2 GOS Defect Injection Methods 81

5.3 GOS Impact on Rectangular Fin Transistors 82

5.4 GOS Impact on Trapezoidal Fins 88

5.5 Hydrodynamic vs. Drift-Diffusion Transport Models 91

5.6 GOS Impact on Triangular Fins 92

5.7 GOS Impact on Double-fin FinFET 103

5.8 GOS Defect Model 104

5.9 HSPICE Model Parameter Extraction 110

5.9.1 HSPICE Model Parameters 110

5.9.2 Model Verification 116

5.10 Summary 117

Chapter 6: Circuit Level Simulation and Fault Modeling 118

6.1 Simulation Strategy 118

6.2 Inverter 120

6.3 AND-OR-Inverter (AOI) 124

6.3.1 Defect in NMOS Transistor M1 126

6.3.1.1 Discussion 137

6.3.2 Defect in NMOS Transistor M2 139

6.3.2.1 Discussion 145

6.3.3 Defect in PMOS Transistors M5 and M7 145
6.4 3:1 Multiplexer (MUX) 146
6.4.1 GOS Defect in Transistor M9 149
6.4.1.1 Discussion 154
6.5 Summary 155

Chapter 7: Results and Discussion 156
7.1 Transient Results 156
7.2 DC Results 161
7.3 Discussion 162

Chapter 8: Conclusion and Future Work 165
8.1 Contributions 168
8.2 Future Work 168
8.3 Publications 169

APPENDICES 170
Appendix A PFinFET Characteristic Curves 170
Appendix B AOI with Tleak Defect in Transistor M1 172
Appendix C GOS in PMOS Transistors 173

REFERENCES 176

List of Tables

Table 4.1 Bulk FinFET design parameters 62

Table 4.2 Local parameter extraction procedure [69] 65

Table 4.3 Noise margin fault classification for GOS defect scenarios 71

Table 4.4 Delay fault categories for GOS defect scenarios 72

Table 5.1 Electrical characteristics of the bulk FinFET 76

Table 5.2 Electrical characteristics of the single-fin and double-fin NFinFETs with $Wtop = 5\ nm$ 80

Table 5.3 Simulation Platform Specification 91

Table 5.4 Total CPU time of Device simulations with different carrier transport models 92

Table 5.5 Saturation current in double-fin vs single-fin FinFET 103

Table 5.6 Error of the compact model currents 108

Table 5.7 The inverter threshold voltage for logic transition 116

Table 6.1 DC characteristic of the defect-free inverter 121

Table 6.2 Noise margins of CUT with different defect scenarios in transistor M1 121

Table 6.3 DC characteristics of CUT with different defect scenarios in transistor M1 122

Table 6.4 The input patterns that detect GD short defect in M1 125

Table 6.5 AOI input combinations applied for GOS in transistor M1 126

Table 6.6 The transition and propagation delays of the defect-free AOI, Case 1 127

Table 6.7 Delay fault categories for transistor M1, Case 1 129

Table 6.8 Delay fault categories for transistor M1, Case 2 133

Table 6.9 The AOI input combinations applied for GOS in transistor M2 139

Table 6.10 Delay fault categories for transistor M2, Case 1 141

Table 6.11 Delay fault categories for transistor M2, Case 2 143

Table 6.12 The input patterns that detect GD short defect in M9 148

Table 6.13 The input combinations considered for detecting the defect in M9 149

Table 6.14 The propagation delays of the defect-free MUX for each input combination 149

Table 6.15 Delay fault categories for MUX, Case 1 150

Table 6.16 Delay fault categories for MUX, Case 2 152

Table 7.1 Simulation results of the INV, transistor M1 156

Table 7.2 Simulation results of the AOI, transistor M1 157

Table 7.3 Simulation results of the AOI, transistor M2 159

Table 7.4 Simulation results of the MUX, transistor M9 160

Table 7.5 GOS impact on IDDQ 161

Table 7.6 GOS impact on the noise margins 162

Table 7.7 Comparison of GOS vs CAT-based models' impact on DC behavior 163

Table 7.8 Comparison of GOS vs CAT-based models' impact on transient behavior 163

Table 7.9 Number of patterns required for defect detection 164

List of Tables in Appendices

Table C.1 AOI input combinations applied for GOS in transistor M5 173

Table C.2 AOI input combinations applied for GOS in transistor M7 174

List of Figures

Figure 1.1 MOSFET relative power and delay in technology nodes [1] 21

Figure 1.2 Cross sections of a FinFET at (a) Fin cut, and (b) Gate cut [3] 22

Figure 2.1 (a) FinFET 3D structure, (b) Top and cross-sectional view [9] 27

Figure 2.2 Sidewall Image Transfer (SIT) used in the fin patterning process [16] 31

Figure 2.3 (a) Fin formation, (b) STI, and (c) dielectric deposition steps [18] 32

Figure 2.4 Direct tunneling through $SiO2$ versus a high-k dielectric [20] 33

Figure 2.5 The processing steps including poly-Si deposition and spacer formation [18] 34

Figure 2.6 S/D epitaxy and doping profile in NMOS (a) and (c), in PMOS (b) and (d) [18] 35

Figure 2.7 PMOS (a) Metal-gate, (b) CESL deposition, (c) contact formation, (d) ILD deposition, (e) via formation, and (f) complete FinFET structure [18] 36

Figure 2.8 Different types of defects and failures that lead to yield loss [25] 38

Figure 2.9 Traditional fault categories [38] 41

Figure 2.10 An inverter with open defect on the drain of the PMOS transistor [44] 43

Figure 2.11 (a) CAT library view generation, (b) CAT-ATPG flow [5] 48

Figure 3.1 An NMOS with a pinhole in the dielectric layer [6] 52

Figure 3.2 (a) Bi-dimensional model of a defect-free MOSFET, (b) GOS defect model [6] 53

Figure 3.3 Linear non-split model [63] 53

Figure 3.4 Non-split GOS defect models in MOSFET [61] 54

Figure 3.5 Gate oxide short model in [30] 55

Figure 3.6 Schematic representation of sample netlist with inserted defect types [5] 57

Figure 3.7 Leakage and drive strength defects in single-fin and multi-fin FinFETs [5] 58

Figure 4.1 Top view of NFinFET for the Synopsys' generic process [66] 60

Figure 4.2 Bulk NFinFET 3D structure [18] 62

Figure 4.3 Double-fin defect-free FinFET structure with $Wtop = 5\ nm$ [18] 63

Figure 4.4 (a) Bulk NFinFET with rectangular fin shape, with injected $8\ nm \times 8\ nm$ pinhole at the (b) fin top center (Top-C), (c) upper sidewall of the fin (Up-C), (d) middle sidewall (Mid-C), (e)

lower sidewall (Low-C), (f) fin top close to drain (Top-D), and (g) fin top close to source (Top-S) 67
Figure 4.5 Pinhole injection scenarios at the a) Up-D, b) Up-C, c) Up-S, d) Mid-D, e) Mid-C, f) Mid-S, g) Low-D, h) Low-C, and i) Low-S 68
Figure 4.6 Simulation setup for defect characterization 69
Figure 4.7 The voltage transfer characteristics and noise margins [73] 70
Figure 5.1 (a) IdVg, and (b) IdVd curves of bulk NFinFET with different fin top widths 75
Figure 5.2 Total gate capacitance of FinFETs at (a) zero bias, (b) saturation bias 77
Figure 5.3 Defect-free FinFET capacitances at (a) zero bias, (b) saturation bias 78
Figure 5.4 Defect-free FinFET capacitances at maximum gate voltage 79
Figure 5.5 (a) IdVg, (b) IdVd characteristic curves of the double-fin FinFET 79
Figure 5.6 Double-fin FinFET capacitances 80
Figure 5.7 The impact of Channel-Filled (CF) vs. Gate-Filled (GF) pinholes on IdVd characteristics 81
Figure 5.8 Electron densities at saturation bias at the fin center of (a) defect-free, and defective transistors with (b) 2 nm × 2 nm, (c) 4 nm × 4 nm, and (d) 8 nm × 8 nm pinholes 82
Figure 5.9 IdVd of rectangular FinFET with (a) 2 nm × 2 nm, (b) 8 nm × 8 nm Top-C pinholes 83
Figure 5.10 The leakage electron current density at the fin center of the defect-free FinFET 84
Figure 5.11 The leakage electron current density of the FinFET with an 8 nm × 8 nm pinhole on 84
Figure 5.12 Electron current densities along the fin length in (a) defect-free FinFET, and FinFETs with 8 nm × 8 nm pinhole on the, (b) Top-C, (c) Up-C, (d) Mid-C, and (e) Low-C of the fin in linear bias. 85
Figure 5.13 IdVd curves of FinFETs with (a) 2 nm × 2 nm, (b) 8 nm × 8 nm Mid-C pinholes 86
Figure 5.14 (a) Leakage, (b) Saturation currents of the rectangular fin FinFET with different scenarios of sidewall pinholes comparing to the top pinhole 87
Figure 5.15 (a) Leakage, (b) Saturation currents of rectangular fin FinFET with top pinholes close to the source (Top-S), at the fin center (Top-C), and close to the drain (Top-D) 88
Figure 5.16 IdVd curves of FinFETs with (a) 4 nm × 4 nm, (b) 8 nm × 8 nm sidewall pinholes 89

Figure 5.17 15 (a) Leakage, (b) Saturation currents of trapezoidal fin FinFET showing the impact of pinhole size and location.....90

Figure 5.18 Leakage, (b) Saturation currents of trapezoidal fin FinFET with pinholes close to the source (Top-S), at the fin center (Top-C), and close to the drain (Top-D).....90

Figure 5.19 IdVg curves of the (a) defect-free, (b) defective FinFET with 4 nm × 4 nm pinhole with different carrier transport models.....91

Figure 5.20 Leakage current densities at the fin center of (a) defect-free, and (b) defective FinFETs, saturation current densities at the fin center of (c) defect-free, and (d) defective transistors.....93

Figure 5.21 (a) IdVg, (b) IdVd curves of the FinFET with 2 nm × 2 nm Top-C pinhole vs. defect-free.....94

Figure 5.22 IdVd curves of FinFETs with (a) 2 nm × 2 nm, (b) 8 nm × 8 nm Mid-C pinholes 95

Figure 5.23 IdVd curves of FinFETs with (a) 8 nm × 8 nm sidewall center pinhole (a) moving along the fin height, (b) moving along the fin length.....96

Figure 5.24 (a) Leakage, (b) Saturation currents of triangular fin with pinholes on the center of fin top and sidewall locations.....97

Figure 5.25 (a) Leakage, (b) Saturation currents of triangular fin with different defect scenarios close to source/drain 98

Figure 5.26 (a) IdVg curves, (b) IgVg curves of FinFETs with 8 nm × 8 nm middle sidewall pinhole, moving close to drain and source at linear bias 99

Figure 5.27 (a) IdVg curves, (b) IgVg curves of FinFETs with 8 nm × 8 nm middle sidewall pinhole, moving close to drain and source at saturation bias.....100

Figure 5.28 (a) The impact of 2 nm × 2 nm pinhole on (a) Cgd, (b) Cgs capacitances.....101

Figure 5.29 The impact of 4 nm × 4 nm pinhole on (a) Cgd, (b) Cgs capacitances 102

Figure 5.30 The impact of 8 nm × 8 nm pinhole on (a) Cgd, (b) Cgs capacitances 102

Figure 5.31 (a) IdVg, (b) IdVd curves of double-fin FinFET with 8 nm × 8 nm sidewall pinhole..103

Figure 5.32 GOS Defect model extraction flow diagram 104

Figure 5.33 (a) The difference in (a) gate currents, (b) drain currents of a FinFET with 8 nm × 8 nm Mid-C pinhole and the defect-free FinFET at various bias voltages.....105

Figure 5.34 Preliminary GOS compact model 106

Figure 5.35 (a) IdVg, (b) IgVg curves of the compact model vs. defective transistor at linear and saturation bias 107

Figure 5.36 CV characteristics of the defective and defect-free FinFETs at (a) linear, (b) saturation bias 108

Figure 5.37 The difference in the (a) gate-source capacitances, (b) gate-drain capacitances of a FinFET with $8\,\mathrm{nm} \times 8\,\mathrm{nm}$ Mid-C pinhole and the defect-free FinFET at various bias voltages 109

Figure 5.38 The proposed GOS compact model for FinFET 109

Figure 5.39 Output resistance at $Vd = 0.05\ V$ and different overdrive voltages 111

Figure 5.40 Threshold voltage roll-off due to (a) SCE, (b) DIBL effects 112

Figure 5.41 Optimized CggVgs characteristics at $Vd = 0.05\ V$ 113

Figure 5.42 Optimized (a) IdVg, (b) GmVg graphs at $Vd = 0.05\ V$ 113

Figure 5.43 Optimized (a) IdVg, (b) GmVg graphs at $Vd = 0.8\ V$ 114

Figure 5.44 Optimized (a) IdVd, (b) GdsVd graphs at $Vgs = 0.4, 0.6, 0.8\ V$ 115

Figure 5.45 Optimized CggVgs characteristics at $Vd = 0.8\ V$ 115

Figure 5.46 (a) VTC, (b) Transient response of the defect-free inverter from TCAD and HSPICE . 116

Figure 6.1 FinFET fault model generation flow diagram 119

Figure 6.2 Inverter circuit diagram 120

Figure 6.3 VTC of the (a) defect-free (b) defective inverter with Up-C4x4 pinhole 121

Figure 6.4 Effective current of transistor M1 123

Figure 6.5 Percentage of falling delay variations for different pinholes in transistor M1 123

Figure 6.6 4-input AOI circuit 125

Figure 6.7 Transient response of (a) AOI, (b) load inverter, Case 1 127

Figure 6.8 Percentage of falling delay variations for different pinholes in transistor M1, Case 1 128

Figure 6.9 AOI with (a) GS short, (b) GD short defect in transistor M1 129

Figure 6.10 VTC of defective AOI with (a) GS short, (b) GD short defect on transistor M1, Case 1 130

Figure 6.11 Transient Response of (a) AOI, (b) the load inverter with GS short on M1, Case 1 130

Figure 6.12 Percentage of delay variations in AOI with GS short in transistor M1, Case 1 131

Figure 6.13 Transient Response of (a) AOI, (b) the load inverter with GD short on M1, Case 1 131

Figure 6.14 Percentage of delay variations in AOI with GD short in transistor M1, Case 1.....132

Figure 6.15 VTC of AOI, Case 2 133

Figure 6.16 The percentage of rising delay variation for different pinholes in transistor M1, Case 3 135

Figure 6.17 Percentage of delay variations in AOI with GS short in transistor M1, Case 3 135

Figure 6.18 Equivalent RC model of the AOI.....137

Figure 6.19 Transient response of (a) AOI with pinholes in transistor M2, (b) load inverter, Case 1 140

Figure 6.20 Percentage of falling delay variations for different pinholes in transistor M2, Case 1 ...140

Figure 6.21 Percentage of delay variations in AOI with GS short in transistor M2, Case 1 141

Figure 6.22 Percentage of delay variations in AOI with GD short in transistor M2, Case 1.....142

Figure 6.23 Schematic diagram of 3:1 MUX.....146

Figure 6.24 MUX consisting of two 4-input AOIs and the required inverters.....147

Figure 6.25 Percentage of the falling delay variations in the MUX with GOS defect, Case 2 152

Figure 6.26 Percentage of the rising delay variations in the MUX with GOS defects, Case 3.....154

List of Figures in Appendices

Figure A.1 Optimized CggVgs characteristics at $Vd = -0.05$ V 170

Figure A.2 Optimized (a) IdVg, (b) GmVg graphs at $Vd = -0.05$ V.....170

Figure A.3 Optimized (a) IdVg, (b) GmVg graphs at $Vd = -0.8$ V 171

Figure A.4 Optimized (a) IdVd, (b) GdsVd graphs at $Vgs = -0.4, -0.6, -0.8$ V.....171

Figure A.5 Optimized CggVgs characteristics at $Vd = -0.8$ V.....171

Figure B.1 VTC of defective AOI with Tleak defect on M1, Case 1 172

Figure B.2 Transient response of (a) AOI with Tleak defect on M1, (b) the load inverter, Case 1....172

Figure C.1 Percentage of delay variations in AOI with GOS defect in transistor M5, Case 1 173

Figure C.2 Percentage of delay variations in AOI with GD short in transistor M5, Case 1 173

Figure C.3 Percentage of delay variations in AOI with GS short in transistor M5, Case 1.....174

Figure C.4 Percentage of delay variations in AOI with GOS defect in transistor M7, Case 3 174

Figure C.5 Percentage of delay variations in AOI with GD short in transistor M7, Case 3 175

Figure C.6 Percentage of delay variations in AOI with GS short in transistor M7, Case 3.....175

List of Acronyms

3D	Three Dimensional
AOI	AND OR Inverter
ATPG	Automatic Test Pattern Generation
BIST	Built-In Self Test
BSIM-CMG	Berkeley Short Channel IGFET Model-Common Multi Gate
CAT	Cell Aware ATPG
CD	Critical Dimension
CESL	Contact Etch Stop Liner
CMP	Chemical Mechanical Polish
CPU	Central Processing Unit
CUT	Circuit Under Test
DD	Drift Diffusion
DIBL	Drain-Induced Barrier Lowering
DOS	Density of States
DP	Double Patterning
DPPM	Defective Parts Per Million
EOT	Equivalent Oxide Thickness
GD	Gate Drain Short
GE	Gate Exhaustive Testing
GIDL	Gate Induced Drain Leakage
GOS	Gate Oxide Short
GS	Gate Source Short
HD	Hydrodynamic
IC	Integrated Circuit
IG	Independent Gate
ILD	Inter Layer Dielectric
LELE	Litho-Etch-Litho-Etch
LER	Line Edge Roughness
LWR	Line Width Roughness
MOL	Middle of Line
MUX	Multiplexer
PTS	Punch Through Stopper
RMG	Replacement Metal Gate
S	Subthreshold slope
SCE	Short Channel Effects
SDD	Small Delay Defects
SDE	Source Drain Extensions
SIT	Sidewall Image Transfer
SoC	System on Chip
SOI	Silicon on Insulator
SRAM	Static Random-Access Memory
STI	Shallow Trench Isolation
SWR	Sidewall Surface Roughness
TCL	Tool Command Language
TDF	Transition Delay Faults
Tdrive	Transistor Drive Strength

TG	Tied Gate
Tleak	Transistor Leakage
VTC	Voltage Transfer Characteristics

Chapter 1: Introduction

According to Moore's Law, the number of transistor count in every square inch has doubled every 2 years. It has required the gate pitch (the technology node) to continue scaling by almost 0.7 every two years [1]. Scaling down the supply voltage and transistor dimensions has resulted in reduced dynamic or switching power and improved performance in consecutive technology nodes, as shown in Figure 1.1. However, all the components of the static power have dramatically increased due to the increase in the subthreshold leakage (I_{off}), the gate leakage (I_{Gate}), and the junction leakage (I_{Junc}) currents. For instance, the gate tunneling current from the gate poly-Si through the dielectric layer to the channel increases with the reduction of dielectric thickness. The subthreshold leakage current has also increased with the channel length reduction.

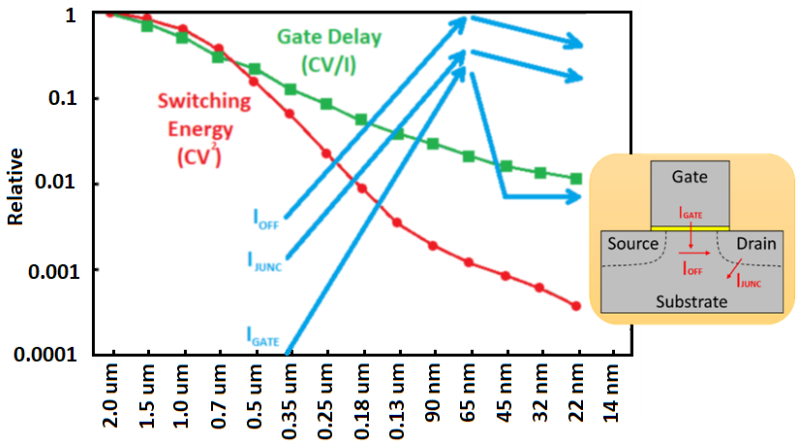

Figure 1.1 MOSFET relative power and delay in technology nodes [1]

Innovative techniques such as strain engineering at 90 nm node and introducing high-k dielectric with metal gate at 45 nm node has improved the transistor performance by increasing the transistor driving current and reducing the gate leakage [2]. However, the challenges due to short channel effects (SCE) and increased leakage have led to the innovative design of the FinFET transistor as a 3D structure in which the channel is

extended outside the substrate, as a fin, and the gate is wrapped around the channel. Figure 1.2 shows the cross section of a 22 nm node FinFET transistor at fin and gate cuts [3]. The novel 3D structure of FinFET, with improved performance and reduced leakage, has made it a promising replacement for planar MOSFET beyond 22 nm technology node. Today, FinFET is widely used in both CPU and SoC products, from high-performance to low-power applications [1].

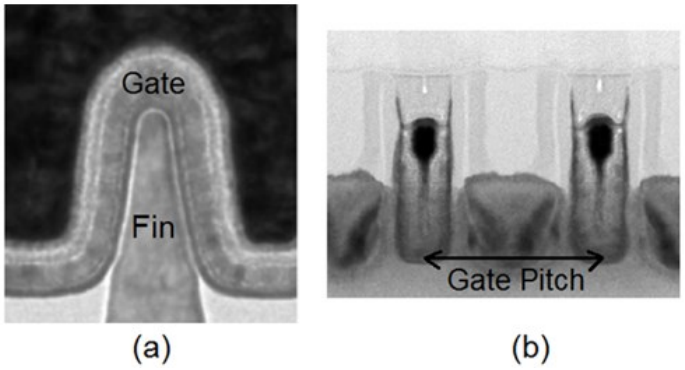

Figure 1.2 Cross sections of a FinFET at (a) Fin cut, and (b) Gate cut [3]

The Integrated circuit (IC) testing has moved from functional testing in early days toward structural testing due to the dramatic increase in the number of transistors in ICs. The manufacturing defects that impact the circuit reliability have become more complex in nanometer technologies. Fault models representing the impact of defects at the circuit level are used in Automatic Test Pattern Generation (ATPG) tools to generate test patterns. Accurate fault models help improve the quality of physical failure analysis and generate more effective diagnostic patterns to determine possible defect locations [14]. With the growing number of applications requiring close to zero defective parts per million (DPPM), the quality of testing ICs has become paramount [4]. Achieving adequate test quality, increasing the defect coverage, and reducing the test time and cost are some of the critical challenges in testing.

1.1 Motivation

Traditional test methods are primarily based on models representing the behavior of the circuit structures with faults injected only at the primary inputs and outputs. These models used to include stuck-at, sequential and transition faults, and later extended to include layout-aware, and bridging between interconnects faults. However, ATPG algorithms failed to detect defects within the circuit structures. Recently, Cell-Aware Test (CAT) technique has been developed by Mentor Graphics and Synopsys to aid in the generation of the test patterns for advanced planar CMOS technologies [5]. It uses the layout information of the standard cells, including parasitic elements, and performs extensive analog simulations for each defect type to find the test patterns that can detect the defect. The CAT methodology has significantly improved the defect detection rate in CMOS technologies due to its defect-based approach in modeling the intra-cell defects missed by traditional test methods.

In FinFET technology, the manufacturing defects have become more complex due to the 3D nature of FinFET transistor and reduced critical dimensions. Although the CAT methodology is extended to FinFET testing, its 2D layout-based approach cannot accurately develop the defect models in the 3D transistor structure [4]. Even after CAT screening, a large number of test escapes has been reported from the industry [4], [5]. Therefore, there is a need for better detection of FinFET defects.

Gate-Oxide-Short (GOS) is one of the dominant defects [6] that is complex to analyze and model, especially in FinFETs, due to its nonlinear impact on the transistor characteristics. In CAT methodology, there are two types of transistor-related defects, the drive strength

defect and the short between transistor terminals, all modeled with resistive elements in extensive circuit simulations of the standard cells [5]. FinFET is a complex 3D structure, and hence, representing its defective behavior by simple linear models will not reflect the true nature of GOS defect and may not lead to the generation of realistic test patterns.

The goal of this research is to propose a novel methodology to develop defect-based fault models for a specific process technology that accurately reflects the behavior of the defective transistor. The effectiveness of the developed fault models in generating more accurate test patterns with higher defect detection rate needs to be addressed.

1.2 Thesis Objective

The main objective to this work is to propose a novel methodology in defect-based fault modeling in FinFET technology. The existing CAT methodology is a layout-based approach that suffers from utilizing simple passive models for transistor-related defects. A solution to this problem is to construct a 3D structure of the FinFET based on a Synopsys template and to introduce the GOS defect at the device structure level to develop accurate defect models. This can be performed in an environment such as Synopsys' Sentaurus TCAD.

The second objective is to apply the extracted defect models to complex circuit structures, such as 4-input AOI and 3:1 MUX and develop accurate fault models. The effectiveness of the proposed defect-based fault modeling approach has to be addressed in comparison to the current CAT methodology.

1.3 Thesis Organization

Chapter two provides a brief background of FinFET structure and its manufacturing steps, as well as the ATPG techniques and their challenges in FinFETs. In Chapter 3, a comprehensive literature review of FinFET defects with emphasis on GOS defect and fault modeling in CMOS and FinFET is presented. The FinFET device models, and tools used in this research, the defect injection strategy, and the defect characterization framework are explained in Chapter 4. The device-level simulations that lead to developing GOS defect models are presented in Chapter 5. The circuit-level simulations for developing fault models of complex gates are covered in Chapter 6. In Chapter 7, a summary of results and discussion to demonstrate the effectiveness of the proposed methodology is presented. Chapter 8 concludes the thesis with a recapitulation and provides recommendation for future work.

Chapter 2: Background

As the focus of the thesis is on GOS defect analysis and modeling, this chapter provides a basic description of the FinFET device structure and a brief overview of its manufacturing process. Testability of circuit built by such an advance and complex process is becoming a major challenge to achieve high level of confidence in product integrity. A description of the processing defects is presented here as well as the fault models that manifest these defects at the higher level of design abstraction. A general overview of design for testability and ATPG techniques are also discussed. This is followed by introducing a more recent test approach referred to as the CAT [5].

2.1 FinFET Device and Manufacturing Process

MOSFET technology has scaled down to 22 nm node by applying innovative techniques such as strain engineering and integrating high-k dielectrics at the gate stack. The improved performance and reduced switching power of these technologies came at the cost of more severe short channel effects and increased leakage currents [7]. FinFET technology has overcome many of these problems due to offering three-dimensional (3D) device structure and hence, is considered now as the mainstream VLSI technology beyond the 22 nm node [8].

2.1.1 FinFET Structure

Innovation in device engineering that led to double-gate and tri-gate MOSFETs has been a solution to the challenges resulted from reducing transistor dimensions in planar CMOS. In the new transistor structure, the channel is vertically extended outside the

substrate (usually named as fin) and the gate is wrapped around the channel, as shown in Figure 2.1.

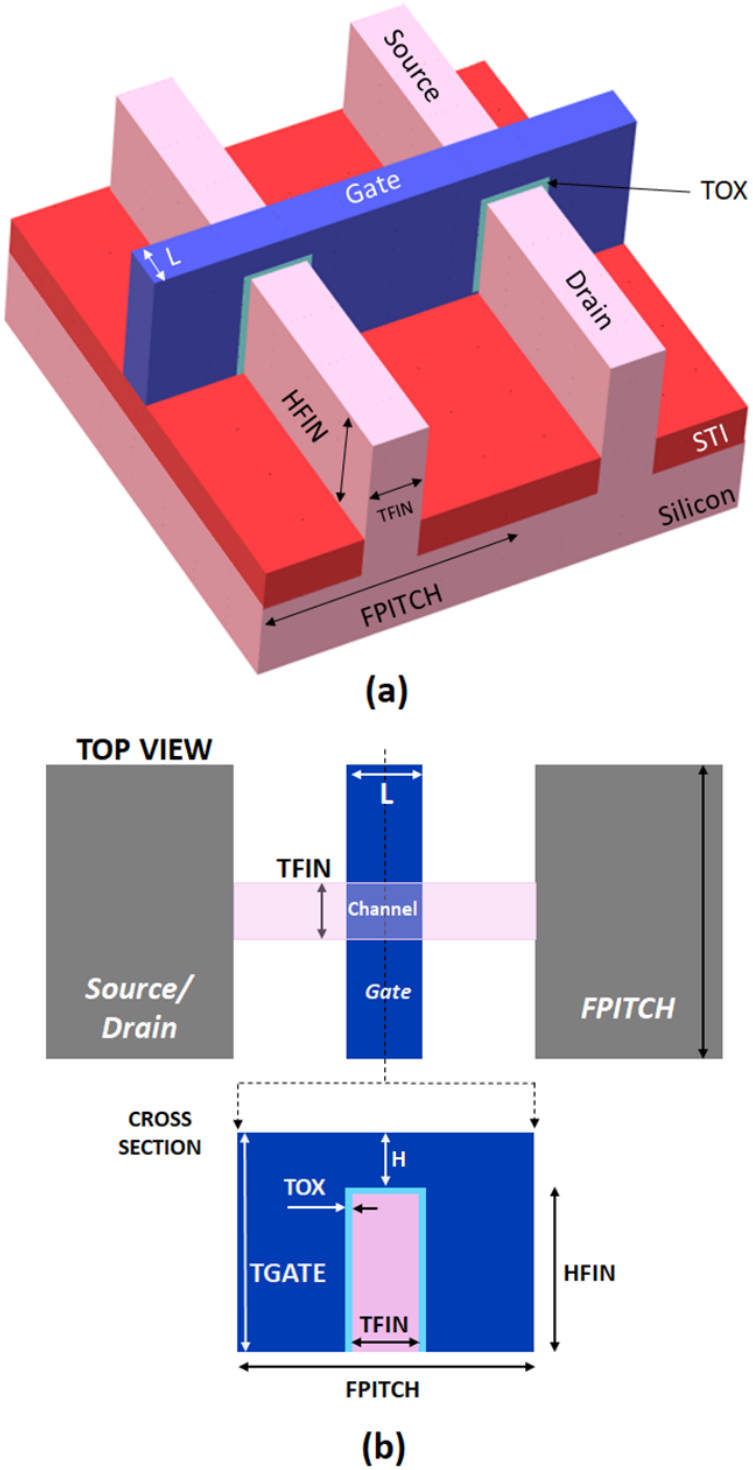

Figure 2.1 (a) FinFET 3D structure, (b) Top and cross-sectional view [9]

The source and drain are also vertical structures appearing as the fins of the gate, hence the transistor is named FinFET. TFIN is the fin width, H is the gate height above the fin, and HFIN represents the fin height. The channel is formed below the gate and its sidewalls. The effective transistor width is obtained from Equation 2.1,

$$W_{eff} = 2 * HFIN + TFIN \tag{2.1}$$

in which, HFIN and TFIN are the fin height and width, respectively. Since the transistor width has a discrete nature and cannot be altered, larger transistors are realized by having arrays of fins. For example, the FinFET in the figure has two fins and supports twice the current of a single-fin FinFET.

The gate length is denoted by L in the figure. The careful choice of L/TFIN is important in downgrading SCEs, specifically reducing the Drain-Induced Barrier Lowering (DIBL) and subthreshold slope (S) [9]. TOX and TGATE in the figure are the gate-oxide thickness and the gate height, respectively. FPITCH is the fin pitch, the minimum allowable distance between two consecutive fins. Another critical dimension (CD) is the gate pitch, the minimum allowable distance between two neighboring gates. Note that in the technology nodes beyond 32 nm, the gate length is not scaled down by the scaling factor, and the gate pitch reduction represents the scaling trend.

FinFETs are realized as double-gate or triple-gate structures. The latter structure, also called tied-gate in the literature, is same as in Figure 2.1. In the former case, also referred to as independent gate, a thick insulator is deposited on the fin top and the channel is formed on the fin sidewalls where the front gate and back gate exist. Therefore, the effective width of the transistor in Equation 2.1 does not include the second term.

FinFET exhibits lower leakage current and improved electrostatic integrity compared to planar MOSFET because of exerting better gate control over the channel [9], [10]. Moreover, random dopant fluctuations in FinFETs are lower due to undoped or lightly doped fins [6]. FinFETs provide higher drain currents at lower supply voltages; hence, making better performance with lower power consumption achievable. From the circuit design perspective, FinFETs consume lower static power due to the reduced leakage currents [1].

FinFETs are categorized into silicon-on-insulator (SOI), and bulk-Si types. SOI FinFET is constructed over a buried-oxide layer and has lower leakage and higher speed, while bulk FinFET has lower process cost, no floating-body effect, and better heat dissipation [11]. Although this research is conducted on the bulk FinFET, the methodology is applicable to SOI structure as well.

2.1.2 FinFET DC Behavior

FinFET operates in a similar manner as MOSFET device. The Shockley model expresses the drain current for long-channel MOSFETs [12]. However, in short-channel devices that extensive electric field in the vicinity of drain region results in the carrier velocity saturation, α-power law provides a simple short-channel model as shown in Equation 2.2 [13],

$$I_D = \begin{cases} 0 & V_{GS} < V_T \quad \text{Cutoff} \\ \frac{I_{Dsat}}{V_{Dsat}} V_{DS} & V_{GS} \geq V_T\,,\ V_{DS} < V_{Dsat} \quad \text{Linear} \\ I_{Dsat} & V_{GS} \geq V_T\,,\ V_{DS} > V_{Dsat} \quad \text{Saturation} \end{cases} \tag{2.2}$$

where

$$I_{Dsat} = P_c \frac{\beta}{2} V_{GT}^{\alpha} \quad , \quad V_{Dsat} = P_v V_{GT}^{\alpha/2}$$

$$\beta = \mu_n C_{ox} \frac{W}{L} \quad , \quad V_{GT} = V_{GS} - V_T \quad , \quad C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} \tag{2.3}$$

and α is the velocity saturation index, P_c and P_v are parameters that can be determined by I-V plot curve fitting, μ_n is the electron mobility, C_{ox} is the gate-oxide capacitance per unit area, ϵ_{ox} is the permittivity of SiO_2 and t_{ox} is the oxide thickness.

2.1.3 FinFET Manufacturing Steps

The manufacturing process of advanced technologies has become more complex due to the reduced critical dimensions and increased number of processing steps. The conventional flow of a replacement metal-gate (RMG) bulk FinFET fabrication process consists of the major steps that are summarized in [18] and briefly explained in the following.

a) Fin patterning

Fin patterning is performed by subsequent lithography and etch pattern transfer steps. Double patterning (DP) techniques are used to increase the resolution of optical lithography beyond the 32 nm technology node [15]. DP techniques are of two primary classes, litho-etch-litho-etch (LELE) and sidewall image transfer (SIT) or self-aligned double patterning technique. In the LELE technique, each exposure has half of the layout pattern, and they are interdigitated to form the final platform. Since the two masks are optically isolated, the mask replacement and alignment increase the overlay error and result in pitch variations or pitch walking. In SIT, half of the layout pattern is deposited and patterned as mandrel. A mask layer deposition is followed by anisotropic etch to form the spacers. By etching the

mandrel, the fins are formed with a pitch half of that of the mandrel. Figure 2.2 shows the major SIT steps in the fin patterning process [16].

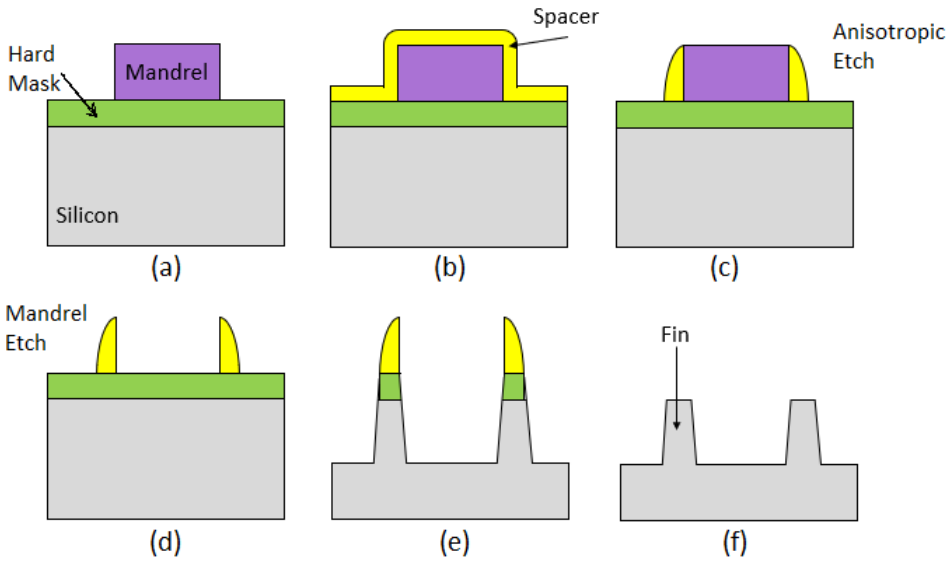

Figure 2.2 Sidewall Image Transfer (SIT) used in the fin patterning process [16]

The fin width is controlled by the spacer thickness and the etch selectivity [16]. The lithography and etch damages are important sources of variability in the fin shape that induce different types of defects to the FinFET.

b) STI and high-k dielectric deposition

In the bulk FinFET, leakage current might flow below the channel; therefore, after the fin patterning step, it is necessary to form a punch-through stopper (PTS) or channel stop region by appropriate ion implantation [17]. The fin height is controlled by the PTS layer position and shallow trench isolation (STI) that is deposited to isolate the adjacent transistors. In Sentaurus TCAD environment, Synopsys has provided a template for simulating the 3D FinFET structure generation. The simulation outcome for the mentioned processing steps are shown for a quarter of fin in Figure 2.3.

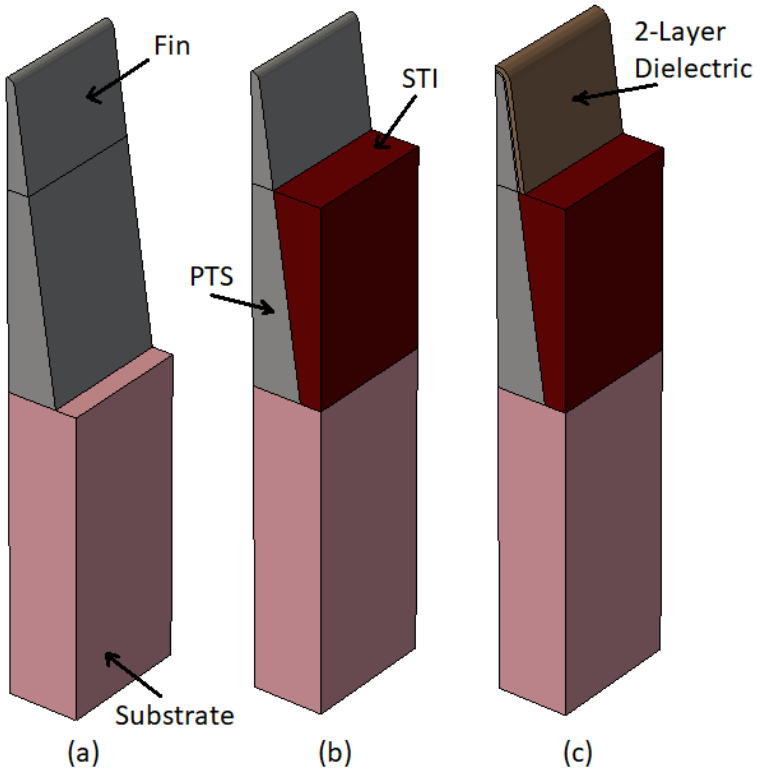

Figure 2.3 (a) Fin formation, (b) STI, and (c) dielectric deposition steps [18]

In some process steps, Hydrogen annealing is performed to reduce the surface roughness and enhance the carrier mobility and device performance. For example, during the fin patterning process, the line-edge-roughness (LER) due to the lithography and the sidewall surface roughness (SWR) due to the etching are reduced by annealing [19], [20]. However, LER and SWR are still important sources of variability in FinFETs.

Scaling down the traditional gate dielectric, SiO_2, thickness below 2 nm has increased the gate leakage due to direct tunneling of electrons [20]. The gate-oxide capacitance is inversely proportional to the dielectric thickness, T_{ox}, as shown in Equation 2.4,

$$C_{ox} = {\varepsilon_0 KA}/{T_{ox}} \tag{2.4}$$

where ε_0 is the free-space permittivity, K is the relative permittivity, and A is the surface area. A solution to reduce the gate tunneling current without affecting the gate capacitance

is incorporating a material such as HfO_2 that has a dielectric constant of 25 compared to 3.9 for SiO_2. It would result in the Equivalent Oxide Thickness (EOT) presented in Equation 2.5,

$$EOT = T_{HiK} * {}^{3.9}/_{K} \tag{2.5}$$

where T_{HiK} is the high-k dielectric thickness. Figure 2.4 shows how using a physically thicker layer of a high-k dielectric reduces the tunneling current [20].

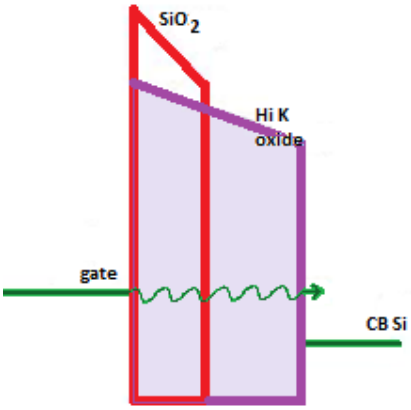

Figure 2.4 Direct tunneling through SiO_2 versus a high-k dielectric [20]

Although HfO_2 is widely used as the gate insulator, it is deposited over an interfacial layer of silicon dioxide. The Si- SiO_2 interface provides better electrical interface since it has lower defect concentration and trapped charges. Moreover, it prevents carrier mobility degradation in the channel compared to using the high-k dielectric alone [20].

c) Poly-silicon gate and spacer formation

Following the gate dielectric deposition, the dummy poly-Si gate and the spacer are deposited over it, as shown in Figure 2.5 (a) and (b). The poly-Si gate will later be removed and replaced by metal gate. The width of the spacer defines the underlap of the S/D

extension below the gate-oxide layer. The unwanted part of the gate dielectric over the S/D region is removed as shown in Figure 2.5 (c).

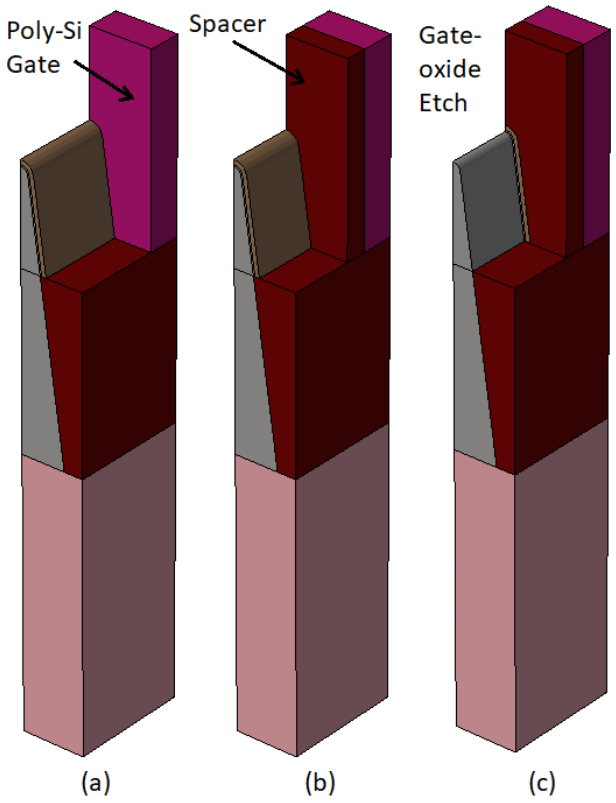

Figure 2.5 The processing steps including poly-Si deposition and spacer formation [18]

d) Source/Drain epitaxy

The Source/Drain (S/D) epitaxial regions are formed and stress engineered by adding SiGe for PFinFET and SiC for NFinFET which apply longitudinal compressive and tensile stress to the channel respectively. The source/drain extensions (SDE) are covered by spacers with low-k material to minimize the parasitic capacitance between the gate and S/D regions [21]. Figure 2.6 (a) and (b) show the resulted structures after the S/D epitaxy for NMOS and PMOS transistors. Figure 2.6 (c) and (d) show the doping profiles in the NMOS

and PMOS final structures, respectively. The raised S/D regions in the PMOS provide lower parasitic resistances and improved performance.

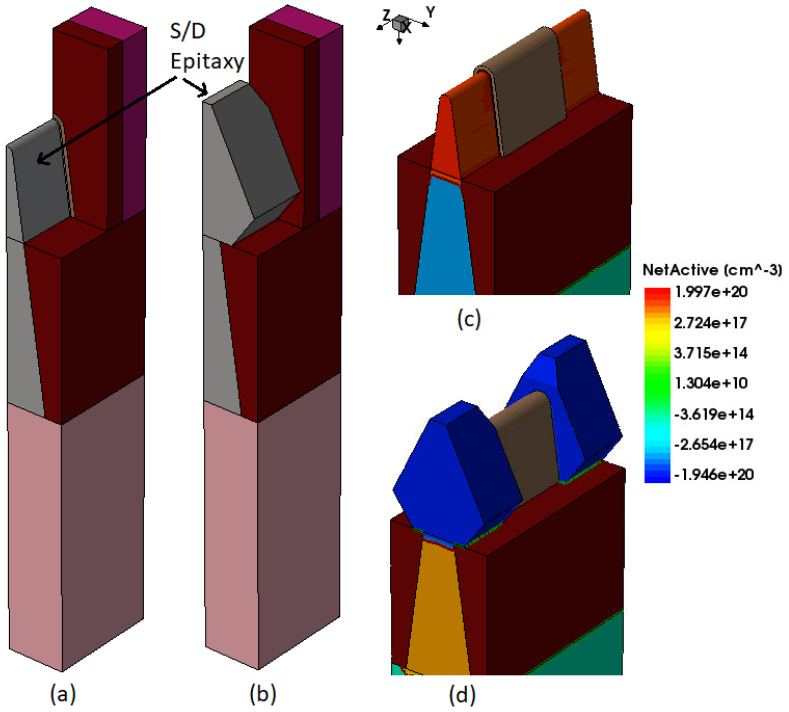

Figure 2.6 S/D epitaxy and doping profile in NMOS (a) and (c), in PMOS (b) and (d) [18]

e) Ploy-Si gate removal and metal gate formation

The gate-last process allows the metal to be deposited after high temperature S/D epitaxial formation and enhances the channel strain as well. Figure 2.7 (a) shows the PMOS structure with the replaced metal gate.

f) CESL and contact formation

Following the metal gate formation, the contact-etch-stop-liner (CESL) is deposited over the metal gate and spacers which applies stress to the channel and improves the saturation current. Finally, the contacts are wrapped around the fin to minimize the contact resistance, and the inter-layer dielectric (ILD) with low-k material is used to isolate trench-

shaped contacts. Figure 2.7 shows these process steps and the final structure after two reflection transformation along the fin width and length.

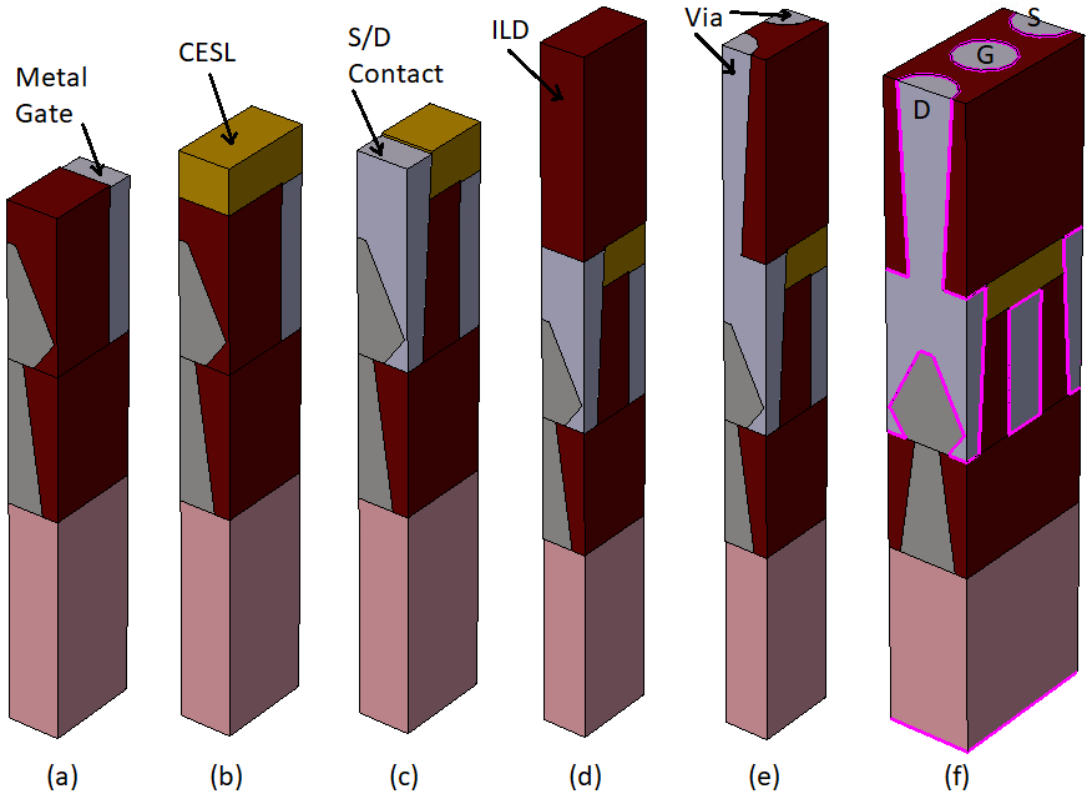

Figure 2.7 PMOS (a) Metal-gate, (b) CESL deposition, (c) contact formation, (d) ILD deposition, (e) via formation, and (f) complete FinFET structure [18]

Due to the 3D nature of FinFET and its complicated manufacturing process, the number and complexity of defects inside standard circuit structures have increased. The next section covers the defect types and the fault models that represent them at the circuit level.

2.2 Manufacturing Defects

Advancement in semiconductor manufacturing process has led to relentless increase in the number of transistors on integrated circuits (ICs). The complexity of the designs and reduced critical dimensions have caused more manufacturing defects to occur

in nanometer technologies, specially FinFETs. This section provides an overview of the defect types and those important in FinFETs that are further referred to in this thesis.

2.2.1 Defects' Classification

Defects are physical imperfections in the device that may happen due to failures or process variations during the manufacturing process [22]. Accordingly, failures can be categorized into two major groups: process-induced and operation-induced failures [23]. The operation-induced failures are usually time-dependent and happen in the metal interconnects, such as electromigration, or in the transistor, such as gate-oxide breakdown and hot-carrier injection. Process-induced failures such as photolithographic misalignments might also occur in interconnects, which might result in missing contacts, or in transistors that may cause variations in the critical dimensions such as gate length and fin width.

Defects are also classified as random and systematic defects. Random dopant fluctuation is an example of randomly distributed defects. In a systematic defect, however, the failure source is predictable, such as mask misalignments that cause fin patterning or metal patterning problems, or chemical mechanical polish (CMP)-related defects [24]. Defects might cause functional failures, or parametric failures where the circuit does not meet the desired specifications, such as timing constraints and power consumption [22]. Defects result in the reduction of process yield that is defined as the percentage of acceptable parts [25]. Figure 2.8 shows the sources of yield loss due to different types of defects and failures.

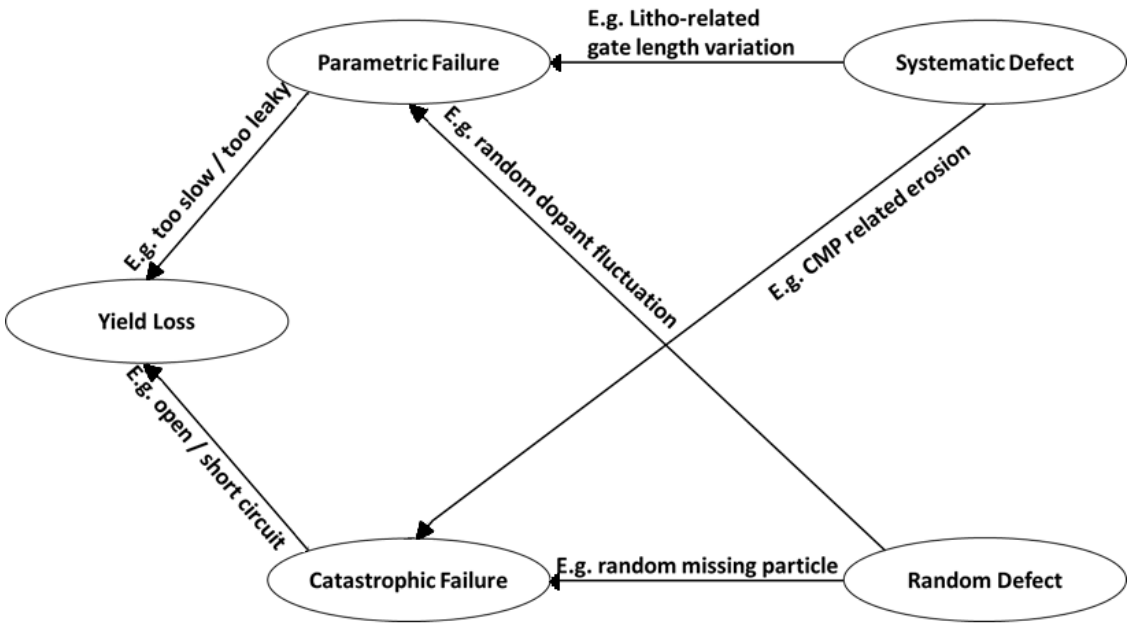

Figure 2.8 Different types of defects and failures that lead to yield loss [25]

Although operation-induced defects produce reliability issues, they are not within the scope of this research, as the focus is mainly on the process-induced defects that affect the transistor behavior. Opens on transistor terminals, short or bridging between the terminals, and gate to channel short are some of the representations of physical imperfections in transistors.

2.2.2 FinFET Device Defects

FinFET has significant structural differences with planar CMOS technology that cause new types of defects to emerge, and some common defects to exhibit different behaviors than in the planar technology. The following are the major classes of FinFET defects [26].

a) Open Defect in Transistor Terminals

In older planar CMOS technology nodes, an open defect in drain or source terminal results in a high impedance state at the transistor output and the faulty gate shows relatively

slow change in potential due to leakage. However, in nanometer technologies with reduced node capacitances, the impact of leakage current is more significant. The subthreshold leakage current of the defective logic gate and the gate tunneling current of the fanout might change the output voltage rapidly [27]. Therefore, the detection of open drain/source in modern technologies is more challenging.

b) Bridging defects

Bridging defect is a short between any two or more signals. Resistive short between metal interconnects, short between transistor terminals, short between a terminal and power supply are examples of bridging defects [28].

c) Fin Open

Fin open occurs due to extremely small size of the fins, existence of airborne particles, or over-etching during the manufacturing process. Fin open reduces the FinFET driving current and can result in small delay defects [29].

d) Gate oxide short (GOS)

Gate oxide short is one of the dominant defects that is complex to analyze and model due to the 3D structure of FinFET. GOS defects are formed due to airborne particles, mask damages, and over-etching of the dummy poly-Si gate that result in the lack of oxide layers in various spots of the thin gate oxide [30]. Also, fin line edge roughness coupled with non-conformal deposition of the gate-oxide layer creates a potential for catastrophic failure causing breakdown, which results in large tunneling leakage or direct current path between the gate and the channel region [31].

e) Defects Caused by Process variations

Some of the major sources of process variations are lithography misalignment, over-etching, or under-etching issues that affect the fin geometry; random dopant fluctuations that affect threshold voltage, and variations in the gate oxide thickness [32]. Also, the line-edge-roughness and line-width-roughness (LWR) have considerable impact on the fin shape. The fin width and height influence the FinFET performance and reliability as the geometrical shape of the fin affects the transistor saturation current, the leakage and subthreshold slope [33-37]. Some of these process variations such as LER and LWR might directly lead to GOS defect which is in the focus of this research.

2.3 Fault Modeling

The physical imperfections or defects can be modeled by circuit elements such as resistors, capacitors, and current sources that can be used in the circuit simulations. The electrical impact of the defect on the circuit behavior is modeled as a fault [22]. The defect can be considered on logic, transistor, layout or physical structure level of abstraction. To characterize the circuit behavior in the presence of defects, the circuit response to the input stimuli must be compared with that of the defect-free circuit. A set of threshold values are also required to determine the fault occurrence [38]. Any deviation from the defect-free response that lies below the defined threshold is considered as non-detectable fault. According to [38], traditionally faults are classified as shown in Figure 2.9. Hard defects cause incorrect function of the circuit and are modeled as logical faults such as stuck-at fault. When the defect creates a short circuit path between the supply and ground it increases the circuit steady-state current and can be detected by I_{DDQ} test [39]. Soft defects

cause performance degradation that are modeled as delay faults, noise margin reduction, and I_{DDQ}.

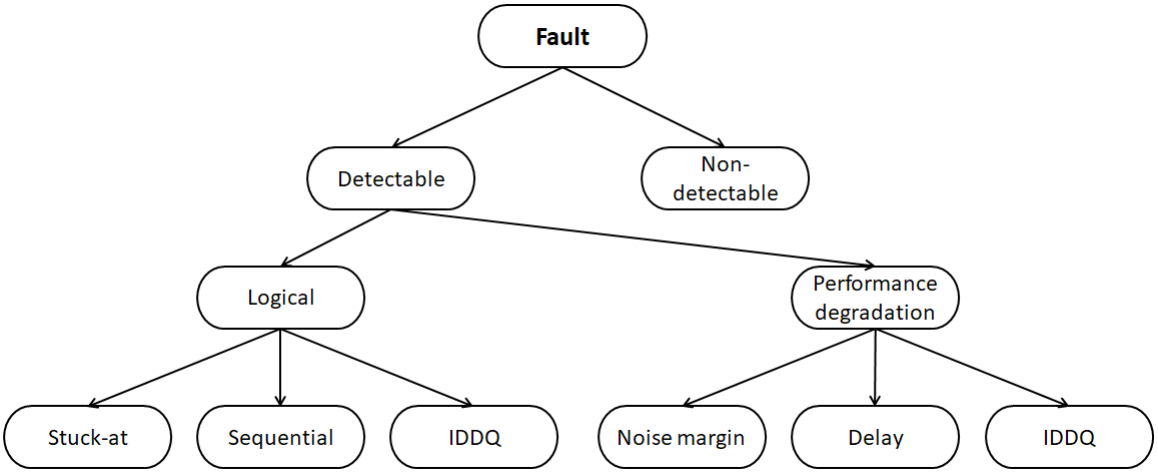

Figure 2.9 Traditional fault categories [38]

Fault models related to different levels of abstraction are described below.

a) Logic level

One of the most commonly used fault models at the logic level is stuck-at fault. Stuck-at fault is traditionally used for test generation since it is technology independent and could model defects such as a short between the gate input and power supply. However, there are many bridging defects that cannot be detected by stuck-at fault test patterns. For example, a bridge might cause a logic gate output to become an undefined value in the mid-rail range, which is modeled as functional change. Moreover, when a logic gate output is floating due to an open defect, it could maintain its previous value for a long time and the defect is modeled as a sequential fault [40].

Short defects affect the quiescent current through the supply and might be detected by I_{DQQ} or σI_{DDQ} test [41]. Due to the increase of leakage currents in nanometer technologies, it has become more difficult to distinguish the defect impact on the current from the process

variation and scaling impacts. The σI_{DDQ} test improves the defect detection by building process-aware leakage current models of the defect-free circuit and removing their impact from the current measurements [41].

Defects might also cause timing failures that are modeled as delay faults. Therefore, a defect can manifest itself as multiple faults depending on its impact on the circuit behavior. Delay faults are categorized into two major classes of gate delay fault and path delay fault. Gate delay fault model is a localized model that associates slow-to-rise and slow-to-fall delays as transition delay faults (TDF) for any gate, while path delay fault considers cumulative delays of the gates along a path [42]. A transition delay test requires at least two at-speed clock cycles, and the defect on a sensitized path is detected when the resulted delay exceeds the slack of the path, that is the difference between the active edge of the test clock cycle and the propagation delay of the path. Small delay defects (SDDs) introduce small delays that should be detected on the longest testable paths passing through the defect, since they are extremely timing sensitive [43], and may not be detected on critical paths.

b) Transistor level

At a lower level of abstraction, transistor level, the impact of defect can be analyzed more accurately. For example, in the inverter gate in Figure 2.10, an open defect on the drain of a PMOS transistor that is modeled as a transistor stuck-open fault is considered [44]. It can be detected by applying two consecutive test vectors ($AB = 11 \rightarrow 10$), where the first input pattern initializes the node and the second pattern excites the fault. The output of the faulty gate is at high-impedance and its voltage is traditionally considered to remain constant.

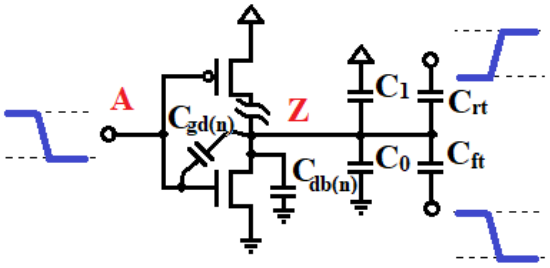

Figure 2.10 An inverter with open defect on the drain of the PMOS transistor [44]

However, in nanometer technologies the impact of transistor leakage currents and downstream parasitic capacitances including the line and load capacitances is non-negligible. Their influence might increase the output voltage above the threshold of logic 0 in this case and result in a test escape. The analysis of the capacitances, as the dominant factor, leads to obtaining the minimum line length that corresponds to the capacitances satisfying this condition [44].

c) Layout level

Due to the complexity of modern ICs, the use of standard library cells has been increased in the circuit design. The traditional fault models used for test pattern generation are successful in representing the defects on the inputs and outputs of the cells as well as the interconnects. However, many test escapes have been reported from the industry that majority of their root cause are the defects inside the library cells [4]. The substantial need for developing fault models based on the manufacturing process outcome and the physical nature of defect has led to the new CAT methodology that has recently been proposed [5]. The CAT methodology uses the layout and extracted parasitic data for each library cell for generating defect-based fault models. The layout data provides accurate information about the potential bridging and open defects inside the cell that can be used in extensive analog simulations to develop accurate fault models.

The next section covers major test types, ATPG techniques, and the advantages of cell-aware test in MOSFET technologies.

2.4 Design for Testability and ATPG Techniques

With substantial advances in microelectronic circuits, testing procedure has become as important as the design process in the product development. The test procedure evaluates the manufactured products to ensure their correct operation over a range of defined operating conditions, passing them to the end-users as "good devices" [45]. Depending on the type of circuit under test, tests are classified into three categories of digital, analog, and mixed signal [46]. Tests could also be classified as functional, parametric, and structural testing. Functional testing is performed to evaluate the correct functionality of the circuit in different operational modes, which is very time consuming and costly in complex digital circuits. Parametric testing measures specific current or voltage levels at the outputs in response to DC or AC input signals. I_{DDQ} is an example of DC parametric testing [45]. Structural testing is undertaken to detect the defects resulted from manufacturing process, assuming the design is verified to be correct [45]. It uses appropriate fault models representing the impact of fabrication defects to excite and propagate the fault effect to the outputs or observation points.

2.4.1 Design for Testability

During the last decades, design and test interactions has led to the development of design for testability techniques such as scan test and logic built-in self-test (BIST). Scan test improves the observability and controllability of sequential circuits and buried nodes

by adding scan cells that launch a test pattern and capture the response. The circuit operation is modified to choose between the normal mode and the test mode. In the test mode, test patterns are shifted into scan cells and the results are recorded. Full scan turns the sequential test generation to a combinational test, though with area overhead, I/O pin increase, and degraded performance [47]. Large designs are partitioned to combinational logic blocks based on the design and test constraints. The partitions are later optimized to maximize the yield/area [48]. The optimization is performed by clustering smaller blocks, using the existing redundancies, and portioning large blocks [48].

In recent years, digital ICs have been widely used in safety-critical applications such as automotive that require on-chip self-test to enhance the reliability of the system. Logic BIST techniques add circuits that generate test patterns and analyze the outputs of the system, with the cost of area overhead. In some applications, a more efficient solution might be the hybrid BIST, where a part of test data is stored on the tester and on-chip hardware generates a reduced set of test patterns. With the growing amount of test data in larger designs, test pattern compression techniques are used to reduce the amount of test data on the tester and the hardware overhead at the same time [47]. Although design for testability techniques help improve the test quality and ease the test pattern generation, ATPG algorithms need more realistic fault models to efficiently reduce the size of test data and improve the test quality. The traditional ATPG algorithms and new approaches including cell-aware test are covered in the next section.

2.4.2 Overview of ATPG Techniques

With the growing number of transistors and I/O pins in digital systems, test patterns cannot be manually generated. Earlier algorithms inserted single stuck-at faults at different nodes to find test vectors that could detect the faults. However, considering a set of test vectors that detect all single stuck-at faults of a circuit, it might happen that a fault masks another specific fault and it could not be detected by any of the test vectors [49]. Therefore, multiple-detect, or N-detect, ATPG algorithms were proposed in which every single stuck-at fault is detected by at least N unique tests. As expected, multiple-detect tests increased the probability of detecting hard-to-detect defects [50].

Physical-aware N-detect ATPG further improved the test quality since it used the localized nature of defects [51]. It considers a large test set and a neighborhood of radius r on the layout for each stuck-at fault and applies specific rules to include proper signal lines in the neighborhood. The test procedure selects the tests with specific number of unique neighborhood states that detect the targeted fault. However, there are challenges in the selection of number of neighborhood states and the increase in the test set size.

Another ATPG method used is known as gate exhaustive testing (GE), which applies all possible input combinations to the circuit input to find the response at primary outputs or scan cells [52]. One of the major problems with gate exhaustive testing in complex circuits is the increased test set size due to the large number of circuit inputs. Although the results in [52] show that gate exhaustive test has higher coverage than N-detect test with N varying from 2 to 15, the test generation time of gate exhaustive method is 3 to 200 times larger than the N-detect tests. Moreover, some input combinations might not be observable when there is no sensitized path from that specific gate to the observation points such as circuit

outputs or scan cells. This problem might be resolved by partitioning the circuit into segments with small input numbers; however, finding the optimum partitions would be a challenge by itself. Also, the defects that need a sequence of two or more test patterns could not be detected by this method.

2.4.3 Cell-Aware Test (CAT)

As integrated circuit complexity has grown and standard library cells have been widely used in circuit design, the traditional fault models such as stuck-at and delay faults don't provide the required fault coverage, since they only consider faults on the cell inputs, outputs and interconnects [5]. Moreover, the manufacturing complexities in new technologies have increased the number of defects inside the standard cells that could not be detected by traditional ATPG tools. Therefore, the defect-based cell-aware test methodology has been proposed in [5], that uses the cell layout data to extract a SPICE netlist that includes the parasitic resistance and capacitances as well. Different types of cell-internal defects such as open, short, transistor drive strength, transistor leakage, and port bridge/open are simulated by inserting representative resistors in the extracted netlist. Extensive analog simulations are performed on the circuit considering different values for each injected resistor at a time to find all input combinations that are useful for detecting that specific defect. The defect is detected by determining the output voltage at specific strobe times. If the difference between the output voltage and the defect-free value exceeds a defined threshold value, e.g. 60% of the supply voltage, the input pattern is saved as a test vector that detects the specific defect. For example, in the case of short defects the CAT procedure replaces every parasitic capacitor between two nets with a resistor and

decreases its value in a specified range and performs extensive analog simulations to find all output erroneous values and build a short defect list. Similarly, for possible open defects, the parasitic resistors at each line segment are increased in repeated simulations. The single-pattern (i.e. static) inputs that detect each specific short defect and two-pattern (i.e. delay) inputs that usually target open defects are also used in developing the defect matrix. The exhaustive defect matrix is then optimized to obtain the CAT library view that will be used by the ATPG algorithm to generate appropriate test patterns. Figure 2.11 shows the cell-aware fault model generation and how it is integrated in the ATPG flow.

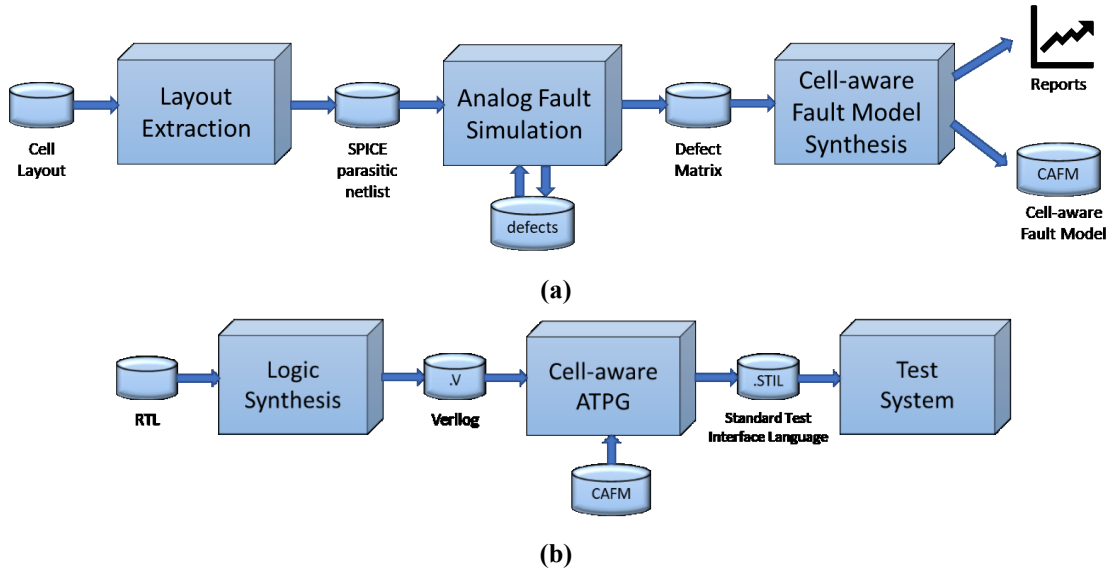

Figure 2.11 (a) CAT library view generation, (b) CAT-ATPG flow [5]

The CAT static and delay tests are proven to detect more hard-to-detect defects than the traditional ATPG and reduce the defect rate in production tests in nanometer technologies [5]. For example, the high-volume production test results of 800000 AMD 32 nm processors show that by adding 7300 additional CAT patterns to have a total of 16900 patterns, the defect rate is reduced to 885 DPPM [53]. Even with the same number of test patterns, the CAT methodology improves the defect coverage comparing to the stuck-at

and transition delay patterns. Moreover, by applying additional CAT patterns, up to 4% improvement in the defect coverage is achieved [53].

In [54], combinational library cells from a 65 nm MOSFET technology are tested for all port stuck-at faults and internal bridge defects with both gate exhaustive and cell-aware tests. Considering a 4-input multiplexer as an example, the largest number of test patterns required in the GE test for this 8-input cell is reported to be more than 2 times of the CAT pattern count. The evaluation of both test methods on 10 circuits have also shown that the number of GE test patterns is, on average, 4 times the test patterns of the cell-aware test. The exponential increase of GE test pattern count with the number of inputs in large designs would substantially increase the test time and hence, the production cost. The other important advantage of the CAT methodology is its improved defect coverage compared to the GE test, which in the mentioned study is an average of 1.8 times [54].

In addition, the CAT methodology is also used in failure diagnosis to identify the defect location with improved resolution compared to the traditional layout-aware diagnosis [55], [4].

2.5 Summary

This chapter provides the required background information regarding the FinFET device and different types of defects and fault models, as well as a brief overview of the traditional ATPG techniques including the defect-based approach of CAT methodology. In the next chapter, the current research in GOS defect and fault modeling and the related test issues are presented.

Chapter 3: Literature Review

In this chapter, a literature review of the previous work in GOS defect modeling as well as the fault models in cell-aware test are presented. The recent work on FinFET defects is described, as well as the GOS mechanism and modeling in MOSFET. The previous research in gate oxide short defect in FinFET are reviewed. This is followed by the description of fault models in CAT methodology and their shortcomings in FinFET technology.

3.1 FinFET Defects

Bhoj et. al, investigated the leakage-delay behavior of simple logic gates consisting of double-gate FinFETs [56]. Different logic styles are simulated considering a cut on the back-gate of the FinFET. Since the back-gate bias has strong influence on the FinFET threshold voltage, it affects the delay-leakage characteristics of the gate. In the case that coupling capacitance of the neighboring lines on the floating back-gate is dominant, the variation in the back-gate voltage causes transition delays, and exponential increase in leakage for a range of back-gate bias. When the layout induced capacitance is negligible, the impact of front gate or S/D induced capacitance dominates based on the fin thickness and gate-drain underlap size. However, the proposed approach cannot be extended to the case where any of the mentioned capacitances are not negligible.

Mesalles et. al, investigated the behavior open defects on various locations of the gate of multi-fin and multi-finger FinFETs considering the cell parasitic capacitances [40]. It is shown that some defect locations exhibit different behavior than in MOSFET technology due to partial disconnection of a few fins. However, the detectability of gate open decreases

in multi-fin structure compared to the single-fin FinFET and is exacerbated in multi-finger transistors as the number of other conducting paths between the output and the rails increases. An accurate defect characterization for open defects depends on the technology specifications, the cell and interconnect layouts, and the test speed which is not in the scope of this research.

Bridging defects in FinFET-based logic cells such as 2-input NAND gate are analyzed in the literature [57], considering resistive short between middle-of-line (MOL) interconnection layers and multi-fin, multi-finger FinFETs. It is shown that in some defect locations the critical resistance has strong dependency on the number of fins/fingers, whereas in other locations there is weak or no related dependency. A metric is defined as bridge defect criticality to identify the short defects that have higher probability of occurrence and lower critical resistance. Based on this defect classification, new test strategies should be designed that are not in the scope of this thesis.

Fin open impact on logic gate delay behavior is investigated in [30]. It is shown that in multi-fin FinFET as the number of broken fins increases the transition delay becomes large enough to be detected by two-pattern tests [30]. However, when the number of broken fins is small, specifically in multi-fin structures timing-aware ATPG techniques are required to tackle the issue of small delay defects in test pattern generation [58], [59].

Although GOS has been widely investigated in planar MOSFET [32], [60], [61], very limited research addressed the FinFET technology. A review of what is available in the literature is presented in the next sections.

3.2 GOS Modeling in MOSFET

Gate oxide short defect creates a low resistance path from the transistor gate to the channel. It can be considered as a pinhole in the gate dielectric layer that provides an undesired current path to the channel. If the pinhole is in the oxide overlap area between the gate and source/drain, it can be modeled as a gate-drain or gate-source short [61]. If the pinhole occurs in the oxide above the channel, it is known as gate-to-channel short or GOS. GOS is a dominant defect [6], because of the continuous reduction in the transistor's critical dimensions and its dielectric thickness. This defect is also complex and hard to analyze since it modifies the transistor characteristics in different manners depending on its size and location [6]. Therefore, it is important to develop accurate defect models that can be used in obtaining reliable fault models and practical test patterns. Figure 3.1 shows an NMOS transistor with a pinhole in the oxide layer.

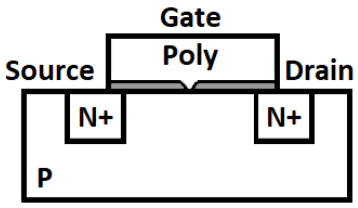

Figure 3.1 An NMOS with a pinhole in the dielectric layer [6]

There are two types of GOS model for MOSFETs, split and non-split [63]. In the split model, a bidimensional array of transistors are used as an electrical model of a defect-free transistor, Figure 3.2 (a). One side of the array represents the source, the other side of the array represents the drain and all transistor gates are connected to form a common gate terminal. The pinhole is modeled as a resistor, R_{GOS}, that is connected between the common gate and one of the internal nodes of the bidimensional array [61], [6], as shown in Figure 3.2 (b).

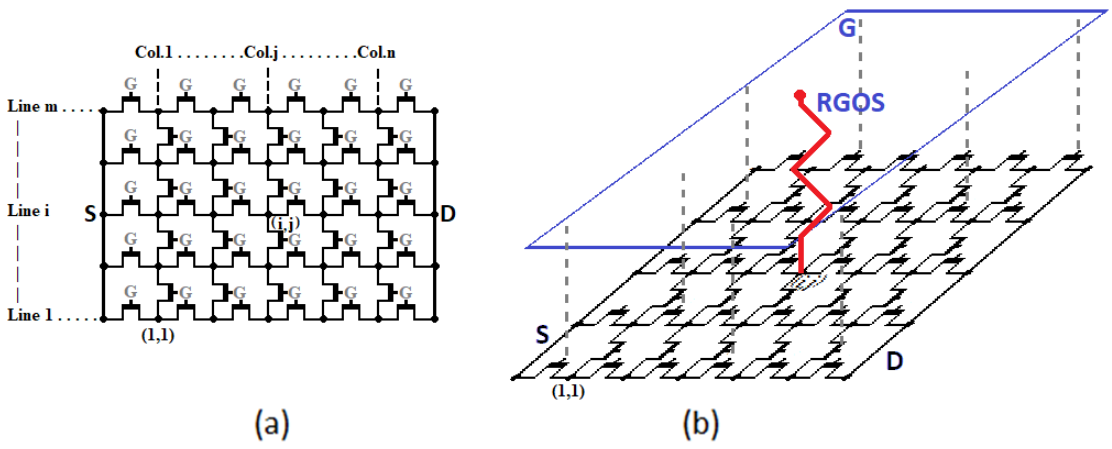

Figure 3.2 (a) Bi-dimensional model of a defect-free MOSFET, (b) GOS defect model [6]

Assuming all transistors in the array to be minimum-sized, the equivalent transistor that the defect is applied to, Figure 3.2(b), has larger effective width and length. Therefore, the split model cannot accurately represent a minimum-size defective transistor [61].

The non-split models are categorized in two groups, linear and non-linear [63]. Figure 3.3 shows the simple linear non-split model that represents the ohmic behavior of the pinhole in the oxide overlap region between the gate and source or drain. The GOS is modeled by non-split linear model in CAT methodology.

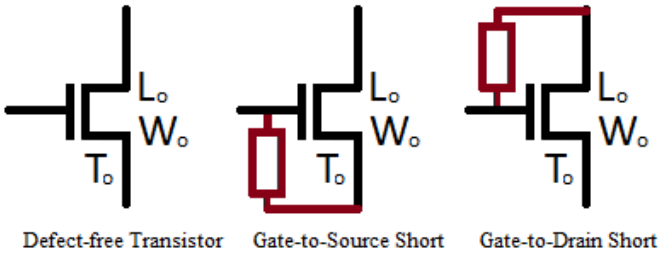

Figure 3.3 Linear non-split model [63]

In [32], authors have proposed the nonlinear non-split model shown in Figure 3.4 (a). Since the defective MOSFET has a lower drain current, the defect-free transistor is modified to have a smaller width, and two additional transistors are used to model the nonlinear current impact of the GOS. The resistive effect of the defect is modeled by adding a resistance

between the common gate and the common drain of the additional transistors. In this model, all transistor lengths are larger than minimum-size due to fitting issues which increases the gate capacitance compared to the defective MOSFET [61].

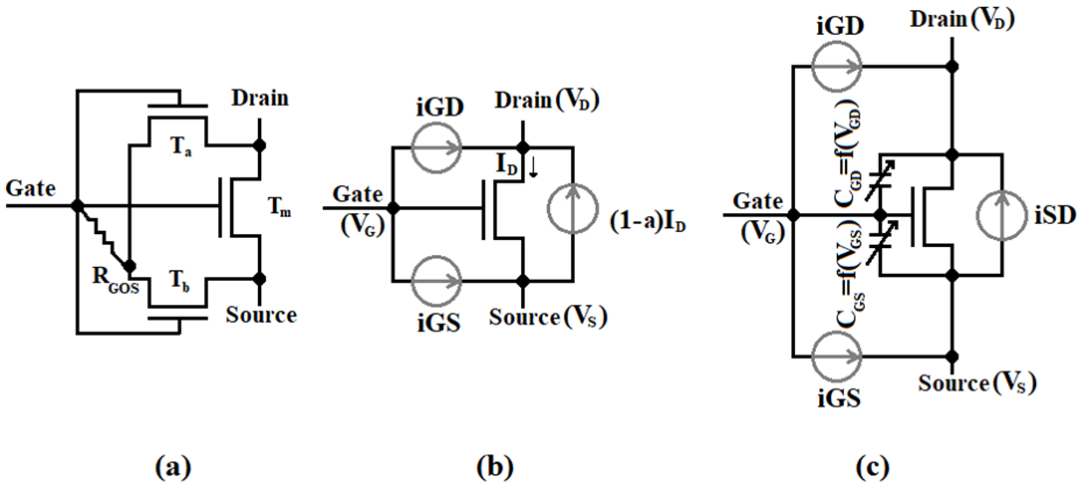

Figure 3.4 Non-split GOS defect models in MOSFET [61]

In [60], the defective MOSFET is modeled by a defect-free transistor and three current sources, as depicted in Figure 3.4 (b). Two voltage-dependent nonlinear current sources, i_{GS} and i_{GD}, model the impact of the defect on the gate current, and the third current source models the saturation drain current reduction by using a fitting parameter, α, in the current source with the value of $(1-\alpha)I_D$.

It is shown that for both models in Figure 3.4 (a) and (b), the I_DV_D curves deviate from the GOS-impacted results in TCAD for small V_G values [61]. Also, the transient response of an inverter with either of these models on NMOS deviates from the GOS results in the output level and slope due to the differences in the capacitances [61].

In [61], three voltage-dependent nonlinear current sources along with two capacitors are utilized, as shown in Figure 3.4 (c), to improve the transient behavior. The current sources are represented by polynomial functions, i_{GS} with a second-order, and i_{DS} and i_{GD} with

third-order polynomials. The DC and transient simulations have shown better matching results with a GOS-impacted MOSFET.

3.3 Previous Work on GOS Defect in FinFET and Shortcomings

The current research in this area consists of only two publications that are covered in this section.

In [30], authors have simulated different FinFET defects such as fin open, stuck-on, and GOS at the circuit level. The GOS is modeled by a 3×2 array of MOSFET transistors as shown in Figure 3.5.

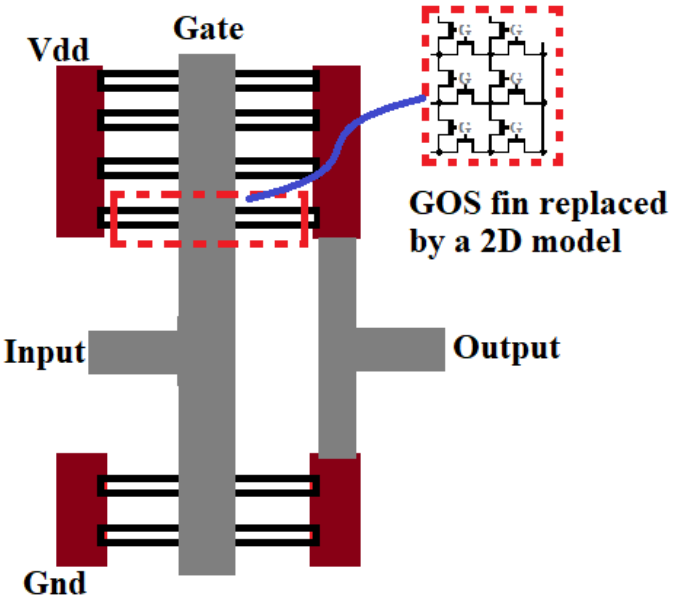

Figure 3.5 Gate oxide short model in [30]

Simple logic gates such as inverter and NAND gate are simulated with this GOS model representing a defective NFinFET. The DC and transient simulations are performed to measure the impact of the number of defective fins on the output voltage and transition delays.

The lumped-element model used in this publication is same as the bi-dimensional model for MOSFETs that is applied in relatively old planar technology nodes. This approach cannot accurately represent the GOS impact in the new technology generations, especially in the 3D structure FinFET. Although increasing the number of transistors in the array would increase the accuracy, it results in relatively larger equivalent transistors. However, these array transistors are difficult to model, also it becomes impractical to model the GOS defect in minimum-sized transistors.

In [64], the GOS defect in independent-gate (IG) and tied-gate (TG) FinFETs in a 25 nm SOI technology is simulated in the TCAD environment. The defect is represented by a 5 nm pinhole injected on the fin sidewall. It is shown that in the IG FinFET with a pinhole in the front-gate, the gate leakage current flows throw the pinhole to the drain and results in negative drain current at low drain bias. While in the FinFET with a pinhole in the back-gate, the gate leakage current is very small since the back-gate is connected to ground. In the tied-gate FinFET, the drain saturation current of the defective transistor is slightly reduced, and the negative I_D at linear bias is same as in MOSFET.

The literature [64] has not yet investigated the GOS impact on the transient behavior of FinFET. It has applied the GOS defect on different transistors in FinFET SRAM designs in the TCAD environment to propose new test methods. It has not extracted defect models that could be used in circuit-level simulations. The literature has not yet proposed an accurate GOS defect model in FinFETs, nor considered the effect of pinhole size and location on the FinFET behavior. The impact of GOS defects on complex gates used in logical designs has not yet been studied.

3.4 FinFET Fault Models in CAT Methodology

In the CAT flow, for each standard cell, a SPICE transistor netlist including parasitic resistances and capacitances is extracted from the layout data. The cell-internal defects are modeled as stuck-at or transition delay faults based on extensive DC and transient simulations. The corresponding test patterns are used to build a defect matrix that is later optimized to generate the cell-aware fault model.

Figure 3.6 shows the supported defect types in a sample netlist [5]. These defect types are Open, Bridge, transistor leakage (Tleak), transistor drive strength (Tdrive), Port Open and Port Bridge [5]. The defects are modeled by simple passive components such as resistors, and the defects related to transistors are limited to bridging between transistor terminals and the drive strength defect.

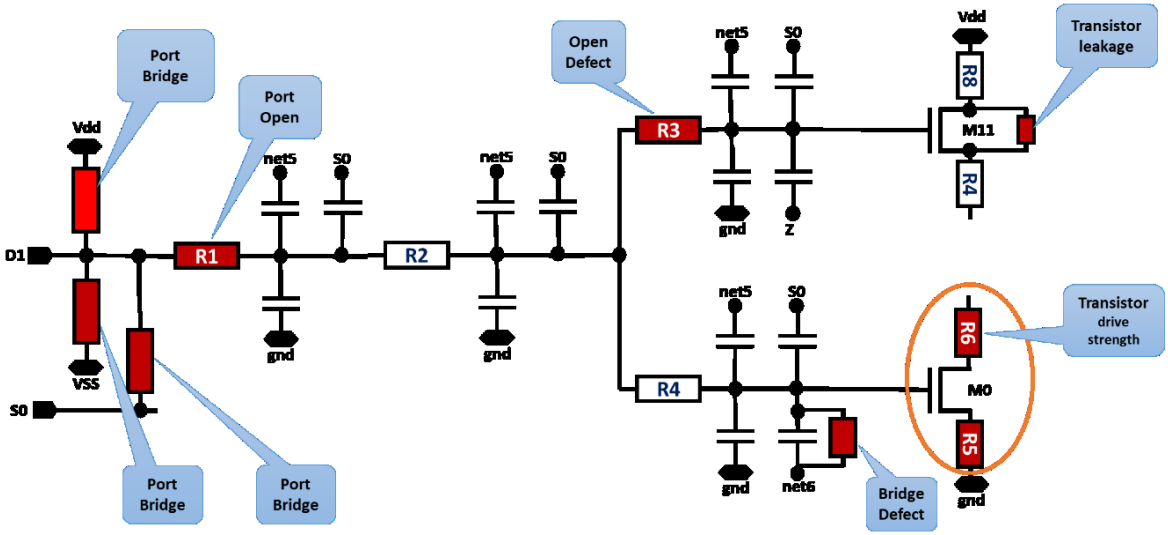

Figure 3.6 Schematic representation of sample netlist with inserted defect types [5]

A complex defect such as the GOS is modeled with either a gate-source or gate-drain short resistance. However, none of these models can represent the nonlinear impact of the GOS on the drain current.

The CAT methodology is reported to be applicable to FinFET technology in a similar manner as planar CMOS [5]. A multi-fin structure is considered as parallel transistors and the mentioned defects are modeled with additional resistors on any of the fins, as shown in Figure 3.7.

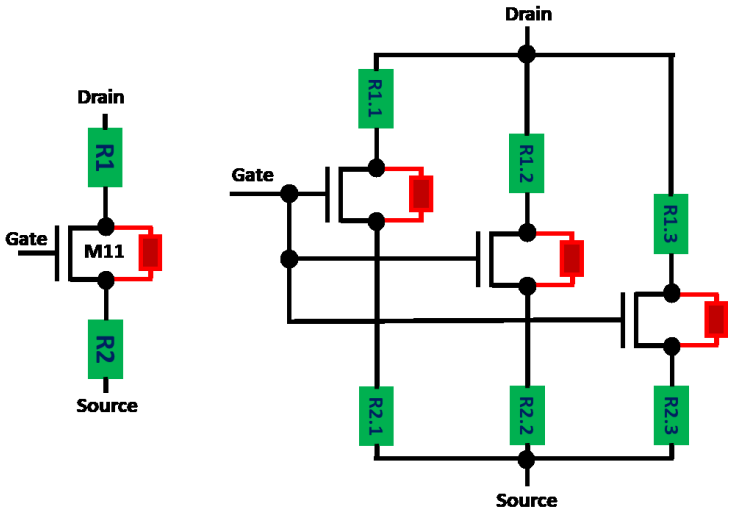

Figure 3.7 Leakage and drive strength defects in single-fin and multi-fin FinFETs [5]

However, applying the CAT methodology to FinFETs has some limitations. Since the layout information used for MOSFET technology is two dimensional, it doesn't accurately reflect the 3D nature of FinFET transistor especially in the presence of a complex defect such as gate oxide short.

Despite the success of the cell-aware test in detecting bridging and open defects inside standard library cells, significant functional test failures have been reported by the industry even after CAT screening [4], [65]. Therefore, in the FinFET technology there is a need for development of more accurate fault models for complex transistor-related defects inside standard cells [4].

In this thesis, a comprehensive investigation of GOS impact on FinFET devices is performed by considering various defect injection scenarios on FinFETs with different fin

shapes from rectangular to triangular form. It characterizes the DC and dynamic behavior of a FinFET in presence of GOS defect to develop accurate GOS defect models for each scenario. These models are then utilized in circuit-level simulations to obtain appropriate fault models.

3.5 Summary

In this chapter, the previous work on the GOS defect and FinFET fault models in CAT methodology were presented. The GOS in FinFETs cannot be represented by simple passive components due to the nonlinear impact of a defect on transistor behavior. The CAT methodology has significantly improved the defect coverage related to the defects inside standard cells, however, considerable test escapes are reported.

In the next chapter, the FinFET 3D structure in TCAD environment and our GOS defect injection strategies are presented. The device models, tools, and the defect characterization setup that is utilized in developing GOS fault models are also explained.

Chapter 4: FinFET Defect Injection and Characterization Methodology

In this research, the GOS defect is simulated at the device level using Sentaurus TCAD tools. The GOS defect is introduced to the FinFET 3D structure as a pinhole. The first section of this chapter presents the FinFET 3D structure, as well as the tools and models required for simulating the FinFET at the device and circuit levels. The GOS defect injection scenarios that are considered in developing the defect models are covered in the second section. Finally, the characterization setup for device- and circuit-level simulations are described in the last section.

4.1 FinFET Device Models and Tools

Recently, Synopsys has provided 3D bulk FinFET templates [18], from which the so called 14/16 nm FinFET template is utilized for generating defect models in this research. Although these FinFET templates are generic, their design characteristics and dimensions are representative of the current technology. The top view of the NFinFET for the Synopsys' generic process is shown in Figure 4.1.

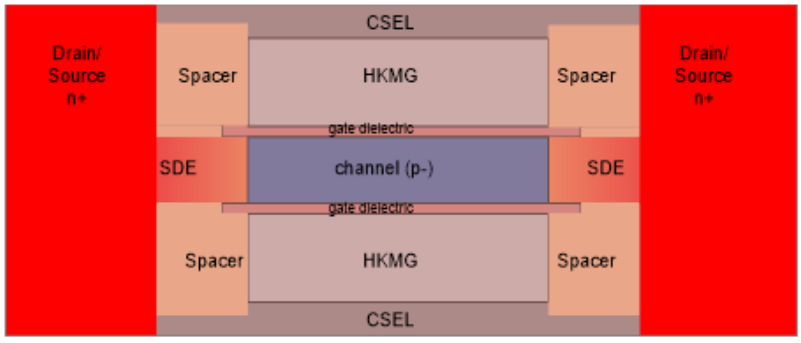

Figure 4.1 Top view of NFinFET for the Synopsys' generic process [66]

The important processing steps and physical models that are used in the TCAD simulations of the Synopsys' template are described in the following subsections.

4.1.1 Sentaurus Process Models

The 3D structure of the bulk NFinFET for the Synopsys' generic process is presented in Figure 4.2. The replacement metal-gate bulk FinFET process consists of fin formation step followed by punch-through stopper ion implantation to prevent leakage current below the channel [17]. The shallow trench isolation (STI) oxide is deposited to isolate the adjacent transistors. A two-layer high-k dielectric is deposited over the fin top and sidewalls, which is covered by a dummy poly-Si gate. This dummy gate is removed and replaced by the metal gate after the spacer formation. The gate-last process allows the metal gate to be deposited after high temperature S/D epitaxial formation and enhances the channel strain as well. The drain and source are heavily-doped (n+) epitaxial regions that are stress-engineered by adding SiGe for PFinFET, and SiC for NFinFET, which applies longitudinal compressive and tensile stress to the channel respectively [18]. The source/drain extensions are covered by spacers with low-k material to minimize the parasitic capacitance between the gate and S/D regions [21]. The contact-etch-stop-liner that is deposited over the metal gate and spacers applies stress to the channel and improves the saturation current. Finally, the contacts are wrapped around the fin to minimize the contact resistance, and the inter-layer dielectric with low-k material is used to isolate trench-shaped contacts. It should be noted that the mechanical stress simulation and mesh refinements are also performed at different processing steps [18], [67].

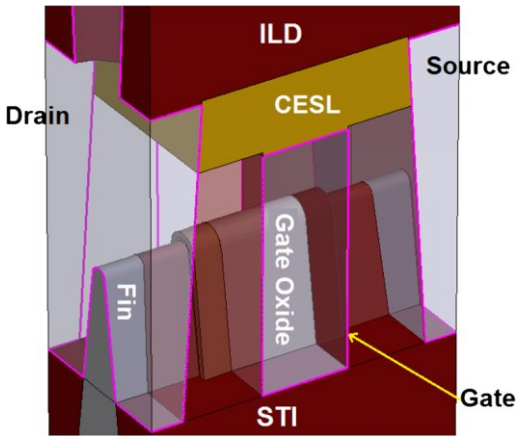

Figure 4.2 Bulk NFinFET 3D structure [18]

The design parameters of the bulk FinFET are listed in Table 4.1. The fin height is 35 nm with corner rounding of 2.5 nm radius at the top. The channel length is 25 nm, which is covered by two-layer gate dielectric, consisting of SiO_2 and HfO_2, with the overall thickness of 2.3 nm. The device has the feature to be expanded to multi-fin structure with a fin pitch of 48 nm. The bottom width of the fin is 15 nm, and it is possible to realize different fin geometries, as rectangular or trapezoidal, by selecting one of the optional values for the fin top width: $W_{top} = 15, 10,$ or 5 nm.

Table 4.1 Bulk FinFET design parameters

Parameter	**Val.**	**Parameter**	**Val.**
Fin height, H (nm)	35	Fin pitch (nm)	48
Fin width, W_{bottom} (nm)	15	Poly pitch (nm)	90
Gate oxide thickness, Tox (nm)	2.3	Channel length, L (nm)	25
Channel doping, Boron (/cm^3)	$2e^{18}$	Metal gate work function (eV)	4.6
S/D doping, Phosphorus (/cm^3)	$2e^{20}$		

Figure 4.3 shows the double-fin FinFET structure with the specifications mentioned in Table 4.1 and simulated by Sentaurus process.

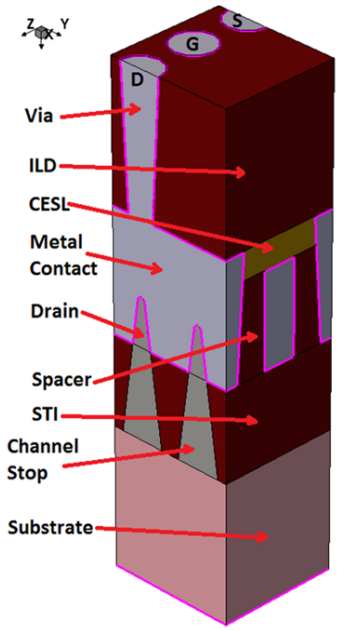

Figure 4.3 Double-fin defect-free FinFET structure with $W_{top} = 5$ nm [18]

In this research, the GOS defect is introduced on different fin shapes to study its impact on the FinFET behavior. However, the defect model is developed based on the triangular fin with rounded corners, $W_{top} = 5$ nm, since it is widely used in high-volume manufacturing of FinFETs [8].

4.1.2 Sentaurus Device Models

The electrical characteristics of the bulk FinFET are obtained with the aid of Sentaurus Device simulations, considering the following major physical models: the thin-layer mobility model, and inversion and accumulation layers mobility models that consider the crystal orientation on the fin top and sidewalls; the bandgap narrowing effect due to high impurity concentrations that is considered in effective intrinsic density calculations; the stress effect on the band structure, on density of states (DOS), and on mobility enhancement; different scattering mechanisms that degrade the mobility in high-k metal-

gate structure; the effect of band-to-band tunneling in modeling the leakage; and the drift-diffusion (DD) model with adjusted velocity saturation [18], [68]. All these physical models are used to solve the Poisson and carrier continuity equations and obtain the device electrical characteristics [18].

There are two commonly used models for the carrier transport in the device simulations, drift-diffusion and hydrodynamic (HD). Since there is no empirical data for the GOS-impacted FinFET available to us, both models are considered in this research. The HD model is more accurate in reproducing velocity overshoot and it considers the impact of lattice and carrier temperature variations on the carrier transport [68]. Therefore, it is worth investigating the FinFET behavior in presence of GOS with this model in the first part of this work.

However, HD model suffers from extremely degraded convergence and longer simulation time than DD, as presented in the next chapter. Moreover, the DD model with adjusted velocity saturation is the carrier transport model that is supported by Synopsys for FinFET templates [18]. Therefore, in the second part of this work the DD model is used in all device simulations performed for developing the defect models.

4.1.3 HSPICE Model Parameters

There are two major parameter extraction procedures for the Berkeley Short-channel IGFET Model - Common Multi-Gate (BSIM-CMG) that lead to either one global set of parameters for transistors with different channel lengths and widths, or a local parameter set for a transistor with a specific channel length [69]. Since the GOS defect model is developed based on the Synopsys' FinFET template with $W_{top} = 5$ nm and $L_g =$

25 nm, local parameter extraction is performed for both NMOS and PMOS transistors. The parameter extraction procedure is summarized in Table 4.2. During the initialization step the model and process control parameters are set. The capacitance optimization is performed using $C_{gg}V_{gs}$ curve (total gate capacitance vs gate-source voltage) at linear bias. First, a long-channel device is considered to obtain an estimate of the mentioned parameters in Step 2 of Table 4.2, then the parameters are refined for the short-channel device.

Table 4.2 Local parameter extraction procedure [69]

Step	Data Used	Bias	Description
1	-	-	Initialize process and model control parameters
2	$C_{gg}V_{gs}$	$V_{ds} = 50$ mV	Extract parameters related to quantum mechanical effects
3	$I_{ds}V_{gs}$, g_m	$V_{ds} = 50$ mV	Extract parameters related to subthreshold slope and mobility
4	$I_{ds}V_{gs}$, g_m	$V_{ds} = 0.8$ V	Extract parameters related to DIBL and velocity saturation
5	$I_{ds}V_{ds}$, g_{ds}	Various V_{gs}	Extract parameters related to output conductance
6	$C_{gg}V_{gs}$	$V_{ds} = 0.8$ V	Extract parameters related to channel length modulation for capacitance

In Step 3, parameters related to the subthreshold slope at low drain bias are extracted using the I_dV_g (drain current vs gate-source voltage) graphs with logarithmic scale. The mobility-related parameters are extracted from the simultaneous observation of the I_dV_g and G_mV_g (transconductance vs. gate-source voltage) graphs through the refinement process. In Step 4 of Table 4.2, the parameters related to the DIBL effect, and velocity saturation at $V_d = 0.8$ V are extracted. The initial estimates of the velocity saturation related parameters are obtained from the I_dV_g and G_mV_g graphs and are later refined from the I_dV_d characteristics. In Step 5, the I_dV_d curves and the output conductance curves at different V_{gs} values are used to extract the related parameters. Iterations between this step and the previous one results in optimized results. Therefore, the impact of each parameter variation on the

characteristics should be observed simultaneously. Finally, in Step 6, the gate capacitance at the saturation bias is used for extracting the related parameters that affect C_{gg} and the graph curvature at high V_{gs} values [69].

The extracted HSPICE model parameters are utilized in circuit-level simulations within the defect characterization setup to develop appropriate fault models and obtain the test patterns that can detect the GOS defect.

4.2 GOS Defect Injection Strategy

The GOS defect is introduced as a pinhole with a specific size on the gate dielectric. In the literature, the pinhole shapes of cube [60] or cylinder [64] with the side or diameter size of $2.5\ \text{nm}$ to $10\ \text{nm}$ are reported. Therefore, in this research cuboid pinholes of $2\ \text{nm} \times 2\ \text{nm}$, $4\ \text{nm} \times 4\ \text{nm}$, and $8\ \text{nm} \times 8\ \text{nm}$ are simulated on distinct locations of the fin. It should be noted that $1\ \text{nm} \times 1\ \text{nm}$ pinhole was initially simulated; however, due to its very close results to those of $2\ \text{nm} \times 2\ \text{nm}$ it is not reported. The pinhole is simulated by removing a small cuboid of the gate dielectric and filling it with the gate material.

4.2.1 GOS in Rectangular and Trapezoidal Fins

The complete FinFET structure with the rectangular fin shape is presented in Figure 4.4. A small pinhole is inserted on the fin top (Top), sidewall, and along the channel. Since the fin height is $35\ \text{nm}$ and the fin corners have roundings, considering a maximum size of $8\ \text{nm} \times 8\ \text{nm}$ pinhole, the defect center is chosen to be located at depth of $10\ \text{nm}$ (Up), $17.5\ \text{nm}$ (Mid), and $25\ \text{nm}$ (Low), as shown in Figure 4.4 (b) to (e). The defect might also happen along the channel length, closer to drain or source, located $8\ \text{nm}$ apart from the

center as depicted in Figure 4.4 (f) and (g). The extension letters D, C, and S stand for closer to drain, center, and closer to source, respectively.

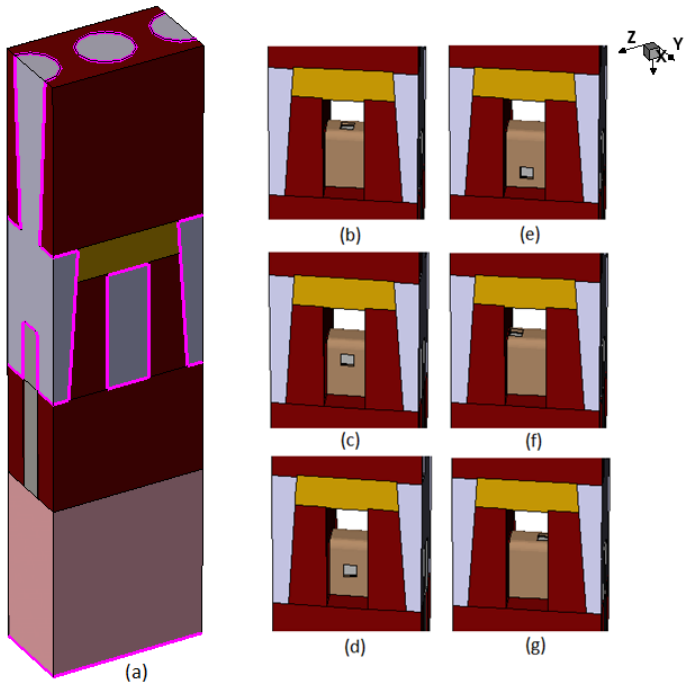

Figure 4.4 (a) Bulk NFinFET with rectangular fin shape, with injected 8 nm × 8 nm pinhole at the (b) fin top center (Top-C), (c) upper sidewall of the fin (Up-C), (d) middle sidewall (Mid-C), (e) lower sidewall (Low-C), (f) fin top close to drain (Top-D), and (g) fin top close to source (Top-S)

In a similar manner, the GOS defects are introduced to the trapezoidal fin with $W_{top} = 10$ nm. The transistor behavior is also studied under the same pinhole scenarios. It is expected to observe different defect impacts on the sidewalls of the trapezoidal fin due to the difference in the surface crystal orientation.

4.2.2 GOS in Triangular Fins

Triangular fins are more practical to realize in the manufacturing process since having non-vertical sidewalls results in easier ion implantation of S/D extensions, easier etching of the gate spacer, and better trench filling [72]. Due to these advantages, Intel

brought this fin geometry to high-volume production at the 22 nm node for the first time in 2011 [8]. Moreover, in fins with smaller top widths the gate has better control over the channel that results in lower subthreshold slopes. Therefore, in this work the GOS defect models are obtained for the FinFET with triangular fin shape ($W_{top} = 5$ nm).

Since the GOS defect might occur at any location of the fin sidewall, the defect injection scenarios are extended to the areas closer to the source or the drain on the fin sidewall. Figure 4.5 shows different injection scenarios of the sidewall pinhole.

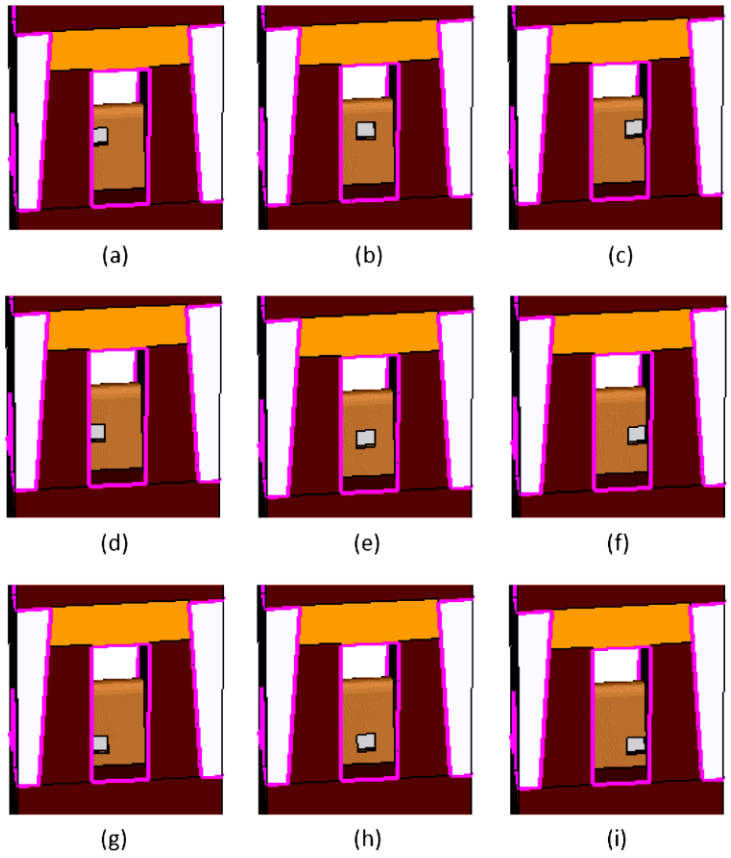

Figure 4.5 Pinhole injection scenarios at the a) Up-D, b) Up-C, c) Up-S, d) Mid-D, e) Mid-C, f) Mid-S, g) Low-D, h) Low-C, and i) Low-S

The simulation conditions for simulating the effect of GOS defect in the FinFET structure and the defect characterization setup are explained in the next section.

4.3 Defect Characterization Setup

The experiment setup for simulating the GOS defect involves the introduction of a pinhole with a specific size in the gate dielectric. The pinhole provides a low-resistance path from the gate to the channel and affects the device behavior at all bias conditions. In the device level simulations, the transistor DC behavior is examined at three different bias situations: (a) $V_{DD} = 0.05$ V to ensure the operation at the linear region, (b) $V_{DD} = 0.8$ V for the saturation region, and (c) $V_{DD} = 0.8$ V and gate voltage of 0.0 V for operating at the cut-off mode with maximum leakage. The dynamic behavior of the transistor is also studied at different bias voltages by performing small signal AC analysis at the frequency of 100 MHz in the mixed mode to extract the FinFET capacitances.

In the circuit-level simulations, the experimental setup shown in Figure 4.6 is used to examine the circuit under test (CUT). It consists of single-fin inverters used as input driver and output unit load. The supply voltage of the driver and CUT is separate from the rest of the circuit to consider the impact of leakage due to GOS on the driver stage. Also, the DC current of the driver and the CUT are measured at both low and high CUT output states as I_{dqL} and I_{dqH}, respectively.

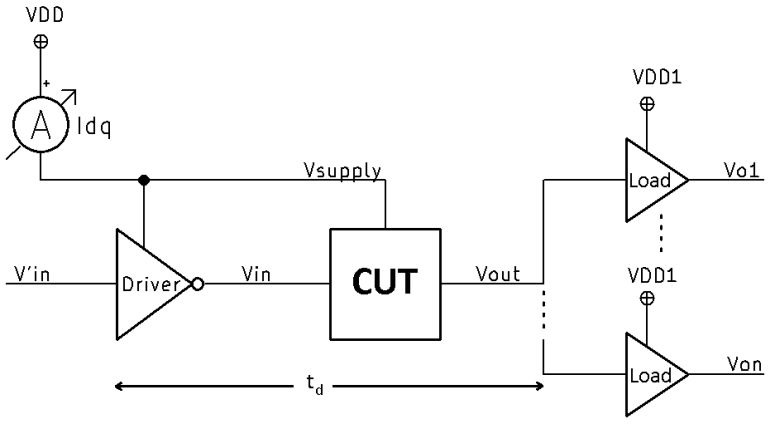

Figure 4.6 Simulation setup for defect characterization

The Voltage Transfer Characteristics (VTC) is used to obtain the noise margin as shown in Figure 4.7. For an inverter, two critical points on the VTC where the slope of the curve is equal to -1, are indicated by V_{IL} and V_{IH}. V_{IL} is the maximum input voltage that generates a valid logic 1 state at the output, and V_{IH} is the minimum input voltage that produces a valid logic 0 at the output. V_{OH} is the largest output voltage where the VTC slope is equal to -1, and V_{OL} is the smallest output voltage that the VTC slope is equal to -1.

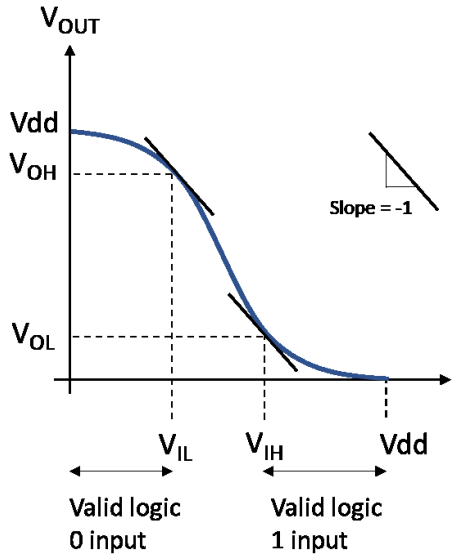

Figure 4.7 The voltage transfer characteristics and noise margins [73]

Considering two cascaded inverters such as M and N, the output voltage of inverter M must reside in the valid range to be recognized as a proper input. The noise margins are obtained from Equation 4.1,

$$NM_L = V_{IL} - V_{OL} \tag{4.1.a}$$

$$NM_H = V_{OH} - V_{IH}. \tag{4.1.b}$$

The percentage of the noise margin reduction in the defective logic gate compared to a defect-free one is obtained from Equation 4.2,

$$\%\Delta_{NM} = \frac{NM_{defective} - NM_{def-free}}{NM_{def-free}} \times 100. \tag{4.2}$$

The threshold values that are considered in the defect characterization process [38], are presented in Table 4.3. The absolute values of the noise margins are used in the table. Depending on the circuit under test and the injected pinhole scenario, the noise margin that is reduced due to the defect impact, either NM_H or NM_L, is used in fault modeling.

Table 4.3 Noise margin fault classification for GOS defect scenarios

Condition (%)	Fault
$20 < \Delta_{NML} \leq 50$	VTC_L1
$\Delta_{NML} > 50$	VTC_L2
$20 < \Delta_{NMH} \leq 50$	VTC_H1
$\Delta_{NMH} > 50$	VTC_H2

In an inverter, the effective drive current (I_{eff}) provides an estimate of the propagation delays, even in sub-nanometer FinFET technologies [74]. The effective current is obtained from Equation 4.3,

$$I_{eff} = \frac{1}{2}(I_H + I_L) \tag{4.3}$$

in which $I_L = I_d(V_{gs} = 0.5V_{DD},\ V_{ds} = V_{DD})$ and $I_H = I_d(V_{gs} = V_{DD},\ V_{ds} = 0.5V_{DD})$. The approximate propagation delay (τ_{pd}) for the inverter output to reach $0.5V_{DD}$ is obtained from Equation 4.4,

$$\tau_{pd} = \frac{C_{out}V_{DD}}{2I_{eff}} \tag{4.4}$$

where C_{out} is the total capacitance at the CUT output.

The transient simulations are performed by applying a piecewise linear pulse with rise/fall time of 1 ps and pulse width of 110 ps. The loading effect of the CUT on the driver stage is considered by measuring the delay from the driver's input to the CUT output, as shown with t_d in Figure 4.6. The number of load inverters is varied from 1 to 5 to examine the impact of the load capacitance on the delay specifications. T_{rise} and T_{fall} are the time

periods for the CUT output to change from 10% to 90% of its maximum swing when rising and falling respectively. $\mathrm{PD_{rise}}$ and $\mathrm{PD_{fall}}$ are the propagation delays between 50% of the driver's input to the same signal level at the CUT output, respectively. The percentage of the delay change with respect to the defect-free case, calculated from Equation 4.5, used for fault modeling is

$$\%\Delta_{\mathrm{PD-rise}} = \frac{\mathrm{PD_{rise-defective}} - \mathrm{PD_{rise-deffree}}}{\mathrm{PD_{rise-deffree}}} \times 100. \quad (4.5)$$

The delay faults are categorized by comparing the percentage of delay variations to the threshold values in Table 4.4. It should be noted that the propagation delays may decrease or increase due to the GOS defect, hence the (+) sign is used for the latter case.

Table 4.4 Delay fault categories for GOS defect scenarios

Condition (%)	Fault
$-50 < \Delta_{\mathrm{PD-rise}} \leq -20$	PD_R1
$\Delta_{\mathrm{PD-rise}} \leq -50$	PD_R2
$20 < \mathrm{PD_{rise}} \leq 50$	PD_R1(+)
$50 \leq \Delta_{\mathrm{PD-rise}}$	PD_R2(+)
$-50 < \Delta_{\mathrm{PD-fall}} \leq -20$	PD_F1
$\Delta_{\mathrm{PD-fall}} \leq -50$	PD_F2
$20 < \Delta_{\mathrm{PD-fall}} \leq 50$	PD_F1(+)
$50 \leq \Delta_{\mathrm{PD-fall}}$	PD_F2(+)

The mentioned metrics are used to construct the fault models for the defective devices. The combinational circuits that are used to develop the GOS fault models include INV, 4-input AOI, and 3:1 MUX. The simulations utilize the HSPICE model parameters that are extracted for the Synopsys FinFET template based on the procedure defined in [69].

4.4 Summary

In this chapter, the 3D FinFET template provided by Synopsys is investigated including the 14nm FinFET generic process and device models. These models are used in

TCAD device simulations to develop the GOS defect models. To perform circuit-level simulations, HSPICE device model parameters of the FinFET are extracted based on the extraction procedure in the BSIM-CMG compact model. The GOS defect injection scenarios on various fin shapes that are implemented in this research are covered. Finally, the defect characterization setup used for obtaining the appropriate fault models for the circuits under test is explained.

Next chapter presents the TCAD simulation results for the defect-free and defective FinFETs in various pinhole scenarios that lead to constructing the GOS defect models.

Chapter 5: Device Level Simulation and Defect Modeling

The GOS defect is introduced as a pinhole at the gate oxide of the 3D structure of the FinFET in Sentaurus TCAD Process. The procedure is performed by placing these pinholes at various locations of different fin shapes including rectangular, trapezoidal, and triangular ones. The Sentaurus Device is appropriately modified to assess the ramification of the defect injection on the DC and transient behavior of the transistor.

The first section covers the defect-free transistor behavior with different fin geometries. In the second and third sections, behavior of the transistors with rectangular and trapezoidal fin geometries are studied in the presence of the pinholes. The impact of defect on the DC behavior of these FinFETs is investigated considering the HD transport model. This is followed by a comparison between the results obtained by incorporating HD and DD models. Due to the advantages of DD model, it is applied in the device simulations of the triangular fin with extended pinhole scenarios. Next, the dynamic behavior of the FinFET in the presence of the defects is also considered in developing accurate GOS defect model. Also, the HSPICE model parameters are extracted from TCAD FinFET template, and the transistor's DC and transient behavior are compared with TCAD results. Finally, the GOS compact model is applied to an NFinFET in the inverter circuit and its behavior is compared with Sentaurus results for verification of the models.

5.1 Defect-Free FinFET

In this work, single-fin and double-fin FinFETs are simulated and their DC characteristics and capacitances are extracted. Due to the importance of investigating

various defect sizes and locations in both device and circuit level simulations, further research on double-fin was recommended as a future work.

The 3D FinFET process is simulated by Sentaurus Process tool, generating the devices structure with aforesaid specifications in Table 4.1.

5.1.1 Single-fin Defect-Free FinFET

Sentaurus Device simulates the FinFET electrical characteristics and generates the I_dV_g curves, shown in Figure 5.1 (a), and the I_dV_d curves, shown in Figure 5.1 (b). The graphs are plotted for different fin shapes with $W_{bot} = 15$ nm and $W_{top} = 15, 10,$ and 5 nm, which are referred to as Wt15, Wt10, and Wt5, respectively.

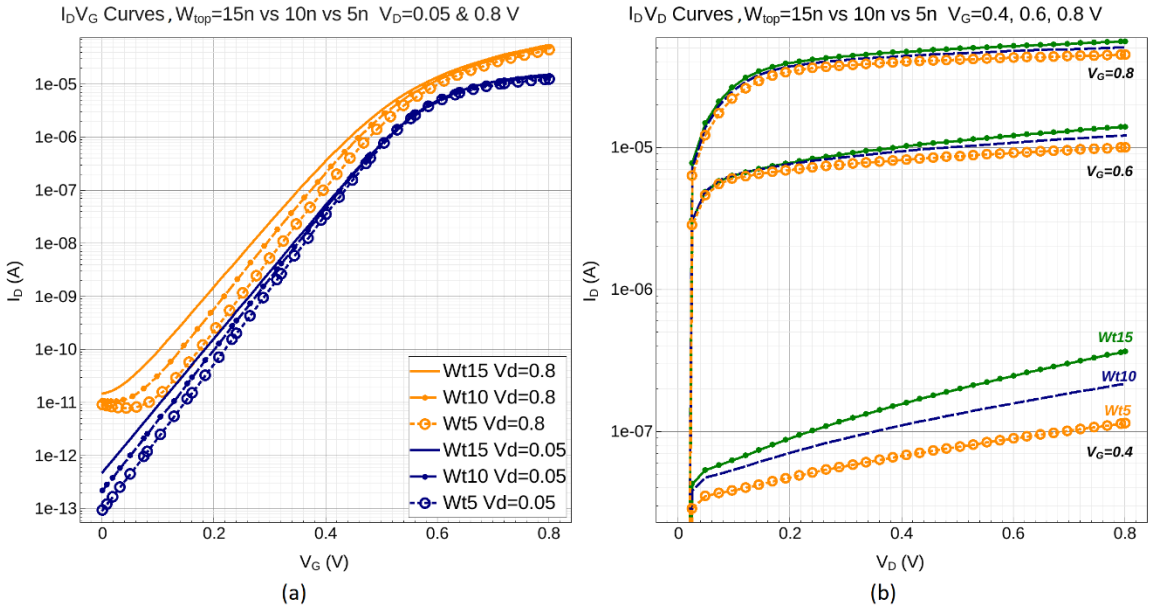

Figure 5.1 (a) I_dV_g, and (b) I_dV_d curves of bulk NFinFET with different fin top widths

Results show that as the fin shape changes from a rectangular to trapezoidal form, the saturation current, I_{on}, degrades due to the reduction of the fin effective width and thin-layer mobility degradation effect. For the leakage current, I_{off}, as the fin shape becomes

closer to the triangular form ($W_{top} = 5$ nm), the control of gate over the fin depth increases, therefore, lower leakage current and improved DIBL effect and smaller subthreshold slopes (S) are obtained with the triangular fin shape. It is observed that the triangular fin ($W_{top} = 5$ nm) has a higher I_{on}/I_{off} ratio, and hence, higher performance. The electrical characteristics of the bulk FinFETs with various W_{top} values are summarized in Table 5.1.

Table 5.1 Electrical characteristics of the bulk FinFET

$\mathbf{W_{top}}$ (nm)	15	10	5
$\mathbf{I_{on}}$ (μA)	55.83	50.90	45.11
$\mathbf{I_{off}}$ (pA)	15.00	10.94	9.34
$\mathbf{V_{th}}$ (V)	0.525	0.518	0.515
S (mV/dec)	79.62	75.18	72.69

There are different methods of threshold voltage extraction at linear and saturation bias conditions in MOSFET. Linear extrapolation methods [70], constant current method [70], and derivative of g_m/I_d [71] are examples of linear methods, while finding the maximum slope of $I_d^{0.5}V_g$ [70] is one of the saturation-bias methods. In constant-current method, the gate voltage at which the drain current reaches a typical value of $W/L * 10^{-7}$ is considered as the threshold voltage. Although this method may not be accurate due its dependence on the channel width (W) and length (L), it is widely used in the initial steps of the HSPICE model parameter extraction procedures due to its simplicity [69]. In the transconductance (g_m) linear method, the gate voltage (V_g) at which $g_m V_g$ is maximum is found. Then, a tangent line is drawn at that specific point of the $I_d V_g$ curve to extract the V_g-axis intercept as the threshold voltage by linear extrapolation [70]. The g_m/I_d method finds the gate voltage at which the gate-channel capacitance is maximum, which is insensitive to the

mobility degradation and series resistance issues [71]. Although this method is reported in the literature as the most precise one [71], it may be sensitive to the simulation step-size. The saturation-bias methods such as $I_d^{0.5}V_g$ suffers from DIBL effect. In overall, the g_m method, which is widely used in extraction algorithms in Sentaurus tools [68], is utilized through our work.

The transient behavior of the defect-free FinFET is studied through AC analysis. Sentaurus Device tool extracts transistor capacitances by small-signal analysis at a specific frequency and bias conditions in mixed-mode simulations. Figure 5.2 shows the gate capacitance at two bias conditions of $V_d = 0.0$ V and $V_d = 0.8$ V at a frequency of 100 MHz.

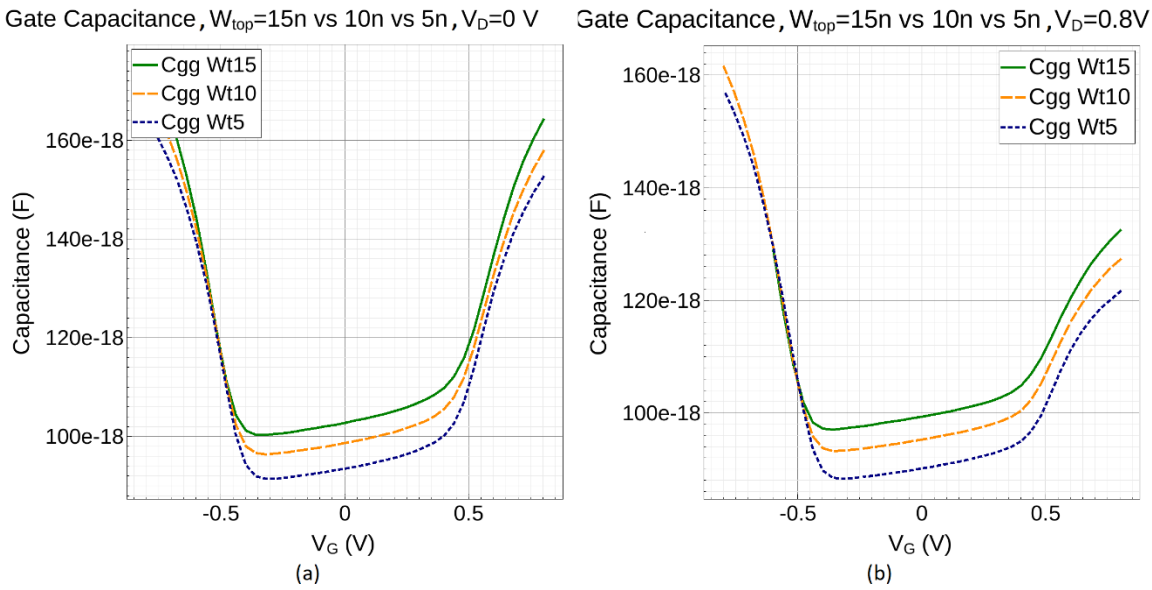

Figure 5.2 Total gate capacitance of FinFETs at (a) zero bias, (b) saturation bias

Since the triangular fin with the top width of 5 nm is later used for GOS defect modeling, all FinFET capacitances of this geometry are extracted. Figure 5.3 shows the defect-free FinFET capacitances at the mentioned bias conditions.

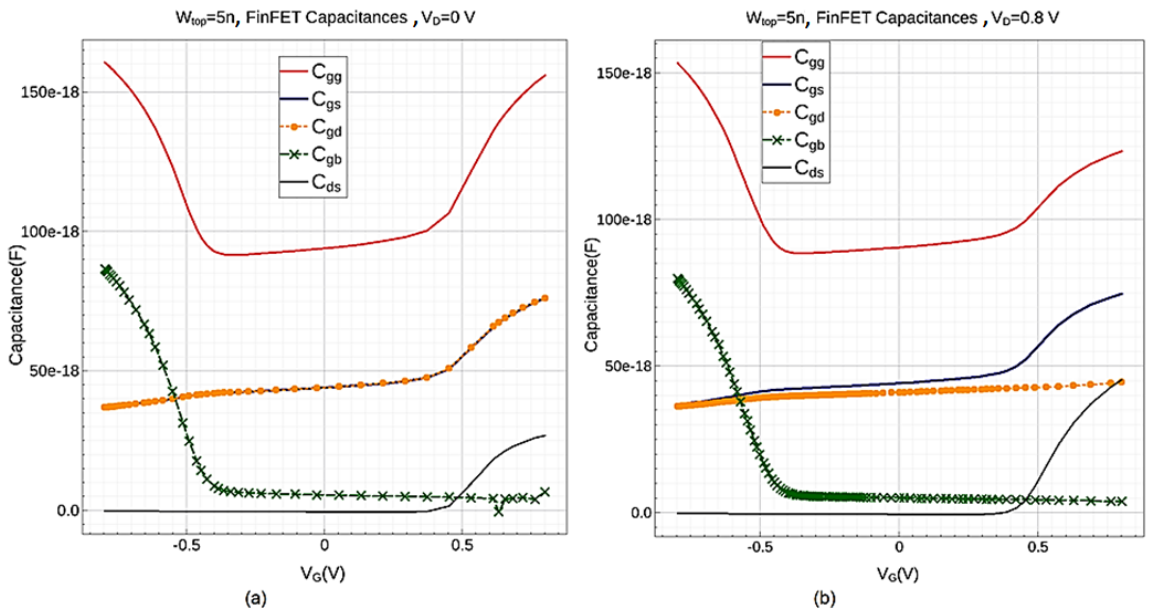

Figure 5.3 Defect-free FinFET capacitances at (a) zero bias, (b) saturation bias

In the above figure, C_{ds} is the drain-source capacitance, C_{gd}, C_{gs}, C_{gb}, C_{gg} are the gate-drain, gate-source, gate-bulk, and the total gate capacitance respectively. In the zero-bias condition, C_{gd} and C_{gs} capacitances are almost the same as in the planar MOSFET. C_{gg} is equal to the sum of C_{gs}, C_{gd}, and C_{gb} capacitances. The C_{gb} capacitance of the FinFET in the assumed bias 0.0 V and 0.8 V is negligible. Therefore, the zero-bias capacitance C_{gg} is almost twice the C_{gs} and C_{gd} at strong inversion. While at $V_d = 0.8$ V, C_{gd} becomes almost 2/3 the of zero bias C_{gd} value [75]. Figure 5.4 shows the defect-free FinFET capacitance variations with the drain voltage at $V_g = 0.8$ V.

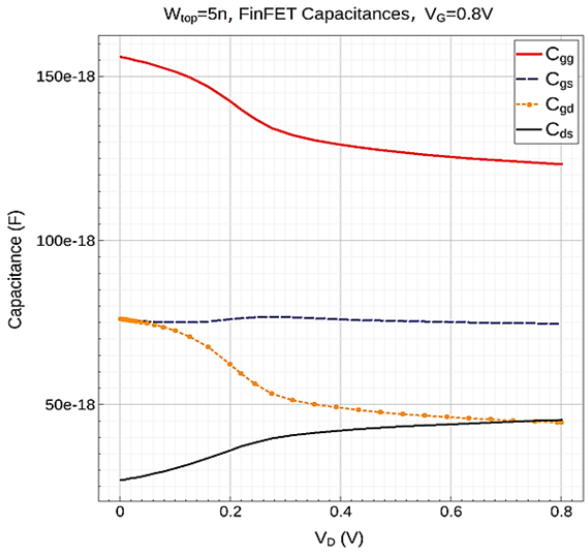

Figure 5.4 Defect-free FinFET capacitances at maximum gate voltage

5.1.2 Double-fin Defect-Free FinFET

The characteristic curves of the double-fin FinFET at different bias conditions are compared with the single-fin FinFET in Figure 5.5. It shows that the drain current almost doubles in double-fin structure due to the increase in transistor effective width.

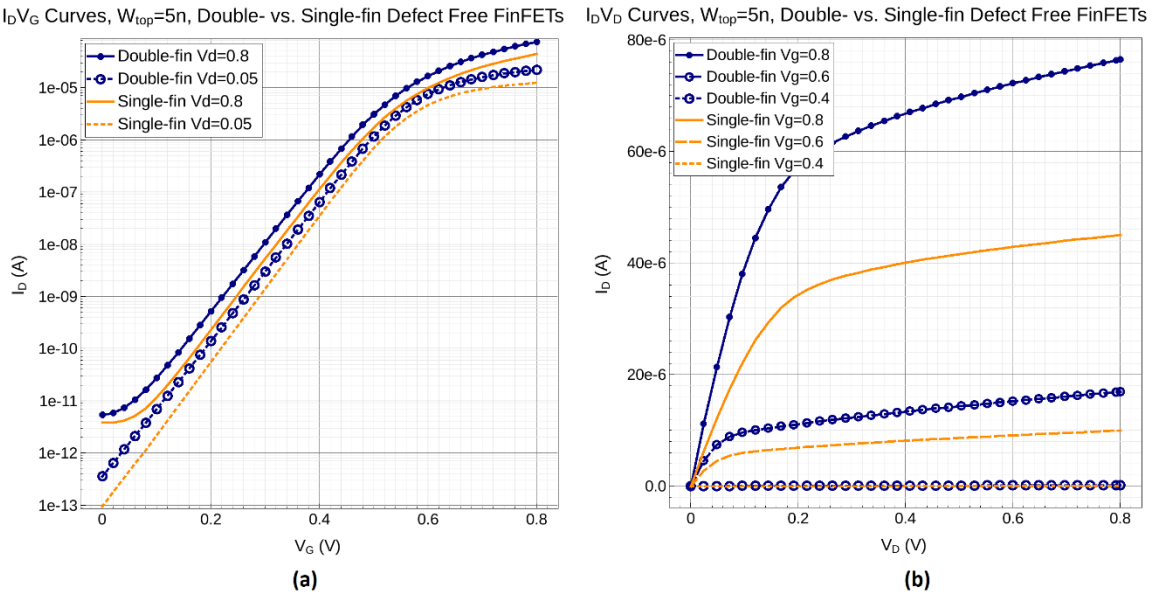

Figure 5.5 (a) I_dV_g, (b) I_dV_d characteristic curves of the double-fin FinFET

The electrical characteristics of both transistors are summarized in Table 5.2. It shows that the saturation drain current has increased due to the increase in the effective width of the transistor. However, it is not twice the saturation current of the single-fin transistor, due to the effect of parasitic resistances at S/D regions of the double-fin structure.

Table 5.2 Electrical characteristics of the single-fin and double-fin NFinFETs with W_{top} = 5 nm

# Fins	1	2
I_{on} (μA)	49.97	71.61
I_{off} (pA)	3.83	6.98
V_{th} (V)	0.49	0.50
S (mV/dec)	72.54	72.97

The capacitances of the double-fin structure are almost twice those of the single-fin device, as plotted in Figure 5.6.

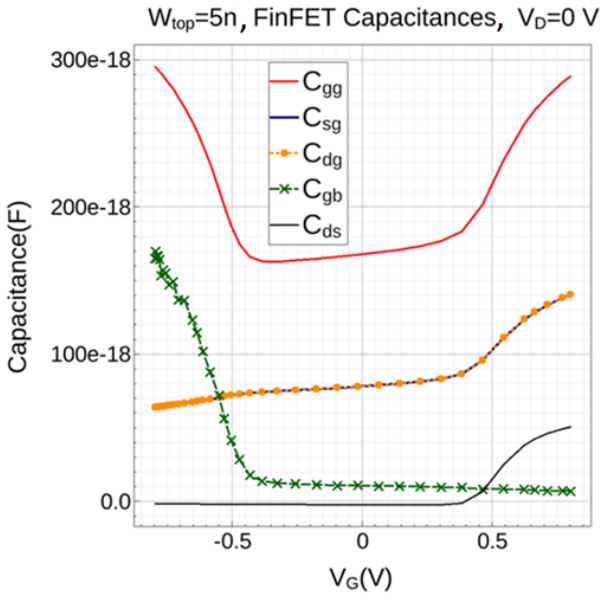

Figure 5.6 Double-fin FinFET capacitances

The defect-free FinFET results will be used as a reference for comparison with the characteristics of each defect scenario to characterize the defective transistor behavior.

In the following section, the realization methods of the GOS defect by injecting a pinhole and subsequent filling material are discussed.

5.2 GOS Defect Injection Methods

The Sentaurus Process is modified in a manner that introduces the defect as a pinhole on the gate dielectric. The pinhole is simulated by removing a small cuboid of the dielectric and filling it with an appropriate material. From the literature, the pinhole is either filled with the channel material [76], or the metal-gate material [61], [64]. Accordingly, both options are investigated by obtaining the I_dV_d characteristics of the defective FinFET with 4 nm × 4 nm pinhole, as plotted in Figure 5.7. It shows the pinhole filling material has a minimal impact, about 6% , on the transistor saturation current. Therefore, in this work it is chosen to fill the pinhole with the metal-gate material [64].

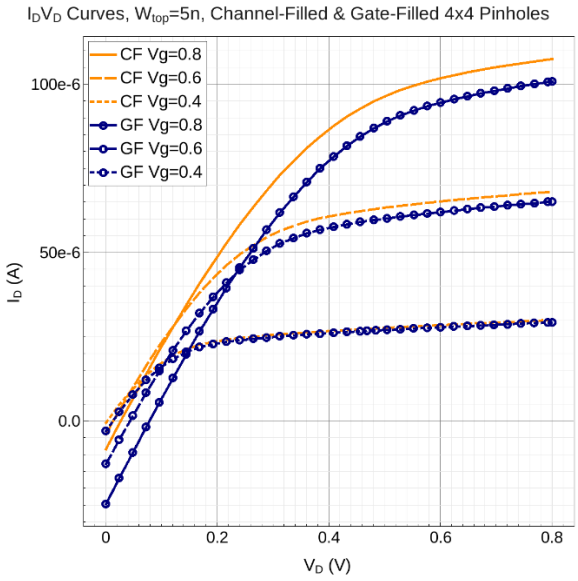

Figure 5.7 The impact of Channel-Filled (CF) vs. Gate-Filled (GF) pinholes on I_dV_d characteristics

The GOS impact on rectangular fin transistors is presented in the following section.

5.3 GOS Impact on Rectangular Fin Transistors

The DC behavior of the rectangular fin-shape FinFET is investigated at the saturation and linear bias conditions. Based on the defect injection scenarios discussed in Section 4.2.1, a pinhole is introduced at top of a rectangular fin as shown in Figure 5.8. When a small pinhole is applied on the fin top, it changes the electron density and mobility in a manner that increases the saturation current. Looking at the electron density of the defect-free transistor in Figure 5.8 (a), it confirms that the saturation current is maximum below the fin top surface. When a 2 nm × 2 nm pinhole is placed at the fin top, as observed in Figure 5.8 (b), the holes flowing directly to the channel cause the electron density to increase below the entire fin surface, specifically at the fin top. Figure 5.8 (c) and (d) display the electron density of the defective transistors with 4 nm × 4 nm and 8 nm × 8 nm pinholes, respectively. As the pinhole size increases, the larger number of holes entering the channel cause the electron density to decline, which causes a reduction in the saturation current.

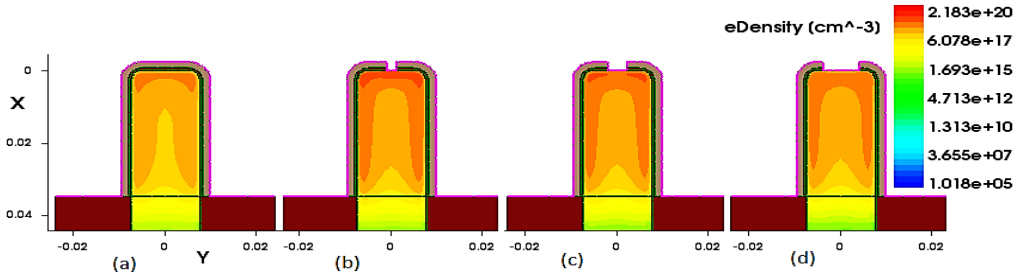

Figure 5.8 Electron densities at saturation bias at the fin center of (a) defect-free, and defective transistors with (b) 2 nm × 2 nm, (c) 4 nm × 4 nm, and (d) 8 nm × 8 nm pinholes

The drain currents of 2 nm × 2 nm and 8 nm × 8 nm pinhole cases are plotted in Figure 5.9 (a) and (b) along with the defect-free transistor results. The orange lines show the defect-free I_dV_d curves at different gate voltages, while the blue lines belong to defective transistors. It is observed that at the maximum gate voltage, the drain saturation current in

Figure 5.9 (a), with 2 nm × 2 nm pinhole, has increased by about 45% with respect to that of defect-free transistor. While in Figure 5.9 (b), with 8 nm × 8 nm pinhole, it is 14% below the defect-free drain current.

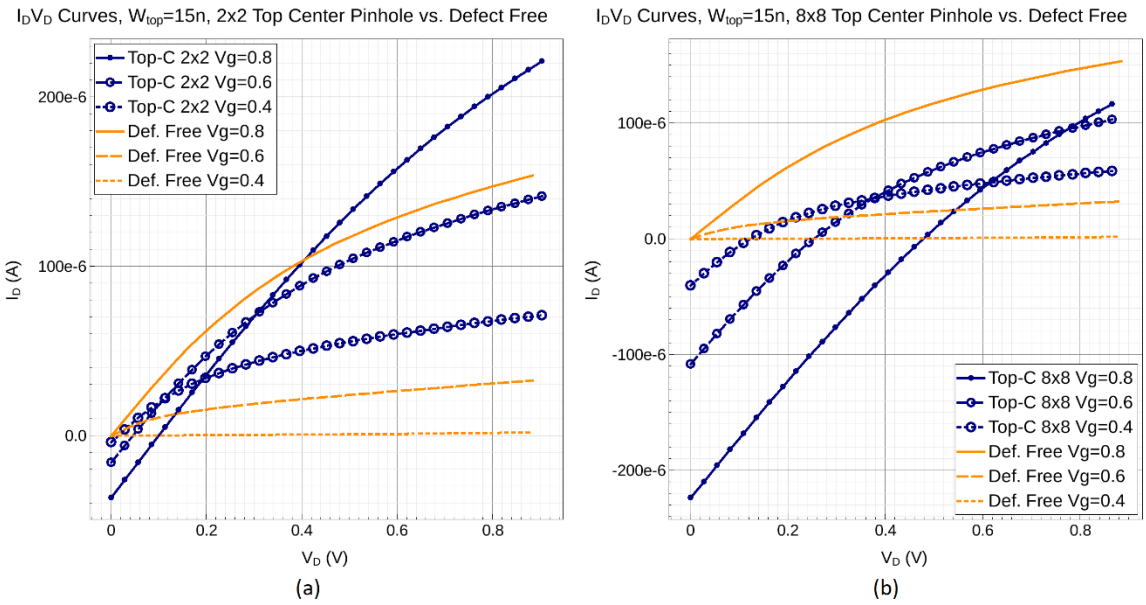

Figure 5.9 I_dV_d of rectangular FinFET with (a) 2 nm × 2 nm, (b) 8 nm × 8 nm Top-C pinholes

The leakage current in the defect-free FinFET flows in the fin depth due to the reduced control of the gate on that region, as shown in Figure 5.10. However, applying a pinhole at different locations of the fin increases the leakage current by several orders of magnitude with its maximum density occurring below the fin top surface. Figure 5.10 shows the absolute value of the electron leakage current density at the fin center of the defective transistor with 8 nm × 8 nm pinhole applied at the top and sidewall center of the rectangular fin.

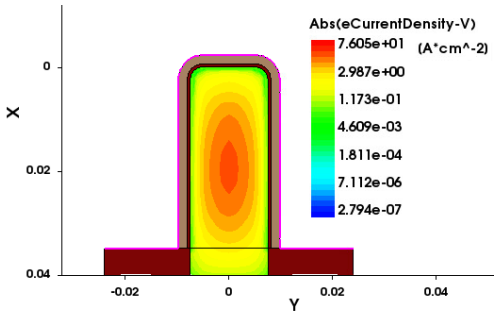

Figure 5.10 The leakage electron current density at the fin center of the defect-free FinFET

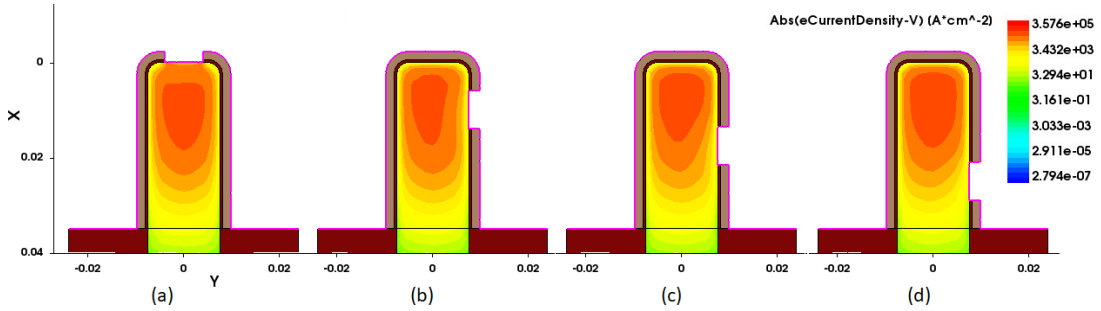

Figure 5.11 The leakage electron current density of the FinFET with an 8 nm × 8 nm pinhole on (a) Top-C, (b) Up-C, (c) Mid-C, and (d) Low-C of the fin

At low drain voltages, the large number of holes injected into the fin top causes the electrons move from drain and source to the gate which results in negative drain current. The magnitude of the negative drain current is exponentially related to the gate voltage, as observed in Figure 5.9. To study the effect of pinhole location on drain current in the linear bias, the electron current density along the fin length is plotted for a specific pinhole size in Figure 5.12. Figure 5.12 (a) shows in defect-free device, the electron current density at the drain side (specified by D) is positive in the linear bias. However, as a pinhole is injected on the fin top, Figure 5.12 (b), the direction of electron flow changes toward the gate which results in negative drain currents observed in Figure 5.9. In all other cases that

the pinhole is applied on the fin sidewall, Figure 5.12 (c) to (e), the electron current density at drain remains positive.

The characteristic curves of the defective FinFET with sidewall pinholes are plotted in Figure 5.13. They show that in the linear bias the drain current is positive, and the saturation current of defective transistor has increased about 50% with respect to the defect-free transistor at the maximum gate voltage. Comparing Figure 5.13 (a) with (b) shows that increasing the pinhole size from $2\,\text{nm} \times 2\,\text{nm}$ to $8\,\text{nm} \times 8\,\text{nm}$ does not affect the characteristics significantly. It has also been observed that all the sidewall pinhole locations produce the same characteristics; therefore, just the Mid-C pinhole results are plotted in Figure 5.13.

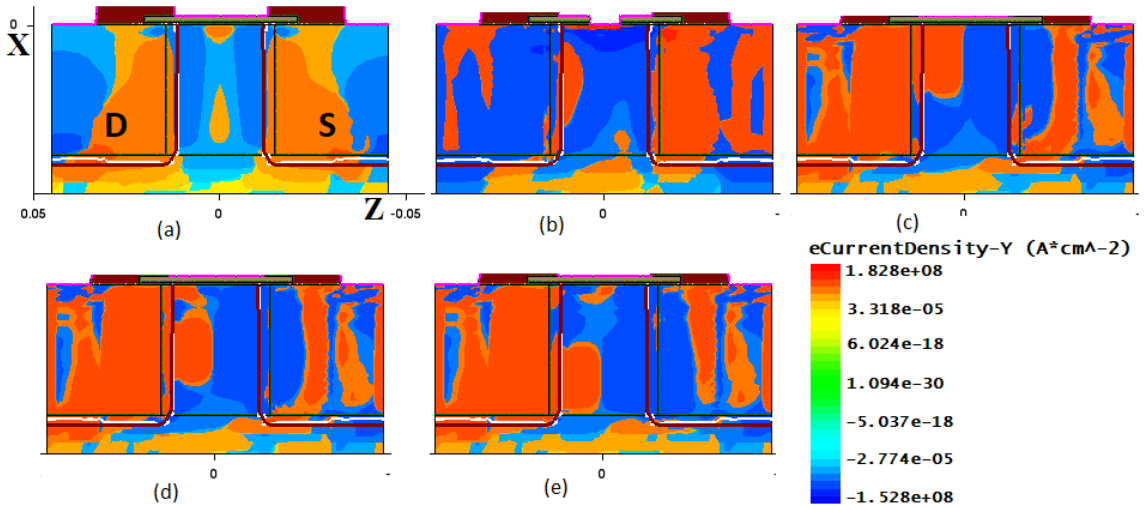

Figure 5.12 Electron current densities along the fin length in (a) defect-free FinFET, and FinFETs with 8 nm × 8 nm pinhole on the, (b) Top-C, (c) Up-C, (d) Mid-C, and (e) Low-C of the fin in linear bias.

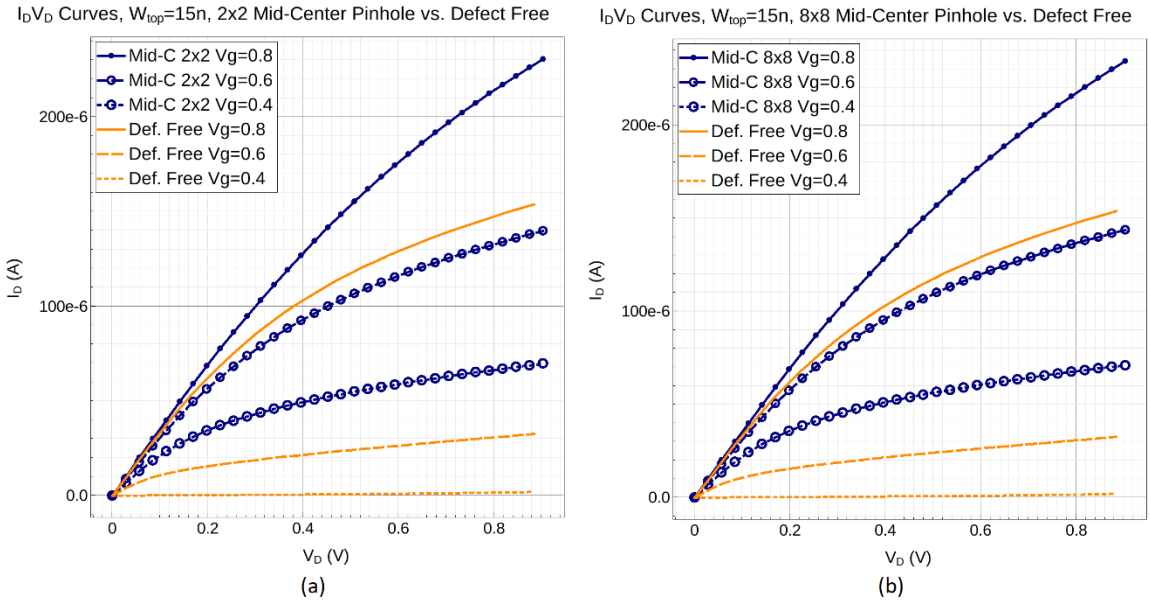

Figure 5.13 I_dV_d curves of FinFETs with (a) 2 nm × 2 nm, (b) 8 nm × 8 nm Mid-C pinholes

The leakage and saturation currents for different pinhole scenarios are plotted in Figure 5.14. Every group indicates a specific pinhole size in nanometer scale, and each bar in a group shows the location of the defect on the rectangular fin. The top pinhole has the highest leakage current at the maximum pinhole size. As for the sidewall pinholes, the leakage increases with the pinhole location moving from the upper sidewall to the lower side since the gate control over the channel is reduced. The defect-free transistor has a leakage current of 72.31 pA, that is almost zero on the scale of the plot. The saturation currents are compared with the defect-free case in Figure 5.14 (b). The solid line at 157.9 uA indicates the saturation current of the defect-free rectangular fin. As previously mentioned, the sidewall pinhole size and location do not affect the saturation current considerably, while for the top pinhole, the current is reduced significantly as the pinhole size increases.

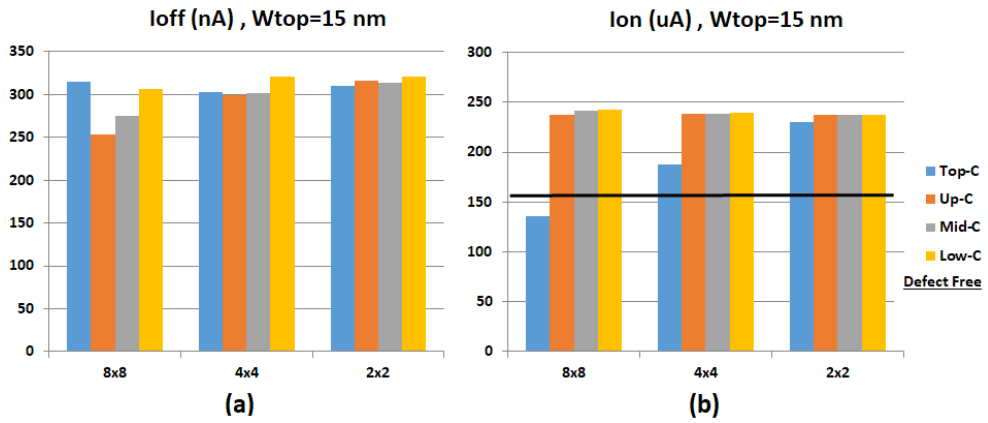

Figure 5.14 (a) Leakage, (b) Saturation currents of the rectangular fin FinFET with different scenarios of sidewall pinholes comparing to the top pinhole

To investigate the impact of pinhole along the channel length, the defect is applied close to the source and close to the drain at the top of the fin, as shown in Figure 4.4 (f) and (g). It is observed that in these cases, the saturation current does not change considerably with the location of the defect especially at small pinhole sizes. However, the leakage current increases dramatically as the pinhole moves closer to the source, due to its strong impact on electron current density in that area. Figure 5.15 shows the results for 4 nm × 4 nm, and 2 nm × 2 nm pinholes along the fin length. It shows that for 4 nm × 4 nm pinhole, as the pinhole moves closer to the source (Top-S), the leakage current almost doubles with respect to the top center case. It should be noted that for 8 nm × 8 nm pinhole, which is not plotted in Figure 5.15, the leakage current increases to 7.25 uA.

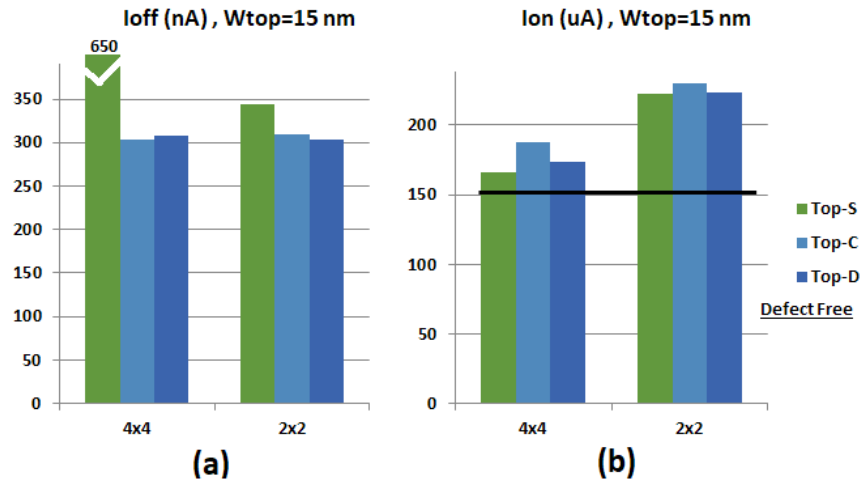

Figure 5.15 (a) Leakage, (b) Saturation currents of rectangular fin FinFET with top pinholes close to the source (Top-S), at the fin center (Top-C), and close to the drain (Top-D)

5.4 GOS Impact on Trapezoidal Fins

GOS defects are introduced in the trapezoidal fins in the same manner as those in the rectangular fins. The fin thickness at the top is selected to be 10 nm. From simulations, it has been noticed that the trapezoidal fin shaped transistor with top pinhole manifests similar behavior as that of a rectangular geometry. However, in the case of sidewall pinholes due to different surface crystal orientation of the sidewalls, the transistor behavior differs from what has been presented in Figure 5.13. Figure 5.16 shows the effect of the size and location of the side pinhole on I_dV_d curves at maximum gate voltage. As the pinhole moves towards the fin top, the saturation current decreases since it flows mainly below the fin top surface, and as the defect occurs closer to that region the injected holes weaken the channel.

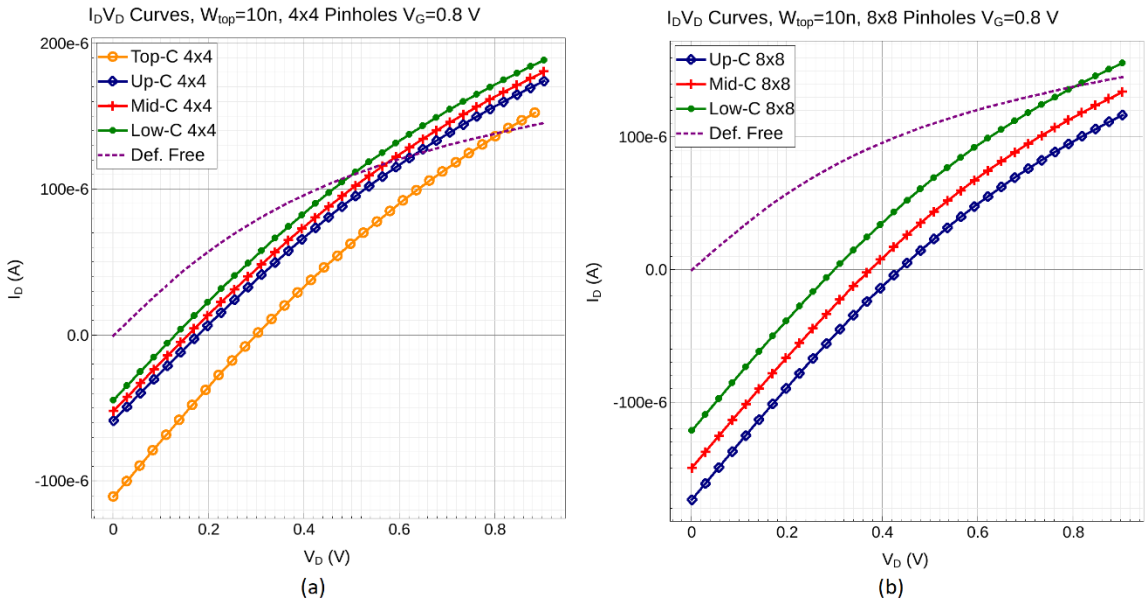

Figure 5.16 I_dV_d curves of FinFETs with (a) 4 nm × 4 nm, (b) 8 nm × 8 nm sidewall pinholes

The leakage and saturation currents of different scenarios are plotted in Figure 5.17. It should be noted that 8 nm × 8 nm pinhole on the fin top, with the top width of 10 nm, is not a practical scenario, hence ignored. The results show that lower side pinhole has larger leakage current than other sidewall locations due to weak control of the gate over the channel at the fin bottom. The 2 nm × 2 nm pinhole saturation current is almost the same at all locations, same as in the rectangular geometry. While, for other sizes the saturation current decreases as the pinhole becomes larger or moves toward the fin top since electron and hole mobilities weaken with these changes. The defect-free leakage current is 19.14 pA, almost zero at the plot scale, while the saturation current is 144.6 uA, shown by the solid line on Figure 5.17 (b).

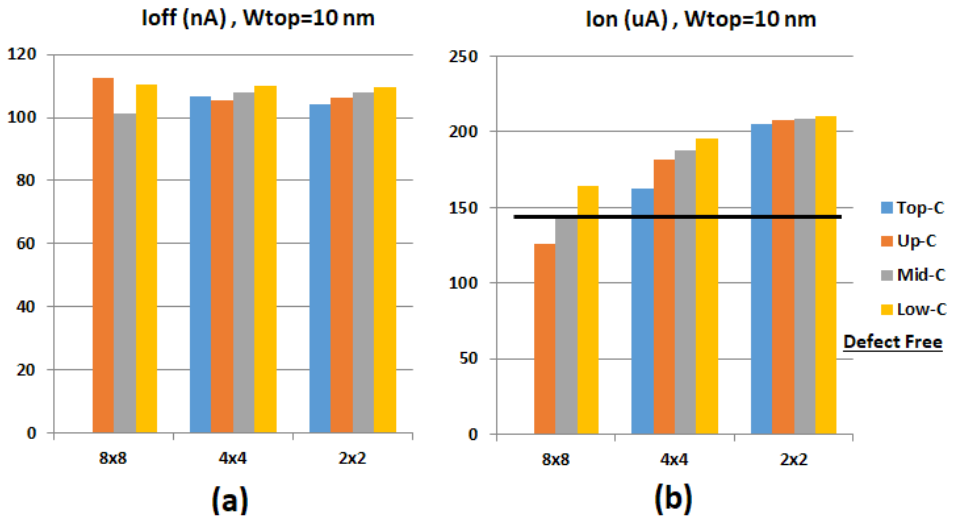

Figure 5.17 15 (a) Leakage, (b) Saturation currents of trapezoidal fin FinFET showing the impact of pinhole size and location

Moving the pinhole along the channel length of the trapezoidal fin affects the saturation and leakage currents in a similar manner as in the rectangular fin, though with smaller magnitude. The simulation results plotted in Figure 5.18 show that pinholes close to source (Top-S) increase the leakage current significantly.

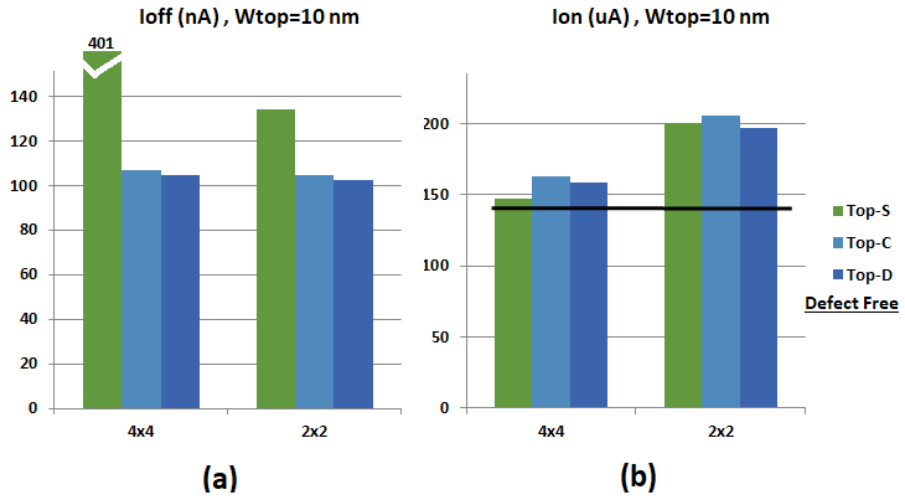

Figure 5.18 Leakage, (b) Saturation currents of trapezoidal fin FinFET with pinholes close to the source (Top-S), at the fin center (Top-C), and close to the drain (Top-D)

5.5 Hydrodynamic vs. Drift-Diffusion Transport Models

The defect-free FinFET template provided by Synopsys is based on the drift-diffusion model with adjusted velocity saturation. To compare it with hydrodynamic model, the transistor characteristics are simulated with both carrier transport models at saturation bias. The I_dV_g characteristics of the defect-free triangular FinFET are plotted in logarithmic scale in Figure 5.19 (a).

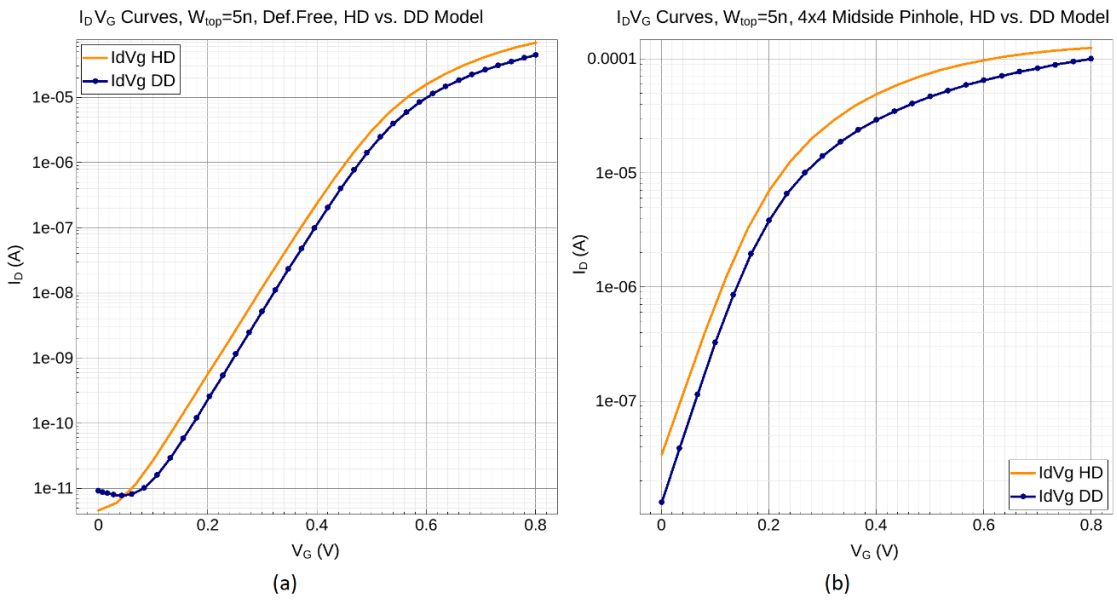

Figure 5.19 I_dV_g curves of the (a) defect-free, (b) defective FinFET with 4 nm × 4 nm pinhole with different carrier transport models

A specific defect scenario, for example 4 nm × 4 nm pinhole on the middle sidewall of the triangular fin, is also simulated with both models. The simulations show both models result in similar FinFET behavior, though with different magnitude of currents. The platform specification used for these simulations are listed in Table 5.3.

Table 5.3 Simulation Platform Specification

CPU	Intel® Xeon® E5-2687W @ 3.1 GHz
Number of Cores	32
Memory	252.4 Gigabytes
Operating System	Linux CentOS release 6.10

The Process simulation CPU time ranges from 45 minutes to 2 hours, which mostly consists of the structure and mesh generation time and the dopant diffusion simulation time. The Device simulation time varies in a wider range due to the complexity of the physical models, the mesh spacing, and the convergence behavior of the system of equations to be solved at each mesh point. The Device simulation times for the above-mentioned scenarios are provided in Table 5.4 for further comparison.

Table 5.4 Total CPU time of Device simulations with different carrier transport models

	Total CPU time (hh: mm)	
Model	**Defect-Free**	**Defective**
HD	49: 35	64: 32
DD	6: 29	9: 12

Although HD model considers the impact of lattice and temperature variations in the Device simulations, which might be more accurate for GOS-impacted transistors, it suffers from degraded convergence and longer simulation times. Therefore, the DD model is used in all the subsequent device simulations and GOS defect model development.

5.6 GOS Impact on Triangular Fins

In triangular FinFET, the pinhole injection scenarios are extended to the fin sidewall closer to source and drain areas in addition to the previous locations, as shown in Figure 4.5. Moreover, the transient behavior of FinFET in the presence of GOS defect is investigated in this section.

First, a 2 nm × 2 nm pinhole is introduced on the fin top. Figure 5.20 shows the absolute value of the electron leakage current density at the fin center. The leakage current in the defect-free FinFET flows in the fin depth due to the reduced control of the gate on that region, as shown in Figure 5.20 (a). However, applying a 2 nm × 2 nm pinhole on the fin

top increases the leakage by three orders of magnitude with its maximum density occurring below the fin top surface, as observed in Figure 5.20 (b). At saturation bias, the depletion regions of the top and sidewall gates merge into each other and result in volume inversion effect [77], as observed in Figure 5.20 (c). When a 2 nm × 2 nm pinhole is placed at the fin top, the holes are flowing directly to the channel, changing the electron density and mobility in a manner that increases the saturation current. Figure 5.20 (d) shows 3 times increase in the electron current density in the presence of the defect.

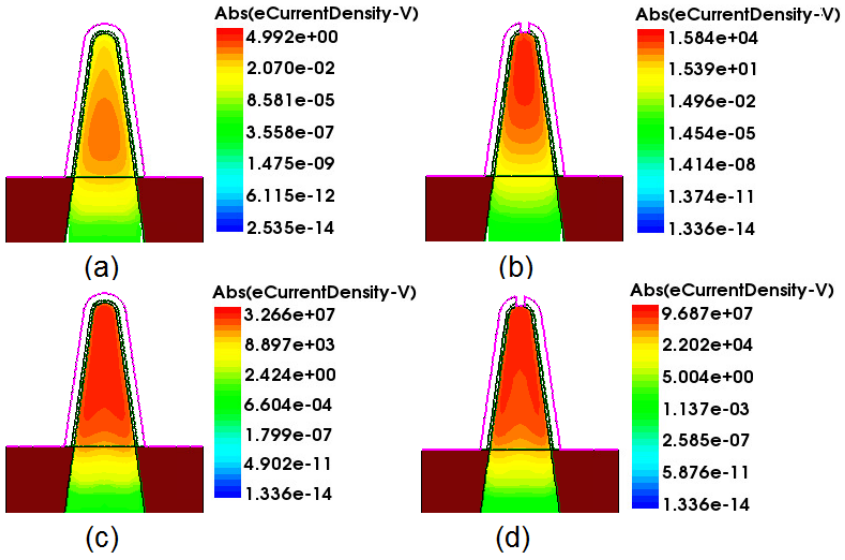

Figure 5.20 Leakage current densities at the fin center of (a) defect-free, and (b) defective FinFETs, saturation current densities at the fin center of (c) defect-free, and (d) defective transistors

The impact of top pinhole on DC characteristics is same as in the previous geometries but with different current magnitudes. Figure 5.21 (a) shows the I_dV_g curves at linear and saturation bias in logarithmic scale. It is observed that at the maximum gate voltage, the drain saturation current is doubled with respect to that of the defect-free transistor. This is because, recalling Figure 5.8, the pinhole causes an increase in the electron density of the channel, especially below the fin top and next to the sidewalls. The transistor leakage current has increased more than three orders of magnitude in the presence of the defect.

The I_dV_d curves at gate voltages of 0.4 V, 0.6 V, and 0.8 V are plotted in Figure 5.21 (b). At low drain voltages, the large number of holes injected into the fin top causes the electrons to move from drain and source to the gate, which results in negative drain current.

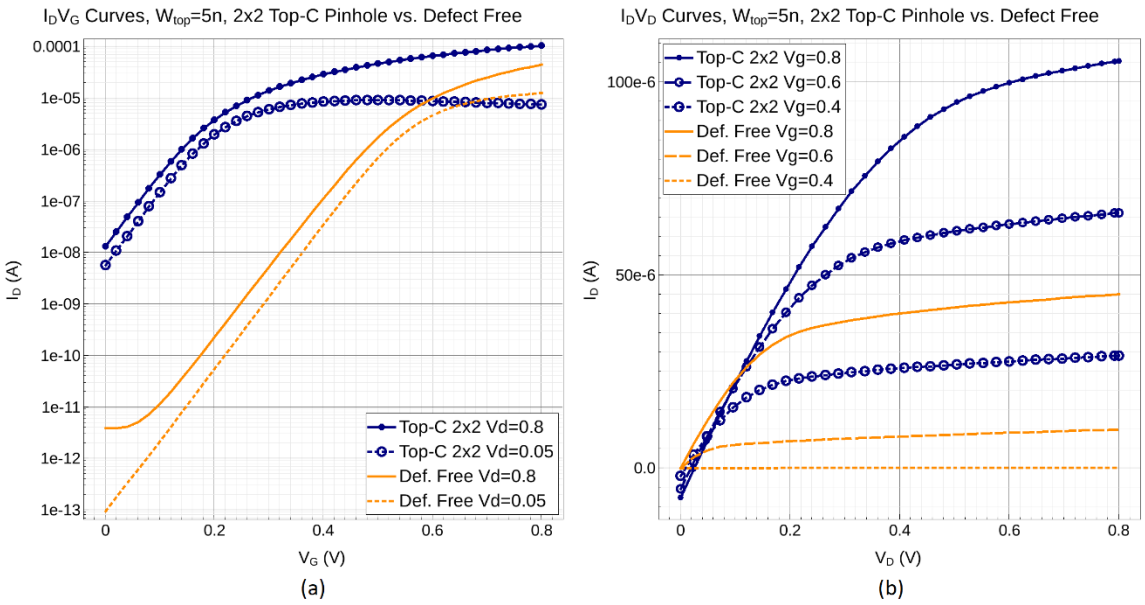

Figure 5.21 (a) I_dV_g, (b) I_dV_d curves of the FinFET with 2 nm × 2 nm Top-C pinhole vs. defect-free

In the next step, the impact of pinhole on the fin sidewall is investigated considering the defect scenarios in Figure 4.5 with defect size changing from 2 nm × 2 nm to 8 nm × 8 nm. First, a pinhole on the midside center of the sidewall is simulated, Mid-C in Figure 4.5 (e). The DC characteristic curves plotted in Figure 5.22 show that the pinhole doubles the saturation current, though for larger pinhole sizes the amount of increase is slightly lower. Comparing Figure 5.22 (a) with (b) shows that increasing the pinhole size significantly increases the magnitude of negative drain current at low drain bias, since large number of holes are injected to the channel causing more electrons to move from drain and source to the gate. It has also been observed that the magnitude of the negative drain current is exponentially related to the gate voltage.

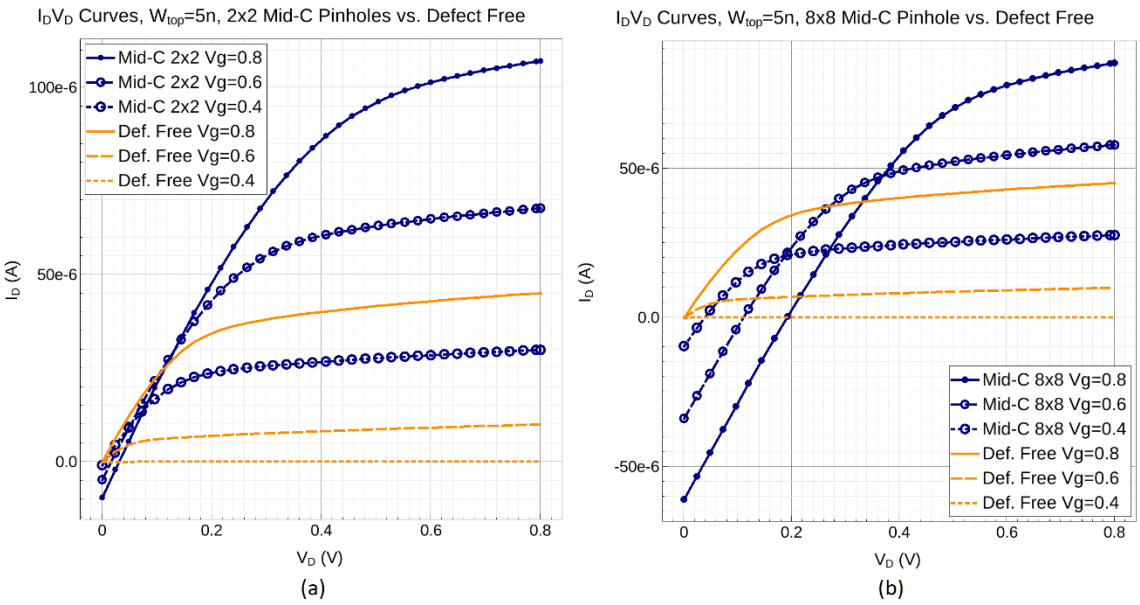

Figure 5.22 I_dV_d curves of FinFETs with (a) 2 nm × 2 nm, (b) 8 nm × 8 nm Mid-C pinholes

Figure 5.23 shows the effect of the location of the sidewall pinhole on I_dV_d curves at maximum gate voltage. At high drain voltage, as the pinhole moves toward the upper sidewall, Figure 5.23 (a), the saturation current decreases as it flows mainly below the fin top surface. The defect closer to that region weakens the channel due to the injected holes. In Figure 5.23 (b), the drain current variations when the pinhole moves closer to source and closer to drain are plotted for a pinhole on the middle sidewall. It is observed that at low drain bias, moving the pinhole close to source, drives more electrons from the drain and results in larger negative drain current.

The impacts of pinhole size and location on the leakage and saturation currents are obtained and plotted in Figure 5.24. The defect-free transistor has a leakage current of 3.83 pA, which is almost zero on the scale of the plot. The 8 nm × 8 nm upside pinhole has the highest leakage current since the maximum current density flows below the fin top surface.

It should be noted that 2 nm × 2 nm pinhole is the only practical defect scenario for the fin top shown in blue bars.

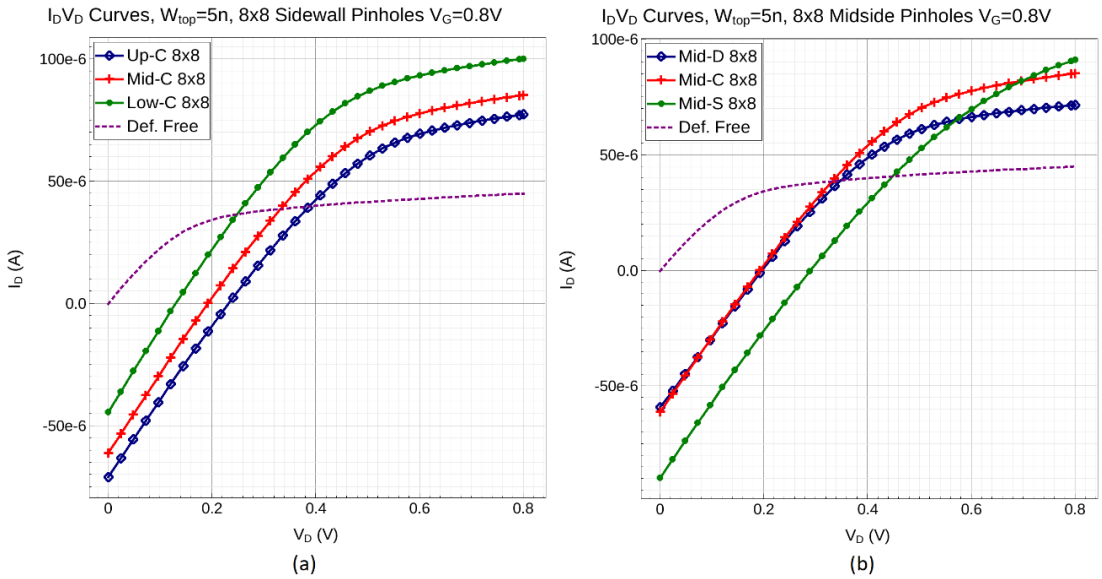

Figure 5.23 I_dV_d curves of FinFETs with (a) 8 nm × 8 nm sidewall center pinhole (a) moving along the fin height, (b) moving along the fin length

The saturation currents are compared with the defect-free case in Figure 5.24 (b). The 2 nm × 2 nm pinhole saturation current is almost the same at all locations. While, for other sizes the saturation current decreases as the pinhole becomes larger or moves toward the fin top, since electron and hole mobilities weaken with these changes. The line at 45 μA indicates the saturation current of the defect-free transistor.

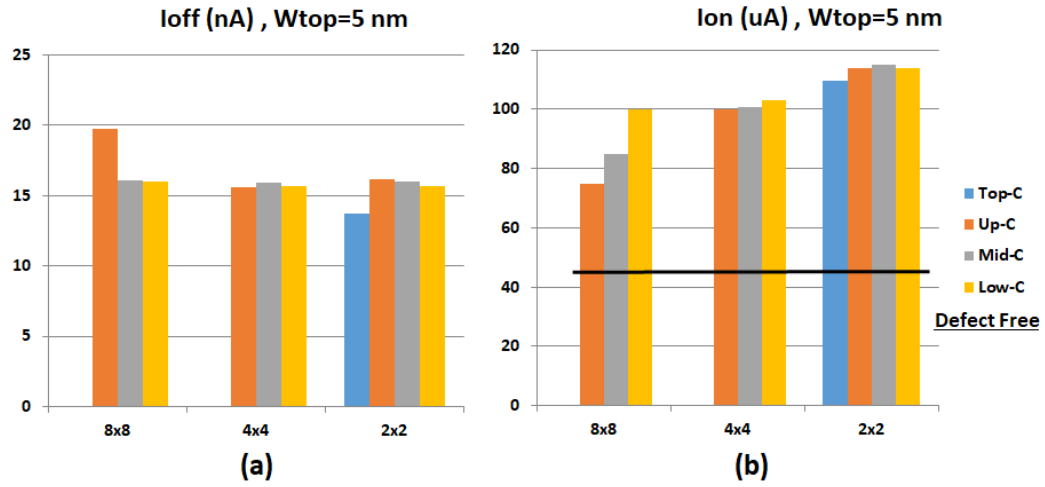

Figure 5.24 (a) Leakage, (b) Saturation currents of triangular fin with pinholes on the center of fin top and sidewall locations

Figure 5.25 shows the impact of the defect on the leakage and saturation currents when an 8 nm × 8 nm pinhole is applied close to the source and close to the drain on the fin sidewall. The groups indicate the location of the pinhole along the channel length, while the bars in each group show the vertical position of the pinhole on the fin sidewall. As previously observed in Figure 5.23 (b), close to source defects result in higher leakage currents among which the upside close to source has the highest leakage current, Figure 5.25 (a). In all GOS-impacted transistors, the leakage current flows below the fin top surface. Therefore, an 8 nm × 8 nm pinhole on the upper sidewall closer to source enhances the electron current density that flows to the drain. At saturation bias, the drain currents for the center and close to source cases are almost similar in values, Figure 5.25 (b). However, under the strong electric field near the drain region, a pinhole close to drain results in lower drain current than other scenarios.

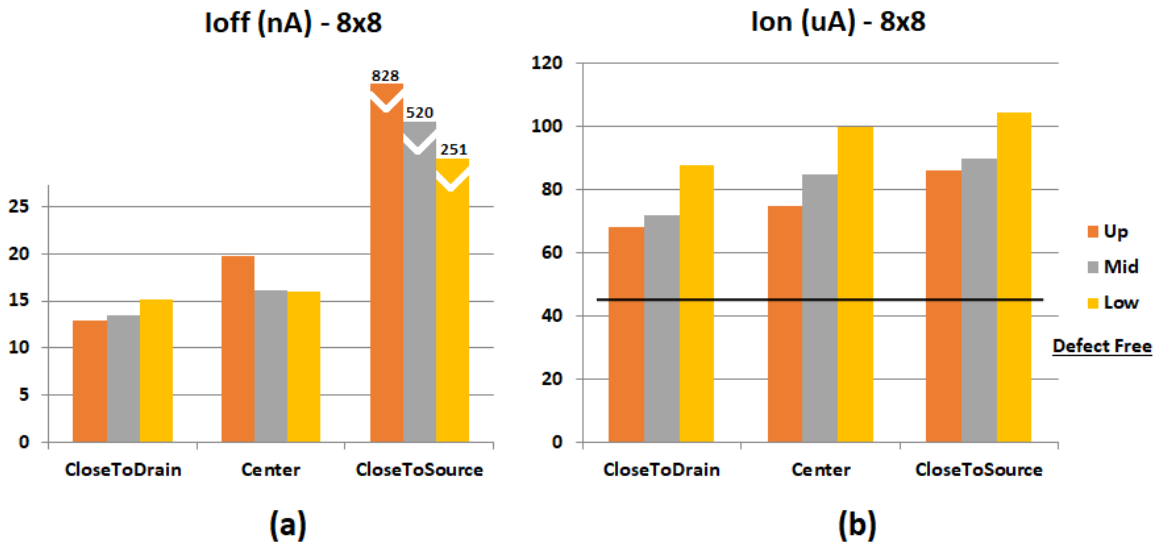

Figure 5.25 (a) Leakage, (b) Saturation currents of triangular fin with different defect scenarios close to source/drain

Other important characteristic curves that show the impact of the defect on the device behavior are the I_gV_g curves, which demonstrate the gate current variations with the gate voltage. To study the impact of pinhole location along the fin length, an $8\,\text{nm} \times 8\,\text{nm}$ pinhole is placed at the middle sidewall of the fin, closer to the source and closer to the drain. These scenarios are simulated at two bias conditions.

Figure 5.26 shows the I_dV_g and I_gV_g curves at linear bias. Although the I_dV_g curves at logarithmic scale such as Figure 5.21 (b) show the leakage, the curves are plotted in linear scale to clearly show the negative drain current, Figure 5.26 (a). A pinhole closer to the source results in the highest current density of electrons flowing from the drain, creating a larger negative drain current similar to Figure 5.23 (b). The gate leakage current exponentially increases with the gate voltage, as shown in Figure 5.26 (b). At low drain bias, the gate leakage current in close to source pinhole scenario is almost equal that in close to drain pinhole, since the source and drain voltages are almost equal.

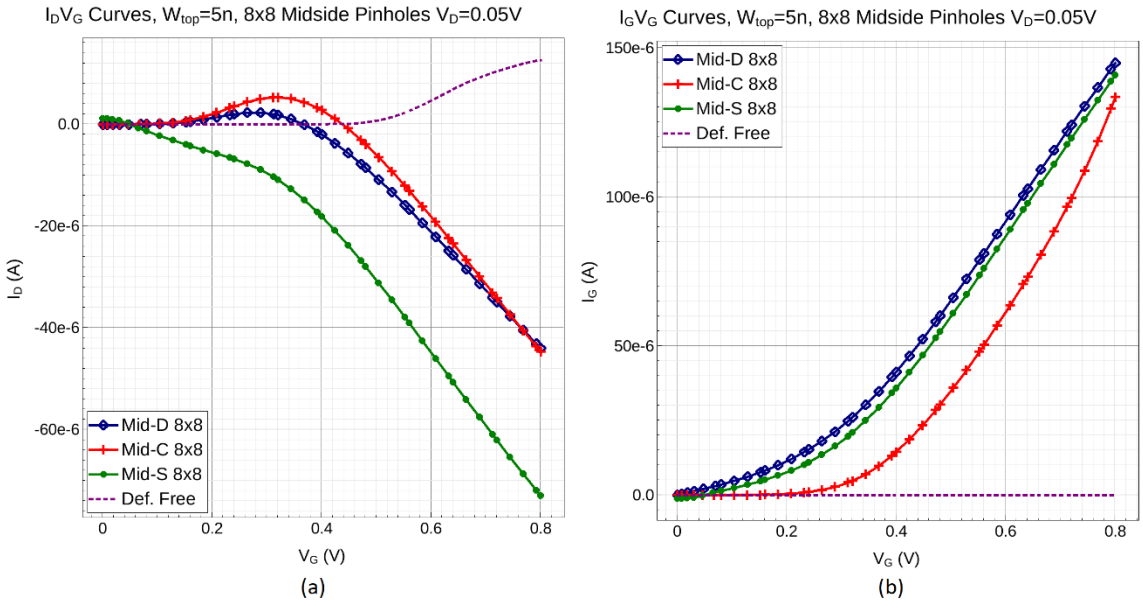

Figure 5.26 (a) I_dV_g curves, (b) I_gV_g curves of FinFETs with 8 nm × 8 nm middle sidewall pinhole, moving close to drain and source at linear bias

The I_dV_g and I_gV_g characteristic curves are also plotted in Figure 5.27 for saturation bias. Figure 5.27 (a) shows that the drain current for close to drain pinhole is lower than the other scenarios. This is because the large number of holes injected into the region under high electric field weakens the channel. The gate leakage current plotted in Figure 5.27 (b) confirms that close to drain pinhole results in the largest flow of holes to the channel.

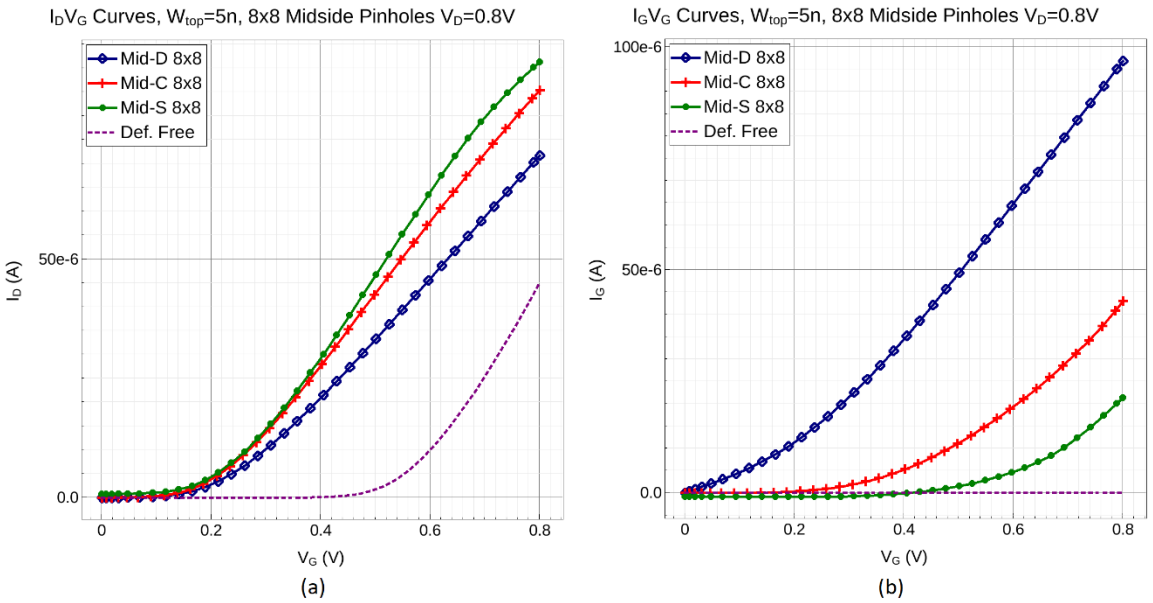

Figure 5.27 (a) I_dV_g curves, (b) I_gV_g curves of FinFETs with 8 nm × 8 nm middle sidewall pinhole, moving close to drain and source at saturation bias

The I_dV_g and I_gV_g characteristic curves are used in developing GOS defect models in Section 5.8.

The impact of 2 nm × 2 nm pinhole, introduced to the fin top and sidewall, on the FinFET gate capacitances at saturation bias are shown in Figure 5.28. There is a maximum of 4 fF increase in C_{gd} due to the defect. The pinhole changes the electron density of the channel and volume inversion occurs at lower gate-source voltages, as shown in the C_{gs} graph. This effect results in the reduction of threshold voltage, which is also observed in the I_dV_g graph, for example in Figure 5.27 (a), where the defective transistors turn on at lower gate voltages. The location of 2 nm × 2 nm pinhole does not affect the gate capacitances.

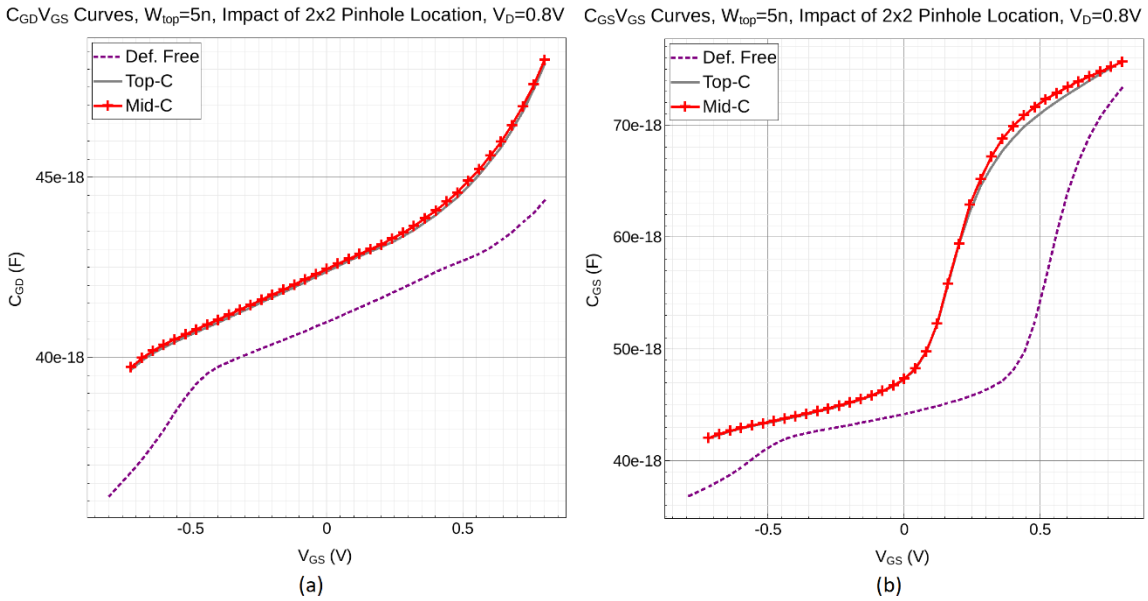

Figure 5.28 (a) The impact of 2 nm × 2 nm pinhole on (a) C_{gd}, (b) C_{gs} capacitances

The gate capacitances of defective transistors with 4 nm × 4 nm pinholes on various locations of the fin sidewall are plotted in Figure 5.29. As the pinhole size increases, the excessive holes injected to the channel reduce the gate-source capacitance at maximum gate voltage compared to the smaller pinhole cases. While the gate-drain capacitance does not have considerable change compared to the that in 2 nm × 2 nm pinhole scenarios since the charge variations does not affect the pinch-off region close to drain.

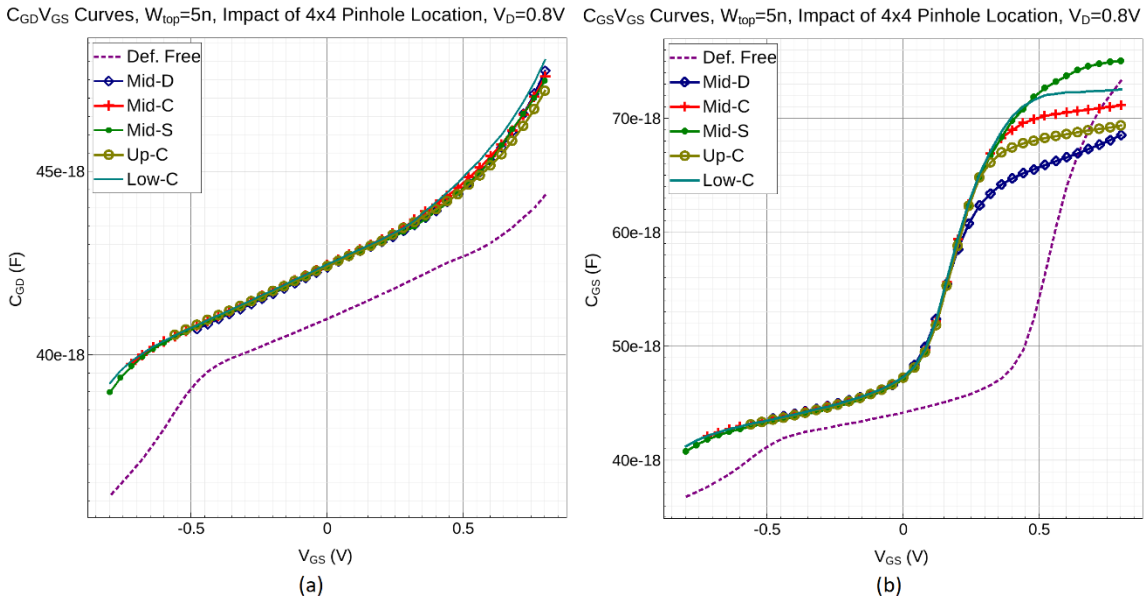

Figure 5.29 The impact of 4 nm × 4 nm pinhole on (a) C_{gd}, (b) C_{gs} capacitances

The close to drain pinhole cause the largest degradation in C_{gs} , in the above figure, at maximum gate voltage compared to other sidewall locations. This reduction is exacerbated in 8 nm × 8 nm pinholes as shown in Figure 5.30 (b). This may be due to the excessive gate leakage current in Mid-D scenario that is depicted in Figure 5.27 (b).

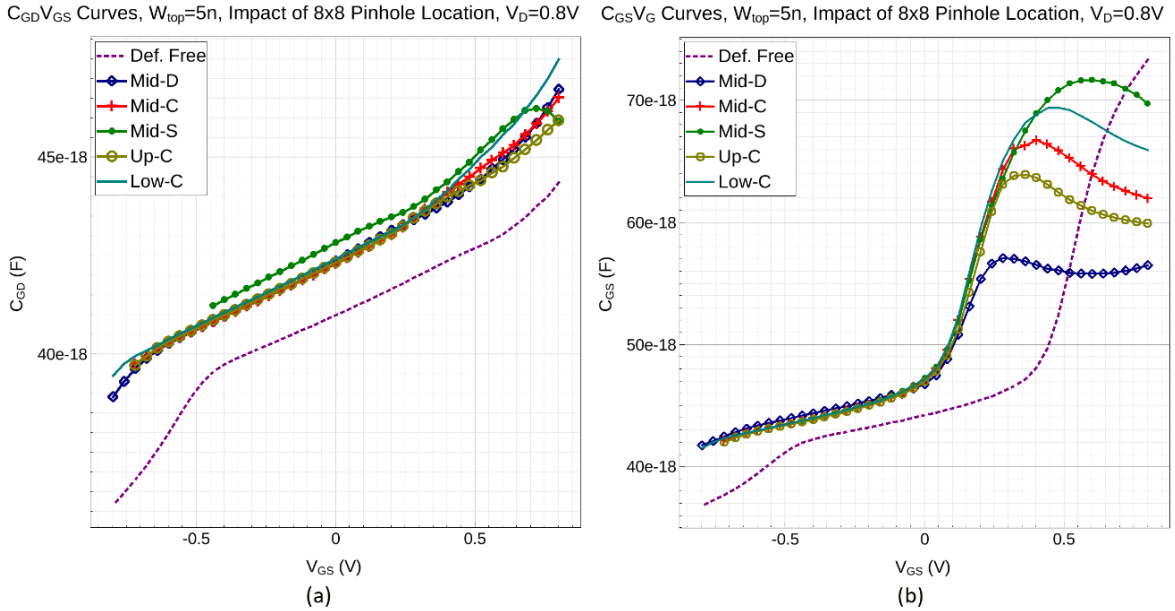

Figure 5.30 The impact of 8 nm × 8 nm pinhole on (a) C_{gd}, (b) C_{gs} capacitances

5.7 GOS Impact on Double-fin FinFET

Before constructing the GOS defect model for single-fin FinFETs, it is important to investigate the impact of defect on a double-fin structure. The 8 nm × 8 nm pinhole is simulated on the inner sidewall of the double-fin transistor shown in Figure 4.3. The I_dV_g and I_dV_d curves of the single-fin and double-fin defective transistors are compared at different bias conditions in Figure 5.31.

The saturation current of the double-fin transistor in presence of GOS defect is almost twice the single-fin FinFET. Considering the defect-free characteristics in Figure 5.5, the impact of defect on the saturation currents are summarized in Table 5.5.

Table 5.5 Saturation current in double-fin vs single-fin FinFET

	I_{on} (μA)	
	Single-fin	**Double-fin**
Def. Free	45.11	76.57
Mid-C8x8	85.39	161.79

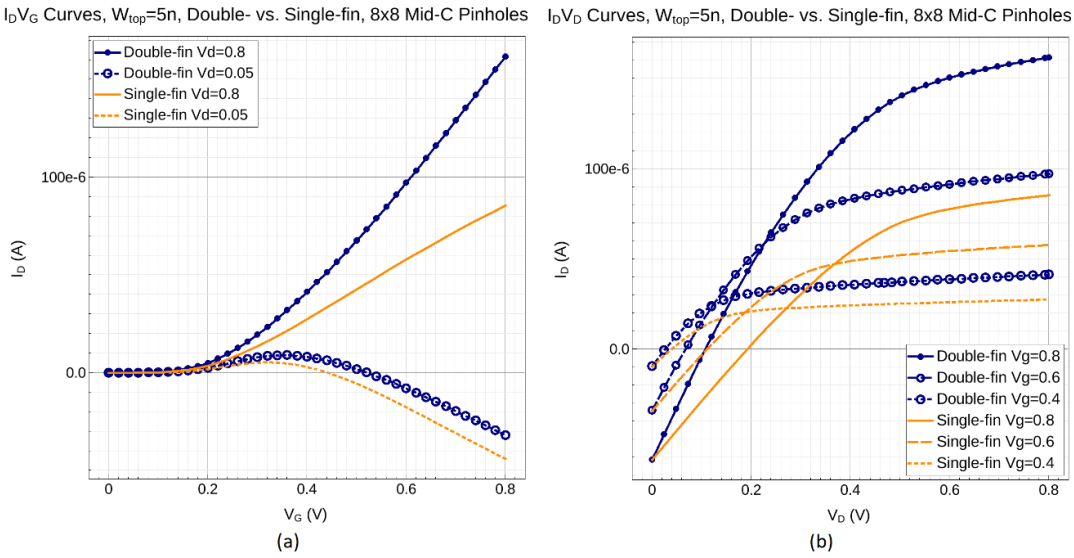

Figure 5.31 (a) I_dV_g, (b) I_dV_d curves of double-fin FinFET with 8 nm × 8 nm sidewall pinhole

The results show that the GOS impact on the saturation current is not weakened as the number of fins is doubled and the proposed defect modeling methodology is extendable to

double-fin transistors. However, in multi-fin FinFETs with large number of fins, the percentage of current variation due to the defect will likely decrease. Further investigation of the GOS impact on double-fin and multi-fin transistors is considered as future work.

5.8 GOS Defect Model

GOS defect is one of the defects that is complex to analyze and model. Therefore, constructing an accurate compact model is crucial in fault modeling and testing applications. Figure 5.32 shows a flow diagram of extracting GOS defect models for each defect size and location by performing DC simulations in TCAD environment. A similar procedure is used to extract the model capacitances by performing AC simulations.

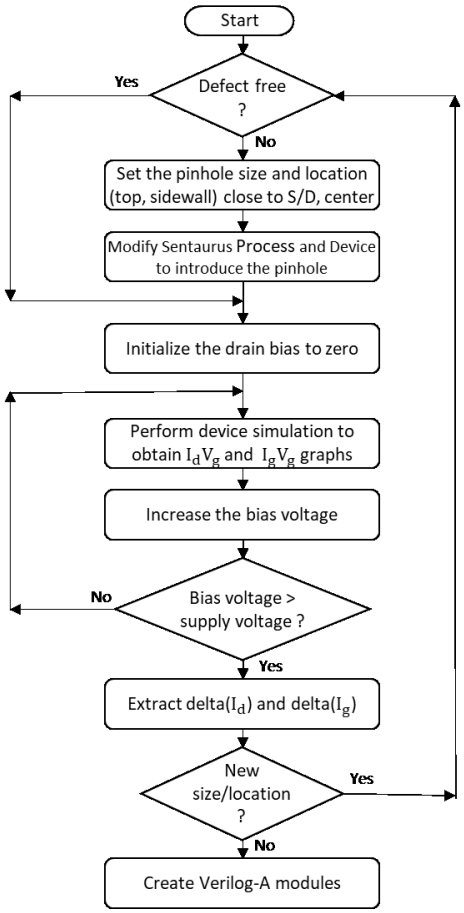

Figure 5.32 GOS Defect model extraction flow diagram

The differences between the DC behaviors of the defective and defect-free FinFETs are extracted, as presented in Equation 5.1,

$$\Delta I_g = I_{g,def} - I_g \tag{5.1a}$$

$$\Delta I_d = I_{d,def} - I_d. \tag{5.1b}$$

The procedure is shown for a specific defect scenario; however, it is extended to other defect sizes and locations. Considering an $8\,\text{nm} \times 8\,\text{nm}$ pinhole on the fin middle sidewall, the difference in currents, ΔI_g and ΔI_d, are obtained at different bias voltages, some of which are plotted in Figure 5.33 (a) and (b) respectively.

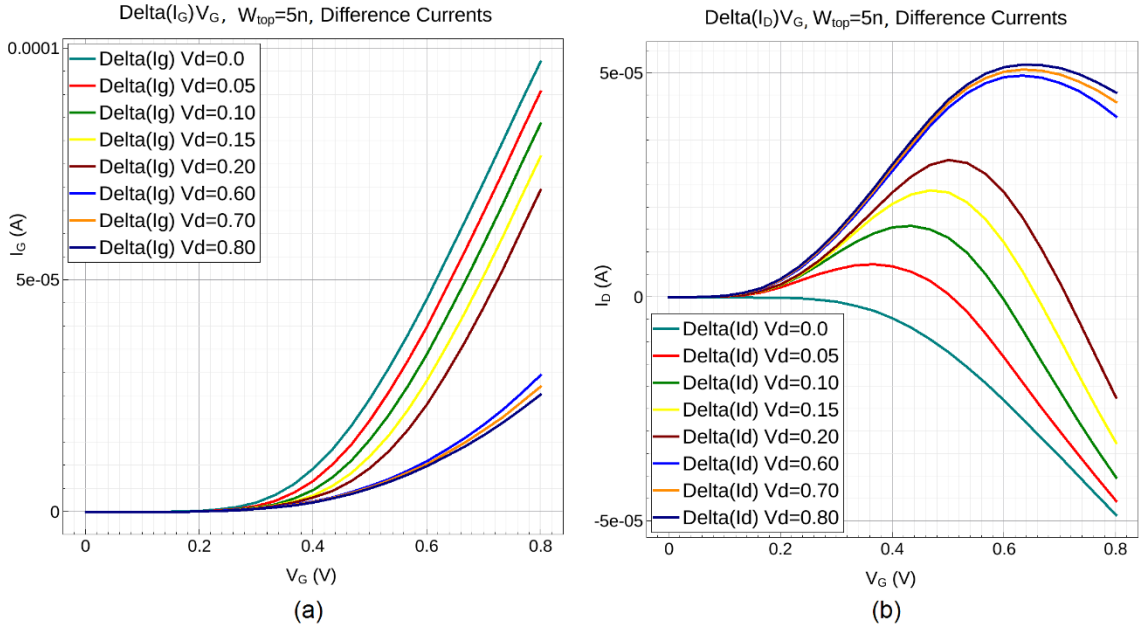

Figure 5.33 (a) The difference in (a) gate currents, (b) drain currents of a FinFET with $8\,\text{nm} \times 8\,\text{nm}$ Mid-C pinhole and the defect-free FinFET at various bias voltages

The difference currents are nonlinear functions of the gate voltage and can be modeled by nonlinear voltage-dependent current sources. However, adding two current sources to the defect-free transistor does not reflect the impact of variations on the drain current as it

happens in the defective device. Therefore, the proposed model consists of three nonlinear current sources depicted as functions f_1 to f_3 in Figure 5.34.

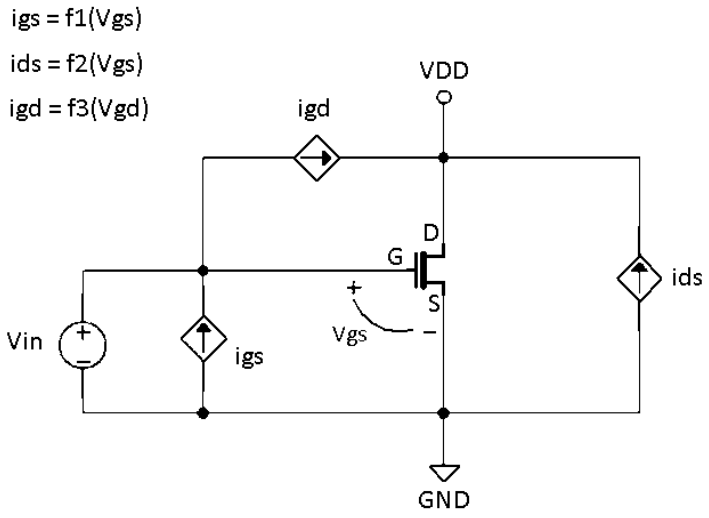

Figure 5.34 Preliminary GOS compact model

The difference currents are obtained from Equation 5.2,

$$\Delta I_g = I_{gs} - I_{gd} \tag{5.2a}$$

$$\Delta I_d = I_{ds} + I_{gd}. \tag{5.2b}$$

The current sources are modeled by analytical functions [60] represented by a family of polynomials with appropriate order in Equations 5.3 to 5.5,

$$I_{gs} = a_2 V_{gs}^2 + a_1 V_{gs} + a_0 \tag{5.3}$$

$$I_{ds} = b_4 V_{gs}^4 + b_3 V_{gs}^3 + b_2 V_{gs}^2 + b_1 V_{gs} + b_0 \tag{5.4}$$

$$I_{gd} = c_3 V_{gd}^3 + c_2 V_{gd}^2 + c_1 V_{gd} + c_0, \tag{5.5}$$

to satisfy Equation 5.2. Although in planar MOSFET the maximum order of these polynomials is 3, realizing the defective FinFET currents require higher orders as stated in Equations 5.4 and 5.5. Since the nonlinear voltage-dependent current source is not available as a built-in model in Sentaurus Device, an interface is provided for user-defined

compact models that are implemented in C++ and linked to the tool. The user-defined model is represented by time-domain equations based on the definitions in [78]. Sentaurus Device solves the system of equations in the form presented in Equation 5.6,

$$\frac{d}{dt}q(t, z(t)) + f(t, z(t)) = 0 \tag{5.6}$$

where z is the vector of electrode voltages $(u1, u2, ..)$, thermal variables, and internal variables $(i1, i2, ..)$; q is the vector consisting of the coefficients of derivative components so called transient right-hand side; and f is the vector of non-derivative coefficients also called DC right-hand side.

Based on the above definitions, the nonlinear current source model is built and linked to the device simulations of the proposed GOS compact model in Figure 5.34. The DC characteristic curves of the compact model are obtained and compared to the defective device at two of the bias voltages as shown in Figure 5.35.

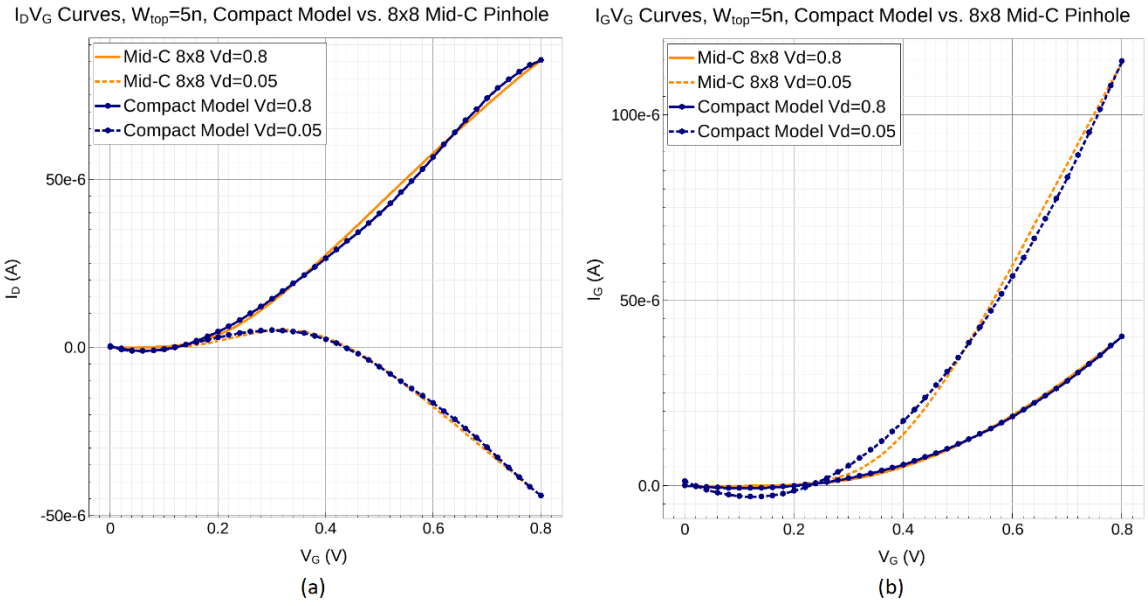

Figure 5.35 (a) I_dV_g, (b) I_gV_g curves of the compact model vs. defective transistor at linear and saturation bias

Considering the drain and gate currents of the defective transistor with 8 nm × 8 nm mid-C pinhole as a reference, the errors of the compact model currents are calculated at the mentioned bias conditions as presented in Table 5.6.

Table 5.6 Error of the compact model currents

	$V_D = 0.05$ V		$V_D = 0.8$ V	
	I_D	I_G	I_D	I_G
μ	$-2.1e-7$	$-6.4e-7$	$-2.2e-7$	$-9.5e-8$
σ	$7.2e-7$	$2.5e-6$	$1.4e-6$	$4.1e-7$

Increasing the order of polynomials in Equations 5.3 to 5.5 would reduce the error, though with computational overhead. This procedure is repeated for all bias conditions with the step of 0.05 V as shown in Figure 5.33; however, smaller steps could be applied as required by the precision of the model.

In order to enhance the transient behavior of the GOS compact model, the gate capacitances of the defect-free and defective FinFETs are extracted at different bias conditions shown in Figure 5.36.

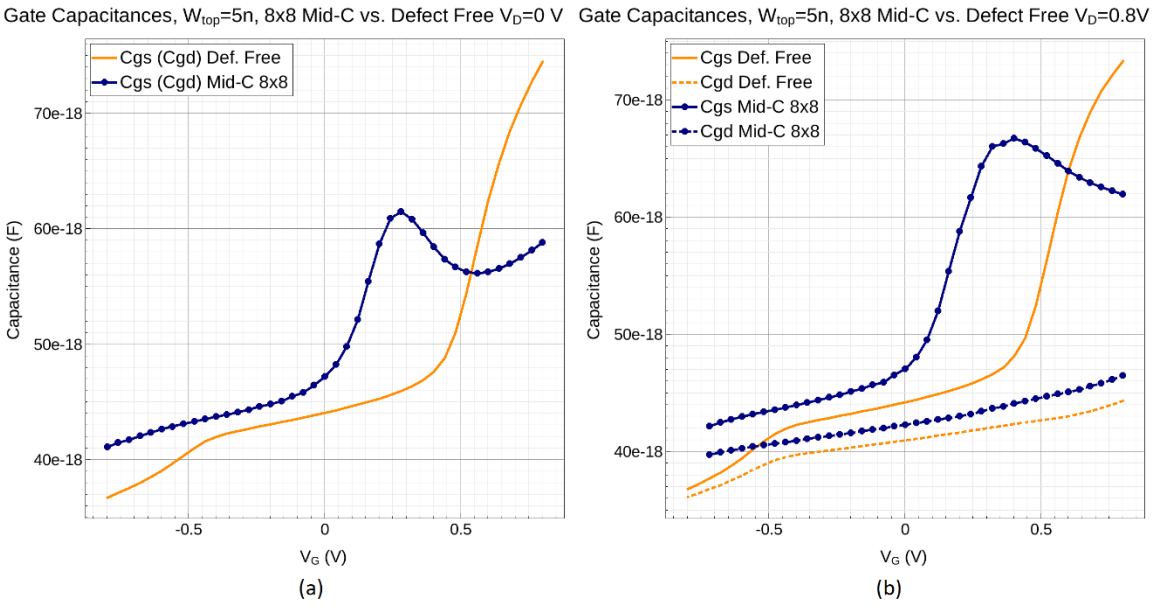

Figure 5.36 CV characteristics of the defective and defect-free FinFETs at (a) linear, (b) saturation bias

The capacitance differences at various bias voltages are obtained, some of them plotted in Figure 5.37. The voltage steps could be increased as required by the precision of the model.

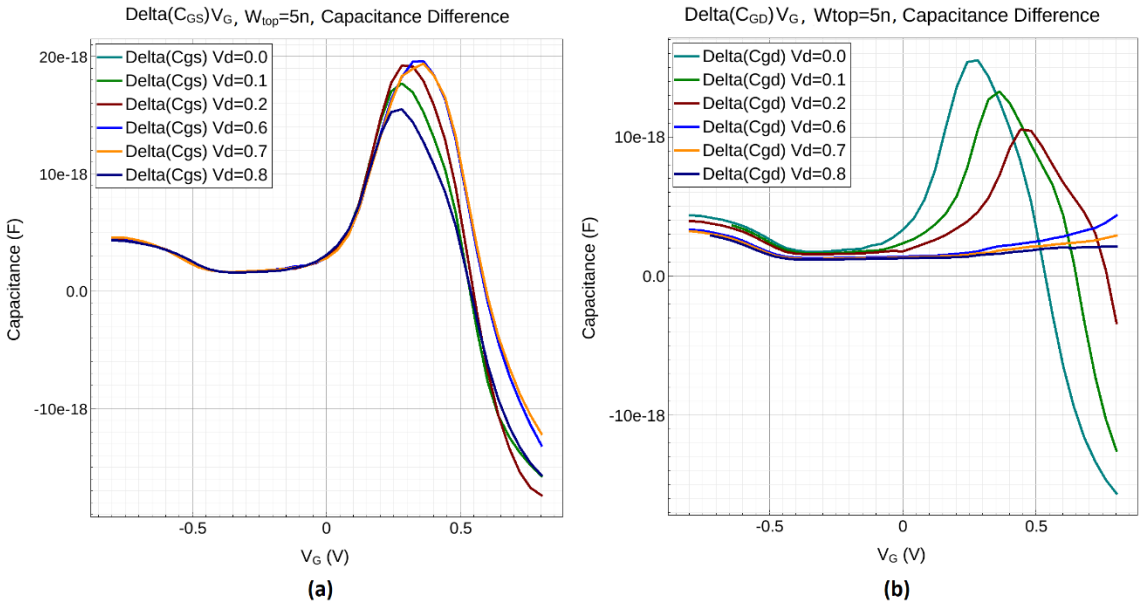

Figure 5.37 The difference in the (a) gate-source capacitances, (b) gate-drain capacitances of a FinFET with 8 nm × 8 nm Mid-C pinhole and the defect-free FinFET at various bias voltages

We have modeled the impact of GOS defect on FinFET dynamic behavior by incorporating nonlinear capacitors, C_{gs} and C_{gd} , represented by functions f_4 and f_5 in our proposed compact model, as shown in Figure 5.38.

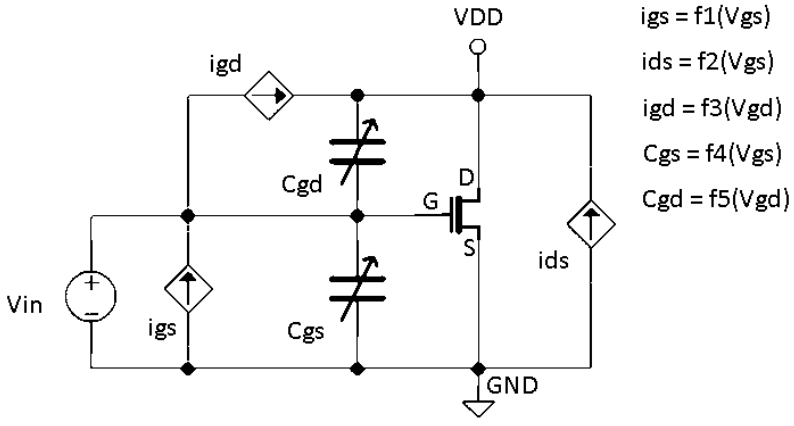

Figure 5.38 The proposed GOS compact model for FinFET

The procedure is repeated for other pinhole sizes and locations in NFinFET at all steps of bias voltages to construct accurate defect models for each scenario. The impact of GOS defect on PFinFET is also investigated by simulating different pinhole scenarios on the triangular fin shape transistor. The defect models are used in circuit-level simulations to obtain appropriate fault models at a higher level of abstraction.

5.9 HSPICE Model Parameter Extraction

To perform circuit level simulations, HSPICE model parameters for n- and p-type FinFETs are required. These parameters are extracted from TCAD simulation results of the triangular FinFET template. The DC characteristic curves of FinFET and the transient behavior of a defect-free inverter are compared with corresponding TCAD results to evaluate their accuracy.

5.9.1 HSPICE Model Parameters

The HSPICE model parameters are extracted based on the local extraction procedure presented in Table 4.2. The initialization process includes setting the model control and process parameters such as $\mathrm{HFIN}, \mathrm{TFIN}, \mathrm{NBODY}, \mathrm{L}, \mathrm{PHIG}$ which are the fin height, fin width, channel doping concentration, channel length, and the gate work-function respectively. According to the Synopsys FinFET-14nm template in Sentaurus process simulations, these parameters are set to $\mathrm{HFIN} = 35\ \mathrm{nm}$, $\mathrm{TFIN} = 15\ \mathrm{nm}$, $\mathrm{NBODY} = 2 \times 10^{18}/\mathrm{cm}^3$. The gate work function is set to the value used in TCAD simulations, $\mathrm{PHIG} = 4.623\ \mathrm{eV}$. According to [69], Matlab scripts are used to calculate the required initial values. Since some of the parameters to be extracted are length-dependent, the

channel length L, is set to different values of 140 nm, 100 nm, 80 nm, 60 nm, 40 nm, and 25 nm. ΔL is the overlap/underlap between the gate and the source/drain diffusions, and in long channel devices it is equal to LINT. To estimate the initial values of LINT and S/D series resistance (RDSW), the $\frac{V_d(\sim 0.05V)}{I_d(V_g,L)}$ is plotted in Figure 5.39. It should be noted that $V_{ov} = V_{gs} - V_{th}$ is the gate overdrive voltage.

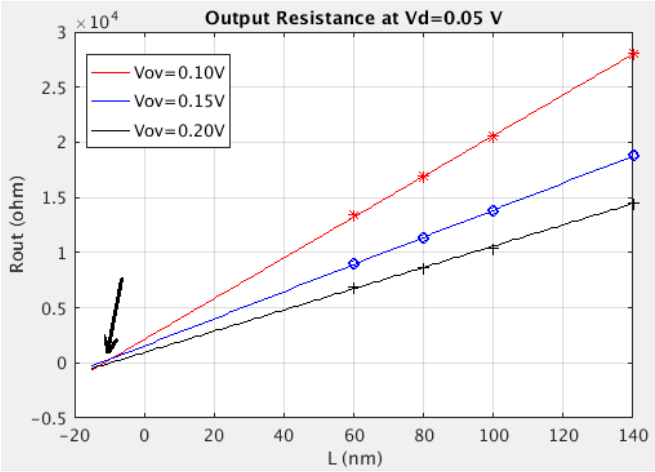

Figure 5.39 Output resistance at $V_d = 0.05$ V and different overdrive voltages

The markers show the output resistance of transistors at different gate overdrive voltages simulated in TCAD environment. The intersection of the fitting lines provides the ΔL, R_{series} [69], that can be used to initialize the mentioned parameters as LINT $= \frac{\Delta L}{2} \approx -5$nm, and RDSW which is equal to $R_{series} \approx 100\Omega$. All the other parameters are set to their default values during the extraction procedure, while at each step some of them are refined accordingly.

To extract the parameters related to V_{th} roll-off, the constant-current method is used in subthreshold region [69]. The threshold voltage roll-off depends on short-channel effects, DIBL, reverse SCE, and temperature [69]. Assuming the nominal temperature to avoid further complications, at low drain bias the parameters related to short-channel effects are

extracted by curve-fitting using the TCAD data. The threshold voltage roll-off based on the extracted parameters is compared with TCAD results in Figure 5.40 (a). A similar procedure is applied to extract the estimated parameters related to DIBL from the saturation region data. Figure 5.40 (b) shows the threshold voltage roll-off due to DIBL effect.

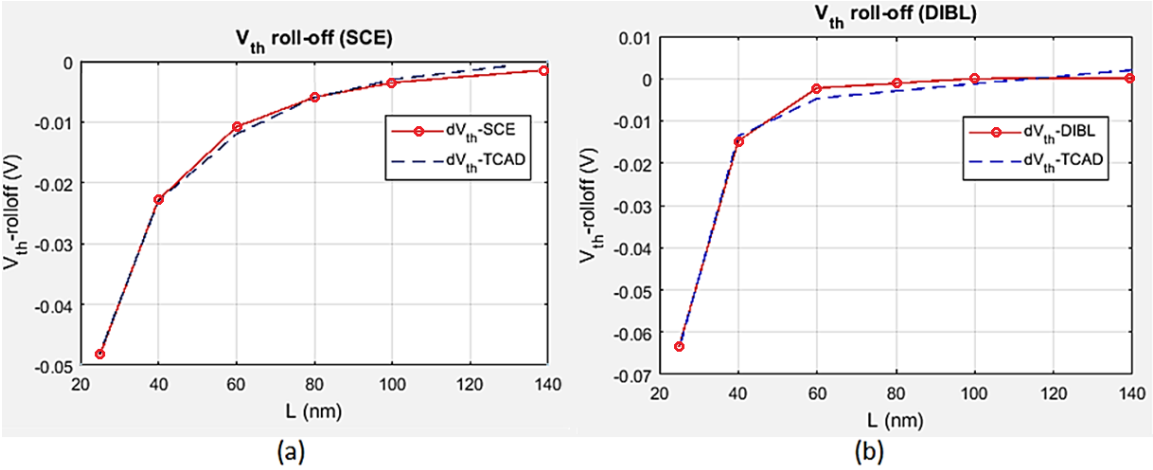

Figure 5.40 Threshold voltage roll-off due to (a) SCE, (b) DIBL effects

The capacitance optimization is performed using $C_{gg}V_{gs}$ curve at linear bias. First, a long channel device capacitance is considered to obtain an estimate of the mentioned parameters in Step 2 of Table 4.2, then the parameters are refined for the short channel device. Recalling from section 4.1.1, the FinFET template has a 2-layer gate dielectric of SiO_2 and HfO_2 with a total thickness of 2.3 nm. While in the extraction procedure, the dielectric constant of SiO_2 is used and the oxide thickness parameters might not represent the physical respective values. The related parameters are EOT, effective oxide thickness, and TOXP, physical oxide thickness, that control the height of the rising edge in the CV characteristics. The quantum mechanical related parameters are also refined for the short channel device. The optimized capacitance graph from HSPICE simulation at low drain bias is plotted in Figure 5.41. The red curve is the HSPICE result and the blue one is from TCAD simulations as the reference.

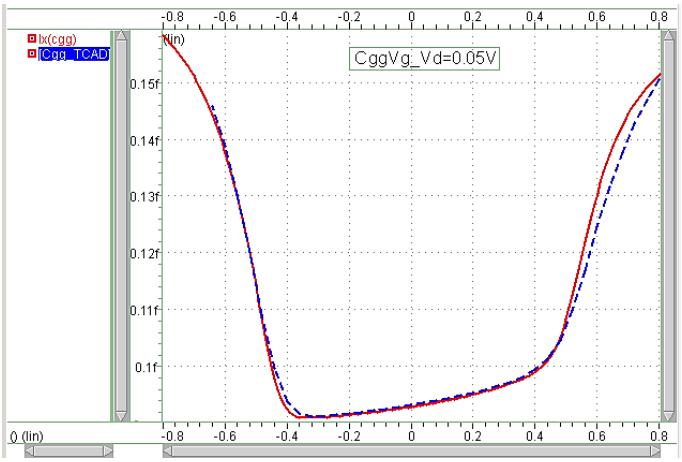

Figure 5.41 Optimized $C_{gg}V_{gs}$ characteristics at $V_d = 0.05$ V

In Step 3 of Table 4.2, the parameters related to the subthreshold slope at low drain bias are optimized using the I_dV_g graphs with logarithmic scale. The initial values of SCE-related parameters are refined as well as those related to the subthreshold leakage and subthreshold slope, S. The optimized I_dV_g curve from HSPICE is plotted in red in Figure 5.42, while the respective curve from TCAD simulation is shown in blue.

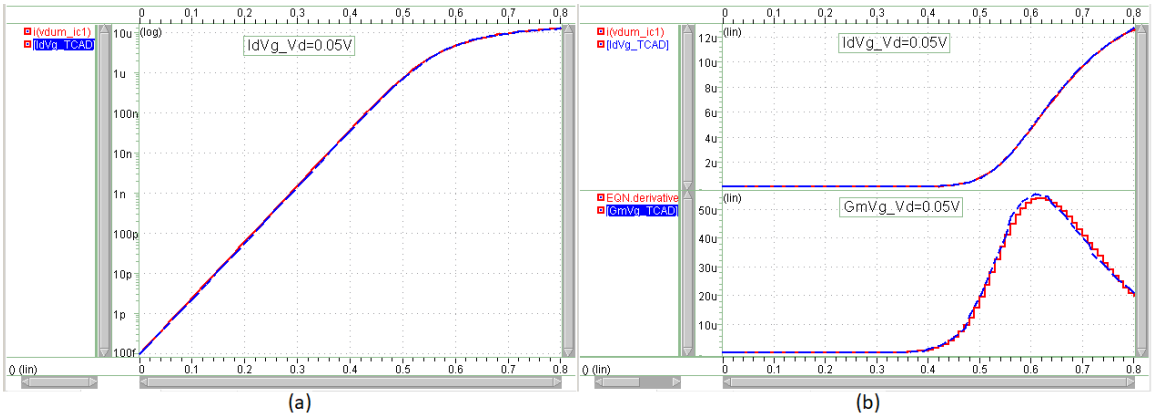

Figure 5.42 Optimized (a) I_dV_g, (b) G_mV_g graphs at $V_d = 0.05$ V

In Step 4 of Table 4.2, the parameters related to DIBL effect, and velocity saturation at $V_d = 0.8$ V are refined. The related parameters are extracted from the simultaneous observation of the I_dV_g and G_mV_g graphs through the refinement process.

On the I_dV_g graph at saturation bias and logarithmic scale, the gate induced drain leakage GIDL-related parameters are extracted by appropriate curve-fitting starting from their default values. The DIBL-related parameters and those affecting the subthreshold slope at high V_{ds} are extracted as well. The optimized I_dV_g graph in logarithmic scale is plotted at saturation bias in Figure 5.43 (a). The initial estimates of the parameters related to velocity saturation are obtained from the I_dV_g and G_mV_g graphs and later refined from the I_dV_d characteristics. The optimized graphs at $V_d = 0.8$ V are plotted on Figure 5.43 (b).

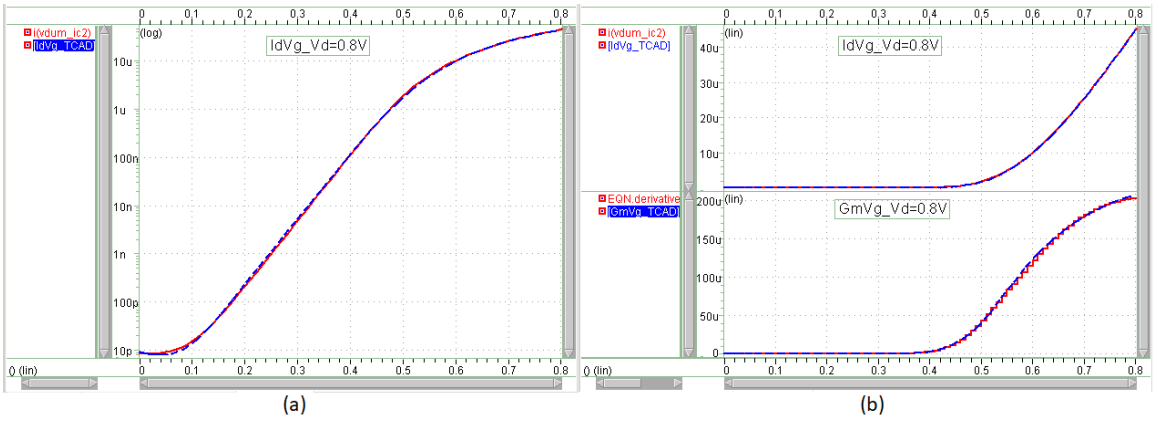

Figure 5.43 Optimized (a) I_dV_g, (b) G_mV_g graphs at $V_d = 0.8$ V

In Step 5 of Table 4.2, the I_dV_d curves and the output conductance curves at different V_{gs} voltages are used to extract the related parameters. Iterations between this step and the previous one results in optimized results. Therefore, the impact of each parameter variation on the characteristics should be observed simultaneously. The parameters that affect the channel length modulation and the slope of the output characteristics are extracted at this step. The variations in these parameters might also affect the G_mV_g and I_dV_g curves and should be considered in the refinements. Figure 5.44 shows the optimized output characteristic curves at three different gate voltages.

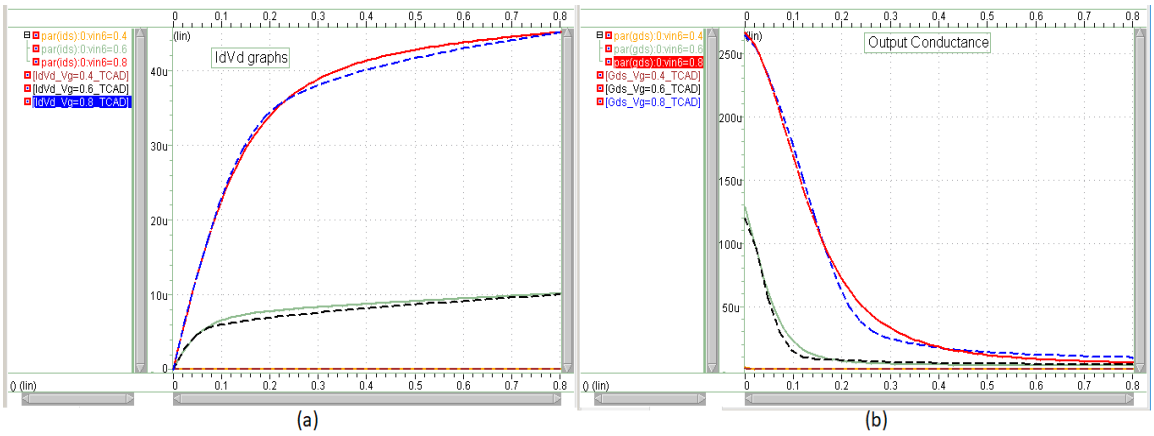

Figure 5.44 Optimized (a) I_dV_d, (b) $G_{ds}V_d$ graphs at $V_{gs} = 0.4, 0.6, 0.8$ V

Finally, in Step 6 of Table 4.2 the gate capacitance at saturation bias is used for extracting the related parameters that affect C_{gg} value and the graph curvature at high V_{gs}. The optimized capacitance at saturation bias is plotted in Figure 5.45. The HSPICE result at maximum V_{gs} has 0.01 fF difference with the TCAD, that might be related to the parameters related to quantum mechanical effects, since the thickness of the charge centroid in the FinFET affects the effective gate capacitance.

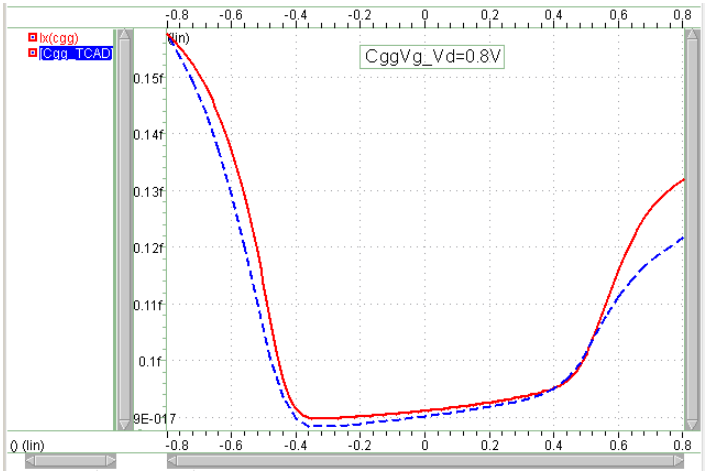

Figure 5.45 Optimized $C_{gg}V_{gs}$ characteristics at $V_d = 0.8$ V

A similar procedure is applied to extract the PFinFET model parameters and the optimized characteristic curves are presented in Appendix A .

5.9.2 Model Verification

To investigate the DC and transient behavior of the extracted model, a defect-free inverter is simulated in both TCAD and HSPICE environments. The load capacitance is 3 fF and the input pulse has a rise/fall time of 10 ps and duration of 150 ps. Figure 5.46 shows the VTCs and the transient responses from TCAD and HSPICE simulations.

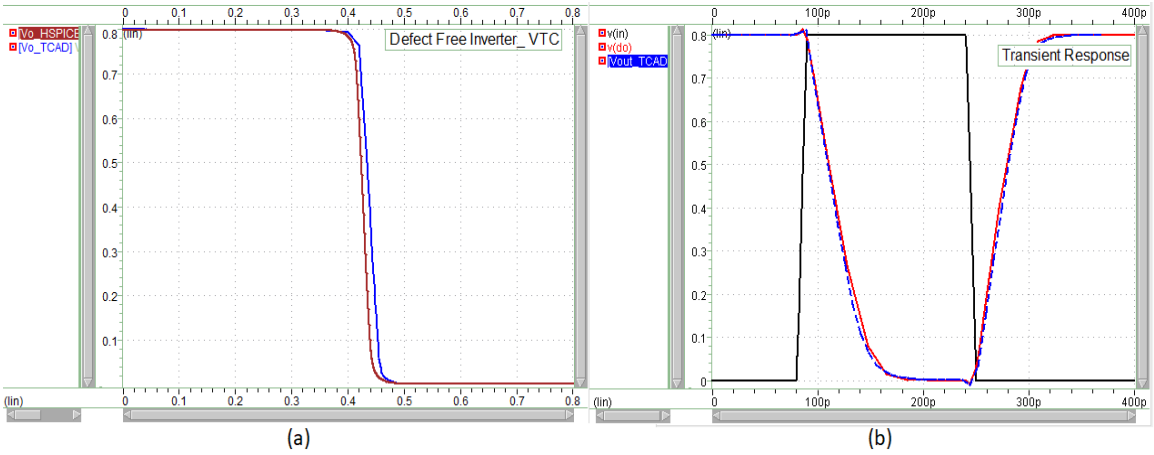

Figure 5.46 (a) VTC, (b) Transient response of the defect-free inverter from TCAD and HSPICE

From the VTC characteristics, the inverter threshold voltage, V_M, at which the logic transition occurs is obtained as shown in Table 5.7. The threshold voltage from HSPICE simulation slightly deviates from the TCAD results with an error of -2.5%.

Table 5.7 The inverter threshold voltage for logic transition

	V_M (V)
TCAD	0.437
HSPICE	0.426
Error	-2.5%

The transient simulation results show that this error has no impact on the inverter dynamic behavior. Therefore, the extracted model parameters can appropriately model the FinFET in further circuit level simulations.

At the next step, the GOS defect models are applied to the transistors in a standard circuit structure such as AOI to investigate the circuit behavior in presence of defect and develop appropriate fault models.

5.10 Summary

This chapter presents the TCAD Sentaurus simulation results for different fin geometries. It covers the impact of various pinhole scenarios on the DC and dynamic behavior of FinFET by obtaining the IV and CV characteristics of the transistor in presence of GOS defect. The results are compared with corresponding characteristics of defect-free transistor to extract the differences in currents and capacitances and develop accurate defect models. Also, the HSPICE model parameters of the FinFET template are extracted. In the next chapter, the GOS compact models are applied on selected transistors in standard circuit structures including Inverter, AOI, and MUX. The results are used to obtain appropriate fault models that represent the impact of defect on circuit behavior and to identify input patterns that can detect the GOS defect.

Chapter 6: Circuit Level Simulation and Fault Modeling

In this chapter, circuit-level simulations are performed with the aid of extracted HSPICE model parameters for triangular-fin FinFET and the developed defect models for various pinhole scenarios. First, the simulation strategy is discussed. Next, the GOS defect models are incorporated to selected transistors in complex gates such as AOI and 3:1 MUX, that have stack of transistors and multiple stage structure, respectively. The circuit behavior in the presence of GOS defects based on our models is studied and compared with the behavior of the same defective circuit using CAT models.

6.1 Simulation Strategy

Among the six defect types considered in the CAT methodology, only the Bridge and Tleak may represent the GOS. With Tleak, any transistor defect that will switch a transistor partially on is modeled as a resistor in parallel with source and drain. A logic input 0 to the defective transistor will not turn off the drain-source path completely and this is in contrast to the GOS behavior. Therefore, it is not included in this chapter. However, the simulation results are available in Appendix B .

The Bridge defect may occur between Gate-Drain (GD) or Gate-Source (GS) which is modeled by various resistor values inserted between these terminals. This scenario is thoroughly investigated in the subsequent sections and are compared to our GOS model. The following GOS defect scenarios are used in the circuit simulations; however, the methodology is applicable to the other pinhole size and locations.

- 2 nm × 2 nm pinhole at the fin Top-C, Top-S, Top-D, Mid-C, Mid-S
- 4 nm × 4 nm pinhole at the fin Up-C, Mid-C, Mid-S, Mid-D, Low-C

- 8 nm × 8 nm pinhole at the fin Up-C, Mid-C, Mid-S, Mid-D, Low-C

The flow diagram in Figure 6.1 shows the important steps in generating the FinFET fault models. For each GOS defect scenario, the defect model is defined as a Verilog-A module. The modules are added to the HSPICE netlist for the specific transistor in the pull-down or pull-up network of the CUT in the experimental setup presented in Section 4.3.

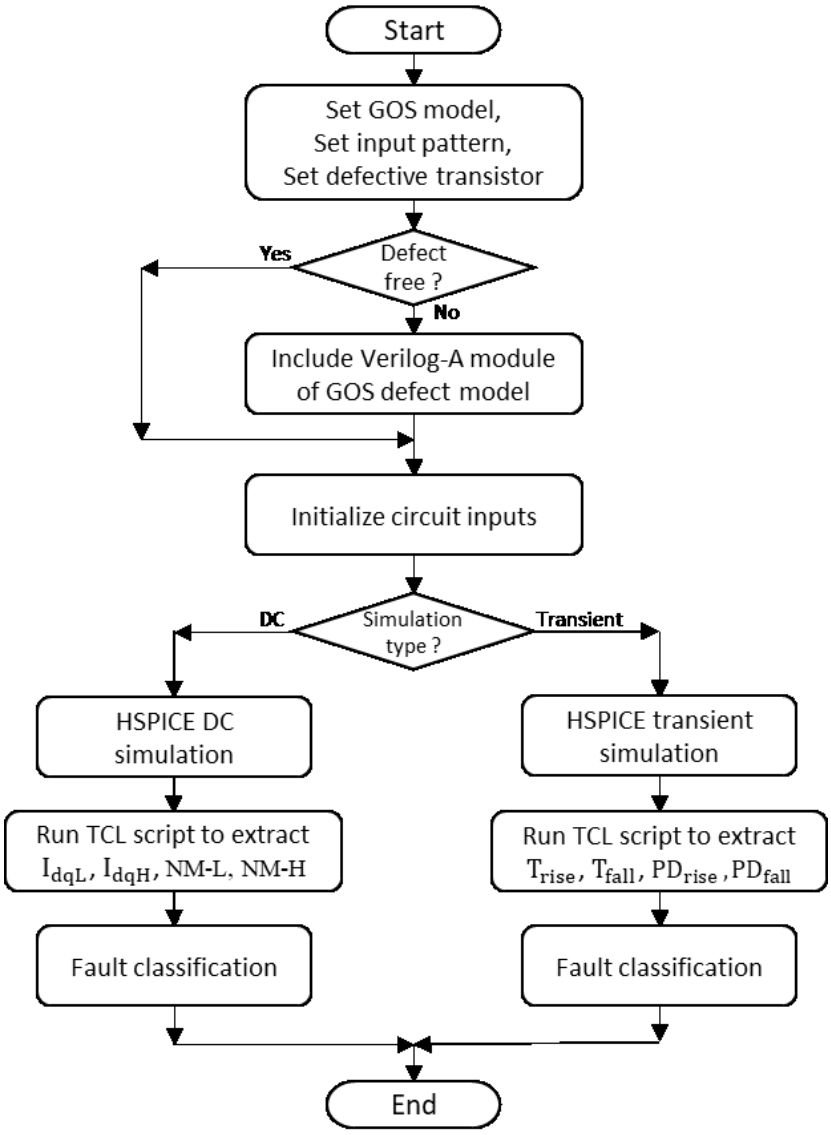

Figure 6.1 FinFET fault model generation flow diagram

The DC and transient analysis are performed on the CUT and the results are used in tool command language (TCL) scripts to extract the required characteristics. The transient

simulations are repeated for up to five unit loads. Although larger load capacitance increases the propagation delays, the percentage of delay variations due to GOS defect are almost the same; hence the results for 3 unit loads are only presented. In fault classification step, if the extracted characteristics deviate from those of the defect-free circuit by more than the specified thresholds, the defect is considered as detectable. The results are compared to the CAT approach to demonstrate the effectiveness of the modeling technique proposed in this thesis.

6.2 Inverter

An inverter with single-fin transistors, shown in Figure 6.2, is considered as the CUT. The defect models of the mentioned pinhole scenarios are applied to the circuit simulations.

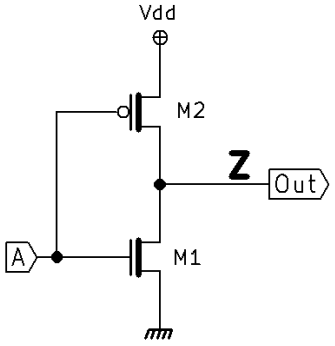

Figure 6.2 Inverter circuit diagram

According to Figure 5.24 and Figure 5.25, the GOS defect in transistor M1, increases the saturation current between 55% to 140% depending on the pinhole size and location. Transistor M1 turns on at lower input voltage, and there is a shift to left in the VTC. Figure 6.3 compares the VTC of the defect-free inverter with that of 4 nm × 4 nm Up-C scenario as example.

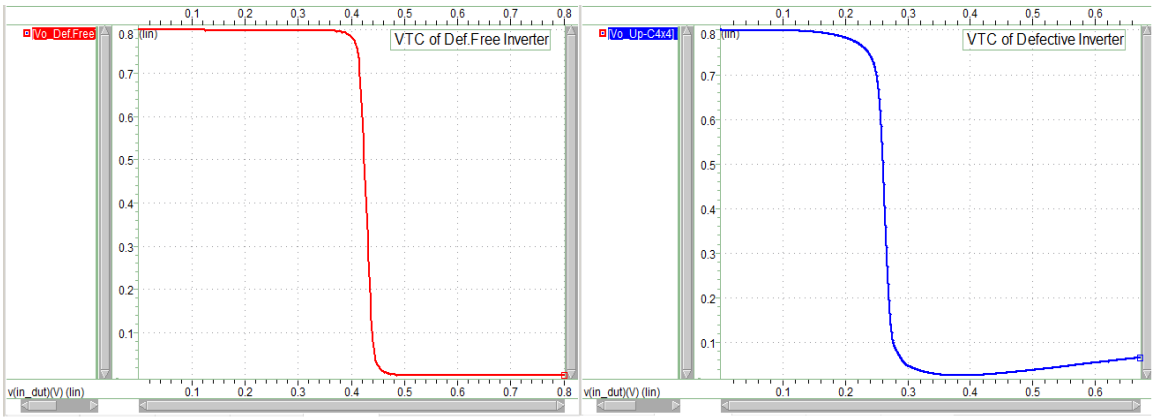

Figure 6.3 VTC of the (a) defect-free (b) defective inverter with Up-C4x4 pinhole

The DC characteristics of the defect-free inverter are summarized in Table 6.1.

Table 6.1 DC characteristic of the defect-free inverter

Scenario	Idq-L (A)	Idq-H (A)	NM-L (V)	NM-H (V)
Def. Free	21.06 p	21.06 p	0.385	0.330

All pinhole scenarios cause significant reduction in the low noise margin due to the shift left in the VTC, as listed in Table 6.2.

Table 6.2 Noise margins of CUT with different defect scenarios in transistor M1

Scenario	NM-L (V)	NM-H (V)
Low-C4x4	-54.69	44.37
Mid-S8x8	-54.79	42.03
Mid-D4x4	-54.82	42.91
Mid-C8x8	-54.85	41.88
Mid-C2x2	-54.90	44.55
Up-C4x4	-54.90	42.88
Top-D2x2	-55.21	44.34
Top-C2x2	-55.21	44.16
Mid-C4x4	-55.31	44.64
Top-S2x2	-55.39	43.28
Mid-S2x2	-55.39	43.16
Low-C8x8	-55.60	45.25
Mid-D8x8	-58.61	38.79
Up-C8x8	-60.54	41.31
Mid-S4x4	-61.60	42.64

The GD short and GS short impact on the DC behavior of the inverter is different from the GOS defects. Considering hard defect resistance values, the GD short results in the output stuck-at input fault. For resistance values that represent soft defect, both noise margins are reduced. The GS short in transistor M1 may result in the input stuck-at 0 fault or reduced voltage swing at the CUT input.

The steady-state currents of the driver and CUT at both output states are listed in Table 6.3. When a high input is applied to the defective transistor, the pinhole provides a direct current path to the channel and I_{dqL} is increased by 6 orders of magnitude. The supply current at high output state also raises by 3 orders of magnitude in the presence of pinholes. The GD or GS short models cannot cause the same current variations as the GOS defect due to their linear impact.

Table 6.3 DC characteristics of CUT with different defect scenarios in transistor M1

Scenario	Idq-L (A)	Idq-H (A)
Def. Free	21.06 p	21.06 p
Top-D2x2	11.07 u	13.17 n
Top-S2x2	11.09 u	15.24 n
Top-C2x2	14.54 u	13.35 n
Mid-C2x2	16.53 u	13.06 n
Mid-S2x2	21.21 u	17.09 n
Low-C4x4	27.32 u	13.27 n
Mid-C4x4	33.02 u	13.04 n
Up-C4x4	36.90 u	14.00 n
Mid-S4x4	37.34 u	44.15 n
Mid-D4x4	39.35 u	12.57 n
Low-C8x8	40.22 u	13.35 n
Mid-C8x8	45.96 u	13.71 n
Mid-S8x8	47.20 u	779.0 n
Up-C8x8	47.75 u	16.93 n
Mid-D8x8	50.12 u	11.25 n

The transient analysis of the inverter is performed by applying a piecewise linear pulse with the rise/fall time of 1 ps to the driver input and finding the delay at the CUT output.

Since the effective drive current provides an estimate of propagation delay, I_{eff} of the transistor with all defect scenarios are extracted, as shown in Figure 6.4.

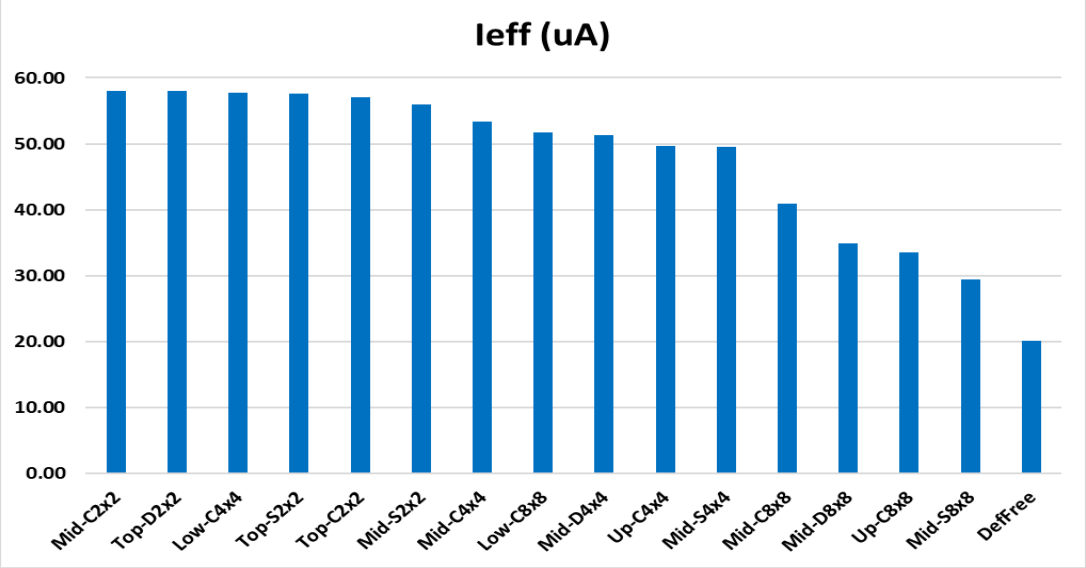

Figure 6.4 Effective current of transistor M1

The 2 nm × 2 nm pinholes cause the largest increase in the drive current and hence, more reduction in the falling delay of the inverter is expected, Figure 6.5. As the pinhole size becomes larger, the increase in I_{eff} becomes smaller.

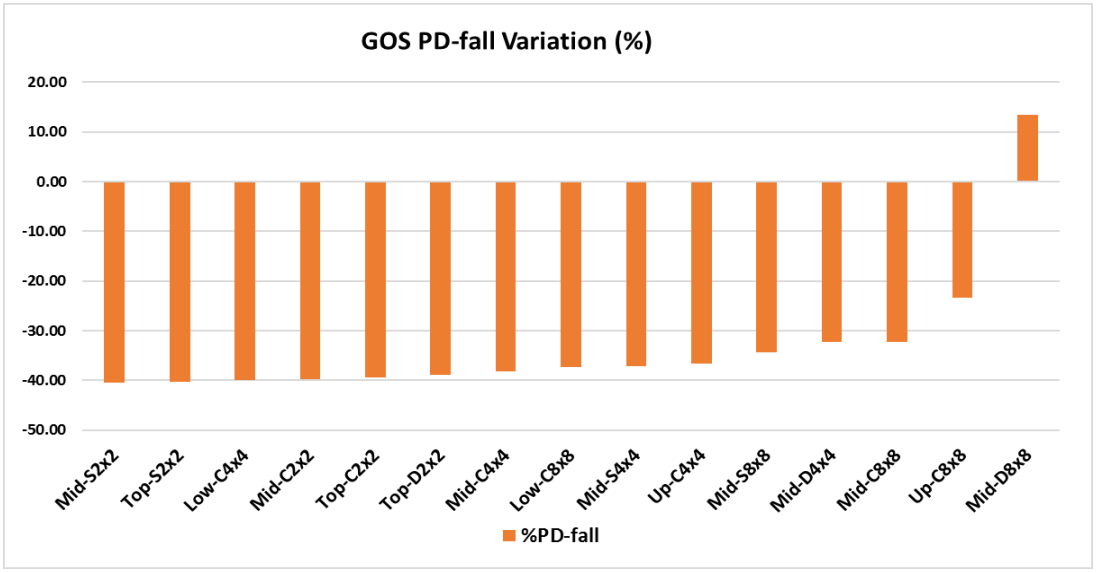

Figure 6.5 Percentage of falling delay variations for different pinholes in transistor M1

When input A is high, GS short with resistance smaller than 9 kΩ keep transistor M1 off and the output is stuck-at 1. For resistance values larger than 11.4 kΩ, the falling delay increases. The GD short results in the output stuck-at input fault for resistance values smaller than 8 kΩ. With resistance values larger than 12 kΩ, the defect does not result in logical faults and causes the falling propagation delay to increase. The results show that GS or GD short cannot model the GOS defect.

In order to reduce clutter in the figures and tables due to the large number of pinhole scenarios and input patterns in complex logic gates, only important ones are included in the following sections.

6.3 AND-OR-Inverter (AOI)

A transistor level representation of the 4-input AOI is shown in Figure 6.6. Since the circuit has symmetric structure, it is sufficient to investigate the impact of defect in Branch 1 of the pull-down network, transistors M1 and M2. Similarly, in the pull-up network, the GOS defect is introduced to transistors M5 and M7.

The input patterns that can detect these defects are obtained using the D Algorithm [79]. Although this algorithm is usually used to detect stuck-at faults at the gate level, the concept is applicable to the transistor-level representation of the AOI with GD and GS short defects.

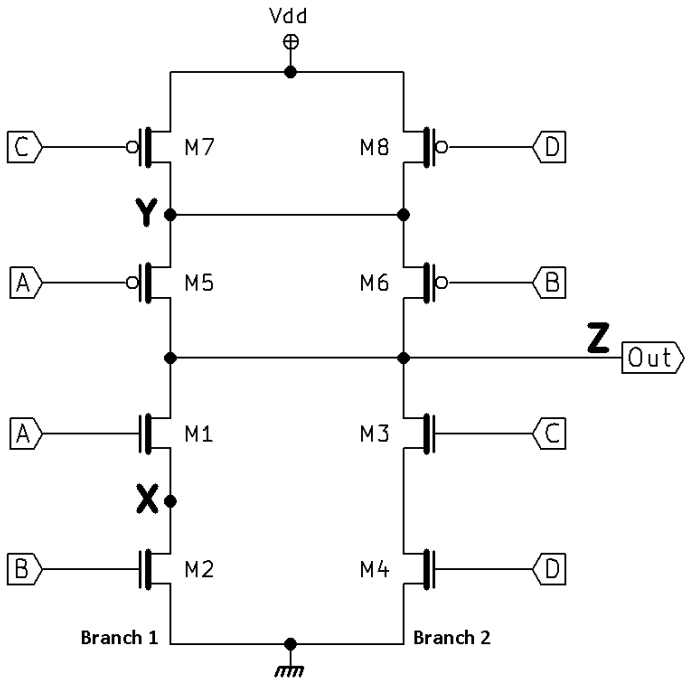

Figure 6.6 4-input AOI circuit

Considering a GD short in transistor M1, the input combinations that activate the faulty behavior should generate a logic 1 at input A. The other inputs should be set in a manner that makes the output zero in the defect-free circuit. There are two possible discharge paths in the pull-down network, therefore, the input patterns should either activate Branch 1, 2 or both. These input combinations are presented in Table 6.4.

Table 6.4 The input patterns that detect GD short defect in M1

A	B	C	D	Z
1	1	0	×	0
1	1	×	0	0
1	0	1	1	0
1	1	1	1	0

The first two patterns can also detect GS short in transistor M1. Since they activate Branch 1 and keep the transistors in Branch 2 off, the defect impact propagates to the output. Considering a GD short in transistor M2, the input combinations that activate the faulty behavior should generate a logic 1 at input B. To propagate the defect impact to the output,

transistor M1 should be on. Therefore, the first two rows of Table 6.4 detect GD short in M2. A similar analysis shows that the same patterns can detect GS short in M2.

6.3.1 Defect in NMOS Transistor M1

To realize the first two rows in Table 6.4, both C and D inputs are considered as 0 and the input pulse is applied to the defective transistor while the other input is held at logic 1, and vice versa, as listed in Cases 1 and 2 of Table 6.5, respectively. In row 3 of the table, the defective transistor is not on the signal path, as the input signal is applied to input C. However, the defective transistor has a high input gate voltage, providing a leakage path to the output; hence, considered as the case of discharge through Branch 2. The last row corresponds to discharging the load capacitance through both branches.

Table 6.5 AOI input combinations applied for GOS in transistor M1

Case	Input Pattern
1	$\mathrm{A}: 0 \rightarrow 1, \mathrm{B} = 1, \mathrm{CD} = 00$
2	$\mathrm{A} = 1, \mathrm{B}: 0 \rightarrow 1, \mathrm{CD} = 00$
3	$\mathrm{AB} = 10, \mathrm{C}: 0 \rightarrow 1, \mathrm{D} = 1$
4	$\mathrm{A} = 1, \mathrm{BC}: 00 \rightarrow 11, \mathrm{D} = 1$

I. Case 1 (A:0->1, B=1, CD=00)

a) GOS Analysis

When the input signal is applied to the defective transistor M1, the GOS leakage reduces the voltage swing of both the driver stage and AOI. The input pattern will cause the AOI behave as an inverter, for which the voltage transfer characteristics is similar to Figure 6.3. The defect reduces the voltage swing of both the driver and AOI. When the driver output is high, which is less than 0.8 V, the AOI output voltage is nonzero due to the GOS leakage current passing through the stacked transistors in this branch. All pinhole

scenarios have the same impact on VTC same as in the inverter and reduce the low noise margin from 55% to 60%.

All GOS defects in transistor M1 increase the supply current I_{dqH} about 3 orders of magnitude and I_{dqL} about 6 orders of magnitude. The close to drain pinholes cause the largest increase in the supply current.

The transient response of AOI considering 3 unit loads is plotted in Figure 6.7 (a). It is observed that the GOS defect results in smaller falling propagation delay and reduced output swing. The CUT output voltage at the low state increases as the defect size increases or when it occurs closer to the source. Since the maximum acceptable voltage for logic 0 input for a FinFET inverter is 0.398 V, the CUT output cannot not be correctly recovered by the load in 8 nm × 8 nm Mid-S scenario at this frequency of operation, as shown in Figure 6.7 (b).

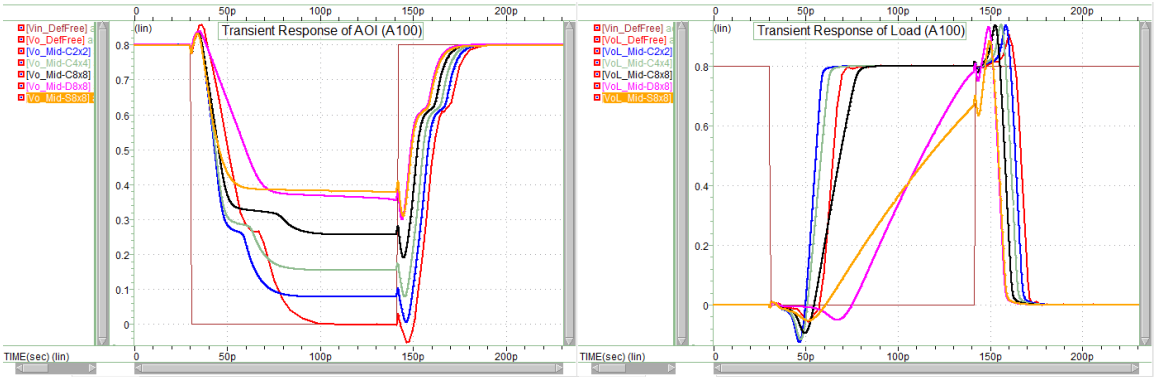

Figure 6.7 Transient response of (a) AOI, (b) load inverter, Case 1

The transition and propagation delays of the defect-free AOI are presented in Table 6.6 for further reference.

Table 6.6 The transition and propagation delays of the defect-free AOI, Case 1

Scenario	PD-rise (ps)	PD-fall (ps)	T-rise (ps)	T-fall (ps)
Def. Free	17.32	23.53	21.87	36.64

Except the 8 nm × 8 nm close to drain pinhole scenarios, the GOS results in the reduction of falling propagation delay ranging from 7% to %37. The percentage of delay variations are plotted in Figure 6.8. Recalling the impact of GOS defect on the transistor IV characteristics, small pinhole results in almost doubling the saturation current comparing to that of defect-free transistor. As the pinhole becomes larger, the amount of increase in the drain current is smaller. Therefore, the percentage of PD-fall reduction of 8 nm × 8 nm Mid-C pinhole is less than 2 nm × 2 nm Mid-C scenario.

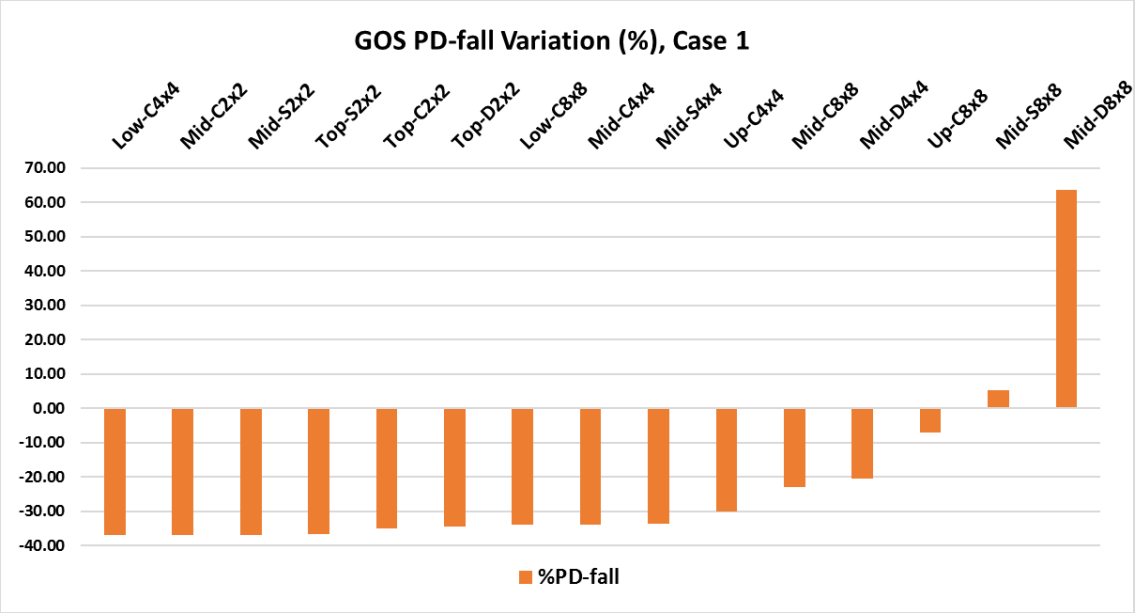

Figure 6.8 Percentage of falling delay variations for different pinholes in transistor M1, Case 1

Since the defect is applied in the pull-down network, the falling delays are considered for fault classification. Also, the output low level voltage and hence the rising delays depend on the input pulse width. Recalling the percentage of delay variations in Table 4.4, and based on the above results, the delay fault categories are listed in Table 6.7.

Table 6.7 Delay fault categories for transistor M1, Case 1

Scenario	Fault
Top-S2x2	PD_F1
Top-D2x2	PD_F1
Top-C2x2	PD_F1
Mid-C2x2	PD_F1
Mid-S2x2	PD_F1, PD_R1
Low-C4x4	PD_F1, PD_R1
Mid-C4x4	PD_F1, PD_R1
Up-C4x4	PD_F1, PD_R1
Mid-S4x4	PD_F1, PD_R1
Mid-D4x4	PD_F1, PD_R1
Low-C8x8	PD_F1, PD_R1
Mid-C8x8	PD_F1, PD_R1
Mid-S8x8	PD_R2
Up-C8x8	PD_R2
Mid-D8x8	PD_F2(+), PD_R2

b) Comparison with CAT-based Modeling

The defect models used in CAT methodology including GS and GD shorts are applied to transistor M1 to examine whether these models reflect the impact of GOS defect on the circuit DC and transient behavior. Figure 6.9 shows the circuit diagram of the AOI with respective defect resistors.

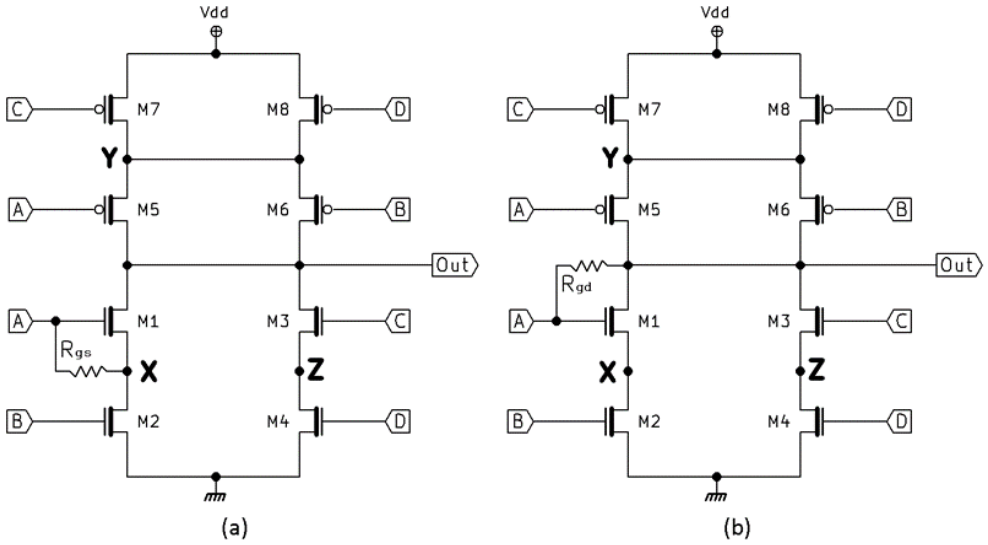

Figure 6.9 AOI with (a) GS short, (b) GD short defect in transistor M1

The GS short may result the output stuck-at 1, or limited output swing as shown in Figure 6.10 (a). It increases the supply current I_{dqL} about 6 orders of magnitude, however, unlike GOS has no impact on I_{dqH}. The impact of GD short on the VTC is also different from GOS scenarios. The defect may cause the output stuck-at input, or reduction in both noise margins, Figure 6.10 (b). Unlike GOS, the GD short increases both I_{dqL} and I_{dqH} about 6 orders of magnitude. Due to the linear nature of GS and GD short defect models, they cannot represent the GOS impact on the circuit DC current behavior.

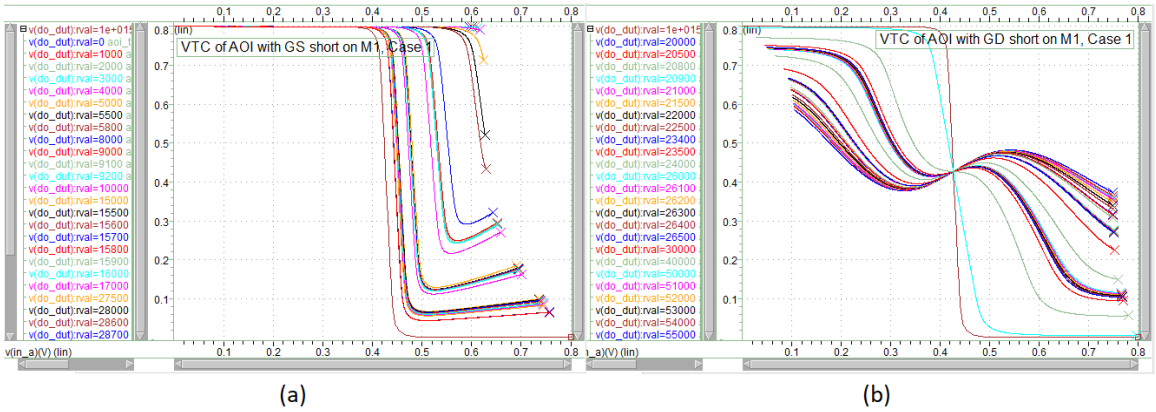

Figure 6.10 VTC of defective AOI with (a) GS short, (b) GD short defect on transistor M1, Case 1

The transient responses of AOI and the load inverter considering GS short defect in transistor M1 are plotted in Figure 6.11.

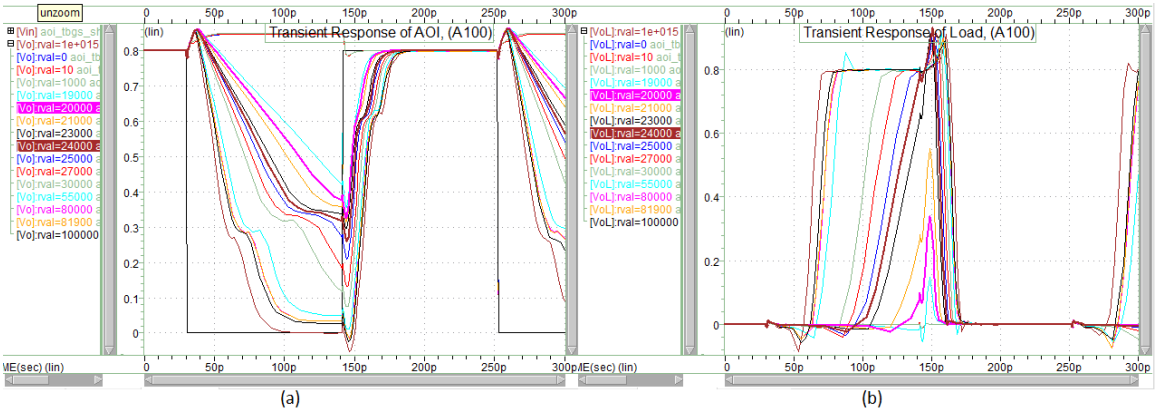

Figure 6.11 Transient Response of (a) AOI, (b) the load inverter with GS short on M1, Case 1

When input A is high, resistance values smaller than 12 kΩ keep transistor M1 off and the output is stuck-at 1. With resistors larger than 19.5 kΩ values (soft defect), the voltage at the internal node X increases, which cause transistor M1 turn on at higher input voltages. Therefore, the falling propagation delay increases significantly, and the output cannot reach its full swing at the low level. The percentage of delay variations for a wide range of defect values are shown in Figure 6.12.

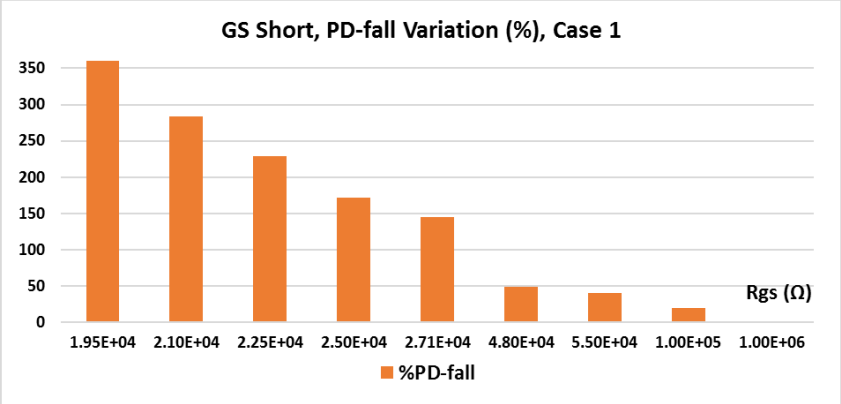

Figure 6.12 Percentage of delay variations in AOI with GS short in transistor M1, Case 1

Next, the transient responses of AOI and the load inverter considering GD short defect in transistor M1 are plotted in Figure 6.13.

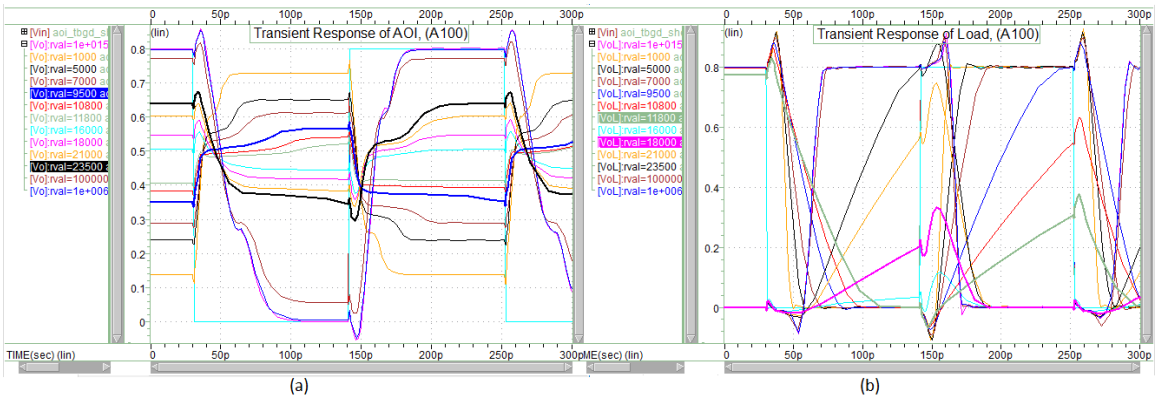

Figure 6.13 Transient Response of (a) AOI, (b) the load inverter with GD short on M1, Case 1

Depending on the GD short resistance value, the output might be stuck-at input for resistance values smaller than 9.5 kΩ. With resistance values larger than 23.5 kΩ, the

defect does not result in logical faults and causes the falling propagation delay to increase. The percentage of delay variations are plotted in Figure 6.14.

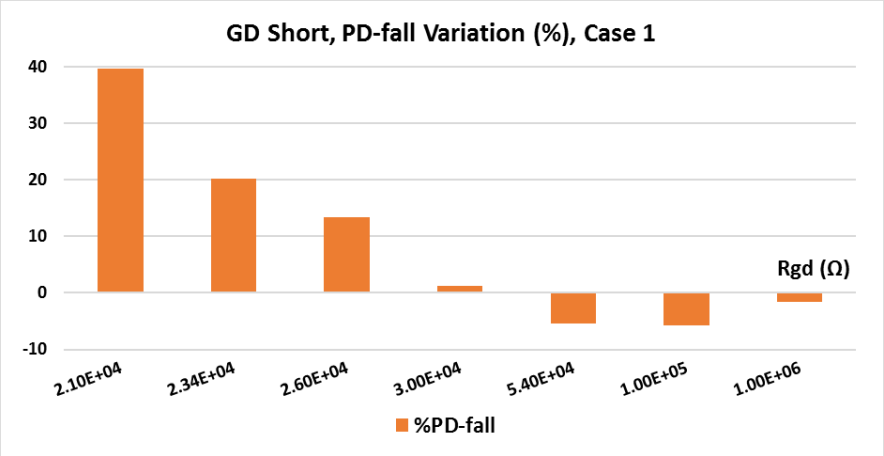

Figure 6.14 Percentage of delay variations in AOI with GD short in transistor M1, Case 1

According to the delay variation results, the GS and GD short can only model the 8 nm × 8 nm Mid-D scenario.

II. Case 2 (A=1, B:0->1, CD=00)

a) GOS Analysis

In this case, the input signal is applied to the defect-free transistor M2. The DC behavior of AOI is different from the previous case since transistor M1 is acting as an active load for M2. When input B transitions to high causing M2 to turn on, the GOS leakage current passing through the channel resistance of M2 sets the internal node X and the output to a nonzero voltage. As the pinhole size increases, the output swing is significantly reduced as shown in Figure 6.15. The GOS defect results in 20% to 80% reduction in the low noise margin, and 50% to 80% reduction in the high noise margin. Since the defect is not on the transistor driven by the input signal, its impact is not reflected in the I_{dq} current measurements.

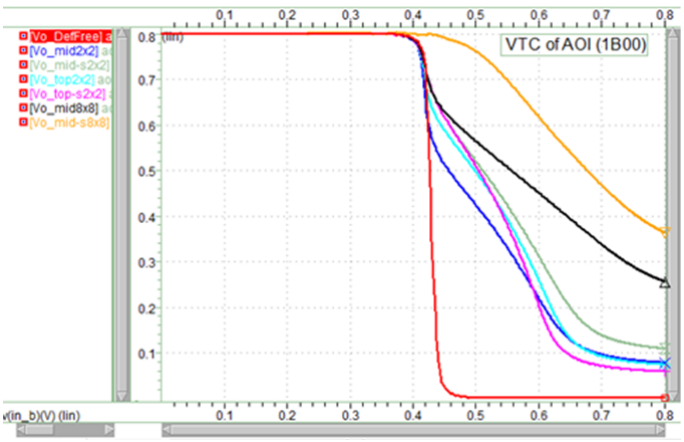

Figure 6.15 VTC of AOI, Case 2

The defect-free propagation delays have a slight increase of up to 5% compared to those in Table 6.6 since the input signal is applied to the lower transistor in the stack. The transient response of defective AOI is similar to Figure 6.7; hence, not repeated.

The GOS defect has the same impact on the delays as with the previous input pattern, though with smaller magnitude of falling propagation delays. It results in 12% to 27% reduction in the falling delay. The delay fault categories are listed in Table 6.8.

Table 6.8 Delay fault categories for transistor M1, Case 2

Scenario	Fault
Top-S2x2	PD_F1, PD_R1
Top-D2x2	PD_F1, PD_R1
Top-C2x2	PD_F1, PD_R1
Mid-C2x2	PD_F1, PD_R1
Mid-S2x2	PD_F1, PD_R1
Low-C4x4	PD_F1, PD_R1
Mid-C4x4	PD_F1, PD_R1
Up-C4x4	PD_F1, PD_R1
Mid-S4x4	PD_F1, PD_R1
Mid-D4x4	PD_R1
Low-C8x8	PD_F1, PD_R1
Mid-C8x8	PD_R1
Mid-S8x8	PD_R2
Up-C8x8	PD_R2
Mid-D8x8	PD_F2(+), PD_R2

b) Comparison with CAT-based Modeling

The GS short defect with resistance values smaller than 12 kΩ keeps transistor M1 off, and the AOI output is stuck-at 1. With resistors larger than 19.5 kΩ values, both transistors gradually turn on when input B increases to 0.8 V. However, the output cannot reach zero voltage due to the resistive connection of the internal node X to the high voltage. The transient behavior of AOI is similar to that in the previous case plotted in Figure 6.11, as well as the percentage of delay variations. The GD short in transistor M1 increases the falling delay similar to the GS short impact, but with smaller magnitude.

The GS or GD short models can only represent the 8 nm × 8 nm Mid-D pinhole scenario.

III. Case 3 (AB=10, C:0->1, D=1)

a) GOS Analysis

In this case, the input pattern that is applied to the AOI will turn transistors M3 and M4 on to discharge the load capacitance, while the input of the defective transistor M1 is high. The VTCs of AOI with different pinhole scenarios are approximately similar to the Figure 6.15. The GOS defect results in 22% to 86% reduction in the low noise margin, and 23% to 64% reduction in the high noise margin. Since the input is applied to the transistor M3, there is no significant changes in the I_{dq} currents.

The discharge of load capacitance occurs through non-defective Branch 2, so the falling delay is not affected by the pinhole. However, by applying a second transition to input C, the rising propagation delay is reduced between 30% to 55% as shown in Figure 6.16.

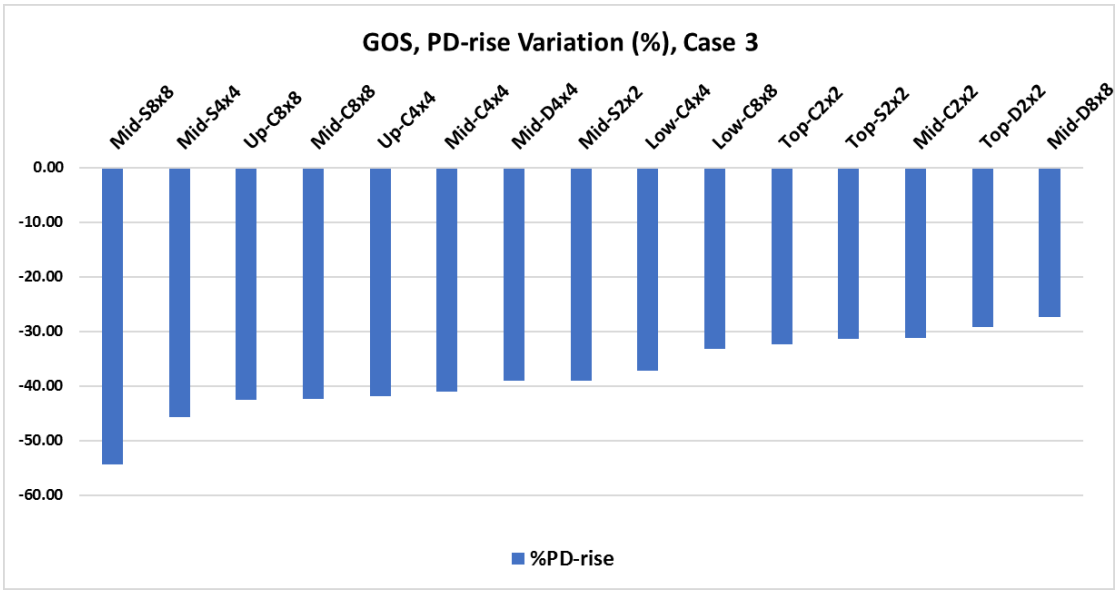

Figure 6.16 The percentage of rising delay variation for different pinholes in transistor M1, Case 3

Except 8 nm × 8 nm Mid-S pinhole that is categorized as PD_R2, all other scenarios are categorized as PD_R1.

b) Comparison with CAT-based Modeling

The transient behavior of AOI for a wide range of GS resistor values shows it has no impact on the falling delays as the discharge occurs via the defect-free branch. However, it affects the rising delays as depicted in Figure 6.17.

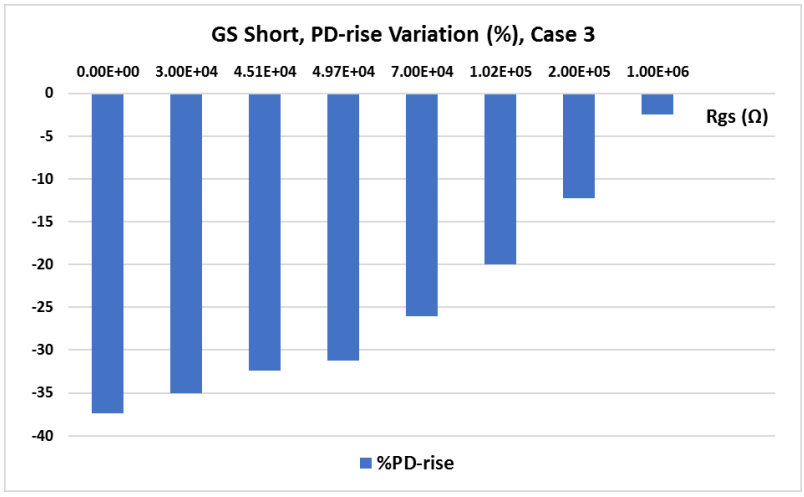

Figure 6.17 Percentage of delay variations in AOI with GS short in transistor M1, Case 3

The GS short defect values that result in the same AOI behavior as GOS defect correspond to the 2 nm × 2 nm Top-C, Top-S, Top-D, Mid-C, 8 nm × 8 nm Mid-D and the Low-C pinhole scenarios.

Unlike the GOS defect, the GD short in transistor M1 increases the falling delay.

IV. Case 4 (A=1, BC:00->11, D=1)

a) GOS Analysis

In this case, the input pattern activates both branches to discharge the load capacitance. The GOS defect results in 15% to 67% reduction in the low noise margin, and 29% to 85% reduction in the high noise margin. Since the input is applied to the transistor M3, there is no significant changes in the I_{dq} currents. When all the AOI inputs are at high voltage, all transistors in the pull-down network conduct. Therefore, in the defect-free circuit the falling propagation delay is reduced by 33% compared to that in Table 6.6, and the rising propagation delay is accordingly increased by 28%. The impact of different pinhole scenarios on the transient response is similar to that in Case 3. Except 8 nm × 8 nm Low-C pinhole that is not detectable and 8 nm × 8 nm Mid-S pinhole that is categorized as PD_R2, all other scenarios are categorized as PD_R1.

b) Comparison with CAT-based Modeling

Both the GS and GD short defects in transistor M1 increase the falling propagation delay that is in contrast to the GOS impact.

6.3.1.1 Discussion

The reduction in the falling propagation delay can be explained by applying Elmore delay model [80], [81] to the AOI circuit which is plotted in Figure 6.18. Each transistor is replaced by a switch and a resistor whose value depends on its channel resistance. The capacitance at each node depends on the source/drain capacitances connected to the node, as well as the corresponding gate-source and gate-drain capacitances for which the equivalent Miller capacitance is used.

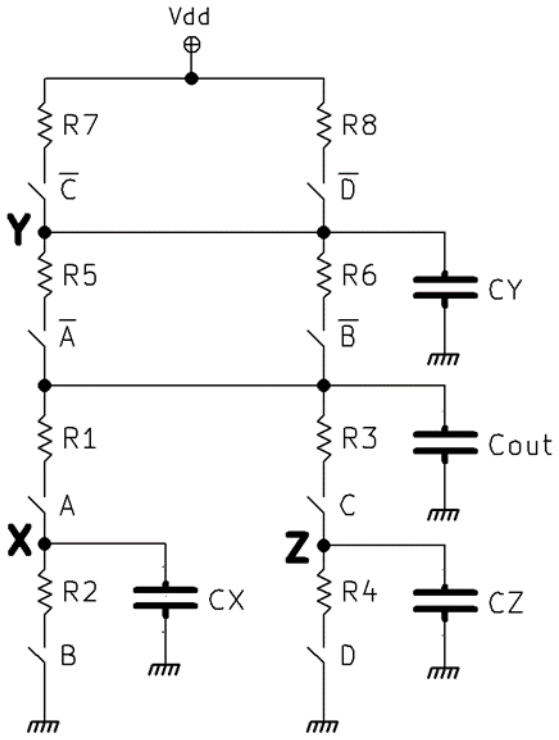

Figure 6.18 Equivalent RC model of the AOI

The Elmore delay in the RC chain is given by Equation 6.1 [81], as

$$\tau_{DN} = \sum_{i=1}^{N} C_i \sum_{j=1}^{N} R_j \qquad (6.1)$$

where N is the number of nodes, C_i is the node capacitance to ground, and R_j is the resistance between consecutive nodes.

Based on the assumed input pattern, switches C and D are open. Therefore, in the pull-down network the internal node capacitance, C_X, and the load capacitance, C_{out}, are those important in the falling propagation delay calculation [81] in Equation 6.2,

$$PD_{fall} = 0.69(R_2C_X + (R_1 + R_2)C_{out}) \tag{6.2}$$

where the node capacitances are calculated from Equations 6.3 and 6.4,

$$C_X = C_{d2} + C_{s1} + 2C_{gd2} + 2C_{gs1} \tag{6.3}$$

$$C_{out} = C_{d1} + 2C_{gd1} + 2C_{d5} + 2 \times 2C_{gd5} + C_{d3} + 2C_{gd3} \tag{6.4}$$

in which C_{di}, C_{si} are the junction capacitances of drain and source, and C_{gsi}, C_{gdi} are the gate-source and gate-drain capacitances of transistor Mi connected to each node, assuming that PMOS transistors have equal capacitances.

Recalling the GOS compact model from previous chapter, it is observed that the C_{gs} and C_{gd} difference capacitances have nonlinear behavior. For example in Figure 5.37, at $V_d = 0.8\,V$ and maximum gate voltage, C_{gs1} of the defective transistor is about 14 fF smaller than the defect-free counterpart, and C_{gd1} is about 4 fF larger than the defect-free one. Therefore, applying GOS defect in transistor M1 would result in smaller C_X and larger C_{out}, which in overall reduces the falling propagation delay. This approximate analysis gives an intuition of how the GOS in transistor M1 has reduced the falling delay. The exception pinhole scenario that increases the falling delay is $8\,nm \times 8\,nm$ Mid-D, in which large number of holes are injected to the region with high electrical field density. The major component of the channel current is the excessive gate leakage current, Figure 5.27 (b), that reduces the gate control over the channel and prevents transistor M1 to turn on.

6.3.2 Defect in NMOS Transistor M2

The input patterns that discharge the load though Branch 1 or both branches are applied to the AOI, as listed in Table 6.9. The input conditions that satisfy the discharge through only Branch 2 is not applicable since the defective transistor would not affect the output node when transistor M1 is off.

Table 6.9 The AOI input combinations applied for GOS in transistor M2

Case	Input Pattern
1	$\mathrm{A}: 0 \rightarrow 1, \mathrm{B} = 1, \mathrm{CD} = 00$
2	$\mathrm{A} = 1, \mathrm{B}: 0 \rightarrow 1, \mathrm{CD} = 00$
3	$\mathrm{A} = 1, \mathrm{BD}: 00 \rightarrow 11, \mathrm{C} = 1$

I. Case 1 (A:0->1, B=1, CD=00)

a) GOS Analysis

The gate of the defective transistor M2 is at high state and the input is applied to M1 through the driver inverter. When input A is low, the output is connected to the supply through the PMOS branch including transistors M7 and M5. While the defective transistor M2 is always on, increasing the input voltage turns M1 on and into linear region. The defect increases the voltage of the internal node X, which causes transistor M1 turn on at a higher input voltage. It causes a shift to right in the VTC and results in 3% to 23% reduction in the low noise margin, and 3% to 29% reduction in the high noise margin. The defect scenarios do not considerably change I_{dq} currents, except the $8\ \mathrm{nm} \times 8\ \mathrm{nm}$ pinholes which increase the low supply current I_{dqL} up to 3 orders of magnitude.

The transient response of the CUT with different pinhole scenarios on the transistor M2 is plotted in Figure 6.19. In $8\ \mathrm{nm} \times 8\ \mathrm{nm}$ Mid-S scenario, increasing the input pulse width to $180\ \mathrm{ps}$ would let the AOI output reach its steady-state value of $0.183\ \mathrm{V}$.

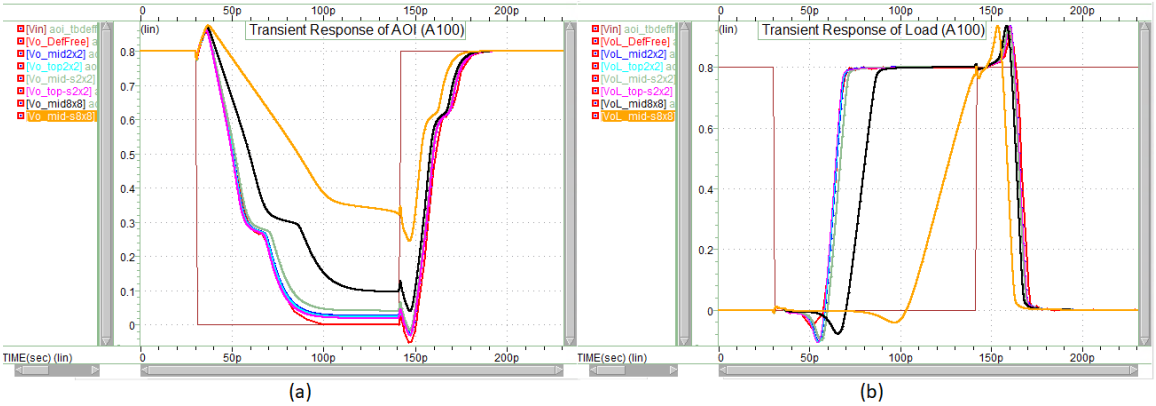

Figure 6.19 Transient response of (a) AOI with pinholes in transistor M2, (b) load inverter, Case 1

The delay characteristics of the defect-free AOI with this input pattern was presented in Table 6.6. By applying the GOS scenarios in transistor M2, the percentage of delay variations are plotted in Figure 6.20.

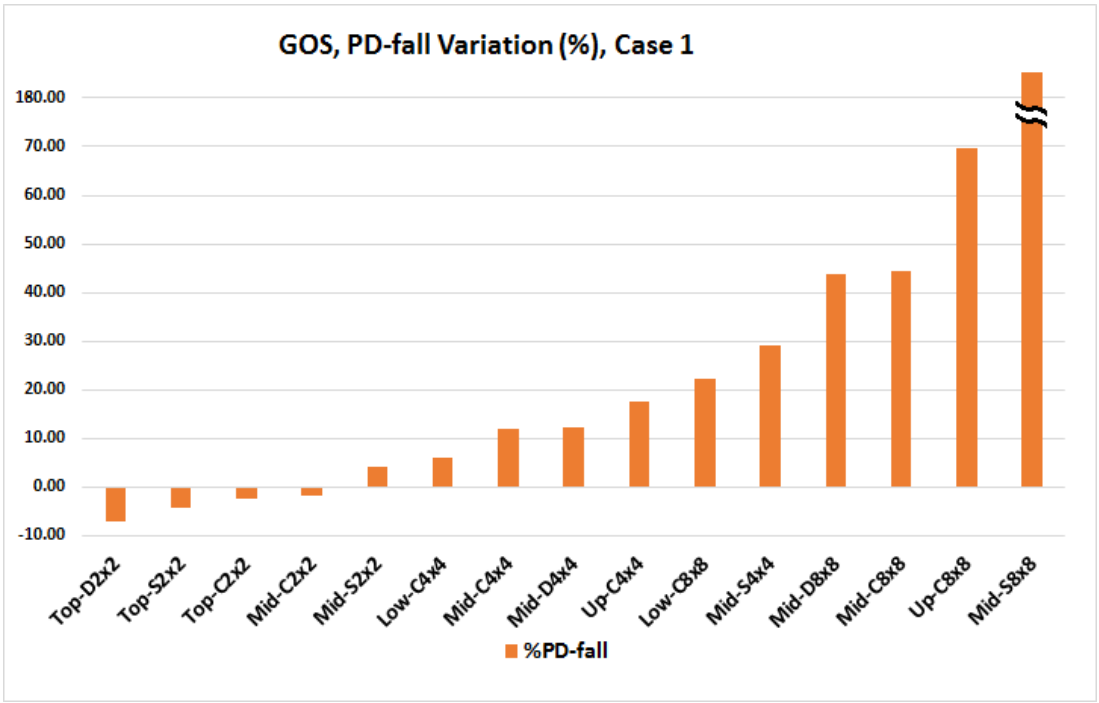

Figure 6.20 Percentage of falling delay variations for different pinholes in transistor M2, Case 1

Based on the above results, the delay fault categories are listed in Table 6.10.

Table 6.10 Delay fault categories for transistor M2, Case 1

Scenario	Fault
Top-S2x2	-
Top-D2x2	-
Top-C2x2	-
Mid-C2x2	-
Mid-S2x2	-
Low-C4x4	-
Mid-C4x4	-
Up-C4x4	-
Mid-S4x4	PD_F1(+)
Mid-D4x4	-
Low-C8x8	PD_F1(+)
Mid-C8x8	PD_F1(+)
Mid-S8x8	PD_F2(+), PD_R1
Up-C8x8	PD_F2(+)
Mid-D8x8	PD_F1(+)

b) Comparison with CAT-based Modeling

The percentage of the delay variations for GS short defects are plotted in Figure 6.21. Although there are resistors that could result in the same falling delay as GOS scenarios, their VTC characteristics is the same as defect-free AOI. This could cause different circuit behavior if the next stage is a pass transistor logic gate.

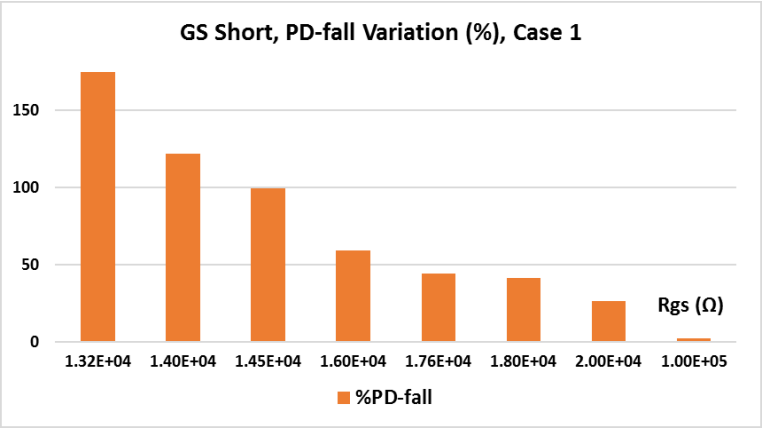

Figure 6.21 Percentage of delay variations in AOI with GS short in transistor M2, Case 1

The GD short resistance directly increases the internal node voltage, V_X , therefore, the AOI output when the pull-down branch is conducting becomes nonzero. However, any GD short resistance that results in the same falling propagation delay as a specific pinhole scenario generates different steady-state output voltage. The output is stuck-at 1 for resistance values smaller than 15.2 kΩ. With resistance values larger than 20.7 kΩ, the defect does not result in logical faults and causes the falling propagation delay to increase. The percentage of the delay variations are plotted in Figure 6.22.

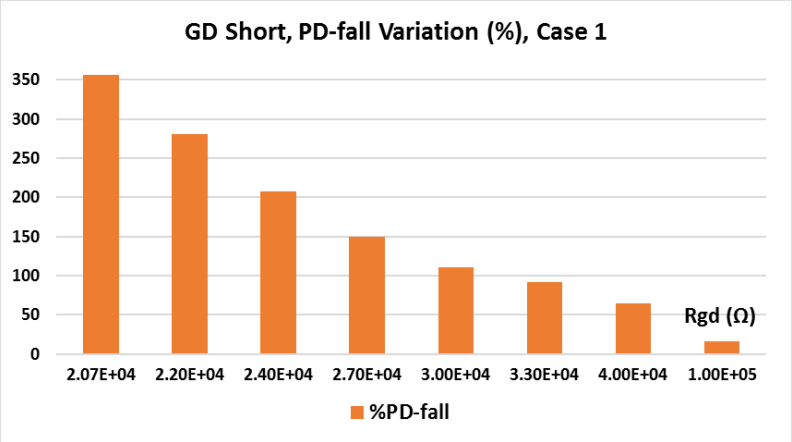

Figure 6.22 Percentage of delay variations in AOI with GD short in transistor M2, Case 1

II. Case 2 (A=1, B:0->1, CD=00)

a) GOS Analysis

The circuit behavior is similar to when the input was applied to the defective transistor M1. When the driver input is low, extensive leakage current passes through the gate of M2 and prevents the driver output to remain at high voltage. The GOS defect results in 50% to 80% reduction in the low noise margin. The defect in transistor M2 increases the supply current I_{dqL} by 6 orders of magnitude and I_{dqH} between 3 to 5 orders.

The transient response of the AOI in presence of GOS defect is similar to Figure 6.19. In the defect-free AOI, applying the input signal to the transistor closer to the output in the pull-down stack results in smaller delays. Despite this, in the defective AOI the falling delay of 8 nm × 8 nm Mid-S scenario, for example, is 40% less than in Case 1. The reason is that in the previous case the defective transistor M2 is always on and extreme GOS leakage increases the internal node voltage and the delays. The delay fault categories are listed in Table 6.11.

Table 6.11 Delay fault categories for transistor M2, Case 2

Scenario	Fault
Top-S2x2	-
Top-D2x2	-
Top-C2x2	-
Mid-C2x2	-
Mid-S2x2	-
Low-C4x4	-
Mid-C4x4	PD_R1
Up-C4x4	PD_R1
Mid-S4x4	PD_R1
Mid-D4x4	PD_R1
Low-C8x8	PD_R1
Mid-C8x8	PD_F1(+), PD_R1
Mid-S8x8	PD_F2(+), PD_R2
Up-C8x8	PD_F1(+), PD_R1
Mid-D8x8	PD_F1(+), PD_R1

b) Comparison with CAT-based Modeling

The GS short prevents input B from reaching the high voltage for resistance values below 8 kΩ. The defect causes the PMOS transistor M6 to conduct and the output is stuck-at 1. Larger resistance values allow M2 to turn on and the PMOS transistor turns off, so the AOI output becomes zero while input B has limited swing. In the transient simulation

of the defective AOI, the falling delay is increased similar to the GS short in transistor M2, Case 1.

The GD short defect in transistor M2 results in output stuck-at 1 with resistance values smaller than 14 kΩ. When input B is low and M2 is off, the conducting PMOS transistor M6 provides a path to charge the load capacitance. As the input voltage increases, and transistor M2 turns on its low drain voltage brings the input voltage down through the GD resistor which holds M6 on. Resistance values larger than 19.8 kΩ increase the internal node voltage, V_X, and the falling propagation delay consequently, same as Figure 6.22.

The results show although GD and GS defect models could generate same falling delays as in 8 nm × 8 nm pinhole scenarios, their output low state have significant difference.

III. Case 3 (A=1, BD:00->11, C=1)

a) GOS Analysis

The input pattern activates both branches to discharge the load capacitance. Since all transistors in the pull-down network conduct when the inputs transition to the high state, it results in smaller equivalent resistance and lower output voltage than Case 2. The GOS defect results in 50% to 65% reduction in the low noise margin. The defect in transistor M2 increases the supply current I_{dqL} by 6 orders of magnitude and I_{dqH} between 3 to 5 orders.

The pinholes have a similar impact on the transient behavior as in Case 1. However, the falling delay increases up to 20%, since both branches in the pull-down network conduct. The 8 nm × 8 nm Mid-S pinhole is the only detectable scenario and is categorized as PD_F1(+) delay fault.

b) Comparison with CAT-based Modeling

The GS short resistor keeps transistor M2 off despite having a high input. Although transistor M2 remains off, the load capacitance is discharged through Branch 2. Therefore, the output transitions to low state with larger falling propagation delay than the defect-free circuit, with an increase of up to 70%. The GD short defect values smaller than 400 Ω cause the AOI output stuck-at 0. As the defect resistance becomes larger than 4 kΩ, the defect increases the internal node voltage and the falling delay will increase up to 30%. Comparing with GOS results, the GS and GD short can only model the 8 nm × 8 nm Mid-S pinhole scenario.

6.3.2.1 Discussion

By approximate delay analysis with Elmore model and recalling the GOS impact on the transistor capacitances, it is observed that the defect in transistor M2 results in smaller gate-source capacitance and larger C_{gd2} at saturation bias. However, C_{gs2} has no impact on the node capacitances in Equations 6.3 and 6.4. Therefore, the delay would increase due to the larger internal node capacitance. The other reason is the non-zero voltage at the source of transistor M1 that increases the threshold voltage of transistor M1 and the falling delay consequently.

6.3.3 Defect in PMOS Transistors M5 and M7

Due to symmetric structure of the AOI circuit, the defect is applied to one of the PMOS branches, Branch 1 including transistors M5 and M7. The load capacitance is charged through any of the branches or both. Hence the input pattern must be applied in a

manner that at least one transistor from each parallel pair conducts. Also, the defective transistor should have 0 input or transition in all cases. Since the AOI circuit has a complementary implementation, the pull-up and pull-down networks are dual [81]. Therefore, a limited number of pinhole scenarios on PMOS transistors are investigated to show that the defect impact is consistent with those in the pull-down network.

- 2 nm × 2 nm pinhole at the fin Top-C, Mid-C
- 8 nm × 8 nm pinhole at the fin Mid-C

The impact of the pinholes on the rising propagation delay are similar to the falling delays in the corresponding cases in the pull-down network. As an example, the simulation results of transistors M5 and M7, with the input applied to the defective transistor, are included in Appendix C .

6.4 3:1 Multiplexer (MUX)

The 3:1 MUX schematic diagram is depicted in Figure 6.23. Each 2:1 MUX is realized with a 4-input AOI.

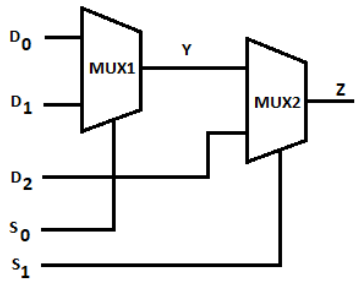

Figure 6.23 Schematic diagram of 3:1 MUX

The simplified output equations of MUX1 and MUX2 are presented in Equations 6.5 and 6.6,

$$Y' = S_0 D_1 + S_0' D_0 \quad (6.5)$$

$$Z' = S_1 D_2' + S_1' Y'. \quad (6.6)$$

The 3:1 MUX is a three-level logic circuit shown in Figure 6.24. In the AOI investigation, it was shown that several GOS defect scenarios in transistor M1 decrease the falling delay. On the other hand, transistor M9 in MUX is in the longest delay path. Since GS and GD shorts can represent the scenarios that increase the falling delay, transistor M9 is selected for the investigation.

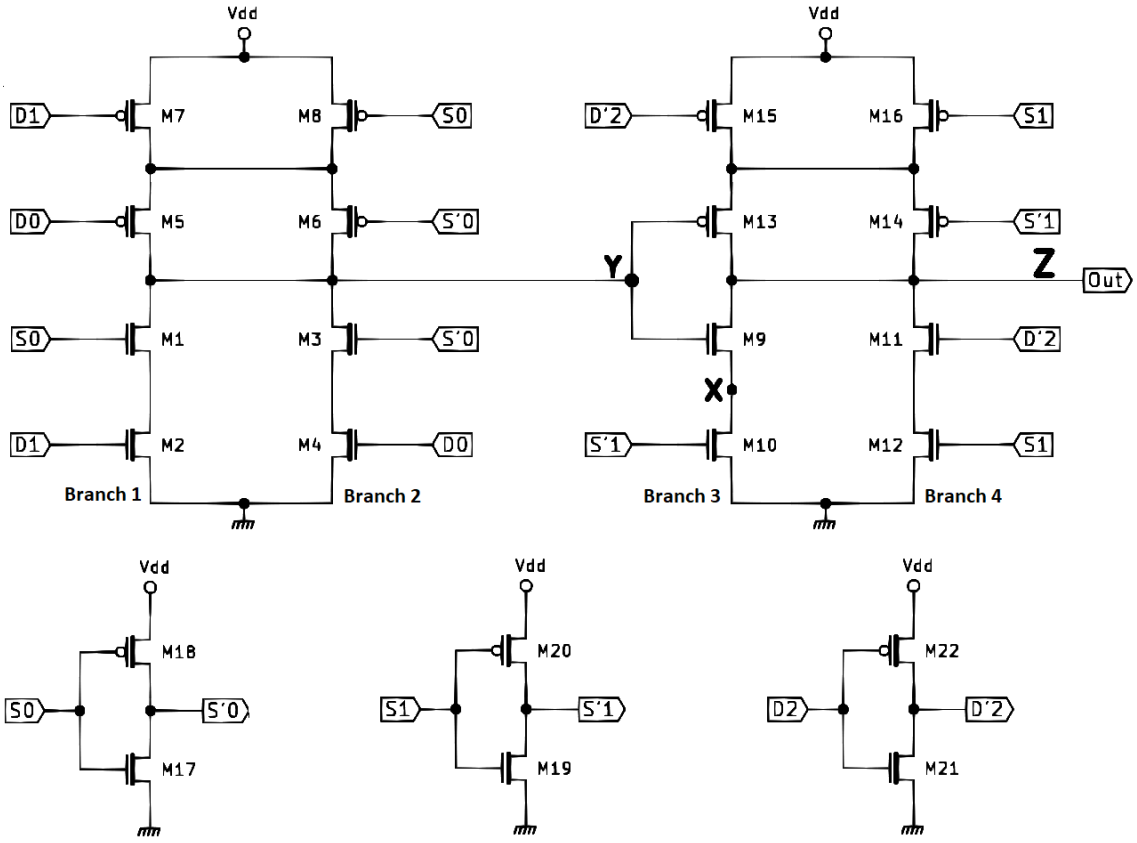

Figure 6.24 MUX consisting of two 4-input AOIs and the required inverters

Considering a GD short in transistor M9, for small resistance values Z is stuck-at Y. Therefore, the input combination that activates the faulty behavior should generate a logic 1 at node Y. If the circuit was defect-free this would have caused the discharge of load capacitance though Branch 3 and the output became zero. However, the GD short defect provides a short path from Y to out and makes it logic 1. The other inputs should be set in

a manner that results in a zero output in the defect-free circuit. The input combinations that lead to a logic 1 at node Y are presented in Table 6.12.

In the first two rows, since Branch 4 has no effect on the output when S_1 is zero, the input D_2' is considered as a don't care. In the last row, the discharge path is through Branch 4. However, the GD short brings the output voltage to the high state.

Table 6.12 The input patterns that detect GD short defect in M9

S_1	S_0	D_2'	D_1	D_0	Z
0	0	×	×	0	0
0	1	×	0	×	0
1	1	1	0	×	0

The GS short in transistor M9 keeps the transistor in the off state. Therefore, to propagate the defect impact to the output, the transistors in Branch 4 must be off and M10 should turn on, same as in the first two rows of the above table. Also, node Y should be at logic 1. Since M10 is on, the GS defect provides a short path to node Y and reduces its voltage. Therefore, the charging path through M13 and M16 to the supply keeps the output at the high state. While, in the defect-free circuit, logic 1 at node Y activates the discharge path through Branch 3 and makes the output zero. The analysis shows that the first two input patterns in Table 6.12 are also used to detect the GS short defect in M9.

In the following section, the defect models for different pinhole scenarios are applied to transistor M9 to investigate the MUX behavior and compare it to the CAT-based GD and GS short defects.

6.4.1 GOS Defect in Transistor M9

In Table 6.12, any of the zero inputs can be considered as the input signal with falling transition. One of the possible scenarios in the first row is having a falling transition in input S_1. The don't care inputs are set to zero. This input combination is listed as Case 1 in Table 6.13. In the second row of Table 6.12, it is assumed that the falling transition is applied to input D_1, and the other inputs are set to the values listed in Case 2. In the last row of Table 6.12, any of the inputs can have a rising transition, from which D_2' is chosen in Case 3. Recalling the defective AOI simulations, Cases 1 and 2 in Table 6.13 cause discharging through Branch 1 with the input signal applied to one of the NMOS transistors and logic 1 to the other. Case 3 results in discharging through Branch 2. The scenario of discharging through both branches cannot be realized in this AOI due to its logic function presented in Equation 6.6.

Table 6.13 The input combinations considered for detecting the defect in M9

Case	S_1	S_0	D_2'	D_1	D_0	Z
1	↓	0	0	0	0	↓
2	0	1	1	↓	1	↓
3	1	1	↑	0	0	↓

Since the propagation delays are pattern-dependent, the rising and falling delays of the defect-free MUX are listed in Table 6.14 for further comparison with the defective circuit behavior.

Table 6.14 The propagation delays of the defect-free MUX for each input combination

Scenario	PD-rise (ps)	PD-fall (ps)	T-rise (ps)	T-fall (ps)
Case 1	29.64	35.01	28.56	37.70
Case 2	36.76	37.17	29.95	39.67
Case 3	30.37	32.91	28.30	40.89

I. Case 1 ($\mathbf{S_1: 1 \rightarrow 0, S_0 D_2' D_1 D_0 = 0000}$)

a) GOS Analysis

The GOS defect results in 25% to 72% reduction in the low noise margin and 58% to 85% reduction in the high noise margin. The pinholes increase the high supply current I_{dqH} about 5 orders of magnitude.

The transient response of the MUX with GOS defect is similar to the AOI response with GOS defect in transistor M1, Case 1. The pinholes on the fin top, middle sidewall, and close to source cause reduction in the falling propagation delay. The percentage of delay variations are similar to the AOI delay variations in Figure 6.8, but with smaller range from -20% to 42%. The delay fault categories are listed in Table 6.15.

Table 6.15 Delay fault categories for MUX, Case 1

Scenario	Fault
Top-S2x2	PD_F1
Top-D2x2	-
Top-C2x2	-
Mid-C2x2	-
Mid-S2x2	PD_R1
Low-C4x4	PD_R1
Mid-C4x4	PD_R1
Up-C4x4	PD_R1
Mid-S4x4	PD_R1
Mid-D4x4	PD_R1
Low-C8x8	PD_R1
Mid-C8x8	PD_R1
Mid-S8x8	PD_R1
Up-C8x8	PD_R1
Mid-D8x8	PD_F1(+)

b) Comparison with CAT-based Modeling

The input pattern makes transistor M10 turn on while node Y is at high state. The GS short with resistance values smaller than 14.0 kΩ causes a hard defect that keeps transistor

M9 off and the output stuck-at 1. While, resistance values larger than 22 kΩ increase the falling propagation delay same as Figure 6.12. The GD short defect results in the output stuck-at 1 fault for resistance values smaller than 14 kΩ, since it provides a short path from node Y to the output. For defect resistance larger than 20.5 kΩ, the circuit behavior is modeled by delay fault.

The results show that a GS or GD short resistance can be found to generate the same falling delay variation as the 8 nm × 8 nm Mid-D pinhole scenario. The other pinhole scenarios that decrease the falling delay cannot be represented by a GS or GD short model.

II. Case 2 ($\mathbf{D_1: 1 \rightarrow 0, S_1\, S_0 D_2' D_0 = 0111}$)

a) GOS Analysis

The GOS defect results in 55% to 64% reduction in the low noise margin. The pinholes increase the high supply current I_{dqH} up to 5 orders of magnitude.

In this case, the input S_0 is at high state and a falling transition in D_1 provides a path to charge node Y to the supply voltage. The toggling signal is applied to the defective transistor, closer to the output. The transient response is similar to the previous case. The percentage of delay variations for all pinhole scenarios are plotted in Figure 6.25 and the delay fault categories are listed in Table 6.16.

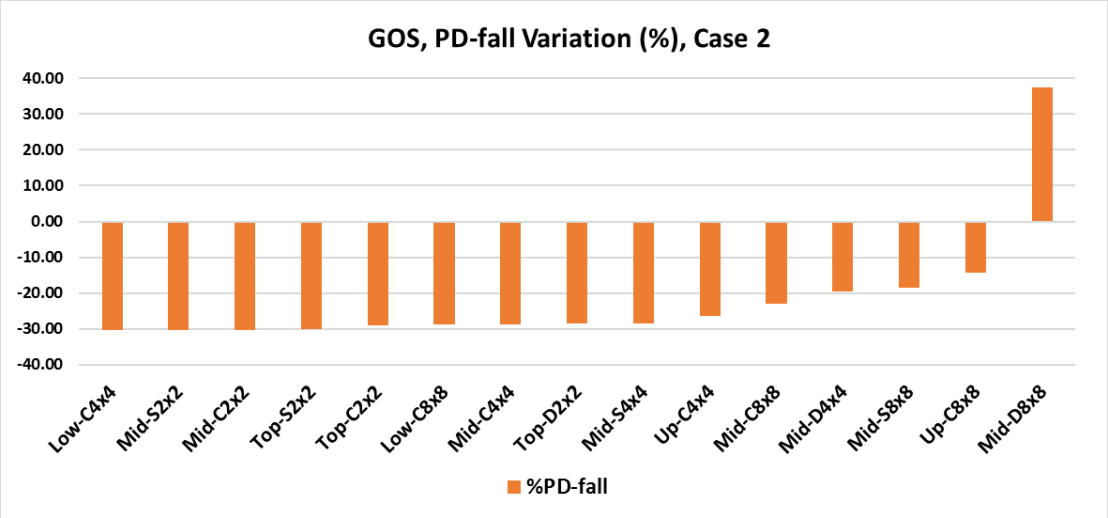

Figure 6.25 Percentage of the falling delay variations in the MUX with GOS defect, Case 2

Table 6.16 Delay fault categories for MUX, Case 2

Scenario	Fault
Top-S2x2	PD_F1
Top-D2x2	PD_F1
Top-C2x2	PD_F1
Mid-C2x2	PD_F1
Mid-S2x2	PD_F1, PD_R1
Low-C4x4	PD_F1, PD_R1
Mid-C4x4	PD_F1, PD_R1
Up-C4x4	PD_F1, PD_R1
Mid-S4x4	PD_F1, PD_R1
Mid-D4x4	PD_R1
Low-C8x8	PD_F1, PD_R1
Mid-C8x8	PD_F1, PD_R1
Mid-S8x8	PD_R1
Up-C8x8	PD_R1
Mid-D8x8	PD_F1(+), PD_R2

b) Comparison with CAT-based Modeling

The GS short with resistance values smaller than 16.0 kΩ keep transistor M9 off and the output is stuck at 1. As the GS resistance increases, transistor M9 turns on. However, the high voltage at the internal node X results in a larger threshold voltage in M9 and

increases the falling propagation delay at the output node, similar to Figure 6.12. The GD short with resistance values smaller than 11.0 kΩ in transistor M9 provides a path from the input signal at node Y to the output node, that is modeled by output stuck-at input. As the GD short resistance increases above 23.0 kΩ, it is represented by the delay fault. The GD short reduces the output swing in both signal level, and has different behavior comparing to the GOS defect.

Except for the pinhole scenarios that increase the falling delay which can be replaced with a GS short resistor (although with different rising delay behavior), the other pinholes have different impact on the MUX transient behavior and cannot be represented by CAT-based models.

III. Case 3 ($\mathbf{D_2'}: \mathbf{0 \rightarrow 1}, \mathbf{S_1 S_0 D_1 D_0 = 1100}$)

a) GOS Analysis

The GOS defect results in 26% to 86% reduction in the low noise margin and 10% to 63% reduction in the high noise margin. The pinholes increase the high supply current I_{dqH} up to 5 orders of magnitude.

The input combination is such that node Y is kept at high voltage, and the changing signal is applied to transistor M11. Therefore, the discharging occurs through Branch 2, while a logic 1 is applied to the defective transistor. Since Branch 2 is conducting in the discharge cycle, the defect in M9 has no impact on the falling transition. However, the leakage through the gate-oxide-short prevents the output to reach zero voltage and affects the rising delay. The percentage of rising delay variations are shown in Figure 6.26.

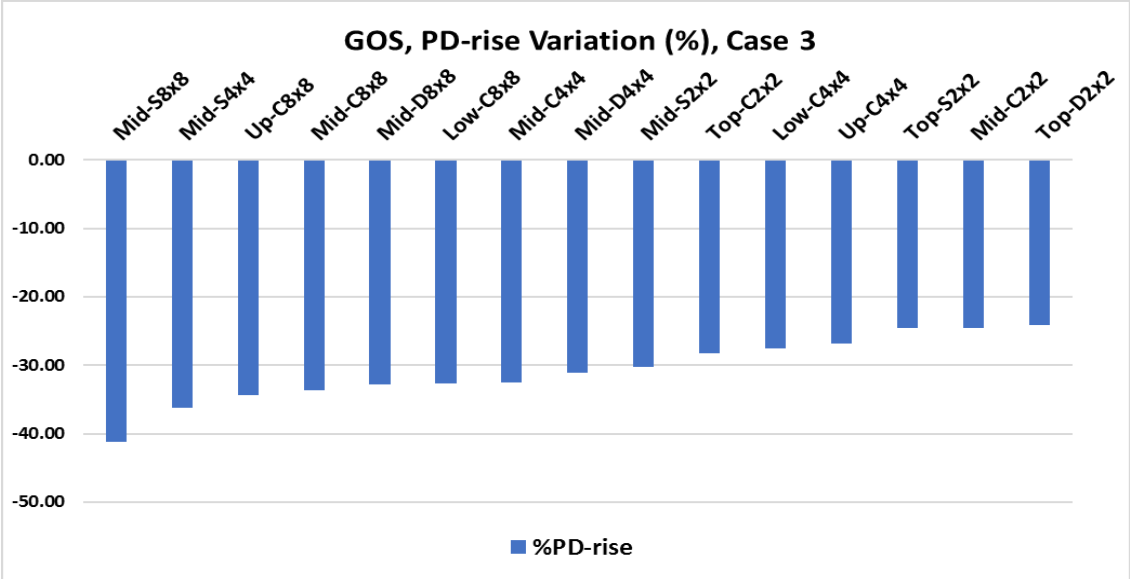

Figure 6.26 Percentage of the rising delay variations in the MUX with GOS defects, Case 3

Based on the above results, all scenarios are categorized as PD_R1.

b) Comparison with CAT-based Modeling

The GS short defect increases the voltage at the internal node and M9 acts as an active load for the discharging branch. The percentage of rising delay variation is similar to Figure 6.17. The results show that the GS short resistor can model 2 nm × 2 nm Top and Mid-C pinholes, and 4 nm × 4 nm Low-C and Up-C scenarios. The GD short defect provides a low resistive path from node Y to the output. For resistance values smaller than 10.0 kΩ, the output is stuck-at 1, while for larger defect resistance values the falling delay increases. Therefore, the circuit behavior is different from the GOS-impacted MUX.

6.4.1.1 Discussion

The GOS defect characterization in the MUX circuit shows that any input combination that cause a rising transition at node Y, for example Case 2, results in significant reduction in the output falling delay. Therefore, that input pattern can detect

most of the pinhole scenarios in transistor M9. When the input signal is applied to the defect-free transistor M10, there is no significant change in the falling delays. Also, when the input signal activates the discharge path through Branch 4 the falling delay is not affected by the defect and a second input transition should be applied to evaluate the output rising delays.

6.5 Summary

In this chapter, the input patterns that can detect different pinhole scenarios are discussed and the impact of GOS defect on Inverter and AOI behavior are studied. The DC analysis of the GOS defect model shows that there are significant changes in the noise margins and I_{dq} currents. It is shown that CAT-based defect models cannot represent the GOS defect in most cases. The GOS defect is also introduced to the selected transistor in a 3:1 MUX. It is shown that despite larger delays on the signal path compared to a single AOI, very few pinhole scenarios can be represented by the CAT-based models.

Based on the simulation results, the comparison of the GOS-impacted CUT with the CAT-based defect models are summarized in the next chapter.

Chapter 7: Results and Discussion

This chapter presents an analysis of the simulation results from the previous chapter to determine the effectiveness of the proposed GOS defect models in fault modeling of complex logic gates in comparison with the existing CAT-based models.

7.1 Transient Results

The simulation results of the applied defects in the inverter, AOI, and MUX are summarized in Table 7.1 to Table 7.4. The test patterns that can detect the GD and GS defects are listed in the first column. The pinhole scenarios that each pattern can detect, as well as those that are missed are listed in separate rows for GOS, GS and GD short. The "Comments" column provides a short description of the defect impact on the delays. If the percentage of delay variations exceed the specific threshold values listed in Table 4.4, the defect is considered as detectable. The important arguments about simulation analysis for each transistor are listed after the corresponding table.

Table 7.1 Simulation results of the INV, transistor M1

Input	Defect Model	Detectable Defect Type[1]		Comments
0→1	GOS	Yes	TS2, TC2, TD2, MC2, MS2, LC4, MS4, MC4, MD4, UC4, LC8, MC8, MS8, UC8	These scenarios except MD8 decrease the falling delay.
		No	MD8	The delay change is not significant.
	GS-CAT	Yes	---	The GS increases the falling delay. The GD has no significant impact in the falling delays.
		No	All	
	GD-CAT	Yes	---	
		No	All	

[1] In this table, the short names of the pinholes are used (e.g. MC4 instead of 4 nm × 4 nm Mid-C).

- The GS short in transistor M1 increases the falling delay and cannot represent any GOS scenarios. The GD short reduces the output low and high voltage swings and the reduction in falling delay is not significant to model any GOS scenarios.

Table 7.2 Simulation results of the AOI, transistor M1

Pattern ABCD	Defect Model	Detectable Defect Type[2]		Comments
Case 1 0100→1100	GOS	Yes	TS2, TC2, TD2, MC2, MS2, LC4, MS4, MC4, MD4, UC4, LC8, MC8, MD8	These scenarios except MD8 decrease the falling delay.
		No	UC8, MS8	The delay changes are not significant.
	GS-CAT	Yes	MD8	GS and GD can model MD8 which causes significant increase in the falling delay.
		No	All others	
	GD-CAT	Yes	MD8	
		No	All others	
Case 2 1000→1100	GOS	Yes	TS2, TC2, TD2, MC2, MS2, LC4, MS4, MC4, UC4, LC8, MD8	These scenarios except MD8 decrease the falling delay.
		No	MD4, MS8, MC8, UC8	The delay changes are not significant.
	GS-CAT	Yes	MD8	GS and GD can model MD8 which causes significant increase in the falling delay.
		No	All others	
	GD-CAT	Yes	MD8	
		No	All others	
Case 3 1001→1011 1011→1001	GOS	Yes	All	No change in falling delays. Additional pattern is applied to use rising delays for detection.
		No	---	
	GS-CAT	Yes	TS2, TC2, TD2, MC2, LC4, LC8, MD8	No change in falling delays. Additional pattern is applied to use rising delays for detection.
		No	All others	
	GD-CAT	Yes	---	Unlike GOS, GD increases the falling delay and cannot model any scenario.
		No	All	
Case 4 1001→1111	GOS	Yes	---	No significant reduction in the falling delays.
		No	All	
	GS-CAT	Yes	---	Unlike GOS, GS and GD increase the falling delay and cannot model any scenario.
		No	All	
	GD-CAT	Yes	---	
		No	All	

[2] In this table, the short names of the pinholes are used (e.g. MC4 instead of 4 nm × 4 nm Mid-C).

- The GS and GD shorts in transistor M1 increase the falling delay. The former increases the voltage of the node X (M1 source) which in turn increases the transistor threshold voltage. The latter provides a resistive path from the driver's pull-up network to the AOI output.
- When the input signal is applied to the GOS-impacted transistor, the delay variations are more significant than those cases in which the transitioning signal is applied to the defect-free transistor. This is due to the GOS impact on the transistor capacitances that makes the transistor turn on faster.
- In Case 3, the non-defective branch discharges the load capacitance; therefore, the GOS in transistor M1 does not affect the falling delay. Using an additional input pattern, the GS short can model a few scenarios. Since GD short increases the falling delay, it cannot represent the circuit behavior with the GOS defect.

Table 7.3 Simulation results of the AOI, transistor M2

Pattern ABCD	Defect Model	Detectable Defect Type[3]		Comments
Case 1 0100→1100	GOS	Yes	MS4, LC8, MS8, MC8, MD8, UC8	These scenarios increase the falling delay.
		No	TS2, TC2, TD2, MC2, MS2, LC4, MC4, MD4, UC4	The delay change is not significant.
	GS-CAT	Yes	MS4, LC8, MS8, MC8, MD8, UC8	GS and GD can model scenarios that cause significant increase in the falling delay. However, the low signal level and the rising delays are different from GOS.
		No	TS2, TC2, TD2, MC2, MS2, LC4, MC4, MD4, UC4	
	GD-CAT	Yes	MS4, LC8, MS8, MC8, MD8, UC8	
		No	TS2, TC2, TD2, MC2, MS2, LC4, MC4, MD4, UC4	
Case 2 1000→1100	GOS	Yes	MS8, MC8, MD8, UC8	These scenarios increase the falling delay.
		No	TS2, TC2, TD2, MC2, MS2, LC4, MS4, MC4, MD4, UC4, LC8	The delay change is not significant.
	GS-CAT	Yes	MS8, MC8, MD8, UC8	GS and GD can model scenarios that cause significant increase in the falling delay. However, their output low signal level is different from GOS.
		No	TS2, TC2, TD2, MC2, MS2, LC4, MS4, MC4, MD4, UC4, LC8	
	GD-CAT	Yes	MS8, MC8, MD8, UC8	
		No	TS2, TC2, TD2, MC2, MS2, LC4, MS4, MC4, MD4, UC4, LC8	
Case 3 1001→1111	GOS	Yes	MS8	Except MS8, the falling delay changes are not significant.
		No	All others	
	GS-CAT	Yes	MS8	
		No	All others	
	GD-CAT	Yes	MS8	
		No	All others	

- All GOS scenarios in transistor M2 increase the capacitance of internal node X. Therefore, the falling delays are increased, and both GS and GD shorts can model some GOS scenarios.

[3] In this table, the short names of the pinholes are used (e.g. MC4 instead of 4 nm × 4 nm Mid-C).

Table 7.4 Simulation results of the MUX, transistor M9

Pattern $S_1S_0D_2'D_1D_0$	Defect Model	Detectable Defect Type[4]		Comments
Case 1 10000→00000	GOS	Yes	TS2, MD8	Except TS2 and MD8, the falling delay changes are not significant.
		No	All others	
	GS-CAT	Yes	MD8	GS and GD can model MD8 which causes significant increase in the falling delay.
		No	All others	
	GD-CAT	Yes	MD8	
		No	All others	
Case 2 01111→01101	GOS	Yes	TS2, TC2, TD2, MC2, MS2, LC4, MS4, MC4, UC4, LC8, MC8, MD8	These scenarios except MD8 decrease the falling delay.
		No	MD4, MS8, UC8	The delay changes are not significant.
	GS-CAT	Yes	MD8	GS can model MD8 which causes significant increase in the falling delay.
		No	All others	
	GD-CAT	Yes	-	GD reduces the swing in both signal levels and cannot model any GOS scenario.
		No	All	
Case 3 11000→11100 11100→11000	GOS	Yes	All	No change in falling delays. Additional pattern is applied to use rising delays for detection.
		No	-	
	GS-CAT	Yes	TS2, TC2, TD2, MC2, LC4, UC4	No change in falling delays. Additional pattern is applied to use rising delays for detection.
		No	All others	
	GD-CAT	Yes	-	Unlike GOS, GD increases the falling delay and cannot model any scenario.
		No	All	

- Since two AOI stages are cascaded in the MUX, the overall delays are increased. Therefore, the impact of GOS on the delay variations is reduced.
- Same as transistor M1 in AOI, the input pattern that applies the toggling signal to the GOS-impacted transistor, detects more scenarios than the patterns in which the transitioning signal is applied to the defect-free transistor.
- Similar to transistor M1 in AOI, the GOS in transistor M9 does not affect the falling delay in Case 3. Using an additional input pattern, the GS short can model a few

[4] In this table, the short names of the pinholes are used (e.g. MC4 instead of 4 nm × 4 nm Mid-C).

pinhole scenarios. Since the GD short increases the falling delay, it cannot represent the GOS defect.

7.2 DC Results

The GS and GD short resistors cause a linear and non-realistic change in the circuit current that cannot represent the GOS defect.

In contrast, the DC simulation results of various GOS defect scenarios show that when the input signal is applied to the defective transistor, there is at least 3 order of magnitude increase in I_{DDQ} currents as listed in Table 7.5. No significant change in the current is represented by NC in the table.

Table 7.5 GOS impact on I_{DDQ}

		$\times 10^{\wedge}$	
		I_{dqL}	I_{dqH}
INV		6	3
AOI	**M1, Case 1**	**6**	**3**
	M1, Case 2	NC	NC
	M1, Case 3	NC	NC
	M1, Case 4	NC	NC
	M2, Case 1	3	NC
	M2, Case 2	**6**	**3**
	M2, Case 3	6	3
MUX	**M9, Case 1**	NC	5
	M9, Case 2	**NC**	**5**
	M9, Case 3	NC	5

The GS and GD short impact on the VTC do not result in the expected voltage transfer characteristics as GOS defect (except a few cases).

In contrast, the GOS defect has consistent behavior in all scenarios and decreases the low noise margin, Table 7.6. The findings about noise margin and I_{DDQ} could be used to improve the GOS fault model.

Table 7.6 GOS impact on the noise margins

		%	
		△NM-L	**△NM-H**
INV		-54 to -61	38 to 45
AOI	**M1, Case 1**	**-55 to -60**	**15 to 40**
	M1, Case 2	-20 to -80	-50 to -80
	M1, Case 3	-22 to -86	-23 to -64
	M1, Case 4	-15 to -67	-29 to -85
	M2, Case 1	-3 to -23	-3 to -29
	M2, Case 2	**-50 to -80**	**45 to 87**
	M2, Case 3	-43 to -87	34 to 64
MUX	**M9, Case 1**	-25 to -72	-58 to -85
	M9, Case 2	**-55 to -64**	**37 to 45**
	M9, Case 3	-26 to -86	-10 to -63

7.3 Discussion

Our analysis has shown that the GOS defect cannot be modeled by stuck-at fault, and its effect on the circuit behavior is modeled with parametric in the AOI, the CAT-based models that might be able to represent the GOS impact on transistors M1 and M2 are GS short in M1 and GD short in M2, as they affect the voltage of internal node X. The DC analysis shows that when the toggling signal is applied to the GOS-impacted transistor, the VTC is shifted to the left which results in the reduction of low noise margin and increase in the high noise margin. This is different from the impacts of GS short on M1 and GD short on M2, as stated in Table 7.7. It is also observed that the GOS defect affects the I_{DDQ} currents different from the CAT-based models.

Table 7.7 Comparison of GOS vs CAT-based models' impact on DC behavior

Transistor	Defect Model	$I_{dqL} \times 10^{\wedge}$	$I_{dqH} \times 10^{\wedge}$	△NM-L (%)	△NM-H (%)
M1 (INV)	**GOS**	**6**	**3**	**-54 to -61**	**38 to 45**
	GS	6	NC	NC	NC
M1 (AOI)	**GOS**	**6**	**3**	**-55 to -60**	**15 to 40**
	GS	6	NC	0 to -43	0 to -38
M2 (AOI)	**GOS**	**6**	**3**	**-50 to -80**	**45 to 87**
	GD	6	6	0 to -64	0 to -82
M9 (MUX)	**GOS**	**NC**	**5**	**-55 to -64**	**37 to 45**
	GS	NC	5	0 to -22	0 to -19

From the transient analysis, our GOS defect models result in fault models that by using them more pinhole scenarios are detected with a smaller number of test patterns, as shown in Table 7.8.

Table 7.8 Comparison of GOS vs CAT-based models' impact on transient behavior

Transistor	Number of Detectable Defects		Percentage of Defect Detection	
	CAT-based Models	GOS Model	CAT-based Models	GOS Model
M1 (INV)	0	14	0	93.3
M1 (AOI)	7	13	46.7	86.7
M2 (AOI)	6	6	40	40
M9 (MUX)	6	12	40	80

For example, from a total of 15 pinhole scenarios in transistor M1 of the AOI, 13 scenarios reduce the falling delay below the threshold by applying 2 consecutive input vectors. While the CAT-based models cannot represent most of the pinhole defect scenarios, the GOS model results in 40% improvement in detecting defects in the upper transistor of the pull-down networks.

The number of required consecutive patterns to achieve the detection rates in the above table are listed in Table 7.9. The CAT-based models can only represent 7 pinhole scenarios by applying 3 consecutive input patterns and considering the rising delay.

Table 7.9 Number of patterns required for defect detection

Transistor	Number of Required Patterns	
	CAT-based Models	GOS Model
M1	3	2
M2	2	2
M9	3	2

In Table 7.2 and Table 7.4, using two input patterns GS and GD shorts can model one pinhole defect (Case 1 and Case 2). In Case 3, by applying additional input pattern more pinhole scenarios can be detected. While, GOS model achieves the defect detection rates in Table 7.8 using two input patterns, and all 15 pinhole scenarios can be detected by applying the 3rd input pattern.

We have shown that in a stack of transistors, that the internal nodes capacitances are important in the propagation delays, simple resistive components cannot accurately model the GOS behavior.

Chapter 8: Conclusion and Future Work

The complexity of FinFET manufacturing process has caused challenges in reliable device testing. Gate oxide short (GOS) is one of the dominant defects that has significant impact on circuit reliability. Defect-based fault modeling is the main approach in CAT methodology which extracts the parasitic data for each library cell. The layout data provides accurate information about the potential bridging and open defects inside the cell. Although the CAT methodology is extended to FinFET testing, its 2D layout-based approach cannot accurately capture the defect models in the 3D transistor structure. The existing methods suffer from applying simplistic resistive defect models for a complex defect such as GOS.

This study is performed, for the first time, at the level of device structure, rather than layout. We have based our analysis on 3D bulk FinFET template provided by Synopsys. This work has investigated the impact of GOS defect on the FinFET characteristics for different locations and sizes of the pinhole on various fin shapes. It is shown that small pinhole doubles the saturation current. As the pinhole size becomes larger, the saturation current degrades due to the large number of charge carriers (holes) injected into the channel. Also, the magnitude of negative drain current at linear bias significantly increases in presence of larger pinhole. Pinhole being closer to the source region on the upper sidewall of the fin results in the highest leakage current. The impact of defect on the transistor capacitances are also captured through small signal analysis at various drain voltages according to the required precision.

The comprehensive investigation of GOS defect has led to the development of a GOS defect model that accurately reflects the behavior of the defective FinFET. To apply the

GOS defect model in circuit-level simulations, the differences in the defective transistor currents and capacitances with those of defect-free device are extracted in appropriate voltage steps to build Verilog-A modules for 15 pinhole scenarios. The other important step was extracting the HSPICE model parameters for the FinFET template based on the BSIM-CMG local extraction procedure. These models are used in the circuit-level simulation of complex logic gates including 4-input AOI and 3:1 MUX.

To compare the proposed fault modeling methodology with the existing approach, the defect models used in the CAT methodology are applied to the circuits under test, as well as the extracted GOS models. Among the transistor defect models used in the CAT methodology, the GD short and GS short are the defects of interest that might be able to represent the GOS behavior. Comprehensive DC and transient analysis are performed using these models with the specific transistors in one of the pull-down and pull-up branches of the CUT due to its symmetric structure. The investigation of GOS impact on the supply current can lead to the improvement of GOS fault models in I_{DDQ} tests. The application of noise margin in test methodologies could be a subject of future research.

In the AOI circuit, the defect is first introduced to the upper transistor in the pull-down stack. It is shown that the two input patterns that activate the defective branch and detect GD or GS defects can only detect one pinhole scenario. This is because most GOS scenarios on this transistor increase the saturation current and reduce the internal node capacitance; and hence, the falling delay. While the GS and GD models result in increased falling delays. The test pattern that activates the defect-free branch requires 3 consecutive input patterns to consider the rising delay for fault detection. In this case, the proposed GOS model detects all pinhole scenarios, and has 50% to 60% improvement compared to

the CAT-based models. Applying additional patterns, as performed in the N-detect approach, that activate both branches in the pull-down network does not improve the defect detection rate.

In the AOI circuit, the defect injection in the lower transistor of the pull-down branch shows that the GS or GD short defects result in the same delay behavior as the GOS model. This is due to the fact that the GOS on the lower transistor increases the internal node voltage and hence, the falling delay, same as the former models.

Finally, the improvement of defect detection is validated by applying the GOS and CAT-based models to a specific transistor in the MUX. Despite larger circuit delays, with the two input patterns that activate the defective branch, the GS or GD models can represent only one pinhole scenario, while 12 scenarios are detectable with the GOS model. By applying additional input and using the rising delay for fault detection, the CAT-based models can only represent 7 pinhole scenarios, while all 15 scenarios are detectable with the GOS model.

Our proposed defect and fault modeling approach contributes to the state-of-the-art CAT methodology by detecting more FinFET-related defects with smaller number of test patterns; and hence, reducing the test time and cost. This methodology could also improve the ability of identifying the location of defect in physical failure analysis. Moreover, our proposed methodology is not limited to standard library cells, as the defect models are developed at the device 3D structure instead of layout. Although the superiority of our proposed approach is shown in complementary MOSFET logic, it could also be applicable to other logic family circuits such as PTL, and other circuit designs such as SRAM. Applying the accurate defect models to any transistor in a design module could be a subject

of future research. Moreover, our proposed approach is applicable to double-fault detection algorithms that analyze the propagation path of single faults to find other single-faults that block their propagation path [81]. The fault pairs that mask each other's propagation paths are considered as uncovered fault pairs, hence, the complexity of finding undetected faults are significantly reduced [81].

8.1 Contributions

Major contributions of this thesis to the area of fault modeling in FinFETs are summarized in the following.

- Proposed a novel methodology for defect and fault modeling in FinFETs
    - Based the model on 3D structure of FinFET
    - Analyzed GOS impact on FinFET behavior
    - Developed accurate GOS defect models
- Applied and evaluated the models, and compared with the state-of-the-art
    - Developed fault models for complex logic gates (AOI and MUX)
    - Compared with CAT and demonstrated superiority

8.2 Future Work

The work presented in this thesis covers the main research objectives. However, the areas that need further research to enhance this work are the following:

- Extending this research to multiple-fin FinFETs.
- Using the TCAD-to-HSPICE extraction tools, to improve the model parameters for various temperatures.

- Investigating extracted models for process and voltage variations.
- Investigating more defect scenarios for P-channel FinFETs.
- Analysis of run time and test time.
- Development of comprehensive fault models.
- Developing 4D models for finding the defect locations.

8.3 Publications

- Dibaj R., Al-Khalili D., Shams M., "Comprehensive investigation of gate oxide short in FinFETs," *Proc. IEEE VLSI Test Symp. (VTS),* pp. 1-6, 2017.
- Dibaj, R., Al-Khalili, D., & Shams, M., "Gate Oxide Short Defect Model in FinFETs," *Journal of Electronic Testing*, vol. 34, no. 3, pp. 351-362, 2018.

Appendices

Appendix A PFinFET Characteristic Curves

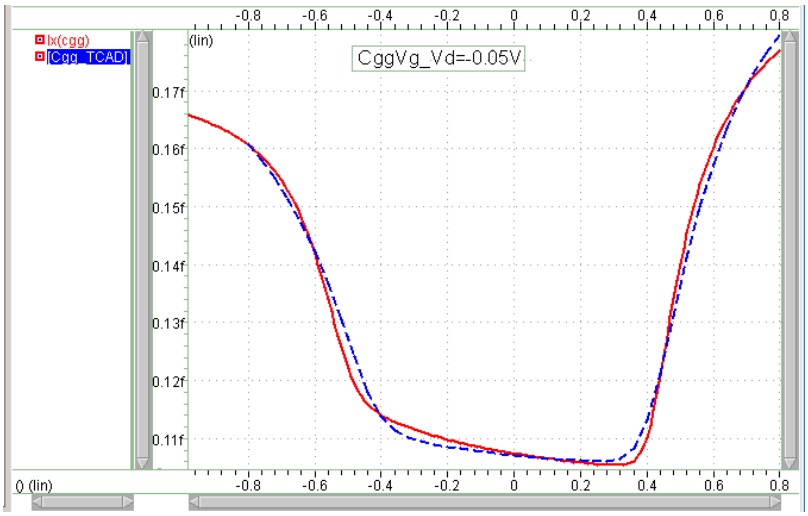

Figure A.1 Optimized $C_{gg}V_{gs}$ characteristics at $V_d = -0.05$ V

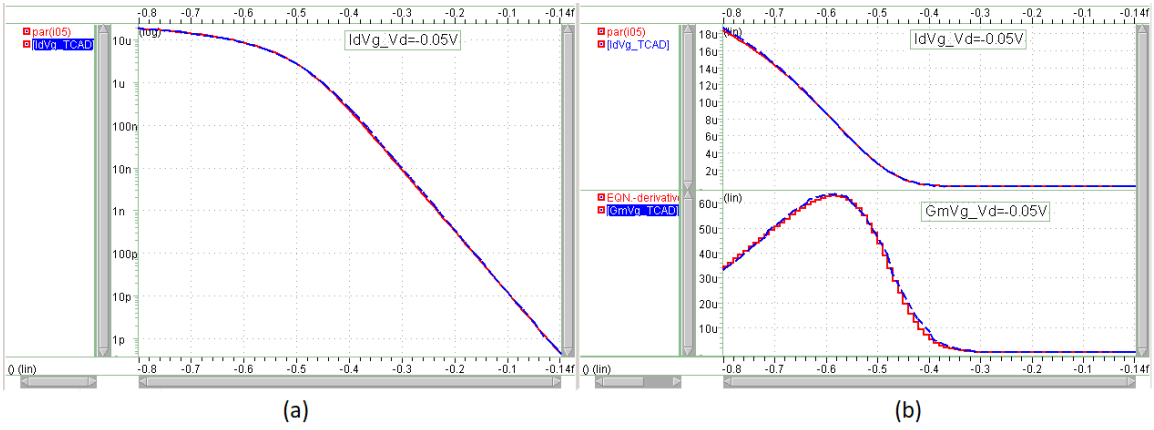

Figure A.2 Optimized (a) I_dV_g, (b) G_mV_g graphs at $V_d = -0.05$ V

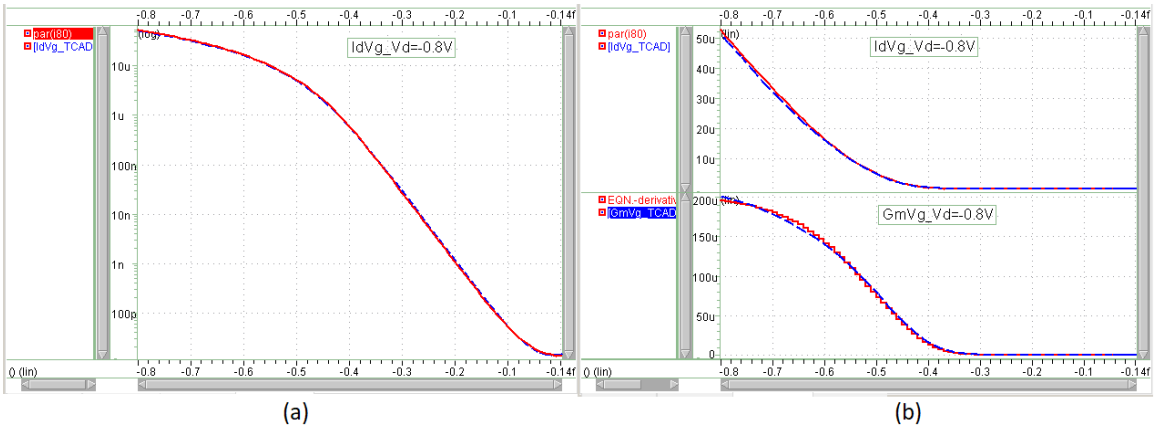

Figure A.3 Optimized (a) I_dV_g, (b) G_mV_g graphs at $V_d = -0.8$ V

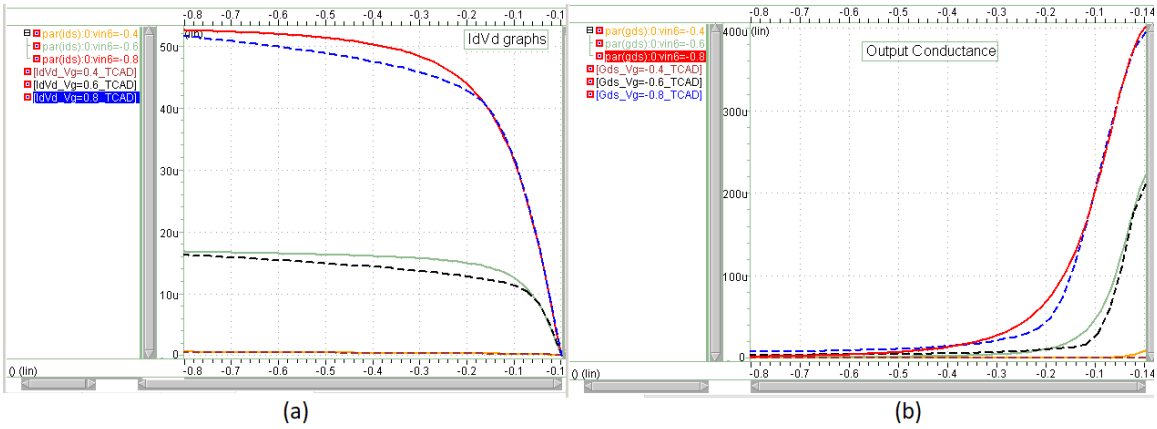

Figure A.4 Optimized (a) I_dV_d, (b) $G_{ds}V_d$ graphs at $V_{gs} = -0.4, -0.6, -0.8$ V

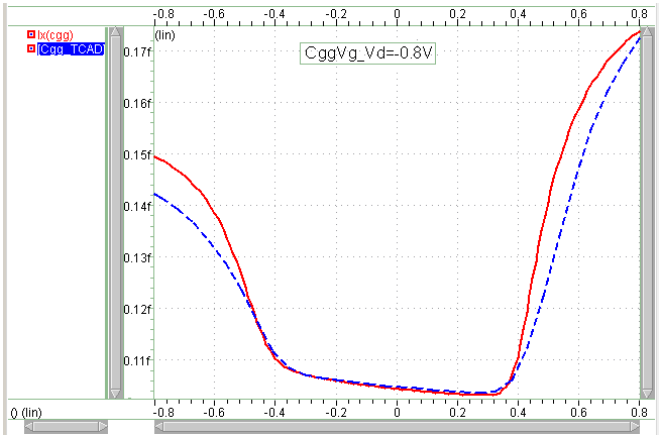

Figure A.5 Optimized $C_{gg}V_{gs}$ characteristics at $V_d = -0.8$ V

Appendix B AOI with Tleak Defect in Transistor M1

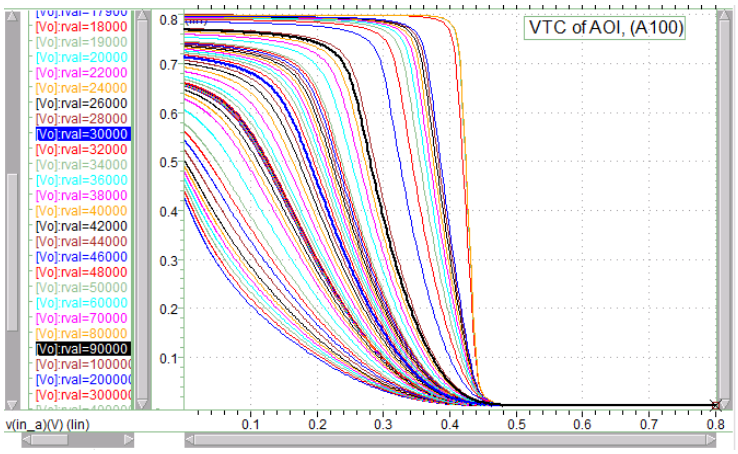

Figure B.1 VTC of defective AOI with Tleak defect on M1, Case 1

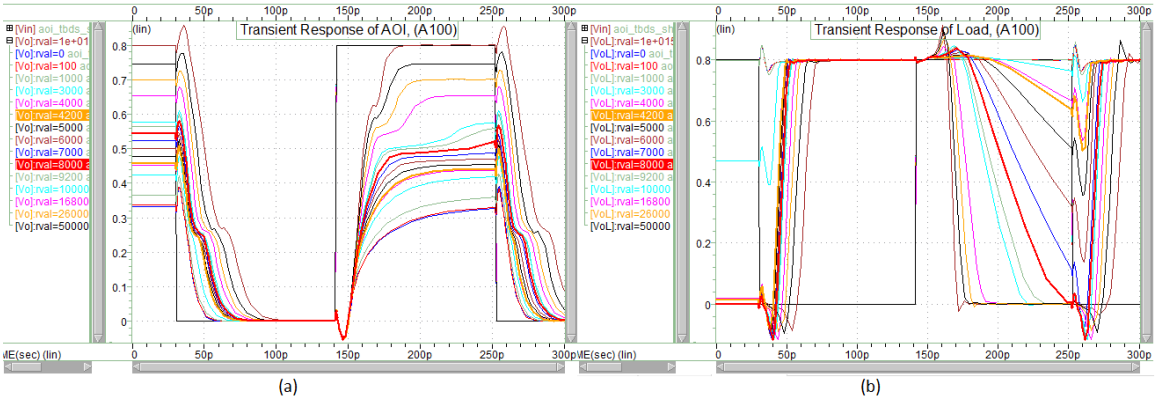

Figure B.2 Transient response of (a) AOI with Tleak defect on M1, (b) the load inverter, Case 1

Appendix C GOS in PMOS Transistors

Table C.1 AOI input combinations applied for GOS in transistor M5

Case	Input Pattern
1	$A: 1 \rightarrow 0, BCD = 101$
2	$AB = 01, CD: 1 \rightarrow 0$
3	$ABCD: 1111 \rightarrow 0000$

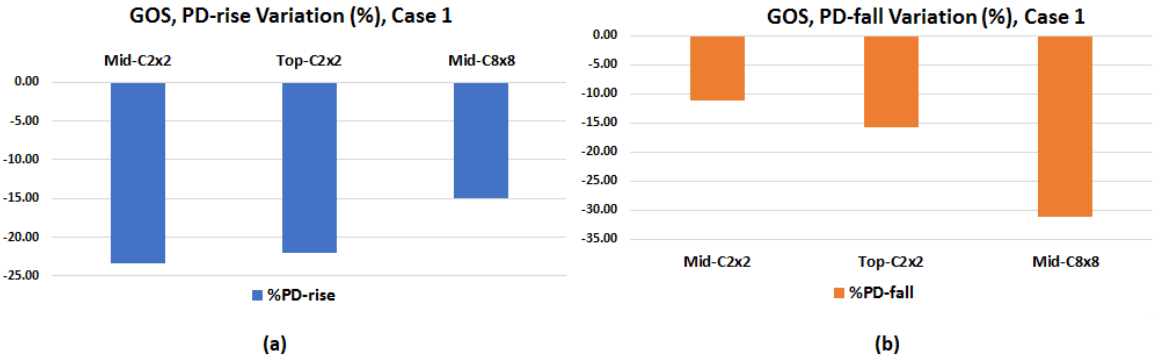

Figure C.1 Percentage of delay variations in AOI with GOS defect in transistor M5, Case 1

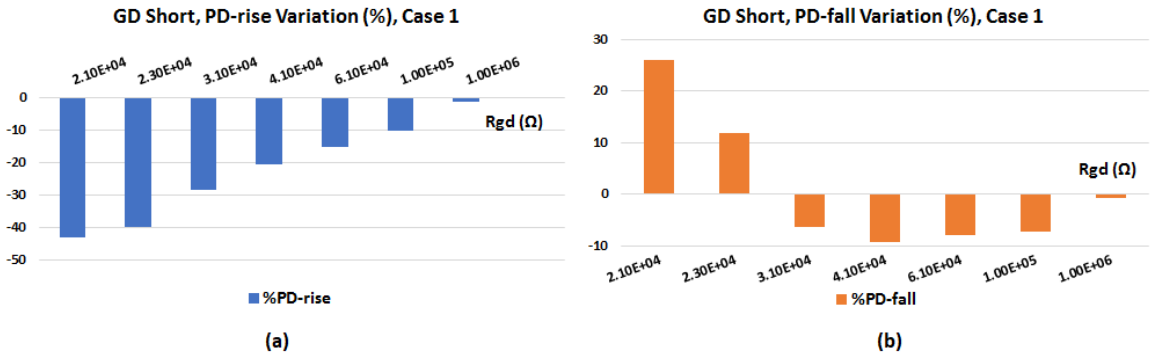

Figure C.2 Percentage of delay variations in AOI with GD short in transistor M5, Case 1

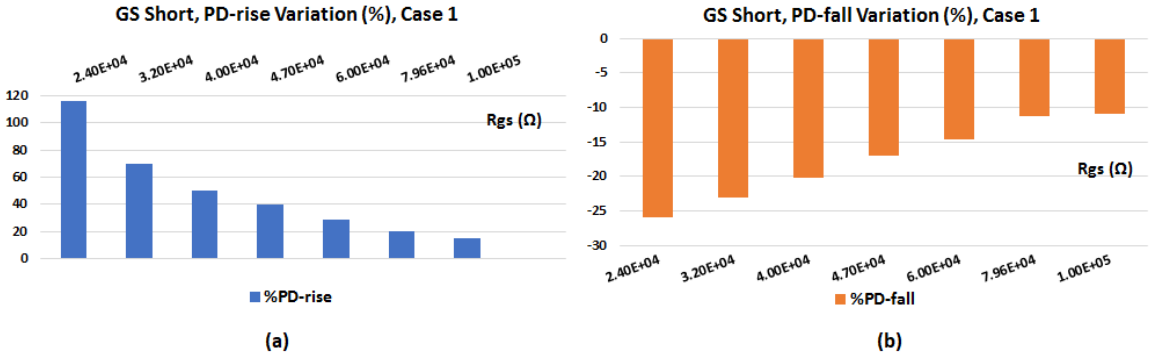

Figure C.3 Percentage of delay variations in AOI with GS short in transistor M5, Case 1

Table C.2 AOI input combinations applied for GOS in transistor M7

Case	Input Pattern
1	$A: 1 \rightarrow 0, BCD = 101$
2	$AB = 01, CD: 1 \rightarrow 0$
3	$ABD: 111 \rightarrow 000, C = 0$

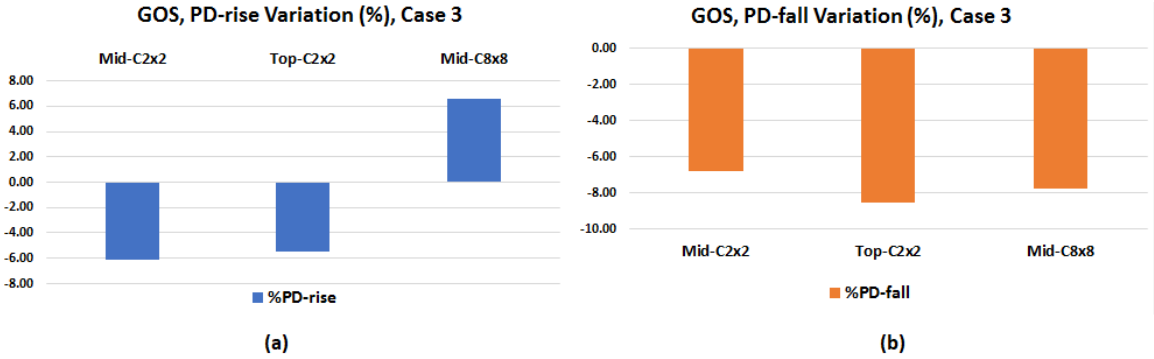

Figure C.4 Percentage of delay variations in AOI with GOS defect in transistor M7, Case 3

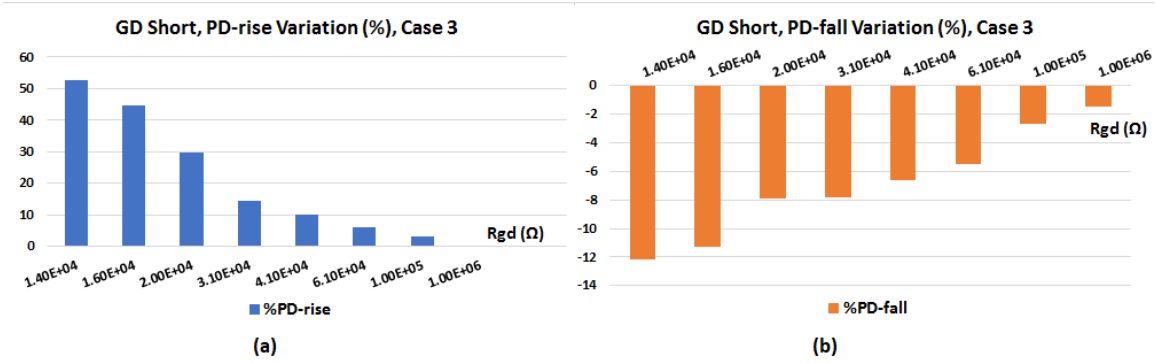

Figure C.5 Percentage of delay variations in AOI with GD short in transistor M7, Case 3

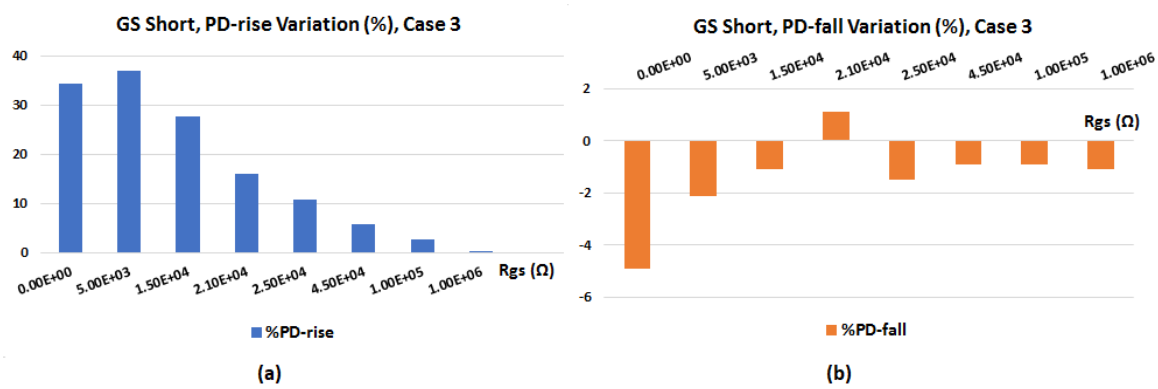

Figure C.6 Percentage of delay variations in AOI with GS short in transistor M7, Case 3

References

[1] Bohr M., "Silicon technology leadership for the mobility era," *Intel developer forum,* 2012.

[2] Packan P., et al., "High Performance 32nm Logic Technology Featuring 2nd Generation High-k+ Metal Gate Transistors," *Electron Devices Meeting (IEDM),* 2009.

[3] Jan C.H., et al., "A 22nm SoC platform technology featuring 3-D tri-gate and high-k/metal gate, optimized for ultra low power, high performance and high density SoC applications," *Electron Devices Meeting (IEDM),* 2012.

[4] Sawicki J., "New Approaches to Improving Quality and Accelerating Yield Ramp for FinFET Technology," *Semicon West,* 2013.

[5] Hapke F., Redemund W., et al., "Cell-Aware Test," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems,* vol. 33, no. 90, pp. 1396-1409, 2014.

[6] Renovell M., Bertrand Y., "Delay testing of MOS transistor with gate oxide short," *Proc. Test Symp.,* p. 168, 2003.

[7] Roy K., Mukhopadhyay S., Mahmoodi-Meimand H., "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits.," *Proc. of the IEEE,* vol. 91, no. 2, pp. 305-327, 2003.

[8] Auth C., Allen C., et al., "A 22nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors," *Proc. IEEE Symp. on. VLSI technology (VLSIT),* pp. 131-132, 2012.

[9] Sinha S., Yeric G., et al., "Exploring sub-20nm FinFET design with predictive technology models," *Proc. IEEE Design Automation Conf. (DAC),* pp. 283-288, 2012.

[10] Nowak E. J., Aller I., et al., "Turning silicon on its edge [double gate CMOS/FinFET technology]," *IEEE Circuits and Devices Magazine,* vol. 20, no. 1, pp. 20-31, 2004.

[11] Zhou H., Song Y., et al. , "Fabrication of Bulk-Si FinFET using CMOS compatible process," *Microelectronic Engineering,* vol. 94, pp. 26-28, 2012.

[12] Ng K. K., Lynch W. T., "The impact of intrinsic series resistance on MOSFET scaling," *IEEE Transactions on Electron Devices,* vol. 34, no. 3, pp. 503-511, 1987.

[13] Sakurai T., Newton A. R., "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," *IEEE Journal of solid-state circuits,* vol. 25, no. 2, pp. 584-594, 1990.

[14] Tang H., Jain A., et al., " Using Cell Aware Diagnostic Patterns to Improve Diagnosis Resolution for Cell Internal Defects," *Proc. IEEE Asian Test Symp. (ATS),* pp. 231-236, 2017.

[15] Liebmann L., Pietromonaco D., Graf M., "Decomposition-aware standard cell design flows to enable double-patterning technology," *Design for Manufacturability through Design-Process Integration, Int. Society for Optics and Photon,* vol. 7974, p. 79740k, 2011.

[16] Park C., Labelle C., et al., "Challenges of VLSI Patterning and Potential Applications of Atomic Layer Etching," *SEMATECH ALE Workshop, Global Foundries,* 2013.

[17] Natarajan S., Agostinelli M., et al., "A 14nm logic technology featuring 2nd generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588μm^2 SRAM cell size," *Proc. IEEE Int. Electron Devices Meeting (IEDM),* pp. 3-7, 2014.

[18] Synopsys Inc., "Three-dimensional simulation of 14/16nm FinFETs with round fin corners and tapered fin shape," *2015.*

[19] Choi Y. K., Chang L., et al., "FinFET process refinements for improved mobility and gate work function engineering," *Proc. IEEE Int. Electron Devices Meeting (IEDM),* pp. 259-262, 2002.

[20] Robertson J., "High dielectric constant oxides," *The European Physical Journal-Applied Physics,* vol. 28, no. 3, pp. 265-291, 2004.

[21] Bohr M., "The evolution of scaling from the homogeneous era to the heterogeneous era," *Proc. IEEE Int. Electron Devices Meeting (IEDM),* pp. 1-1, 2011.

[22] Sachdev M., De Gyvez, J. P., *Defect-oriented testing for nano-metric CMOS VLSI circuits,* Springer Science & Business Media, vol. 34, 2007.

[23] Segura J., Hawkins C. F., *CMOS electronics: how it works, how it fails,* John Wiley & Sons, 2004.

[24] Daino M., Jensen G., et al. , "Line end voids defectivity improvement on 64 pitch Cu wire interconnects of 14 nm technology," *Proc. IEEE Advanced Semiconductor Manufacturing Conf. (ASMC),* pp. 60-64, 2016.

[25] Gupta P., Papadopoulou E., Yield Analysis and Optimization, Handbook of Algorithms for Physical Design Automation, 2008.

[26] Tshagharyan G., Harutyunyan G., et al., "Overview study on fault modeling and test methodology development for FinFET-based memories," *Proc. IEEE East-West Design & Test Symp. (EWDTS),* pp. 1-4, 2015.

[27] Champac V., Hernandez J. V., et al., "Testing of stuck-open faults in nanometer technologies," *IEEE Design & Test of Computers,* vol. 29, no. 4, pp. 80-91, 2012.

[28] Bushnell M., Agrawal V., *Essentials of electronic testing for digital, memory and mixed-signal VLSI circuits,* Springer Science & Business Media, 2004.

[29] Chiang K. Y., Ho Y. H., et al., "Fault simulation and test pattern generation for cross-gate defects in FinFET circuits," *Proc. IEEE Asian Test Symp. (ATS),* pp. 181-186, 2015.

[30] Liu Y., Xu Q., "On modeling faults in FinFET logic circuits," *Proc. IEEE Int. Test Conf. (ITC),* pp. 1-9, 2012.

[31] Foran B., Ives N., Lysaght P., "Tomography of high-k dielectrics on fin-FET sidewalls," *Microscopy and Microanalysis,* vol. 14, no. S2, p. 444, 2008.

[32] Renovell M., Galliere J. M., et al. , "Modeling the random parameters effects in a non-split model of gate oxide short," *Journal of Electronic Testing,* vol. 19, no. 4, pp. 377-386, 2003.

[33] Lee J., Jeong H., et al., "Impact of fin thickness and height on read stability/write ability in tri-gate FinFET based SRAM," *Proc. IEEE Int. SoC Design Conf. (ISOCC),* pp. 657-660, 2012.

[34] Yeh W. K., Zhang W., et al., "Effects of fin width on performance and reliability for N-and P-type FinFETs," *IEEE Int. Conf. on Electron Devices and Solid-State Circuits (EDSSC),* pp. 361-364, 2016.

[35] Oliveira A. V., Agopian P. G. D., et al., "Fin width influence on analog performance of SOI and bulk FinFETs," *Proc. IEEE Int. Caribbean Conf. on Devices, Circuits and Systems (ICCDCS),* pp. 1-4, 2014.

[36] Huang W. T., Li Y., "Electrical characteristic fluctuation of 16-nm-gate trapezoidal bulk FinFET devices with fixed top-fin width induced by random discrete dopants," *Nanoscale research letters,* vol. 10, no. 1, p. 116, 2015.

[37] Sohn C. W., Kang C. Y., et al., "Effect of fin height of tapered FinFETs on the sub-22-nm system on chip (SoC) application using TCAD simulation," *Proc. IEEE Int. Symp. on VLSI Technology, Systems, and Applications (VLSI-TSA),* pp. 1-2, 2013.

[38] Al-Khalili D., Adham S., et al., "Comprehensive defect analysis and defect coverage of CMOS circuits," *Proc. IEEE Int. Symp. on Defect and Fault Tolerance in VLSI Systems,* 1998.

[39] Rajsuman R., " Iddq testing for CMOS VLSI," *Proceedings of the IEEE,* vol. 88, no. 4, pp. 544-568, 2000.

[40] Mesalles F., Villacorta H., et al., "Behavior and test of open-gate defects in FinFET based cells," *Proc. IEEE European Test Symp. (ETS),* pp. 1-6, 2016.

[41] Chang C. L. L., Wen C. H. P., Bhadra J., "Process-variation-aware Iddq diagnosis for nano-scale CMOS designs-the first step," *Proc. Conf. on Design, Automation and Test in Europe, EDA Consortium,* pp. 454-457, 2013.

[42] Savir J., Patil S., "Scan-based transition test," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems,* vol. 12, no. 8, pp. 1232-1241, 1993.

[43] Xiang D., Li J., et al., "Test compaction for small-delay defects using an effective path selection scheme," *ACM Trans. on Design Automation of Electronic Systems (TODAES),* vol. 18, no. 3, p. 44, 2013.

[44] Arumí D., Rodríguez-Montañés R., Figueras J., " Test Escapes of Stuck-Open Faults Caused by Parasitic Capacitances and Leakage Currents," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 24, no. 5, pp. 1739-1748, 2016.

[45] Grout I. A., *Integrated circuit test engineering: modern techniques,* Springer Science & Business Media, 2005.

[46] Jha N. K., Gupta S., *Testing of digital systems,* Cambridge University Press, 2003.

[47] Wang L. T., Wu C. W., Wen X. , *VLSI test principles and architectures: design for testability,* Elsevier, 2006.

[48] Mirza-Aghatbar M., Breuer M. A., et al., " A design flow to maximize yield/area of physical devices via redundancy," *Proc. IEEE Int. Test Conf. (ITC),* pp. 1-10, 2012.

[49] Abramovici M., Breuer M. A., Friedman A. D., *Digital Systems Testing and Testable Design,* Computer Science Press, vol. 2, 1990.

[50] Nelson J. E., Brown J. G., et al., "Multiple-detect ATPG based on physical neighborhoods," *Proc. ACM/IEEE Design Automation Conf.,* pp. 1099-1102, 2006.

[51] Lin Y. T., Poku O., et al., "Evaluating the Effectiveness of Physically-Aware N-Detect Test using Real Silicon," *Proc. IEEE Int. Test Conf. (ITC),* pp. 1-9, 2008.

[52] Cho K. Y., Mitra S., McCluskey E. J., "Gate exhaustive testing," *Proc. IEEE Int. Test Conf.,* 2005.

[53] Hapke F., Reese M., et al., "Cell-aware Production test results from a 32nm notebook processor," *Proc. IEEE Int. Test Conf. (ITC),* pp. 1-9, 2012.

[54] Hapke F., Schloeffel J., et al., "Gate-exhaustive and cell-aware pattern sets for industrial designs," *Proc. IEEE Int. Symp. on VLSI Design, Automation and Test (VLSI-DAT),* pp. 1-4, 2011.

[55] Tang H., Tai T. P., et al., "Diagnosing timing related cell internal defects for FinFET technology," *Proc. IEEE Int. Symp. on VLSI Design, Automation and Test (VLSI-DAT),* pp. 1-4, 2015.

[56] Bhoj Ajay N., Muzaffer O. Simsir, et al., " Fault models for logic circuits in the multigate era," *IEEE Transactions on Nanotechnology,* vol. 11, no. 1, pp. 182-193, 2012.

[57] Forero F., Galliere J. M., et al., "Analysis of short defects in FinFET based logic cells," *Proc. IEEE Latin American Test Symp. (LATS),* pp. 1-6, 2017.

[58] Goel S. K., Chakrabarty K., eds., *Testing for Small-Delay Defects in Nanoscale CMOS Integrated Circuits,* CRC Press, 2013.

[59] Lin A. F., Liao K. Y., et al., "TARGET: Timing-aware gate exhaustive transition ATPG for cell-internal defects," *Proc. IEEE Int. Symp. on VLSI Design, Automation and Test (VLSI-DAT),* pp. 1-4, 2015.

[60] Chehade S., Chehab A., and Kayssi A., "Modeling of gate oxide short defects in MOSFETS," *Int. Design and Test Workshop (IDT), IEEE,* 2009.

[61] Lin C. W., Chao M. C. T., Hsu C. C., "Novel circuit-level model for gate oxide short and its testing method in SRAMs," *IEEE Trans. on VLSI Systems,* vol. 22, no. 6, pp. 1294-1307, 2014.

[62] Sytrzycki M., "Modeling of gate oxide shorts in MOS transistors," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems,* vol. 8, no. 3, pp. 193-202, 1989.

[63] Renovell M., Galliere J. M., et. al., " Modeling gate oxide short defects in CMOS minimum transistors," *IEEE European Test Workshop,* pp. 15-20, 2002.

[64] Lin C. W., Chao M. C. T., Hsu C. C., "Investigation of gate oxide short in FinFETs and the test methods for FinFET SRAMs," *Proc. IEEE VLSI Test Symp. (VTS),* pp. 1-6, 2013.

[65] Singh A. D. , "Cell Aware and stuck-open tests," *Proc. IEEE European Test Symp. (ETS),* pp. 1-6, 2016.

[66] Dibaj R., Al-Khalili D., Shams M., "Comprehensive investigation of gate oxide short in FinFETs," *Proc. IEEE VLSI Test Symp. (VTS),* pp. 1-6, 2017.

[67] Synopsys Inc., Sentaurus™ Process User Guide, Version K, 2015.

[68] Synopsys Inc., Sentaurus™ Device User Guide, Version K, 2015.

[69] Khandelwal S., Duarte J., et al., *BSIM-CMG 110.0. 0–Multi-gate MOSFET compact model technical manual,* Regents University California, Berkeley, 2015.

[70] Ortiz-Conde A., Sanchez F.G., et al., "A review of recent MOSFET threshold voltage extraction methods," *Microelectronics Reliablility,* vol. 42, pp. 583-596, 2002.

[71] Rudenko T., Kilchytska V., et al., "On the MOSFET threshold voltage extraction by transconductance and transconductance-to-current ratio change methods: Part I and II," *IEEE Transaction on Electron Devices,* vol. 58, pp. 4172-4180, 2011.

[72] Brozek T., *Micro-and nanoelectronics: Emerging device challenges and solutions,* CRC Press, 2014.

[73] Sedra A. S., Smith K. C., *Microelectronics circuits,* Oxford University Press, 1998.

[74] Razavieh A., Deng Y., et al., " Effective Drive Current in Scaled FinFET and NSFET CMOS Inverters," *IEEE Device Research Conf. (DRC),* pp. 1-2, 2018.

[75] Chauhan Y. S., Lu D. D., et al., *FinFET Modeling for IC Simulation and Design: Using the BSIM-CMG Standard,* Academic Press, 2015.

[76] Mohammadi H. G., Gaillardon P. E., De Micheli G., "From defect analysis to gate-level fault modeling of controllable-polarity silicon nanowires," *IEEE Trans. on Nanotechnology,* vol. 14, no. 6, pp. 1117-1126, 2015.

[77] Balestra F., Cristoloveanu S., et al., "Double-gate silicon-on-insulator transistor with volume inversion: a new device with greatly enhanced performance," *IEEE Electron Device Letter,* vol. 8, no. 9, pp. 410-412, 1987.

[78] Synopsys Inc., Sentaurus™ Compact Model User Guide, Version K, 2015.

[79] Roth J. P., "Diagnosis of automata failures: A calculus and a method," *IBM journal of Research and Development,* vol. 10, no. 4, pp. 278-291, 1966.

[80] Elmore W. C., "The transient response of damped linear networks with particular regard to wideband amplifiers," *Journal of Applied Physics,* vol. 19, no. 1, pp. 55-63, 1948.

[81] Wang P., Moore C. J., et al., "An ATPG method for double Stuck-At faults by analyzing propagation paths of single faults," *IEEE Trans. on Circuits and Systems*, vol. 65, pp. 1063-1074, 2018.