DISSERTATION

TIMING ANALYSIS OF CMOS LOGIC GATES IN DEEP SUBMICRON VLSI DESIGN

Submitted by

Xueping Jiang

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2005

INFORMATION TO USERS

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

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

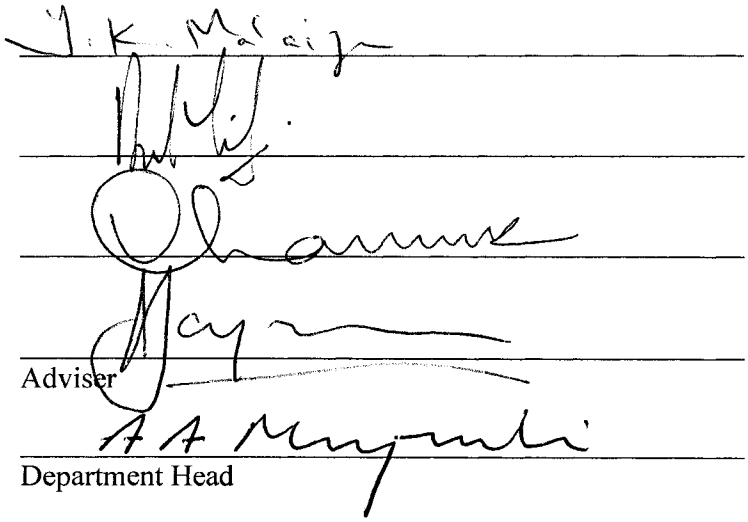
ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

COLORADO STATE UNIVERSITY

DEC. 16, 2004

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY XUEPING JIANG ENTITLED TIMING ANALYSIS OF CMOS LOGIC GATES IN DEEP SUBMICRON VLSI DESIGN BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work



Adviser

Department Head

ABSTRACT OF DISSERTATION

TIMING ANALYSIS OF CMOS LOGIC GATES IN DEEP SUBMICRON VLSI DESIGN

A comprehensive timing analysis of complementary metal-oxide semiconductor (CMOS) logic gates in deep submicron very large scale integration (VLSI) design is carried out. A computationally simple proper deep submicron MOSFET model (PDSMM) is proposed to accurately describe the static I–V characteristics of both long channel and deep submicron metal-oxide semiconductor field effect transistors (MOSFETs). The PDSMM is an extension of Shockley MOSFET model and Shichman-Hodges MOSFET model by incorporating the second-order effects of deep submicron MOSFETs such as charge carrier velocity saturation and channel length modulation. Using the PDSMM we develop propagation delay models for a deep submicron CMOS inverter. The effects of input transition time, gate-to-drain coupling capacitance, parasitic drain diffusion capacitance, and external load capacitance are taken into account. Even in the extreme conditions, the proposed propagation delay model of a deep submicron CMOS inverter can accurately duplicate the propagation delays from SPICE BSIM3.

In order to extend the propagation delay methodologies for a deep submicron CMOS inverter to model single-stage complex CMOS logic gates such as NAND or NOR gate, a generic series-connected MOSFET chain model (SMCM) is proposed to

accurately depict the peculiar static I–V characteristics of a series-connected MOSFET chain with an equal input pattern. The effect of initial states of parasitic capacitances of intermediate nodes of a series-connected MOSFET chain is taken into account by using a modified input suppression technique associated with activation time. We focus on fast initial state (FIS) and slow initial state (SIS), which are the most useful in practice. Using the SMCM and the modified input suppression technique we develop propagation delay models for a single-stage complex CMOS logic gate. We propose a simple input mapping technique to investigate the effect of arbitrary input pattern of a single-stage complex CMOS logic gate. These timing models include initial states of parasitic capacitances and arbitrary input pattern in addition to the factors used in propagation delay models of a CMOS inverter. The comprehensive propagation delay models of a single-stage complex CMOS logic gate can accurately reproduce the propagation delays from SPICE BSIM3.

The proposed comprehensive propagation delay models are suitable to precisely describe transition behaviors of deep submicron CMOS logic gates. They can be exploited in electronic design automation (EDA) flows to implement, verify and optimize deep submicron CMOS VLSI designs.

Xueping Jiang (姜学平)
Department of Electrical and Computer Engineering
Colorado State University
Fort Collins, CO 80523
Spring 2005

ACKNOWLEDGMENTS

I would like to thank the numerous individuals who contributed to the completion of the manuscript. First and foremost, I am particularly indebted to my advisor, Dr. Anura P. Jayasumana for his excellent continuous guidance and instructive discussions throughout the long time-consuming course of this work. I would also like to express my gratitude to the committee members, Dr. V. Chandrasekar, Dr. D. L. Lile, and Dr. Y. K. Malaiya for their advice and support. Also, I am grateful to Dr. H. Temkin who led me into the area of semiconductor devices and materials.

To my parents, Binggui Jiang and Sihua Tong, my brother Min Jiang, and my sisters Xiuli Jiang and Qiaohong Jiang, I would like to express my many thanks for their kindness and encouragement. To my wife Weijun Zhang and my two lovely children Jeffrey Jiang and Jessica Jiang, who support and encourage me in everyway, I give my gratitude.

TABLE OF CONTENTS

Chapter 1	Introduction	1
Chapter 2	Background and Related Researches	5
2.1	Background of CMOS Digital Circuits	5
2.1.1	Dominance of CMOS Technology	5
2.1.2	Moore's Law	7
2.1.3	Digital Design Hierarchy	8
2.1.4	Timing Verification of CMOS Digital Circuits	9
2.2	Related Researches in Timing Methodologies of CMOS Digital Circuits	10
2.3	Numerical Approaches of Timing Analysis	12
2.4	Analytical Approaches of CMOS Timing Models	14
2.4.1	Shockley MOSFET Model	15
2.4.2	Step Input Delay Model with Shockley MOSFET model	17
2.4.3	Delay Models Considering Input Waveform	17
2.4.4	The Simplified Bulk-Charge MOSFET Model	18
2.4.5	Delay Models by Using the Simplified Bulk-Charge MOSFET Model	19
2.4.6	α -Power Law MOSFET Model and Delay Calculation	20
2.4.7	The nth-Power Law MOSFET Model and Timing Models	23
2.5	Resistance-Capacitance (RC) Delay Approaches	24
2.5.1	Elmore Delay	25
2.5.2	Signal Delay in RC Tree Networks and RC Mesh Networks	26
2.5.3	Timing Models for CMOS VLSI circuits with Interconnects	28
2.6	Tabular Delay Approaches with Empirical Delay Equations	30
Chapter 3	Deep Submicron MOSFET Models	44
3.1	Channel Length Modulation	45
3.2	Charge Carrier Velocity Saturation	46
3.3	Derived MOSFET Model with Effect of Carrier Velocity Saturation	49
3.4	Power Index of a Deep Submicron MOSFET	52
3.5	Deep Submicron MOSFET Model (DSMM)	54
3.6	Proper Deep Submicron MOSFET Model (PDSMM)	56
Chapter 4	Timing Models of a Deep Submicron CMOS Inverter	64
4.1	Normalized Variables	66
4.2	CMOS Inverter	66
4.2.1	Differential Equation	66

4.2.2 Capacitances 68
4.2.3 MOSFET Current Models 69
4.3 Timing Analysis of a CMOS Inverter for a Step Input 71
4.3.1 Operation of a CMOS Inverter for a Step Input 72
4.3.2 Timing Analysis of a CMOS Inverter for a Step Input 73
4.3.3 Results and Discussions 77
4.4 Timing Analysis of a CMOS Inverter Modeled with a Dominating MOSFET for a Ramp Input 79
4.4.1 Operation of a Dominating nMOSFET for a Ramp Input 79
4.4.2 Output Voltages of a Dominating nMOSFET for a Ramp Input 80
4.4.3 Results and Discussions 86
4.5 Comprehensive Propagation Delay Analysis of a CMOS Inverter 89
4.5.1 Operation of a CMOS Inverter for a Ramp Input 89
4.5.2 Output Voltages of a CMOS Inverter for a Ramp Input 90
4.5.3 Results and Discussions 107

Chapter 5 A Series-Connected MOSFET Chain with an Equal Input Pattern 134
5.1 Naming Conventions 137
5.2 Static I–V Characteristics 138
5.3 Power Index 140
5.4 Generic Series-Connected MOSFET Chain Model (SMCM) 141
5.5 Modified Input Suppression Technique 144
5.6 Two Extreme Initial States – Slow and Fast 146
5.7 Step-Input Activation Time t_{as} 149
5.8 Ramp-Input Activation Time t_a 150
5.8.1 Large Input Transition Time: $(\tau \geq \tau_u)$ 151
5.8.2 Medium Input Transition Time: $(\tau_l \leq \tau < \tau_u)$ 152
5.8.3 Small Input Transition Time: $(\tau < \tau_l)$ 153
5.8.4 Discussions 154

Chapter 6 Timing Models of a Single-Stage CMOS Logic Gate with an Equal Input Pattern 180
6.1 A Single-Stage CMOS Logic Gate with an Equal Input Pattern 183
6.1.1 Differential Equation 183
6.1.2 Capacitances 184
6.1.3 Current Models of I_p and I_n 185
6.2 Timing Analysis of a Single-stage CMOS Logic Gate for a Step Input 188
6.2.1 Operation of a Single-Stage CMOS Logic Gate for a Step Input 189
6.2.2 Timing Analysis of a Single-Stage CMOS Logic Gate for a Step Input 190
6.2.3 Results and Discussions 195
6.3 Timing Analysis of a Single-Stage CMOS Logic Gate for a Ramp Input without a Trivial Parallel MOSFET Cluster 197
6.3.1 Operation of a Series-Connected nMOSFET Chain for a Ramp Input 198
6.3.2 Output Voltages of a Dominant Series-Connected MOSFET Chain for a Ramp Input without a Trivial Parallel MOSFET Cluster 199

6.3.3 Effect of Initial States of Parasitic Capacitances of Intermediate Nodes .. 207
6.3.4 Results and Discussions 209
6.4 Comprehensive Timing Analysis of a Single-Stage CMOS Logic Gate with an Equal Input Pattern 212
6.4.1 Operation of a Single-Stage CMOS Logic Gate with an Equal Input Pattern 212
6.4.2 Output Voltages of a Complex CMOS Logic Gate with an Equal Input Pattern 214
6.4.3 Effect of Initial States of Parasitic Capacitances of Intermediate Nodes .. 231
6.4.4 Results and Discussions 235
Chapter 7 Timing Model of a Single-Stage CMOS Logic Gate with an Arbitrary Input Pattern 269
7.1 Naming Conventions 271
7.2 Arbitrary Input Pattern 272
7.2.1 Inactive Inputs 273
7.2.2 Active Inputs 273
7.2.3 Partially-Active Inputs 274
7.3 Input Mapping Technique 276
7.4 Results and Discussions 278
Chapter 8 Conclusions 302
References 305
Appendix I Derivation of New DSMM model 317
I.1 Drain-Source Saturation Voltage V_{sat} in DSMM 317
I.2 Drain-Source Saturation Current I_{sat} in DSMM 318
Appendix II First-order Linear Differential Equation, Bernoulli Equation, and Raccati Equation 320
II.1 First-order linear differential equation 320
II.2 Bernoulli equation 321
II.3 Raccati equation 321
Appendix III Cumulative Distribution Function (CDF), *Q-function*, and useful expressions 323
III.1 Cumulative Distribution Function (CDF) and *Q-function* 323
III.2 Useful Expressions 324

LIST OF TABLES

Table 4.1 The PDSMM Parameters of nMOSFET and pMOSFET of a CMOS Inverter.108

Table 5.1 The SMCM Model Parameters for All the Sorts of a Four-MOSFET Series-Connected Chain with an Equal Input Pattern Used in Chapter 5 156

Table 5.2 The Step-Input Activation Time t_{as} for a Four-MOSFET Series-Connected Chain with an Equal Input Pattern 157

Table 6.1 The Speedup of the Proposed Method over SPICE with BSIM3 Model..... 237

Table 6.2 The Coefficients for Calculating the Input-to-Output Coupling Capacitances and the Diffusion Capacitances 238

Table 6.3 The SMCM Model Parameters Used in Chapter 6 239

Table 7.1 Classification of an Individual Input of an Arbitrary Input Pattern..... 280

Table 7.2: Four Examples of Input Mapping Technique for FIS..... 281

Table 7.3: Four Examples of Input Mapping Technique for SIS 282

Table 7.4 Propagation Delays from SPICE and Input Mapping Technique for FIS..... 283

Table 7.5 Propagation Delays from SPICE and Input Mapping Technique for FIS..... 284

Table 7.6 Propagation Delays from SPICE and Input Mapping Technique for FIS..... 285

Table 7.7 Propagation Delays from SPICE and Input Mapping Technique for FIS..... 286

Table 7.8 Propagation Delays from SPICE and Input Mapping Technique for SIS 287

Table 7.9 Propagation Delays from SPICE and Input Mapping Technique for SIS 288

Table 7.10 Propagation Delays from SPICE and Input Mapping Technique for SIS 289

Table 7.11 Propagation Delays from SPICE and Input Mapping Technique for SIS 290

LIST OF FIGURES

Fig. 2.1 A generic CMOS logic gate used in analytical timing approaches. There are two basic circuit components: pMOSFET tree and its dual nMOSFET tree. 32

Fig. 2.2 Static I–V characteristics of a deep submicron MOSFET from Shockley model and BSIM3 model. The parameters of Shockley model are fit to have the same characteristics with SPICE BSIM model at drain-source voltage close to zero in linear region. 33

Fig. 2.3 Saturation drain-source currents of a deep submicron MOSFET when $V_{ds} = V_{dd}$ *from Shockley model and from SPICE BSIM3 model. The parameters of Shockley model are fit to have the same characteristics with SPICE BSIM3 model at small drain-source voltage in linear region.* 34

Fig. 2.4 Comparison of static I–V characteristics of a deep submicron MOSFET from Shockley model and from SPICE BSIM3 model. The parameters of Shockley model are fit to have the same current with SPICE BSIM model when $V_{gs} = V_{dd}$ *and* $V_{ds} = V_{dd}$ *in saturation region.* 35

Fig. 2.5 Saturation drain-source currents of a deep submicron MOSFET when $V_{ds} = V_{dd}$ *from Shockley model and from SPICE BSIM3 model. The parameters of Shockley model are fit to have the same current with SPICE BSIM3 model when* $V_{gs} = V_{dd}$ *and* $V_{ds} = V_{dd}$ *in saturation region.* 36

Fig. 2.6 Static I–V characteristics of a deep submicron from simplified bulk-charge MOSFET model and SPICE BSIM3 model. The parameters of simplified bulk-charge MOSFET model are fit to have the same drain-source current with SPICE BSIM3 model when $V_{gs} = V_{dd}$ *and* $V_{ds} = V_{dd}$ *in saturation region. The static I–V characteristics from simplified bulk-charge MOSFET model where square-law in saturation region still exists deviate from SPICE BSIM3.* 37

Fig. 2.7 Saturation drain-source current of a deep submicron MOSFET when $V_{ds} = V_{dd}$ *from simplified bulk-charge MOSFET model and from SPICE BSIM3 model. The parameters of simplified bulk-charge MOSFET model are fit to have the same current with SPICE BSIM3 model when* $V_{gs} = V_{dd}$ *and* $V_{ds} = V_{dd}$ *in*

saturation region. Obviously, power index of a deep submicron MOSFET deviates from two as in simplified bulk-charge MOSFET model when gate-source voltage increases. 38

Fig. 2.8 Static I–V characteristics of a deep submicron MOSFET from α-power law MOSFET model and from SPICE BSIM3 model. The parameters of α-power law MOSFET model are fit to have the same drain-source current with SPICE BSIM3 model when $V_{gs} = V_{dd}$ *and* $V_{ds} = V_{dd}$ *in saturation region. The static I–V characteristics from α-power law MOSFET model are piecewise linearized....* 39

Fig 2.9 Saturation drain-source current of a deep submicron MOSFET when $V_{ds} = V_{dd}$ *from α-power law MOSFET model and from SPICE BSIM3 model. The parameters of α-power law MOSFET model are fit to have the same current with SPICE BSIM3 model when* $V_{gs} = V_{dd}$ *and* $V_{ds} = V_{dd}$ *in saturation region..* 40

Fig. 2.10 Static I–V characteristics of a deep submicron MOSFET from nth-power law MOSFET model and SPICE BSIM3 model. The parameters of nth-power law MOSFET model are fit to have the same drain-source current with SPICE BSIM3 model when $V_{gs} = V_{dd}$ *and* $V_{ds} = V_{dd}$ *in saturation region. The static I–V characteristics from nth-power law MOSFET model are well matched with SPICE BSIM3 model.* 41

Fig. 2.11 Saturation drain-source current of a deep submicron MOSFET when $V_{ds} = V_{dd}$ *from nth-power law MOSFET model and from SPICE BSIM3 model. The parameters of nth-power law MOSFET model are fit to have the same current with SPICE BSIM3 model when* $V_{gs} = V_{dd}$ *and* $V_{ds} = V_{dd}$ *in saturation region. The nth-power law MOSFET model can accurately reproduce static I–V characteristics of a deep submicron MOSFET.* 42

Fig. 2.12 RC chained line and RC tree network 43

Fig. 3.1 Drifting velocity of surface electrons as a function of channel electric field [22]. Drifting velocity saturates when electric field increases. The critical electric field E_c *is obtained by dividing saturation drifting velocity* V_{sat} *with low-field mobility* μ_0. 58

Fig. 3.2 Mobility of surface electrons as a function of channel electric field derived from (3.3) for data of Fig. 3.1. The mobility of surface electrons decreases as channel electric field increases or drifting velocity of surface electrons saturates. 59

Fig. 3.3 Saturation drain-source currents as a function of gate-source voltage at $V_{ds} = V_{dd}$ *from SPICE BSIM3 model and from the proposed models (DSMM and PDSMM) for a deep submicron nMOSFET (upper line) and pMOSFET (lower*

line), respectively. It shows that the proposed models (DSMM and PDSMM) can accurately reproduce I–V characteristics of a deep submicron MOSFET in saturation region. *60*

Fig. 3.4 Power indices as a function of gate-source voltage for a deep submicron pMOSFET (dashed line) and nMOSFET (solid line), respectively. The vertical dotted lines indicate the corresponding threshold voltages. The power index β *of a deep submicron MOSFET is two when gate-source voltage is equal to threshold voltage and gradually decreases as gate-source voltage increases. The lower limit of power index* β *of a deep submicron MOSFET is one.* *61*

Fig. 3.5 Static I–V characteristics are plotted from deep submicron MOSFET model (DSMM) and SPICE BSIM3 model for a deep submicron MOSFET. The static I–V characteristics from DSMM are well matched with those from SPICE BSIM3 model. *62*

Fig. 3.6 Static I–V characteristics are plotted from proper deep submicron MOSFET model (PDSMM) and SPICE BSIM3 model for a deep submicron MOSFET. The static I–V characteristics from PDSMM are well matched with those from SPICE BSIM3 model. *63*

Fig. 4.1 A CMOS Inverter. The input-output coupling capacitance C_m *and output load capacitance* C_{load} *are included.* *109*

Fig. 4.2 The typical switching trajectory traces of both pMOSFET and nMOSFET of a CMOS inverter for rising ramp inputs. The static I–V characteristics of both pMOSFET and nMOSFET of the CMOS inverter are also plotted. Obviously the nMOSFET drain-source current I_{ds} *dominates the output transition or discharging process in case of rising ramp inputs.* *110*

Fig. 4.3 The typical discharging trajectory traces of both pMOSFET and nMOSFET for a step input. In this case, the pMOSFET can be ignored since there is no current passing through the pMOSFET. There is a jump of drain-source current I_{ds} *at* $t = 0$. *The output voltage has an overshoot at* $t = 0$. *When* $t > 0$, *the discharging trajectory trace of nMOSFET goes along the characteristic curve of the nMOSFET with* $V_{gs} = V_{dd}$ *111*

Fig. 4.4 Output voltage waveforms of a CMOS inverter with different external capacitances of 10fF, 100fF, 200fF, 400fF, and 800fF (from left to right) for a rising step input. The transient voltage waveforms of a step input based on PDSMM in Subsection 4.3.2 are shown in solid lines. As a comparison, the output voltage waveforms of a ramp input with input transition time of 1ps from SPICE BSIM3 are also plotted in dashed lines. *112*

Fig. 4.5 Propagation delay versus external capacitance for a rising step input. The propagation delay derived from PDSMM in Subsection 4.3.2 is indicated as the solid line. The dashed line with circle symbol represents the propagation delay obtained from SPICE BSIM3 for a ramp input of 1ps transition time. As a comparison, the propagation delay of a CMOS inverter with a ramp input of 50ps transition time is shown in dotted line with plus symbol. 113

Fig. 4.6 The typical discharging trajectory traces of an nMOSFET for ramp inputs. The static I–V characteristics of only dominating nMOSFET of the CMOS inverter are also plotted. 114

Fig. 4.7 Five operation regions (R1-R5) of the dominating nMOSFET of a CMOS inverter for rising ramp inputs. The five operation regions are introduced in the text. The ramp input is indicated with the dashed line. The typical discharging output waveforms are also plotted. 115

Fig. 4.8 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (solid lines). Only a dominating nMOSFET denotes the CMOS inverter. Input transition time is 1ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). The calculated output waveforms from PDSMM and the output responses from SPICE BSIM3 agree with each other excellently. 116

Fig. 4.9 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (solid lines). Input transition time is 50ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). Note that for smaller loads, the calculated waveforms are obviously slower than SPICE BSIM3 results due to exaggerated voltage overshoots caused with input-to-output coupling current. In practice, voltage overshoot is lowered by being partially compensated with inverse discharging current through the trivial pMOSFET of a CMOS inverter to power rail. 117

Fig. 4.10 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (solid lines). Input transition time is 100ps. Load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). Voltage overshoots caused with input-to-output coupling capacitance are exaggerated. Comparing with Fig. 4.9, the longer input transition time makes the effect of short-circuit current more significant. The only nMOSFET currents push the calculated waveforms left (or fast). The effect of exaggerated overshoots on propagation delays are partially compensated with no short-circuit currents. 118

Fig. 4.11 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in

Subsection 4.4.2 (solid lines). Input transition time is 200ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). Voltage overshoots caused with input-to-output coupling current are exaggerated. Comparing with Fig. 4.10, the longer input transition time makes the effect of short-circuit current more significant. No short-circuit currents push the calculated waveforms left (or fast). The effect of exaggerated overshoots on propagation delay are almost compensated with no short-circuit currents. *119*

Fig. 4.12 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (solid lines). Input transition time is 400ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). Voltage overshoots are exaggerated. The very slow input transition time makes the effect of no short-circuit current more significant. The uncompensated nMOSFET current discharges the load capacitance much faster. The phenomena make the calculated propagation delays unreliable. *120*

Fig. 4.13 Propagation delays versus external load capacitances for input transition times of 1ps, 10ps, 100ps, 200ps, and 400ps (from bottom to top), respectively. The propagation delays of a CMOS inverter are obtained from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (circles). For small input transition time or large load capacitance, the method proposed in Subsection 4.4.2 is very accurate. However, large input transition time and small load capacitance make a significant error, which cause the expressions in Subsection 4.42 unreliable. 121

Fig. 4.14 Propagation delays versus input transition times with the external load capacitances of 10fF, 100fF, 200fF, 400fF, and 800fF (from bottom to top), respectively. The propagation delays of a CMOS inverter are obtained from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (circles). For small input transition time or large load capacitance, the method proposed in Subsection 4.4.2 is very accurate. However, large input transition time and small load make significant errors, which cause the expressions in Subsection 4.4.2 unreliable. *122*

Fig. 4.15 Correlation of propagation delays from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 with propagation delays from SPICE BSIM3. Only a dominating nMOSFET denotes the CMOS inverter. The two dashed lines are ±5% of propagation delays deviating from SPICE BSIM3. Significant errors are caused with large input transition time and small load capacitance. *123*

Fig. 4.16 Eight operation regions of a CMOS inverter for a rising ramp input. The input ramp is indicated with dashed line. However, the solution of the differential equation for region two is not accurate enough due to simplification of the property equations. The eleven operation regions shown in Fig 4.17 are used in

the comprehensive timing mode. *124*

Fig. 4.17 Eleven operation regions of a CMOS inverter with a rising ramp input for the comprehensive timing model. The input ramp is indicated with dashed line. The dotted lines represent the output responses of rising input ramps. *125*

Fig. 4.18 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. Input transition time is 1ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). *126*

Fig. 4.19 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. Input transition time is 50ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). They match with each other excellently. *127*

Fig. 4.20 The output voltage waveforms of a CMOS inverter from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. The input transition time is 100ps. The external load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). They agree with each other excellently. *128*

Fig. 4.21 The output voltage waveforms of a CMOS inverter from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. The input transition time is 200ps. The external load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). They match with each other excellently. *129*

Fig. 4.22 The output voltage waveforms of a CMOS inverter from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. The input transition time is 400ps. The external load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). *130*

Fig. 4.23 Propagation delay versus external load capacitance. The propagation delays of a CMOS inverter are obtained from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. The input transition times are 1ps, 10ps, 100ps, 200ps, and 400ps, respectively (from bottom to top)..... *131*

Fig. 4.24 Propagation delay versus external load capacitance. The propagation delays of a CMOS inverter are obtained from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. The input transition times are 1ps, 10ps, 100ps, 200ps, and 400ps, respectively (from bottom to top)..... *132*

Fig. 4.25 The correlation of propagation delays from the comprehensive timing model based on PDSMM in Subsection 4.5.2 with propagation delays from SPICE BSIM3. The dashed lines indicate ±5% of propagation delays deviating from

SPICE BSIM3. All the data match with each other excellently. *133*

Fig. 5.1 A four-nMOSFET series-connected chain. The numbering of nMOSFET and its corresponding input starts from the nMOSFET connected to ground as number one. *158*

Fig. 5.2 The static I–V characteristics of 'n41' from SPICE. The square-law of chain current versus input voltage is valid when input voltage is small and close to threshold voltage. However, the square-law fails at input voltage close to operating voltage where the chain current strongly saturates as input voltage increases. *159*

Fig. 5.3 The static I–V characteristics of 'n42' from SPICE. The square-law of chain current versus input voltage is valid when input voltage is small and close to threshold voltage. However, the square-law fails at input voltage close to operating voltage where the effect of chain current saturation is pretty large as input voltage increases. *160*

Fig. 5.4 The static I–V characteristics of 'n43' from SPICE. The square-law of chain current versus input voltage is valid when input voltage is small and close to threshold voltage. However, the square-law fails at input voltage close to operating voltage where the effect of chain current saturation is still large as input voltage increases. *161*

Fig. 5.5 The static I–V characteristics of 'n44' from SPICE. The square-law of chain current versus input voltage is valid when input voltage is small and close to threshold voltage. However, the square-law slightly fails at input voltage close to operating voltage where the effect of chain current saturation is visible as input voltage increases. *162*

Fig. 5.6 The static I–V characteristics of 'n41234' from SPICE. They are similar with the static I–V characteristics of a deep submicron MOSFET. *163*

Fig. 5.7 The dependence of chain current I_n on input voltage V_{in} when output voltage V_{out} is equal to V_{dd} for 'n41', 'n42', 'n43', 'n44', and 'n11' (from top to bottom). The deep submicron nMOSFET of 'n11' has the same equivalent width of the series-connected chains of 'n41', 'n42', 'n43', and 'n44'. We can clearly see the effect of chain current saturation in cases like 'n41' and 'n42' as input voltage increases. The current amplitude of an nMOSFET with the same equivalent width of a series-connected MOSFET chain is smaller than that of the chain currents. In other words, the speed of logic NAND gate is faster than an inverter with the same equivalent width. *164*

Fig 5.8 Power index as a function of input voltage from (3.14) for 'n41', 'n42', 'n43', 'n44', and 'n11', respectively (from bottom to top). Power index of a deep submicron MOSFET of 'n11' ranges between one and two. Power indices of

series-connected MOSFET chains are close to two when input voltage is around threshold voltage. They decrease when input voltage increases. Their values are even lower than one at input voltage close to operating voltage. 165

Fig. 5.9 The static I–V characteristics of 'n41' from SPICE and SMCM. From the figure, SMCM can accurately reproduce the peculiar I–V static characteristics of 'n41'. The chain current strongly saturates as input voltage increases..... 166

Fig. 5.10 The static I–V characteristics of 'n42' from SPICE and SMCM. From the figure, SMCM can accurately reproduce the peculiar I–V static characteristics of 'n42'. The effect of chain current saturation is pretty large as input voltage increases. 167

Fig. 5.11 The static I–V characteristics of 'n43' from SPICE and SMCM. From the figure, SMCM can accurately reproduce the peculiar static I–V characteristics of 'n43'. The effect of chain current saturation is still large as input voltage increases. 168

Fig. 5.12 The static I–V characteristics of 'n44' from SPICE and SMCM. From the figure, SMCM can accurately reproduce the peculiar static I–V characteristics of 'n44'. The effect of chain current saturation is visible as input voltage increases. 169

Fig. 5.13 The static I–V characteristics of 'n41234' from SPICE and SMCM. From the figure, SMCM can accurately reproduce the static I–V characteristics of 'n41234'..... 170

Fig. 5.14 Voltage waveforms of real input and suppressed input of a series-connected MOSFET chain. Before the activation time t_a, the real input is suppressed in the suppressed input. After the activation time t_a, the suppressed input is the same as the real input. 171

Fig. 5.15 Input pattern to generate fast initial state propagation delay (FISPD) of a series-connected nMOSFET chain. The order of turning nMOSFETs off makes parasitic capacitances of intermediate nodes have the minimized initial charges, which give rise to fast ISPD..... 172

Fig. 5. 16 Input pattern to generate slow initial state propagation delay (SISPD) of a series-connected nMOSFET chain. The order of turning nMOSFETs off makes parasitic capacitances of intermediate nodes have the maximized initial charges, which give rise to slow ISPD. 173

Fig. 5.17 Output voltage waveforms of 'n413' due to fast and slow initial states (FIS and SIS). The effect of initial states of intermediate nodes is clearly shown in this figure. The difference of the waveforms of FISPD and SISPD is caused by initial charges of parasitic capacitances of effective intermediate nodes of a series-

connected MOSFET chain. *174*

Fig. 5.18 Chain current waveforms of 'n414' for a step input with of fast and slow initial states (FIS and SIS). The step-input activation times for FIS and SIS are measured from the time when input voltage jumps to the time when chain current reaches the half of steady-state current. *175*

Fig. 5.19 Activation time of 'n413' as a function of input transition time. Given SIS and FIS step-input activation times, the SIS and FIS ramp-input activation times for input transition time are calculated. The charge conservation and SMCM are assumed in the derivation. The difference between SIS and FIS is caused by initial charges of parasitic capacitances of intermediate nodes between $nMOSFET_1$ *and* $nMOSFET_3$*. The difference between FIS and the threshold line is caused by initial charges of parasitic capacitances of intermediate nodes between* $nMOSFET_3$ *and output node.* *176*

Fig. 5.20 Activation time of 'n41234' as a function of input transition time. The difference between SIS and FIS is caused by initial charges of parasitic capacitances of intermediate nodes between $nMOSFET_1$ *and* $nMOSFET_4$*. There is no difference between FIS and the threshold line due to no intermediate nodes between* $nMOSFET_4$ *and output node.* *177*

Fig. 5.21 Activation time of 'n41' as a function of input transition time. There is no difference of activation times of SIS and FIS since there is only one active input in n41. There is no effective intermediate node in 'n41'. *178*

Fig. 5.22 Activation time of 'n44' as a function of input transition time. There is no difference of activation times of SIS and FIS since there is only one active input in n44. There is no effective intermediate node in 'n44'. *179*

Fig. 6.1 A four-input NAND gate. The numbering of an nMOSFET and its corresponding input starts from the nMOSFET connected to ground as number one. Output load capacitance C_{load} *is shown. The possible input-output coupling capacitances of* C_{mn4}, C_{mp1}, C_{mp2}, C_{mp3}, *and* C_{mp4} *are also plotted.* *240*

Fig. 6.2 Output voltage waveforms of a NAND4234 gate with the active inputs of V_{i2}, V_{i3}, *and* V_{i4} *for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. There is a pair of traces for the external load capacitance of 50fF: the left one for the fast initial state (FIS) and the right one for the slow initial state (SIS). We can see an obvious mesa in the voltage waveform for slow initial state (SIS). There is a voltage overshoot, but not a mesa in the fast initial state (FIS). Both phenomena are well explained by the modified input suppression technique in Section 5.5.* *241*

Fig. 6.3 Output voltage waveforms of a NAND41 gate with the active input of V_{i1} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference. 242

Fig. 6.4 Output voltage waveforms of a NAND423 gate with the active inputs of V_{i2}, and V_{i3} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively (from left to right). There is a pair of traces for each external load capacitance: the left one for fast initial state (FIS) and the right one for slow initial state (SIS). A mesa exists in the voltage waveforms for both slow initial state (SIS) and fast initial state (FIS). Both are accurately explained by the modified input suppression technique in Section 5.5. 243

Fig. 6.5 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. Note that the left family of traces is for fast initial state (FIS) and right one for the slow initial state (SIS). For each family the external capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively (from left to right). Note there is a mesa in the voltage waveforms for slow initial state (SIS). No mesa exists in fast initial state (FIS). Both are explained by the modified input suppression technique in Section 5.5. 244

Fig. 6.6 Seven operation regions of a complex NAND434 gate modeled with a series-connected MOSFET chain 'n434' with the active inputs of V_{i3} and V_{i4} for a rising ramp input. The typical discharging traces from SPICE BSIM3 are also plotted. Each pair of traces corresponds to the same external load capacitance for fast initial state (FIS) and slow initial state (SIS). 245

Fig. 6.7 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} for an external load capacitance of 10fF. The NAND42 gate is modeled with a series-connected MOSFET chain 'n42'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference. 246

Fig. 6.8 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} for an

external load capacitance of 100fF. The NAND42 gate is modeled with a series-connected MOSFET chain 'n42'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference. *247*

Fig. 6.9 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} for an external load capacitance of 400fF. The NAND42 gate is modeled with a series-connected MOSFET chain 'n42'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference. *248*

Fig. 6.10 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for an external load capacitance of 10fF. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots are exaggerated with no trivial parallel pMOSFET cluster. *249*

Fig. 6.11 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for an external load capacitance of 100fF. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots are exaggerated with no trivial parallel pMOSFET cluster. *250*

Fig. 6.12 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for an external load capacitance of 400fF. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state

(SIS). The reduction of voltage overshoots is caused with large external load capacitance. *251*

Fig. 6.13 Propagation delays of FIS versus external load capacitances for a NAND42 gate. The NAND42 gate is modeled with a series-connected MOSFET chain 'n42'. The propagation delays from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the propagation delays of SIS are the same. *252*

Fig. 6.14 Propagation delays of FIS versus external load capacitances for a NAND4124 gate. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The propagation delays from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the activation time of FIS of a NAND4124 is zero. *253*

Fig. 6.15 Propagation delays of SIS versus external load capacitances for a NAND4124 gate. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The propagation delays from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). The part of error in propagation delays is caused with the exaggerated voltage overshoots. *254*

Fig. 6.16 The correlation of derived propagation delay of fast initial state (FIS) from SMCM in Section 6.3 with propagation delay from SPICE BSIM3. The NAND4 gate is modeled with a series-connected MOSFET chain. There are 375 simulation points which cover all the 15 sorts of active inputs. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively. *255*

Fig. 6. 17 The correlation of derived propagation delay of slow initial state (SIS) from SMCM in Section 6.3 with propagation delay from SPICE BSIM3. The NAND4 gate is modeled with a series-connected MOSFET chain. There are 375 simulation points which cover all the 15 sorts of active inputs. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively. *256*

Fig. 6.18 Eleven operation regions of a complex NAND4. The typical discharging traces from SPICE BSIM3 are also plotted. Each pair of traces corresponds to the same external load capacitance for fast initial state (FIS) and slow initial state (SIS). *257*

Fig. 6.19 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} for an external load capacitance of 10fF. The voltage waveforms from SMCM in

Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference. *258*

Fig. 6.20 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} *for an external load capacitance of 100fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference.* *259*

Fig. 6. 21 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} *for an external load capacitance of 400fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference.* *260*

Fig. 6. 22 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, *and* V_{i4} *for an external load capacitance of 10fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots, including a mesa for fast input, are well compensated with a trivial parallel pMOSFET cluster.* *261*

Fig. 6. 23 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, *and* V_{i4} *for an external load capacitance of 100fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots, including a mesa for fast input, are well compensated with a trivial parallel pMOSFET cluster.* *262*

Fig. 6.24 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, *and* V_{i4} *for an external load capacitance of 400fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and*

400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots, including a mesa for fast input, are well compensated with a trivial parallel pMOSFET cluster. *263*

Fig. 6.25 Propagation delays of FIS versus external load capacitances for a NAND42 gate. The propagation delays from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the propagation delays of SIS are the same. *264*

Fig. 6.26 Propagation delays of FIS versus external load capacitances for a NAND4124 gate. The propagation delays from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the activation time of FIS of a NAND4124 is zero. *265*

Fig. 6.27 Propagation delays of SIS versus external load capacitances for a NAND4124 gate. The propagation delays from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top)..... *266*

Fig. 6.28 The correlation of proposed propagation delay of fast initial state (FIS) from SMCM in Section 6.4 with propagation delay from SPICE BSIM3. There are 375 simulation points which cover all the 15 sorts of active inputs. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively. *267*

Fig. 6.29 The correlation of proposed propagation delay of slow initial state (SIS) from SMCM in Section 6.4 with propagation delay from SPICE BSIM3. There are 375 simulation points which cover all the 15 sorts of active inputs. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively. *268*

Fig 7.1 Input classification of an arbitrary input pattern in input mapping technique. There are three kinds of inputs – active, partially-active, and inactive. *291*

Fig 7.2 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for FIS. The four dotted lines indicate the real input pattern given in case one of Table 7.2. The dashdot line denotes the mapped input. *292*

Fig 7.3 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for FIS. The four dotted lines indicate the real input pattern given in case two of Table 7.2. The dashdot line denotes the mapped input. *293*

Fig 7.4 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for FIS. The four dotted lines indicate the real input pattern given in case three of Table 7.2. The dashdot line denotes the mapped input. 294

Fig 7.5 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for FIS. The four dotted lines indicate the real input pattern given in case four of Table 7.2. The dashdot line denotes the mapped input. 295

Fig 7.6 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for SIS. The four dotted lines indicate the real input pattern given in case one of Table 7.3. The dashdot line denotes the mapped input. 296

Fig 7.7 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for SIS. The four dotted lines indicate the real input pattern given in case two of Table 7.3. The dashdot line denotes the mapped input. 297

Fig 7.8 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for SIS. The four dotted lines indicate the real input pattern given in case three of Table 7.3. The dashdot line denotes the mapped input. 298

Fig 7.9 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for SIS. The four dotted lines indicate the real input pattern given in case four of Table 7.3. The dashdot line denotes the mapped input. 299

Fig 7.10 The correlation of propagation delays of NAND4 from SPICE BSIM3 and from the input mapping technique for FIS. The real input patterns are given in Tables 7.4-7.7. The external capacitances are 100fF and 400fF, respectively. The dashed lines denote ±10% deviation from SPICE results. 300

Fig 7.11 The correlation of propagation delays of NAND4 from SPICE BSIM3 and from the input mapping technique for SIS. The real input patterns are given in Tables 7.8-7.11. The external capacitances are 100fF and 400fF, respectively. The dashed lines denote ±10% deviation from SPICE results. 301

Fig. III.1 Q-function from new approach and numerical approach. It shows that the new approach is very accurate to describe Q-function. 327

CHAPTER 1 INTRODUCTION

In recent decades, very large-scale integrated circuit (VLSI) industry has experienced astounding revolution. It brings about electronics, computer, communication, and semiconductor to become one of dominant industries with a high growth rate. During the 1990s, semiconductor fabrication technology has rapidly developed into very deep submicron area and the complexity of designing digital circuits has increased dramatically. VLSI digital circuits due to advanced fabrication processing technologies and robust designs are appearing in high performance with large die size and high integration density. Much of this growth has been in complementary metal-oxide semiconductor (CMOS) technology. High performance, high integration density with small feature size, complex functions, low power consumption, low cost, and short design-to-market time are the primary industry requirements. Deep submicron CMOS circuits able to meet these industry requirements have proven to be an ideal technology for the rapid expansion of VLSI digital circuits.

Due to the difficulty of designing deep submicron VLSI CMOS digital circuits as well as the expensive costs of design and fabrication, effective electronic design automation (EDA) tools are necessary to implement, optimize, and verify deep submicron VLSI CMOS designs in every design stage from architecture, logic design, circuit realization to physical implementation. Design verification and optimization have to be used to guarantee that implemented VLSI CMOS circuits achieve their desired functions and performances [43]. One of main tasks in design verifications (DV) is timing verification to ensure that all timing paths meet specified timing constraints [33].

Regarding VLSI logic design, speed and complexity of deep submicron CMOS digital circuits have increased rapidly. Logic designers have to make sure that designed CMOS circuits meet all the functional specifications. Designs are captured with register transfer level (RTL) codes by using hardware description languages (HDL) such as Verilog. EDA tools take care of implementations from RTL codes to layouts with the following steps. First logic synthesis tools generate pre-layout gate-level netlists from RTL codes. These pre-layout gate-level netlists need to do timing verifcation. Once pre-layout gate-level netlists satisfy timing constraints, timing-driving place and route tools (P&R) are used to implement physical designs. Again, post-layout extracted netlists need to meet timing constraints. If designed CMOS digital circuits do not satisfy timing constraints, they have to be optimized or redesigned [94].

Timing verification is one of the most difficult processes in VLSI engineering. Timing verification of a CMOS digital circuit generally includes not only maximum timing verification but minimum timing verification also. Timing constraints [95] of maximum timing verification for every internal timing path in CMOS digital circuits are

$$T_{ckq} + T_{com\_slow} + T_{wire} + T_{setup} + T_{ck\_skew} < T_{ck} \tag{1.1}$$

In (1.1), T_{ckq} is the clock-to-q time of driving register; T_{com_slow} is the slow-corner propagation delay of combination logic circuits in the timing path; T_{wire} is the interconnection wire delay in the timing path; T_{setup} is the setup time of receiving register; T_{ck_skew} is the clock skew; T_{ck} is the period of maximum specified frequency. From (1.1) maximum timing verification of a CMOS digital circuit is associated with the performance of a CMOS digital circuit determined by the longest signal propagation time

of all timing paths in a CMOS digital circuit. In order to achieve the performance of a CMOS digital circuit, it is necessary to optimize the CMOS digit circuit such that propagation delays for all timing paths satisfy the given maximum timing constraints. Timing constraints [95] of minimum timing verification for every internal timing path in CMOS digital circuits are

$$T_{ckq} + T_{com\_fast} + T_{wire} > T_{hold} + T_{ck\_skew} \tag{1.2}$$

In (1.2) T_{com_fast} the fast-corner propagation delay of combination logic circuits in the path; T_{hold} the hold time of receiving register. Minimum timing verification is associated with the hold timing issues determined by the shortest signal propagation delays of all the timing paths in a CMOS digital circuit. It becomes more and more important for deep submicron CMOS digital design due to faster logic gates. In timing verification there are huge amount of various timing paths to be verified in a CMOS VLSI design. Fast and accurate timing approaches to evaluate CMOS logic gates of fast and slow corners are needed for VLSI digital designers to verify and optimize their CMOS designs.

Our objective in this research is to develop simple and accurate analytical timing models for deep submicron CMOS logic gates. Obviously the effects of input transition time and external load capacitance on propagation delays are included in the timing models. Furthermore, the important second-order MOSFET effects such as charge carrier velocity saturation, gate-drain coupling, and channel length modulation are also incorporated into the timing models. In addition to these factors, we take into account the effects of an arbitrary input pattern and initial state of parasitic capacitances of intermediate nodes in the propagation delay models for a single-stage complex CMOS

logic gate. The trade-offs between accuracy and simplicity are addressed. The developed comprehensive timing models can be well suited for VLSI timing verification and optimization of deep submicron CMOS logic gates such as logic level timing verifiers and simulators, logic synthesis and optimization tools, place and route tools, and timing closure tools [33].

The contents of the investigation are arranged as follows. The background and related researches in timing analyses and models for CMOS logic gates are introduced in Chapter 2. The important second-order effects of a deep submicron MOSFET are introduced in Chapter 3. Then the deep submicron MOSFET model (DSMM) and the proper deep submicron MOSFET model (PDSMM) are proposed to accurately describe the static I–V characteristics of a deep submicron MOSFET. In Chapter 4 the propagation delay models for a deep submicron CMOS inverter are developed based on the PDSMM. In Chapter 5 we discuss the characteristic properties of a series-connected MOSFET chain with an equal input pattern. A generic series-connected MOSFET chain model (SMCM) is proposed to accurately describe the peculiar static I–V characteristic of a series-connected chain with an equal input pattern. An input suppression technique is modified to depict the effect of initial states of parasitic capacitances of intermediate nodes on propagation delays. In Chapter 6 the propagation delay models are developed for a single-stage complex CMOS logic gate with an equal input pattern. The effect of an arbitrary input pattern is presented by using an input mapping method in Chapter 7. The summary of this research is given in Chapter 8. Except SPICE simulations, MATLAB is used for calculating and plotting in this research [6] [64].

CHAPTER 2 BACKGROUND AND RELATED RESEARCHES

2.1 BACKGROUND OF CMOS DIGITAL CIRCUITS

2.1.1 Dominance of CMOS Technology

During the recent decades, digital CMOS technology emerged as the major technology for VLSI applications including high performance microprocessors and application specific integrated circuits (ASICs). This trend continues as digital CMOS technology has crowded out less competitive technologies such as bipolar and NMOS [3] [24]. Bipolar digital logic lost the battle in the digital design world mainly because of high power consumption per gate. High power consumption per gate in bipolar technology limits the number of gates, which can be reliably integrated on a single die or package. Digital CMOS technology becomes the most important primarily due to the low static power consumption of a CMOS digital circuit [2] [24] [104]. Aside from small leakage current, power in a CMOS digital circuit is dissipated only during transitions. When a CMOS digital circuit is in a non-switching stable state, virtually no power is consumed if small leakage current is ignored. This is a major improvement for a CMOS digital circuit over bipolar and NMOS circuits, where considerable power is dissipated even in the non-switching stable state. Although dynamic power consumption concerns are rapidly becoming influential in CMOS digital design as well, there does not seem to be a new technology around the corner to alleviate the problem.

Another significant feature of a CMOS digital circuit is its stability of operation [2] [104]. For bipolar and NMOS technologies, small deviations of device characteristics may cause circuits to work improperly. In CMOS technology, small variations in process,

temperature and voltage (PTV) usually do not disturb the operation of a CMOS digital circuit. The noise margin of a CMOS digital circuit is better than bipolar and NMOS since CMOS output voltage can go to power or ground. The static I–V characteristics of a CMOS digital circuit is better than bipolar and NMOS also.

Third, digital CMOS technology has advanced due to the availability of simple scaling laws because of the stability of digital CMOS operation [4] [35] [44] [51] [104]. These scaling laws have served as an outline for the reduction of MOSFET dimensions. By combining these scaling laws with the inherently low power dissipation of CMOS logic gates, extremely high integration density of CMOS logic gates has been made possible.

Finally, the CMOS fabrication technology has proven to be a robust and sophisticated [96]. The recent continual improvements of CMOS manufacture process technology including photolithography equipments enable more MOSFETs with smaller feature size to be fabricated in a larger single die with reasonably good yield. The sophisticated photolithographic technique, simple scaling laws and robust CMOS operation make the feature size of a CMOS digital gate reduced steadily. Reducing the feature size continually improves device and circuit performance. Smaller CMOS logic gates consume less power and take up less space. The number of MOSFETs that can be designed and manufactured in a single IC chip has grown enormously, now approaching more than hundreds of millions.

As a result, CMOS technology has reached the appealing position. CMOS fabrication technique is very flexible, and a single basic CMOS process is easily employed for a wide variety of applications. VLSI CMOS circuits proliferate in a growing

variety of applications with amazing diversity. For example, CMOS applications include (but are certainly not limited to) digital logic, static RAMs, dynamic RAMs, digital signal processing (DSP), high-performance microprocessors (CPUs) and a variety of analog functions. The need to serve this diversity is adding new demands to CMOS technology. In addition to higher and higher frequency, logic functionality of integrated circuits is becoming more complicated than before. CMOS designers integrate more functions onto a single IC chip, which previously were distributed across several IC chips and combined at the board level. Combining the significant improvements of CMOS technology with sophisticated electronic design automation (EDA) tools, the trend of system-on-a-chip (SOC) has emerged.

2.1.2 Moore's Law

Integration density and performance of digital MOS VLSI circuits have dramatically increased for the past decades. In the 1960s, Gordon Moore [68] predicted that the number of transistors that can be integrated on a single die would grow exponentially with time. This predication, called as Moore's law, has proven to be amazingly visionary until now. As can be observed from [58], integration complexity doubles approximately every 1 to 2 years and memory density has increased by more than a thousand fold since 1970. Let's take the microprocessor as an example. From its inception in the early 70s, the microprocessor has grown in performance and complexity at a steady and predictable pace. The barrier of million-transistor per chip was crossed in the later 80s. Now the number of transistors in one high performance microprocessor exceeds one hundred millions. Clock frequencies double every three years and have reached into several gigahertz range. An even more important observation is that, as of

now, these trends have not shown any signs of a slow-down.

The high integration density and small feature size is driven by the sophisticated process technology. New generation of process technology emerges around every three years [44]. Recently the gate lengths of 65nm and 90nm are widely used in high performance microprocessor design. New techniques like copper interconnection greatly improve the chip performance and reduce the power consumption.

As CMOS process technology goes down to deep submicron, second order effects of short channel MOSFETs become important. Shockley MOSFET model that can accurately describe long channel MOSFET characteristics does not predict the deep submicron MOSFETs well. Recently, several new MOSFET models are proposed to describe short channel effects. One of them is well-known Berkeley short-channel IGFET model (BSIM) [2] [34] [35] [53] [63] [66], which is the physics-based, deep-submicron MOSFET model for digital and analog circuit designs developed at the University of California at Berkeley. It is a complicated model with more than one hundred model parameters. It is only used in numerical solution of CAD circuit simulators like SPICE. It is impossible to use BSIM model in analytical analysis in VLSI design.

2.1.3 Digital Design Hierarchy

Instead of earlier individualized device design and optimization approach, CMOS VLSI digital circuits are constructed in a hierarchical way. Chip-level module is partitioned into a collection of modules or blocks, each of which consists of a number of modules, or blocks, or cells on its own. Cells are reused as much as possible to reduce the design effort and to enhance the chances for a first-time-right implementation. In fact this hierarchical approach is the key factor for the success of modern CMOS VLSI digital

circuit design.

The essential concept in hierarchical approach is abstraction, which is the most important in dealing with the complexity issue. At each design level, the internal details of a complex module or block are abstracted away and replaced by a black box view or model. This model of abstraction contains virtually all the information needed to deal with the block at the next higher hierarchical level. For instance, once a circuit designer has implemented a multiplier module, it is considered a black box with known and necessary characteristics for chip-level integration. Since there exists no need to look inside this black box, design complexity at chip-level integration is substantially reduced. The impact of this divide-and-conquer approach is extraordinary for complicated deep submicron VLSI design and verification. Instead of having to deal with countless elements, designers have to consider only a few components, each of which are characterized in performance, function and cost by a small number of parameters.

This hierarchical design approach affects the timing methodologies of digital VLSI design. With good partition of design [79], most of internal critical paths within a block are gate-dominated where the interconnect effects can be ignored. The inter-block critical paths are wire-dominated. The timing analyses and models for these two kinds of paths are different. Most of timing paths in a design are internal gate-dominated paths while a few of wire-dominated paths usually determine the design performance.

2.1.4 Timing Verification of CMOS Digital Circuits

Due to the difficulty of designing high integration density of MOSFETs with small feature size, complex VLSI designs have to be verified. With thousands or millions of transistors, effective electronic design automation (EDA) tools are necessary in every

phase of implementation, verification, and optimization of digital VLSI designs. VLSI circuits have to be designed, evaluated, redesigned and verified by using computer simulations and verifications before designs go to fabrication. Design verification (DV) is used to guarantee that designed circuits perform specified functions. Design verification (DV) includes three main parts: functional design verification (FDV), physical design verification (PDV), and timing verification. Functional design verification (FDV) is to make sure that designed CMOS circuits meet all functional specifications and requirements. The physical design verification (PDV) is to guarantee circuit realizations and implementations to meet physical design constraints such as design rule checking (DRC), layout versus schematic (LVS), etc. Time verification is to ensure that all the timing paths meet specified timing constraints such as (1.1) and (1.2). If an implemented circuit does not satisfy the requirements or constraints, it needs to be redesigned or optimized [94]. Design optimizations driven by performance or power [101] or die size [20] are the important topics in VLSI design [28] [33] [37] [43]. The performance-driven tools are based on timing models of CMOS logic gates. Since there are huge number of timing paths in a VLSI circuit it is impossible to use comparatively slow SPICE in EDA flows, other methods to calculate timing delays of CMOS circuits have to be found and used in EDA flows. Timing models of CMOS logic gates in deep submicron VLSI design are the topic of the paper.

2.2 RELATED RESEARCHES IN TIMING METHODOLOGIES OF CMOS DIGITAL CIRCUITS

Propagation delay is one of the most critical performance parameters in VLSI

CMOS digital circuits. Basically there are several main methods to investigate propagation delay of CMOS digital circuits. The first one is numerical approaches like transistor level circuit simulators such as SPICE [2] [34] [37] [63] [66] [90] [100]. These numerical methods usually provide accurate and detailed transient output response and are suited for small circuit designs such as megacells or IOs or critical path simulations. However, numerical approaches are usually time-consuming and need a lot of hardware supports such as memory. Therefore these slow numerical methods are difficult to directly be incorporated into full chip timing verification or logic level timing simulators when there are over millions of transistors involved.

Analytical approaches of propagation delays, as opposed to numerical approaches, make more evident the functional relationship between design objectives and various device parameters. These analytical expressions allow for the insight of further improvement and optimization. Analytical expressions of timing models of CMOS logic gates are obtained by solving for transient response from differential equations describing CMOS logic gates. One of significant advantages of these analytical approaches of propagation delays is that they usually are faster to calculate propagation delay than numerical approaches by at least two orders of magnitude [31] [70]. These analytical expressions can easily be incorporated in CMOS design flows.

In the third way resistance-capacitance (RC) approaches make the delay calculation much easier and faster with reasonably accuracy. In RC approaches, all the circuit components including MOSFETs are considered to be linear. For instance, MOSFETs are approximated as linear effective resistances. With this assumption, nonlinear CMOS circuits are simplified into linear RC networks. The propagation delays

are estimated through the solution of linear RC networks. RC circuit approaches were used to model the average circuit behaviors since the nonlinear behaviors of MOSFET logic gates are not well represented by linear resistances.

Finally, table lookup techniques provide a direct and efficient way to obtain propagation delays. However, they are based on presimulation data that need lot of time to prepare. Besides that interpolation may give rise to errors [21].

2.3 NUMERICAL APPROACHES OF TIMING ANALYSIS

Numerical approaches provide a way to accurately simulate and verify circuits in VLSI design. One of numerical approaches to circuit design is SPICE circuit simulator [2] [34] [37] [63] [66] [90] [100]. The acronym SPICE stands for Simulation Program with Integrated Circuit Emphasis, reflecting its origin as a method of dealing with the enormous complexity inherent in IC design. The original SPICE core was developed at the University of California at Berkeley in the late 1960s, as a successor to a first attempt at a computer-aided circuit simulator with the quite horrible name of CANCER (Computer Analysis of Nonlinear Circuits, Excluding Radiation) [73].

In SPICE, a circuit is treated in a node/element fashion and depicted as a collection of various elements (e.g., resistor, capacitors, etc.) and numbered or named nodes to which each particular element is connected. Nodal analysis in circuit theory is used in SPICE analysis. SPICE representation of a circuit with n nodes can be thought of as an nxn matrix. If two nodes are connected by some element, the corresponding matrix element is nonzero with the property of that element. By defining state variables such as the nodal voltage at each internal node, and fixing the state variables at external nodes

such as the power supply, ground, etc., the matrix is solved to determine the values of the state variables at the various internal circuit nodes. To reach a solution, specific SPICE model to any circuit element is required. The circuit can be thought of as the main program, while the element models are subroutines that are called when the behavior of an element is evaluated.

SPICE circuit simulator can produce the most detailed and accurate timing information. It has been extensively employed for circuit designers to design and verify small circuits. Basically, each circuit element is described in its SPICE model. With circuit element models and circuit topology depicted in circuit netlist, the accurate waveforms of circuit nodal voltages and branch currents can be obtained from SPICE simulations. The precise timing information can be measured from the relevant transient waveforms. Because of its accuracy and widespread adoption, SPICE has been used as a criterion or standard for circuit performance and other timing analysis methodologies in the integrated circuit industry. The drawback for SPICE simulation is that the analysis requires substantial hardware resources like storage and CPU computation time. Therefore, this numerical method of circuit simulation with accurate results is suitable for comparatively small and specialized circuits like megacells or IOs or critical path simulations but is inconvenient for larger block of circuits. This numerical method is difficult to directly be incorporated into full chip timing verification and logic level timing simulators due to hardware demands. One way to overcome the drawbacks of SPICE is to partition the whole circuits to many small subcircuits (stages) defined by channel connection as in timemill/powermill tools [95]. Then numerical circuit transient response can be obtained for each small circuit with reasonable errors from SPICE

results. This modified numerical method still takes too much time to be incorporated into digital design flow.

2.4 ANALYTICAL APPROACHES OF CMOS TIMING MODELS

In analytical approaches, propagation delay of a CMOS logic gate is determined by solving for analytical transient response from differential equation describing the CMOS logic gate. One of significant advantages of analytical approaches is that they are faster to obtain propagation delays than numerical approaches by at least two orders of magnitude [21] [31] [70] [89]. These analytical delay models can easily be incorporated in CMOS design flows such as logic-level timing simulators, timing verification and optimization, performance-driven synthesis optimization, performance-driven place and routing tools, gate sizing, delay minimization, etc [20] [33].

In general, a CMOS logic gate as shown in Fig. 2.1 is compound of two dual logic components, i.e. pMOSFET and nMOSFET logic tree. The differential equation to describe transient characteristics can be obtained from the application of the Kirchoff's current law at output node [55]. Consequently, The differential equation is

$$\frac{dV_o(t)}{dt} = \frac{C_m}{C_{load} + C_m} \cdot \frac{dV_i(t)}{dt} + \frac{1}{C_{load} + C_m} \cdot (I_p - I_n) \tag{2.1}$$

This is the general differential equation used in analytical approaches. In addition to gate-drain coupling capacitance C_m and load capacitance C_{load} of output node, the transient output response $V_o(t)$ depends on input waveform $V_i(t)$ and drain current models for nMOSFET tree and pMOSFET tree. For simplicity input voltage $V_i(t)$ is usually either a step function or a ramp function in analytical approaches. The drain current models for

MOSFETs should be simple and accurate enough to be used in the analytical solution of (2.1). There are a few popular MOSFET models used in analytical approaches. In the following subsections, we will briefly review some important MOSFET models and their derived propagation delay models.

2.4.1 Shockley MOSFET Model

Shockley [90] first proposed the MOSFET model for long channel MOSFETs. In Shockley MOSFET model, the drain current I_{ds} is expressed as

$$I_{ds} = \begin{cases} 0 & V_{gs} \le V_{th} \\ k \cdot \left(V_{sat} - \frac{V_{ds}}{2} \right) \cdot V_{ds} & V_{gs} > V_{th}, V_{ds} \le V_{sat} \\ \frac{1}{2} \cdot k \cdot {V_{sat}}^2 & V_{gs} > V_{th}, V_{ds} > V_{sat} \end{cases} \tag{2.2}$$

where V_{sat} is the drain-source saturation voltage of a MOSFET given by

$$V_{sat} = V_{gs} - V_{th} \tag{2.3}$$

In (2.2) and (2.3), V_{th} is the threshold voltage of a MOSFET. In (2.2) the coefficient k is the drivability factor or transconductance parameter of a MOSFET, which is

$$k = \mu \cdot \left(\frac{\varepsilon_{sio2}}{t_{sio2}} \right) \cdot \left(\frac{W}{L_{eff}} \right) \tag{2.4}$$

where μ denotes effective carrier mobility, ε_{SiO_2} dielectric constant of gate oxide, t_{SiO_2} gate oxide thickness, W channel width, and L_{eff} effective channel length.

Shockley MOSFET model was developed for long channel MOSFETs [83] [90]. Each parameter of this model has its physical meanings. This simple model is widely used in treating long channel MOSFET circuits analytically. Many derived formulas

based on the model such as CMOS timing model and CMOS short-circuit current or power have been extensively used in long-channel VLSI designs [13] [45] [52] [65] [104] and early versions of SPICE circuit simulator [72].

Shockley MOSFET model cannot accurately reproduce static I–V characteristics of a deep submicron MOSFET since the second-order effects of a short channel MOSFET are ignored in the model. Figs 2.2-2.5 show static characteristics of a deep submicron MOSFET from SPICE BSIM3 model and from Shockley MOSFET model. There are several main discrepancies for deep submicron MOSFET behaviors. The first one is that drain current I_{ds} in saturation mode of a deep submicron MOSFET illustrated in Figs. 2.3 and 2.5 does not show square-law dependence on gate-source voltage V_{gs} as in (2.2). The second is that drain saturation voltage V_{sat} of a deep submicron MOSFET is deviated from the linear expression expected in (2.3), especially when gate-source voltage is large as shown in Fig. 2.4. The third inconsistency is that drain saturation current I_{ds} of a deep submicron MOSFET with a given V_{gs} is much less than the predicted value of Shockley MOSFET model as shown in Figs. 2.2-2.3. The fourth disagreement is that the effect of channel length modulation is not taken into account in Shockley MOSFET model. The main reason for the above first three discrepancies is that Shockley MOSFET model does not include the effect of charge carrier velocity saturation occurring in a deep submicron MOSFET. In other words, the effect of charge carrier velocity saturation in high electric field along channel associated with short channel length has to be considered in deep submicron MOSFET model. Besides that, the channel length modulation should be included in deep submicron MOSFET model also.

2.4.2 *Step Input Delay Model with Shockley MOSFET model*

Burns [13] published the first closed-form delay expression and delay expressions of CMOS inverters. Based on Shockley MOSFET model, Burns derived an analytical formula for the output response of a CMOS inverter driven with a step input. The propagation delay of a CMOS inverter was calculated from the analytical output response. It is well known that the propagation delay of output falling may be written

$$t_{dHLStep} = \frac{C_t}{k_n V_{dd}(1-u_{th})}\left\{\frac{2u_{th}}{1-u_{th}} + \ln\left[4(1-u_{th})-1\right]\right\} \tag{2.5}$$

where k_n is transconductance parameter, C_t is total load capacitance of output node, V_{dd} is operating voltage, and u_{tn} is the normalized threshold voltage $u_{tn} = V_{th} / V_{dd}$.

The propagation delay model given in (2.5) is insufficiently accurate, especially for a deep submicron CMOS inverter, since it does not take into account the influence of the input waveform on the propagation delay. Furthermore, as discussed in Section 2.4.1 Shockley MOSFET model used in Burns's derivation is not accurate enough to describe the static *I–V* characteristics of a deep submicron MOSFET.

2.4.3 *Delay Models Considering Input Waveform*

Hedenstierna and Jeppson [40] presented analytical expressions for the output waveform and propagation delay including the effect of input waveform slope. The influences of the short-circuit current and gate-to-drain coupling capacitance were ignored in their analysis. Shockley MOSFET model was used in their analysis. They separated input ramp into fast input ramp or slow input ramp according to whether the MOSFET is still saturated or not when input voltage ramp reaches operating voltage. For fast input ramp, the total propagation delay is

$$t_{dHL} = t_{dHLStep} + \frac{\tau}{6} \cdot (1 + 2 \cdot u_{th}) \tag{2.6}$$

where $t_{dHLStep}$ is the step-input response delay given in (2.5), τ is the rising time of input voltage ramp, and u_{tn} is the normalized threshold voltage $u_{tn} = V_{th} / V_{dd}$. Equation (2.6) clearly shows that the propagation delay does depend on the input slope. For slow input ramp, the propagation delay can be obtained by

$$t_{dHL} = t\Big|_{V_{out}=0.5V_{dd}} - \frac{\tau}{2} \tag{2.7}$$

where $t\Big|_{V_{out}=0.5V_{dd}}$ is obtained from analytical transient response expressions. Kayssi et al [52] extended these output waveform expressions for the case of exponential input waveform. Park and Soma [77] proposed a linear system to model the switching transitions of CMOS logic and solved the responses using convolution theory.

2.4.4 The Simplified Bulk-Charge MOSFET Model

One of the analytical MOSFET models is the simplified bulk-charge MOSFET model [99], an empirical MOSFET model simplified from LEVEL3 SPICE model. The parameter δ of the simplified bulk-charge MOSFET model is introduced to describe the effect of charge carrier velocity saturation under high electric field. Channel length modulation effect is ignored in this simplified bulk-charge MOSFET model. The drain current I_{ds} of the simplified bulk-charge MOSFET model is expressed as

$$I_{ds} = \begin{cases} 0 & V_{gs} \le V_{th} \\ k \cdot (1+\delta) \cdot \left(V_{sat} - \dfrac{V_{ds}}{2} \right) \cdot V_{ds} & V_{gs} > V_{th}, V_{ds} \le V_{sat} \\ \dfrac{1}{2} k \cdot (1+\delta) \cdot {V_{sat}}^2 & V_{gs} > V_{th}, V_{ds} > V_{sat} \end{cases} \tag{2.8}$$

where V_{th} is the threshold voltage, k is the drivability factor or transconductance parameter, and δ is the Taylor series expansion coefficient of the bulk charge. In (2.8), V_{sat} is the drain-source saturation voltage, which is defined as

$$V_{sat} = \frac{(V_{gs} - V_{th})}{(1+\delta)} \tag{2.9}$$

The simplified bulk-charge MOSFET model is simple so that it is widely used in circuit analysis. However, the simplified bulk-charge MOSFET model cannot accurately reproduce the *I–V* characteristics of a deep submicron MOSFET. Figs. 2.6-2.7 show the static characteristics of a deep submicron MOSFET from SPICE BSIM3 and from simplified bulk-charge MOSFET model. There are two main inconsistencies when the simplified bulk-charge MOSFET model is used to describe deep submicron MOSFET behaviors. The first discrepancy is that the drain current I_{ds} in the saturation mode of a deep submicron MOSFET does not show square-law dependence on gate-source voltage as in (2.8) and (2.9). The other disagreement is that the effect of channel length modulation is ignored in the simplified bulk-charge MOSFET model. The main reason for the first discrepancy is that the drain-source saturation voltage in the simplified bulk-charge MOSFET model is linearly scaled from that of Shockley MOSFET model. Channel length modulation has to be included in deep submicron MOSFET model.

2.4.5 Delay Models by Using the Simplified Bulk-Charge MOSFET Model

Jeppson [45] solved the differential equation describing the discharge of the load capacitance for a rising input ramp considering the currents through pMOSFET and nMOSFET and the coupling capacitance. The simplified bulk-charge MOSFET model was used without considering the effect of channel length modulation. However, in the

case where the pMOSFET is in linear mode, the quadratic term of current through the pMOSFET was neglected. Moreover, it was not mentioned how the integration constant between the linear and the saturation modes of the pMOSFET device is calculated for fast inputs. For slow inputs least-square fitting techniques are used.

Vemuru and Thorbjornsen [102] derived an expression for the output waveform, which includes the previously mentioned quadratic term of the pMOSFET current but ignored the influence of the coupling capacitance. A power series was used to approximate the solution of the differential equation. However, only the first five terms of the series were considered, and a recursion form for the calculation of higher order terms in order to obtain better accuracy was not given. Bisdounis et al. recently published two papers where analytical expressions for output waveforms [7] [8] were derived. The simplified bulk-charge MOSFET model was used in one of them [7]. The gate-to-drain couplig capacitance and the pMOSFET of the inverter were included. Taylor series expansion method was used to approach the solution of nonlinear Riccati equation, which cannot be solved analytically if a particular solution is not known.

2.4.6 α -Power Law MOSFET Model and Delay Calculation

One of well-known MOSFET models to describe deep submicron MOSFET behaviors is α -power law MOSFET model proposed by Sakurai and Newton [84]. They extended the long-channel delay model such as Shockley MOSFET model to include the effect of carrier's velocity saturation of a submicron MOSFET in α -power law MOSFET model. The α -power law model uses power index α to correct those behavior discrepancies in short channel MOSFETs predicted by the square-law dependence of drain-source current on gate-source voltage as in Shockley MOSFET model. The fitting

expression was explained due to the carrier velocity saturation. A full description of the α -power law model is given below:

$$I_d = \begin{cases} 0 & V_{gs} \le V_{th} \\ (I_{d0}{}' / V_{d0}{}')V_{ds} & V_{gs} > V_{th}, V_{ds} \le V_{d0}{}' \\ I_{d0}{}' & V_{gs} > V_{th}, V_{ds} > V_{d0}{}' \end{cases} \tag{2.10}$$

where

$$I_{d0}{}' = I_{d0} \cdot \left(\frac{V_{gs} - V_{th}}{V_{dd} - V_{th}} \right)^{\alpha} = \frac{W}{L_{eff}} P_c (V_{gs} - V_{th})^{\alpha} \tag{2.11}$$

and

$$V_{d0}{}' = V_{d0} \cdot \left(\frac{V_{gs} - V_{th}}{V_{dd} - V_{th}} \right)^{\frac{\alpha}{2}} = P_v (V_{gs} - V_{th})^{\frac{\alpha}{2}} \tag{2.12}$$

where V_{dd} indicates a supply voltage and P_c and P_v are the model parameters. Figs 2.8-2.9 show static characteristics of a deep submicron MOSFET from SPICE BSIM3 and from α -power law MOSFET model. There are several discrepancies when α -power law MOSFET model is to describe deep submicron MOSFET behaviors. The first one is that drain current I_{ds} in saturation mode of a deep submicron MOSFET does not show α -power law dependence on gate-source voltage as in (2.11). The second is that the drain saturation voltage V_{sat} of a deep submicron MOSFET is deviated from (2.12), especially when gate-source voltage is large as illustrated in Fig. 2.8. The third disagreement is the current of α -power law MOSFET in linear mode was described as linear dependence on drain-source voltage. The fourth inconsistency is that the effect of channel length modulation is not taken into account in α -power law MOSFET model.

The α -power law MOSFET model is simple and suitable for circuit analysis.

Useful and simple expressions were derived for the propagation delay, short-circuit power and logic threshold voltage with the model. They presented closed-form expression of propagation delay for CMOS inverter from α-power law MOSFET model [84]. They claimed that their delay model is accurate to within a few percent of SPICE simulations [84]. Both the short-circuit current and the coupling capacitance are neglected in the derivation of the output expressions. By using α-power law MOSFET model, the propagation delay t_{pHL} of falling output of an inverter or the propagation delay t_{pLH} in the charging case is calculated as

$$t_{pHL,pLH} = \left(\frac{1}{2} - \frac{1-u_{th}}{1+\alpha}\right)\cdot\tau + \frac{C_t \cdot V_{dd}}{2 \cdot I_{d0}} \tag{2.13}$$

with $u_{th} = V_{th} / V_{dd}$, where τ is the input transition-time. The output transition-time is given by

$$t_T = \frac{t_{0.9} - t_{0.1}}{0.8} = \frac{C_t \cdot V_{dd}}{I_{d0}} \cdot \left[\frac{0.9}{0.8} + \frac{V_{d0}}{0.8 \cdot V_{dd}} \cdot \ln\left(\frac{10 \cdot V_{d0}}{e \cdot V_{dd}}\right)\right] \tag{2.14}$$

Dutta et al. [29] extended the work [84] of CMOS inverter to the case of very slow slope input or very light load cases. For fast slope input or reasonable heavy load, they used the same expression as that proposed in [84]. For very slow slope input or very light load, they used fitting curves to connect the results obtained from α-power law MOSFET model and from static characteristics. The delay model presented in [31] uses α-power MOS model taking into account short-circuit current of an inverter, but output voltage and currents through both transistors are assumed to be piecewise linear, resulting in inaccurate modeling of the nonlinear MOSFET behavior. Cherkauer and Friedman extended the delay model of (2.12) and (2.13) into CMOS tapered buffer [16]. Bisdounis

et al. [8] used α-power law MOSFET model in their analysis of propagation delay. The analytical expressions for the output waveforms were derived. The gate-to-drain coupling capacitance and the pMOSFET of a CMOS inverter were included. Taylor series expansion method was used to solve the nonlinear differential equation that cannot be solved analytically. Vemuru and Scheinberg [103] exploited α-power MOS model to derive short-circuit power dissipation estimation for CMOS logic gates. The α-power law MOSFET model was extended to enable technology projections [11].

2.4.7 The nth-Power Law MOSFET Model and Timing Models

Another interesting MOSFET model is *nth*-power law MODFET model proposed by Sakurai and Newton [85] [86]. It has been implemented in SPICE3 circuit simulator. The drain current I_{ds} of the proposed model is as follows.

$$I_{ds} = \begin{cases} 0 & V_{gs} \le V_{th} \\ I_{sat} \cdot (1 + \lambda \cdot V_{ds}) \cdot \left(2 - \dfrac{V_{ds}}{V_{sat}}\right) \dfrac{V_{ds}}{V_{sat}} & V_{gs} > V_{th}, V_{ds} \le V_{sat} \\ I_{sat} \cdot (1 + \lambda \cdot V_{ds}) & V_{gs} > V_{th}, V_{ds} > V_{sat} \end{cases} \quad (2.15)$$

where V_{th} is the gate-source threshold voltage, λ the channel length modulation coefficient, V_{sat} the drain-source saturation voltage:

$$V_{sat} = K \cdot (V_{gs} - V_{th})^m \quad (2.16),$$

and I_{sat} is the MOSFET saturation-related current, which is

$$I_{sat} = A \cdot (V_{gs} - V_{th})^n \quad (2.17)$$

In (2.17) A is the MOSFET drivability factor that is proportional with the width of a MOSFET and inversely proportional with the effective length of a MOSFET.

As shown in Figs 2.10-2.11, *nth*-power law MODFET model can accurately express the static *I–V* characteristics of short-channel MOSFETs. Since there are seven parameters in the model it is not easy to do analytical treatments of circuits with this model. Moreover, the drain current I_{ds} depends cubically on drain-source voltage that makes the analytical propagation delays almost impossible.

Sakurai and Newton [85] [86] extended α-power law MOSFET model [84] to *nth*-power law MOSFET model to describe series-connected MOSFET structure (SCMS) as in complex NAND/ NOR gates. They adopted the current expressions that are more similar to the Shockley MOSFET model. The other improvement of the extended model is that the channel length modulation is included. In [85] [86], the influence of the short-circuit current was considered by assuming a suppressed input voltage clamped to the rail until the switching voltage of the inverter was exceeded. This approximation is exact only for some cases of input ramps. Hirata et al. [41] developed a propagation delay model for CMOS inverters considering short-circuit currents with *nth*-power law MOSFET model while complex gates were modeled using the collapsing technique developed in [85]. Rossello and Segura [18] presented a method to compute the propagation delay of complex CMOS gates based on *nth*-power law MOSFET model. A transistor collapsing technique for complex CMOS gates was developed. The short-channel effects, internal coupling capacitances and body effect were included.

2.5 RESISTANCE-CAPACITANCE (RC) DELAY APPROACHES

In resistance-capacitance (RC) approaches, CMOS circuits are approximated with linear or piecewise linear RC networks. These RC approaches for determining the

propagation delay from such RC networks generally fall into two main categories. In one category, the so-called Elmore delay of a RC network is computed as propagation delay of a CMOS circuit [30] [58] [74] [82] [105]. In the other category, the objective is to bound propagation delay time of RC circuits [26] [61] [82] [105] [106] [107]. Since Elmore delay is defined as the first moment of impulse response, it yields only a reasonably good approximation to propagation delay. With comparatively accuracy the simple RC delay approach for obtaining propagation delay is very fast so that it is widely used in timing simulation, timing verification, transistor resizing [20] [57], wiresizing [23], and circuit optimization [74]. The trade-off between the RC approach and the analytical approach is simplicity and speed versus accuracy [33]. In this section, the important theorems of RC delay approaches are introduced.

2.5.1 Elmore Delay

Elmore published an important analytical approximation to RC chain delay [30]. He called the first moment of impulse response the RC chain delay. The definition of RC chain delay is inadequate because it does not define propagation delay in terms of signal threshold such as half of operating voltage. Elmore delay for transient rising output is defined as [30]

$$\tau_d = \int_0^\infty t \cdot \frac{dV}{dt} \cdot dt \tag{2.18}$$

Elmore delay has been used as a measure of propagation delay for RC interconnects and for MOS circuits. Before considering a general RC network, a simple and non-branched RC chain shown in Fig. 2.12 (a) is analyzed. Assume that all capacitances are initially pre-discharged to ground. The step input is applied at source point $[V_{in}(t) = V_{dd} \cdot u(t)]$. The

objective is to find a simple closed-form formula that gives the time for voltage V_i at node i to rise from ground to half of power voltage after step input $V_{in}(t)$ is applied at source point. This is equivalent to deriving the equivalent first-order time-constant of RC networks. Elmore delay at node i can be computed using the following equation:

$$\tau_i = \sum_{k=1}^{N}\left(C_k \cdot \sum_{j=1}^{\min(k,i)} R_j\right) = \sum_{k=1}^{i}\left(R_k \cdot \sum_{j=k}^{N} C_j\right) \tag{2.19}$$

where N is the total number of nodes in RC chain. The first-order time constant τ_i at node i is found as the sum of a number of components. The fraction of contribution of node k on τ_i is determined as the product of the capacitance C_k at node k and the sum of resistances of the common path from source to node i and node k. Equation (2.19) was used to compute Elmore delay of MOS circuits [76]. The effective resistance of each device and the capacitive load seen by each device are used to approximate a MOS circuit by an RC circuit and Elmore delay is computed.

2.5.2 Signal Delay in RC Tree Networks and RC Mesh Networks

In CMOS integrated circuits, signals may propagate between stages with fanouts. A given logic gate may drive several gates, some of them through long wires. The exact calculation of signal delay through such networks is difficult. Rubinstein, Penfield, and Horowitz [82] extended Elmore delay to RC tree networks and derived simple formulas for finding upper and lower bounds of propagation delay. Given a RC tree network shown in Fig. 2.12 (b), assume that all capacitances are initially pre-discharged to ground. The step input is applied at source point $[V_{in}(t) = V_{dd} \cdot u(t)]$. A simple closed-form formula is derived to obtain the time for voltage V_i at node i to rise from ground to half of operating

voltage after step input $V_{in}(t)$ is applied at source point. The extended Elmore delay at node i can be computed using the following equation [82]:

$$\tau_i = \sum_{k=1}^{N} (C_k \cdot R_{ki}) \tag{2.20}$$

where N is the total number of nodes in the RC chain. The resistance R_{ki} is defined as the resistance of the common portion of the unique path between the input and node i and the unique path between the input and node k. It is obvious that RPH Elmore delay of (2.20) is equal to Elmore delay if the RC tree network is a RC chain. Wyatt [105] extended RPH's work to include resistor loops in RC mesh networks, instead of RC tree networks, and tighten the bounds. The resistor network extracted from the original RC mesh is used to represent to generate the resistance matrix. Its relevant elements can be substituted for R_{ki} in the definition above. Lin and Mead [58] redefined Elmore delay to properly handle the effect from initial charges in RC networks. Chu and Horowitz [19] presented a two-pole-one-zero model to improve the accuracy of RC circuits with charge-sharing operations. Horowitz introduced a simple transformation of variables to convert a nonlinear MOS circuit into a pseudo linear network such that the techniques developed for linear RC networks are still applicable while taking into account nonlinear MOSFET characteristic. This nonlinear model becomes quite complicated especially when the input waveform effect is included. As a result from these remarkable contributions, timing simulators and analyzers [76] based on linear RC models have a computational speed close to that of switch-level logic simulators. However, RC delay approaches generally suffer from one major drawback; namely, the over-simplified linear resistance model on nonlinear MOSFET cannot guarantee reliable accuracy compared with the more

complicated circuit simulators [72]. Moreover, all the above-mentioned techniques of propagation delays are based on step-input waveform while the real propagation signal always exhibits a finite-slope waveform in practice. In the timing verifier Crystal [76], Ousterhout used an effective resistance based on the rising time ratio to include the effect of the finite-slope input waveform. Substantial inaccuracies still exist. A generic linear RC delay modeling for digital CMOS circuits was proposed by Deng and Shiau [25]. In order to take advantage of the simplicity of linear RC model and the reliability of empirical model, the proposed propagation delay is determined by the input rising/falling time and the generic delay calculated from the generic linear RC network. RC circuit approaches were used to model the average circuit behaviors since nonlinear behaviors of MOS transistors are not well represented by linear and piecewise resistors. More efficient and general algorithms for computing Elmore delay have been developed [62]. These extended Elmore delay formula have been widely used in the timing analyzers or time verifiers of VLSI digital design. It is a basis to estimate the propagation delay of interconnects, complex transmission gate and MOS networks.

2.5.3 Timing Models for CMOS VLSI circuits with Interconnects

As fabrication technology advances, device sizes get smaller, VLSI circuits get faster, and chip die sizes increase. As a result, on-chip interconnects play an important role in determining the overall performance of VLSI circuits. However, the effect of interconnects is difficult to accurately model due to the following main reasons. The first one is how to extract a reasonable simple interconnect model for propagation delay. The other is that interconnects are driven by nonlinear deep submicron CMOS logic gates. Interconnections not only delay signals, but alter the shape of signals also [55]. Thus,

input waveform shape, RC-interconnect loading and nonlinear deep submicron CMOS logic gates with second-order effects affect circuit performance [54]. The interconnect models depend on the relationship of interconnect length and signal transition time [51]. If signal transition time is much longer than propagation delay of a signal on an interconnection a lumped RC model is valid. One way to obtain the propagation delay of a CMOS logic gate with this kind of interconnection was proposed in [51]. A linear or piecewise-linear model approximates the nonlinear behaviors of a logic gate. Then RC timing approaches are exploited to obtain the propagation delay of RC interconnects driven by a logic gate. Harbour and Drake proposed a procedure for calculating the signal delay in on-chip interconnections modeled by RC network [39]. This method is based on the elimination of all the internal nodes of the RC network, using a first order approximation. The method is valid for calculating the approximate transient response of RC mesh networks for a ramp input. A fundamental limitation of this method is that a nonlinear CMOS logic gate cannot be accurately described with a linear circuit. The way to derive the propagation delay was presented in [38]. There are two steps involved. The first one is to approximate the MOSFET output voltage waveform by analyzing it with an effective capacitance equivalent to that of interconnects. In the second step interconnects driven by the output waveform are efficiently estimated.

If propagation delay of interconnects is longer than signal transition time transmission line model or complicated interconnection model should be used. Eo and Eisenstadt proposed high-speed VLSI interconnect modeling based S-parameter measurements [32]. They claimed the simple RC model of interconnect is not suitable for the highest speed VLSI circuit design. Restle, Jenkins, Deutsh, and Cook [80] showed

that transmission line effects associated with long wires like clock distributions are significant. The approximated analytical solution for transients in the case of a capacitive loaded uniformly distributed RC (URC) line was proposed in [108].

2.6 TABULAR DELAY APPROACHES WITH EMPIRICAL DELAY EQUATIONS

Tabular delay approaches, in general, build a set of look-up tables including propagation delays and output transition times as functions of design parameters, input transition time, and external load capacitances for typical primitives. These look-up tables are generated from either pre-characterization using SPICE simulations or measured data. They inherently include all device parameters and internal parasitic properties. However, any change in device parameters requires the rebuilding of the look-up tables. Tabular delay approaches are not only time and memory space consuming, but incorporate interpolation errors as well [55].

In order to overcome the drawbacks mentioned above, tabular delay approaches combined with empirical delay equations are used to interpolate the the dependence of propgation delay on input transition time and output capacitance. The tabular delay method is combined with the RC tree approach [12]. An approach based on the empirical delay models for fixed gate structures has been reported [5]. Reliable accuracy and extremely efficient speed can be achieved as the delay values are directly estimated, with very minor computations, from the pre-characterization library. The disadvantages of the method are the pre-characterization which is quite time-consuming in order to cover a wide range of different parameter combinations such as transistor sizes, node

capacitances, and input waveforms. Tabular delay approaches combined with polynomial approcimation were proposed in [50].

M. W. Jetton and A. P. Jayasumana proposed the method of determining delay in logic cell models based on tabular delay approaches with delay equations[46]. In it the constants of delay equations are extracted from look-up delay tables. A logic-cell delay model takes nonlinear effects into consideration in determining propagation delay, thus providing improved accuracy as compared to existing models, particularly when input transition times are large. Given a logic cell of the type where propagation delay is a function of input transition time (T_{rf}) and load capacitance (C_l), the method involves choosing some special discrete simulation points associated with propagation delays in [C_l, T_{rf}] space to ensure accuracy for a wide range of T_{rf} and C_l values, determining the constants of empirical delay equations [87] in accordance with the chosen simulation points, and solving for propagation delay on a more accurate basis. Generally propagation delay, T_d, may be determined with the relation $T_d = A + B \cdot C_l + C \cdot T_{rf} + E(C_l, T_{rf})$, where A, B, and C are constants from the simulation points, and $E(C_l, T_{rf})$ represents a correction factor associated with the nonlinear effects. The propagation delay calculated from the method is accurate compared with the results of SPICE BSIM3 model, especially in the case of large load capacitance. The more important is that the simulaiton points are greatly reduced.

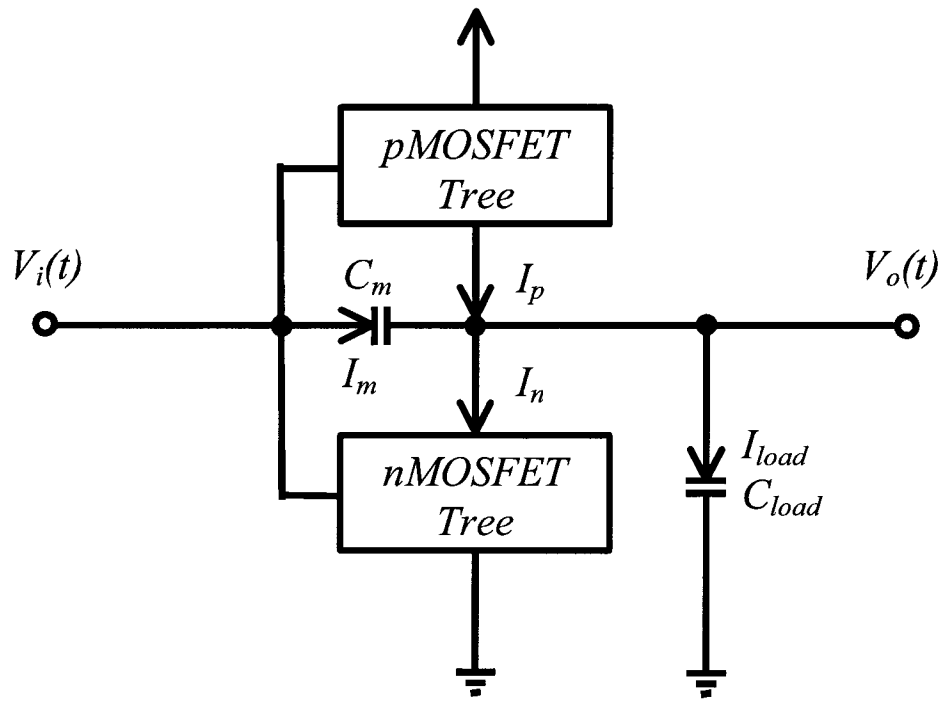


Fig. 2.1 A generic CMOS logic gate used in analytical timing approaches. There are two basic circuit components: pMOSFET tree and its dual nMOSFET tree.

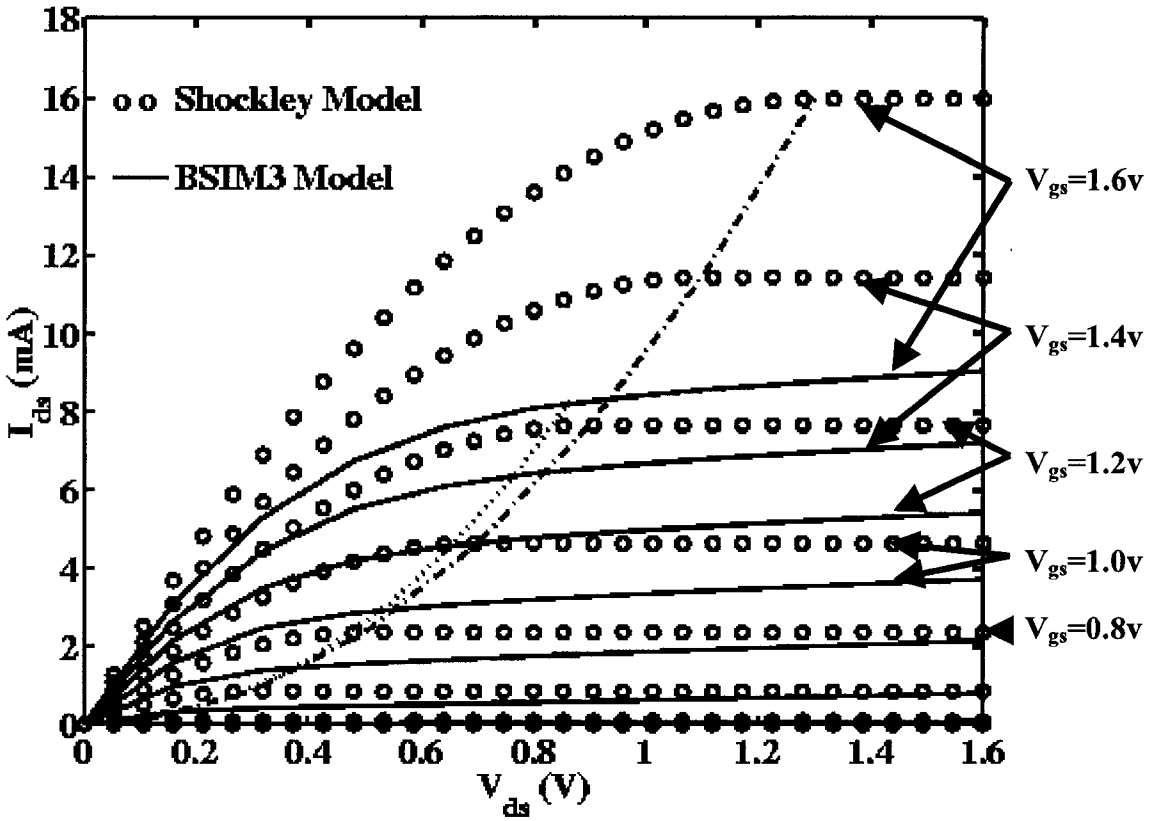


Fig. 2.2 Static I–V characteristics of a deep submicron MOSFET from Shockley model and BSIM3 model. The parameters of Shockley model are fit to have the same characteristics with SPICE BSIM model at drain-source voltage close to zero in linear region.

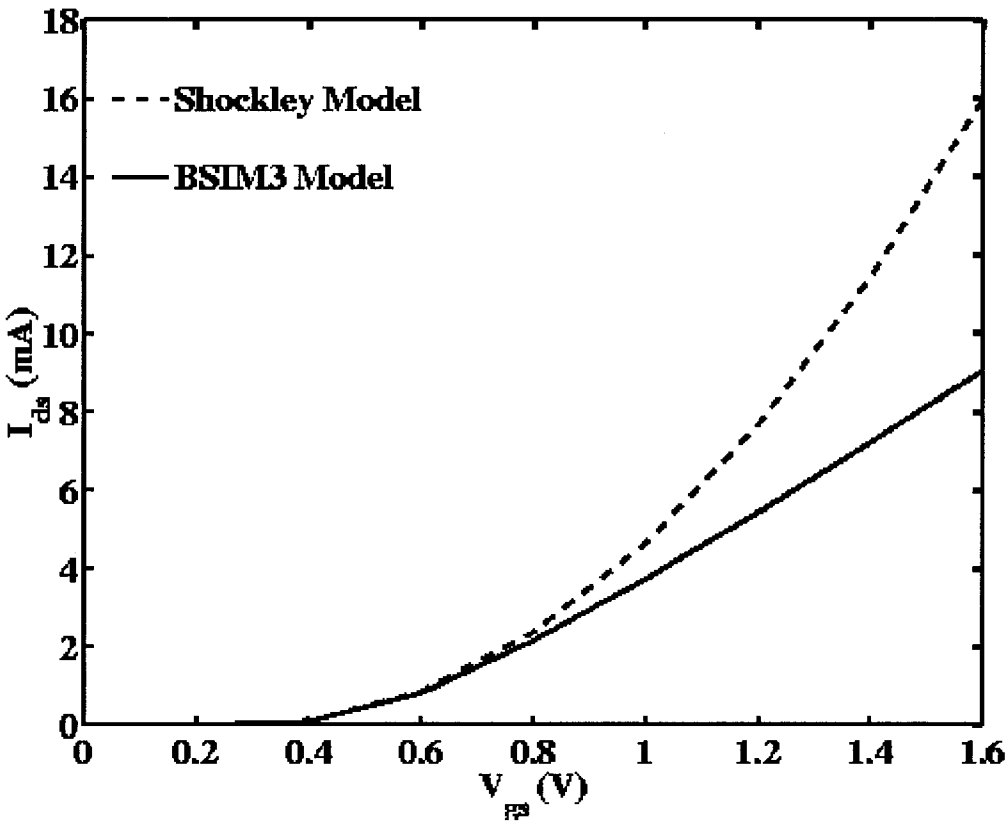


Fig. 2.3 Saturation drain-source currents of a deep submicron MOSFET when $V_{ds} = V_{dd}$ from Shockley model and from SPICE BSIM3 model. The parameters of Shockley model are fit to have the same characteristics with SPICE BSIM3 model at small drain-source voltage in linear region.

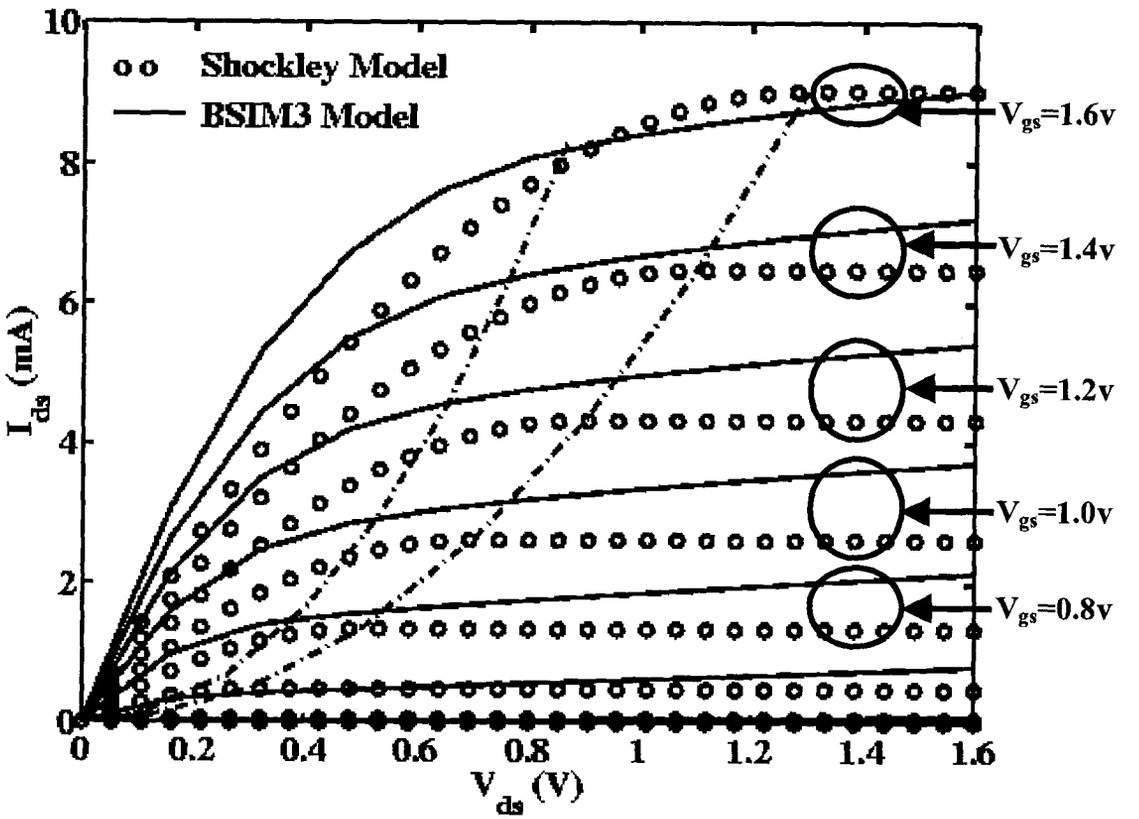


Fig. 2.4 Comparison of static I–V characteristics of a deep submicron MOSFET from Shockley model and from SPICE BSIM3 model. The parameters of Shockley model are fit to have the same current with SPICE BSIM model when $V_{gs} = V_{dd}$ and $V_{ds} = V_{dd}$ in saturation region.

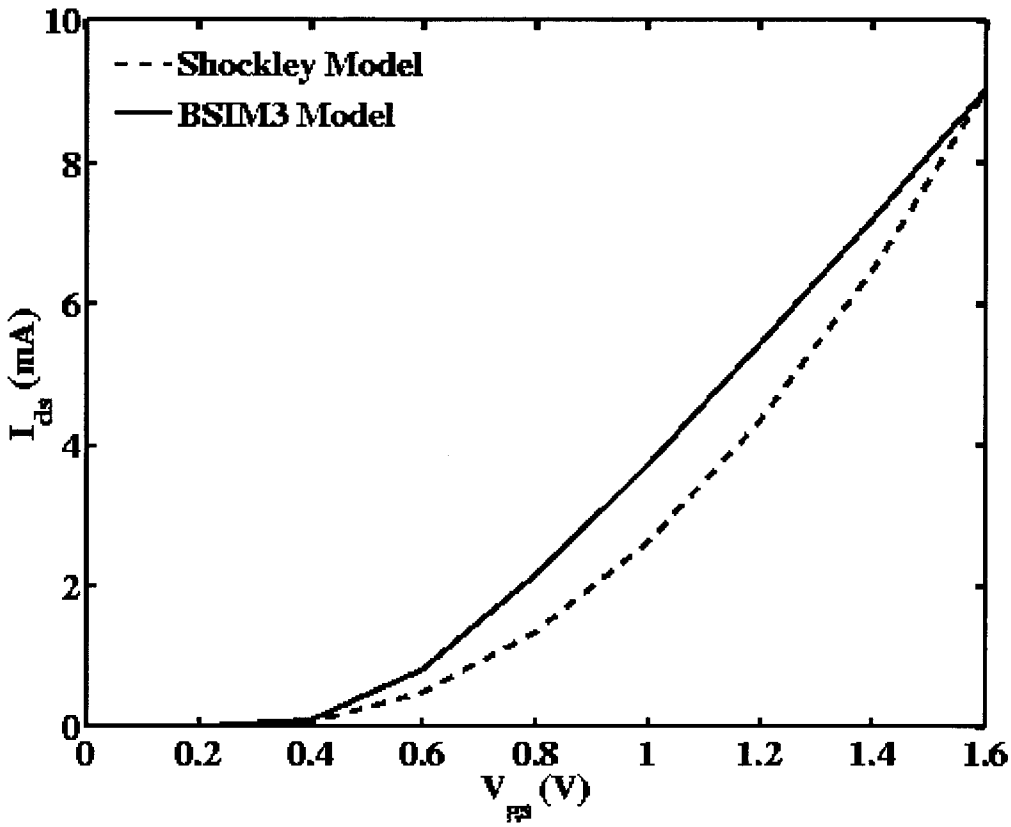


Fig. 2.5 Saturation drain-source currents of a deep submicron MOSFET when $V_{ds} = V_{dd}$ from Shockley model and from SPICE BSIM3 model. The parameters of Shockley model are fit to have the same current with SPICE BSIM3 model when $V_{gs} = V_{dd}$ and $V_{ds} = V_{dd}$ in saturation region.

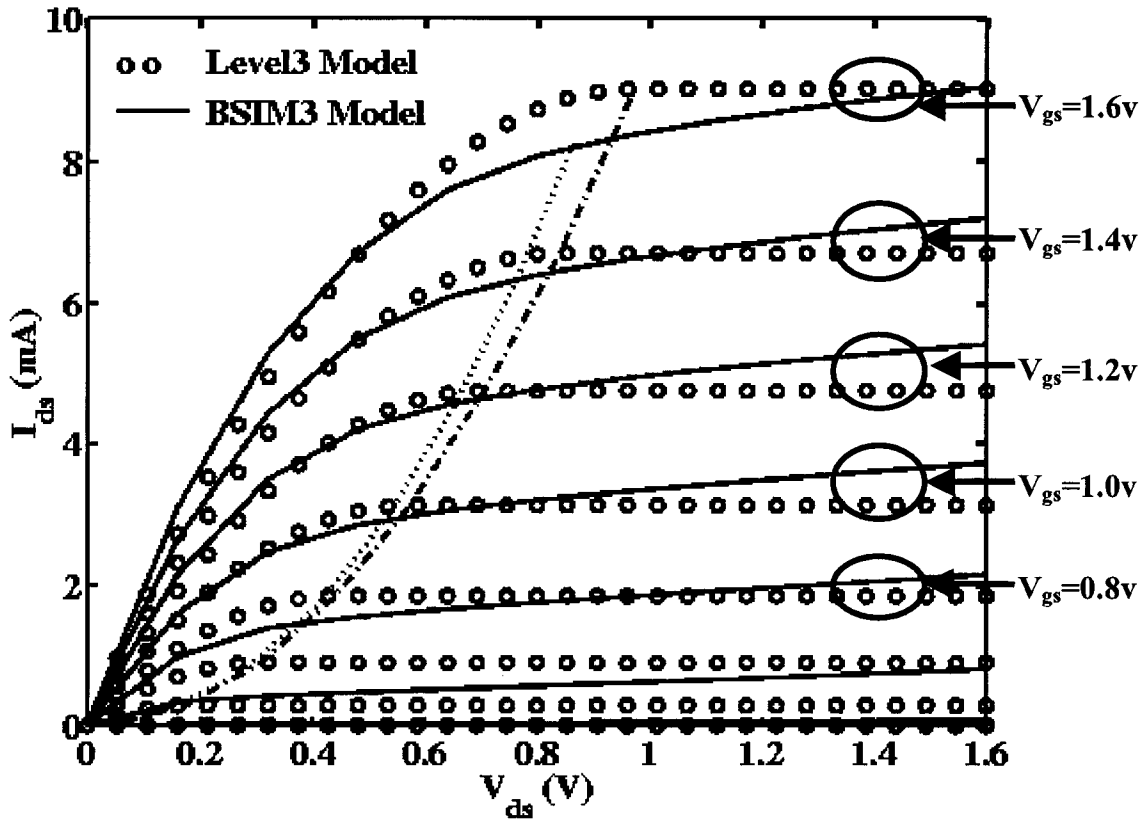


Fig. 2.6 Static I–V characteristics of a deep submicron from simplified bulk-charge MOSFET model and SPICE BSIM3 model. The parameters of simplified bulk-charge MOSFET model are fit to have the same drain-source current with SPICE BSIM3 model when $V_{gs} = V_{dd}$ and $V_{ds} = V_{dd}$ in saturation region. The static I–V characteristics from simplified bulk-charge MOSFET model where square-law in saturation region still exists deviate from SPICE BSIM3.

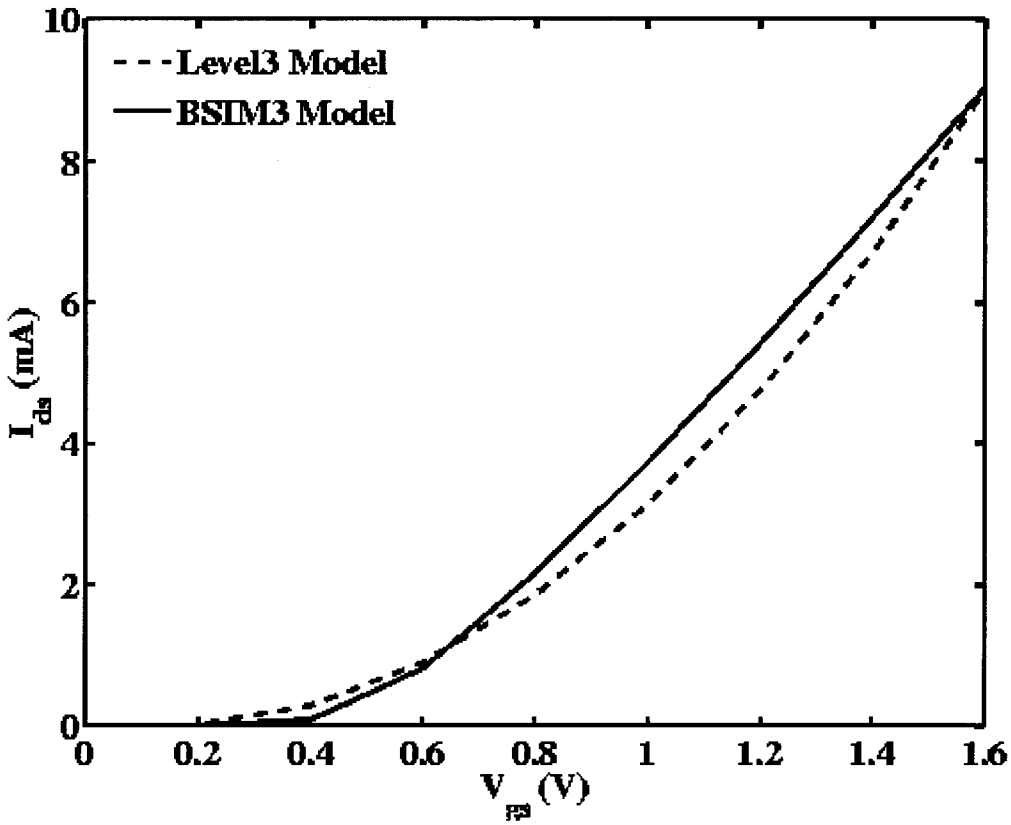


Fig. 2.7 Saturation drain-source current of a deep submicron MOSFET when $V_{ds} = V_{dd}$ from simplified bulk-charge MOSFET model and from SPICE BSIM3 model. The parameters of simplified bulk-charge MOSFET model are fit to have the same current with SPICE BSIM3 model when $V_{gs} = V_{dd}$ and $V_{ds} = V_{dd}$ in saturation region. Obviously, power index of a deep submicron MOSFET deviates from two as in simplified bulk-charge MOSFET model when gate-source voltage increases.

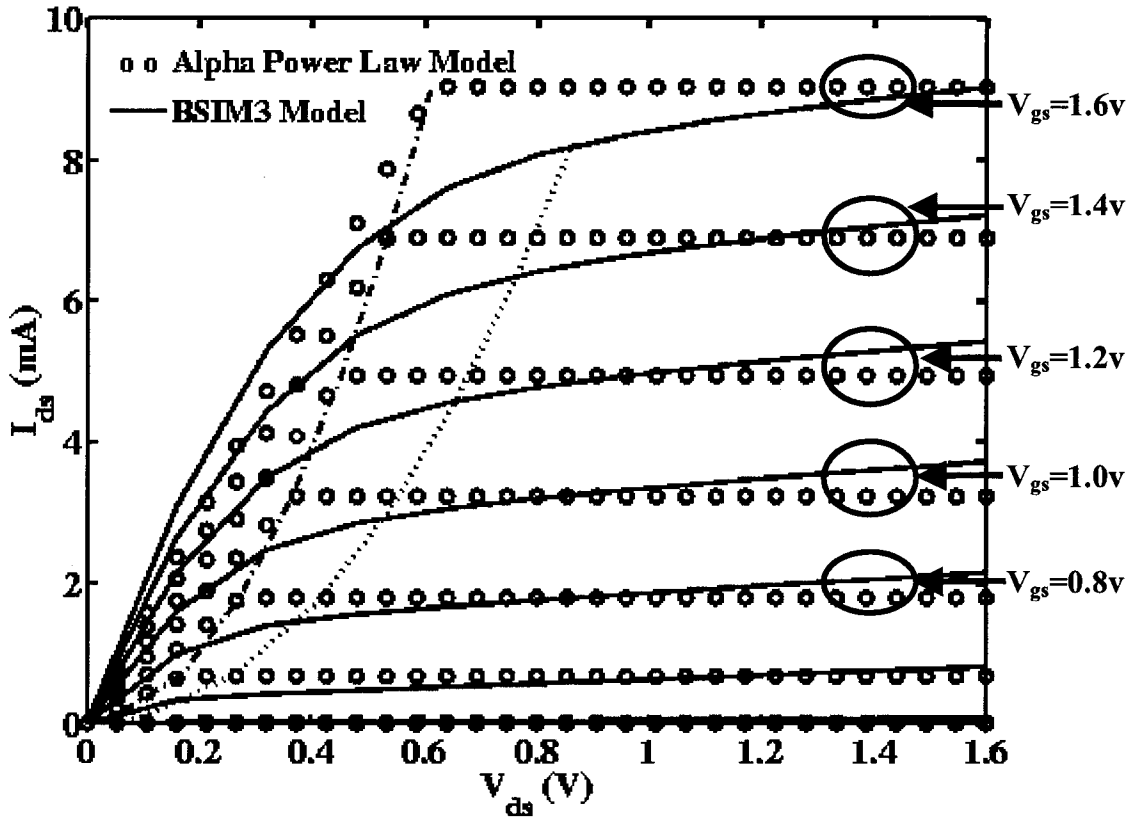


Fig. 2.8 Static I–V characteristics of a deep submicron MOSFET from α-power law MOSFET model and from SPICE BSIM3 model. The parameters of α-power law MOSFET model are fit to have the same drain-source current with SPICE BSIM3 model when $V_{gs} = V_{dd}$ *and* $V_{ds} = V_{dd}$ *in saturation region. The static I–V characteristics from α-power law MOSFET model are piecewise linearized.*

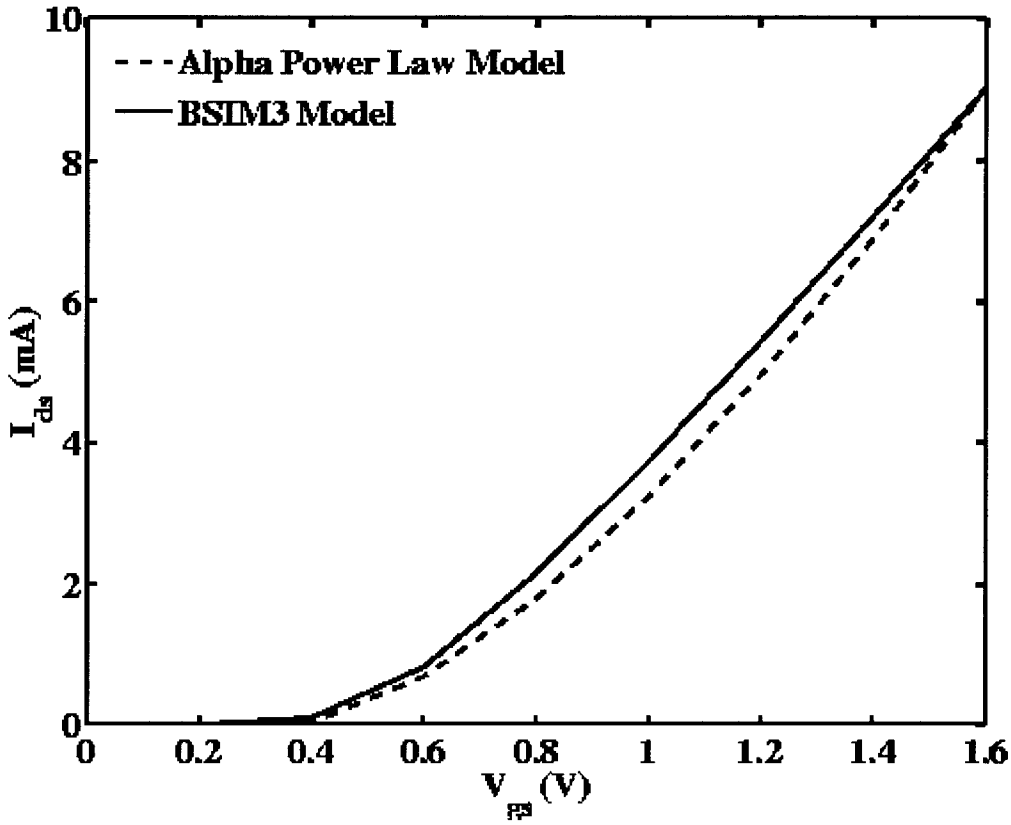


Fig 2.9 Saturation drain-source current of a deep submicron MOSFET when $V_{ds} = V_{dd}$ from α-power law MOSFET model and from SPICE BSIM3 model. The parameters of α-power law MOSFET model are fit to have the same current with SPICE BSIM3 model when $V_{gs} = V_{dd}$ and $V_{ds} = V_{dd}$ in saturation region.

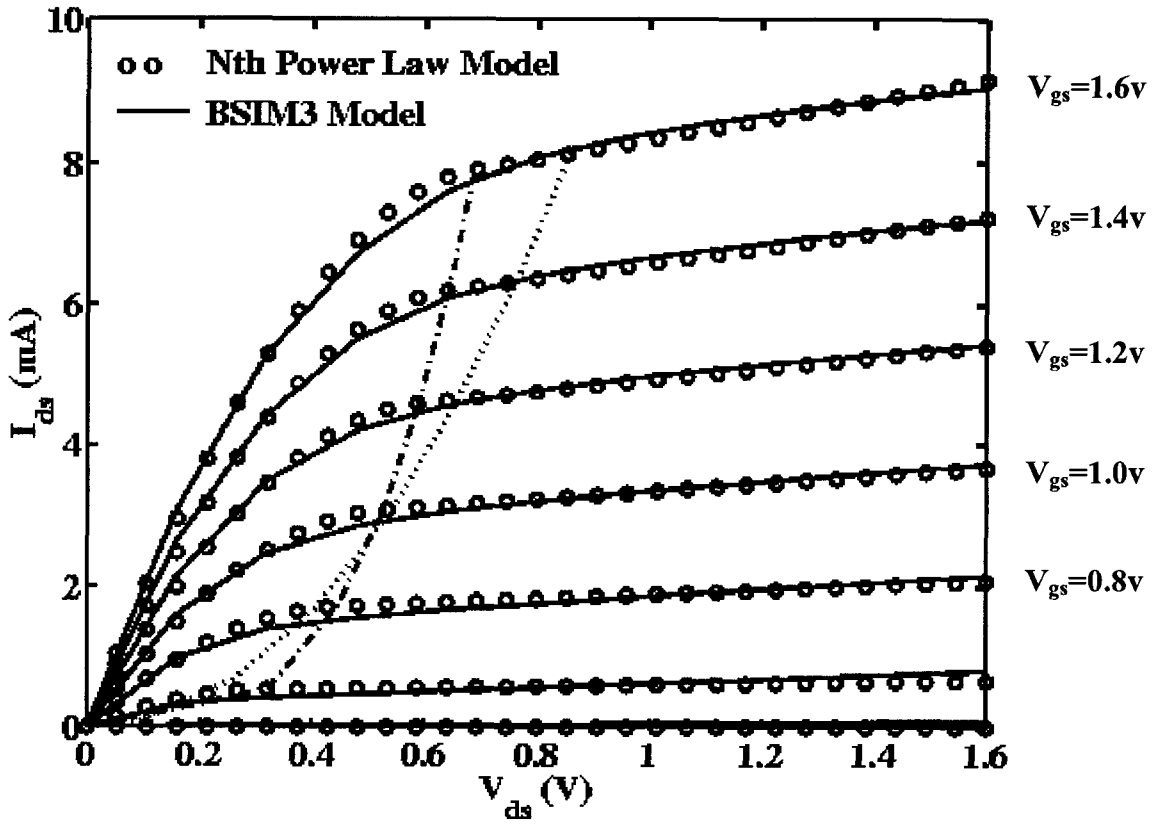


Fig. 2.10 Static I–V characteristics of a deep submicron MOSFET from nth-power law MOSFET model and SPICE BSIM3 model. The parameters of nth-power law MOSFET model are fit to have the same drain-source current with SPICE BSIM3 model when $V_{gs} = V_{dd}$ *and* $V_{ds} = V_{dd}$ *in saturation region. The static I–V characteristics from nth-power law MOSFET model are well matched with SPICE BSIM3 model.*

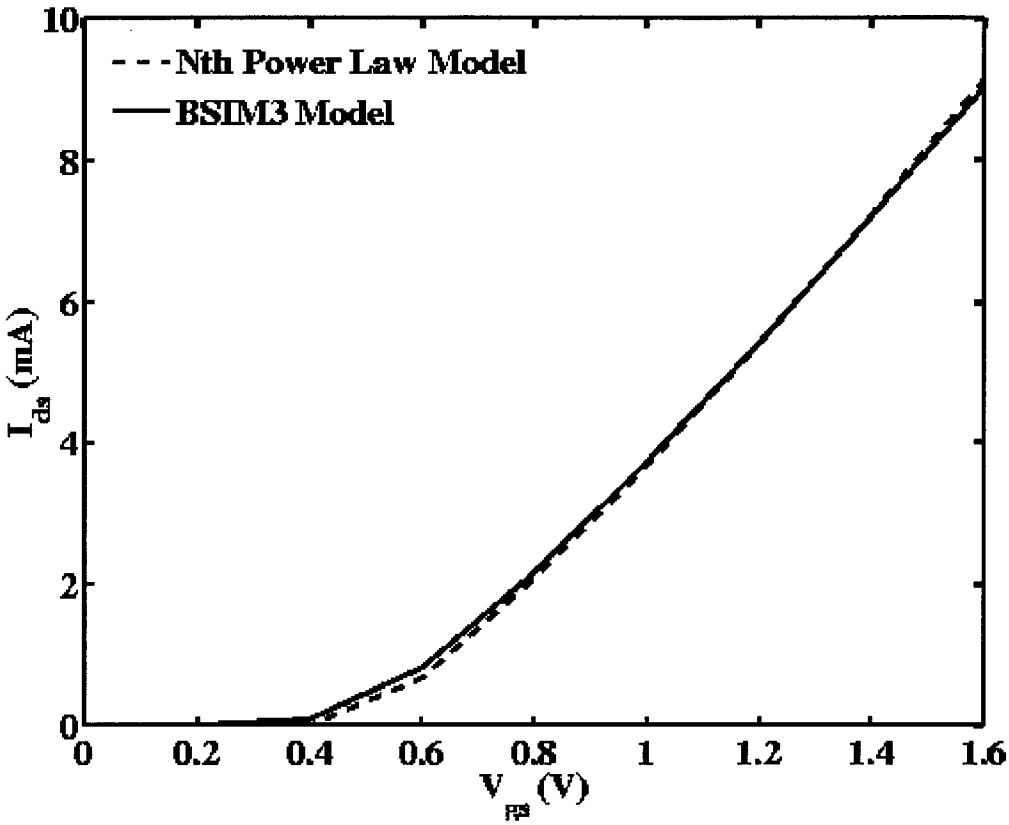


Fig. 2.11 Saturation drain-source current of a deep submicron MOSFET when $V_{ds} = V_{dd}$ from nth-power law MOSFET model and from SPICE BSIM3 model. The parameters of nth-power law MOSFET model are fit to have the same current with SPICE BSIM3 model when $V_{gs} = V_{dd}$ and $V_{ds} = V_{dd}$ in saturation region. The nth-power law MOSFET model can accurately reproduce static I–V characteristics of a deep submicron MOSFET.

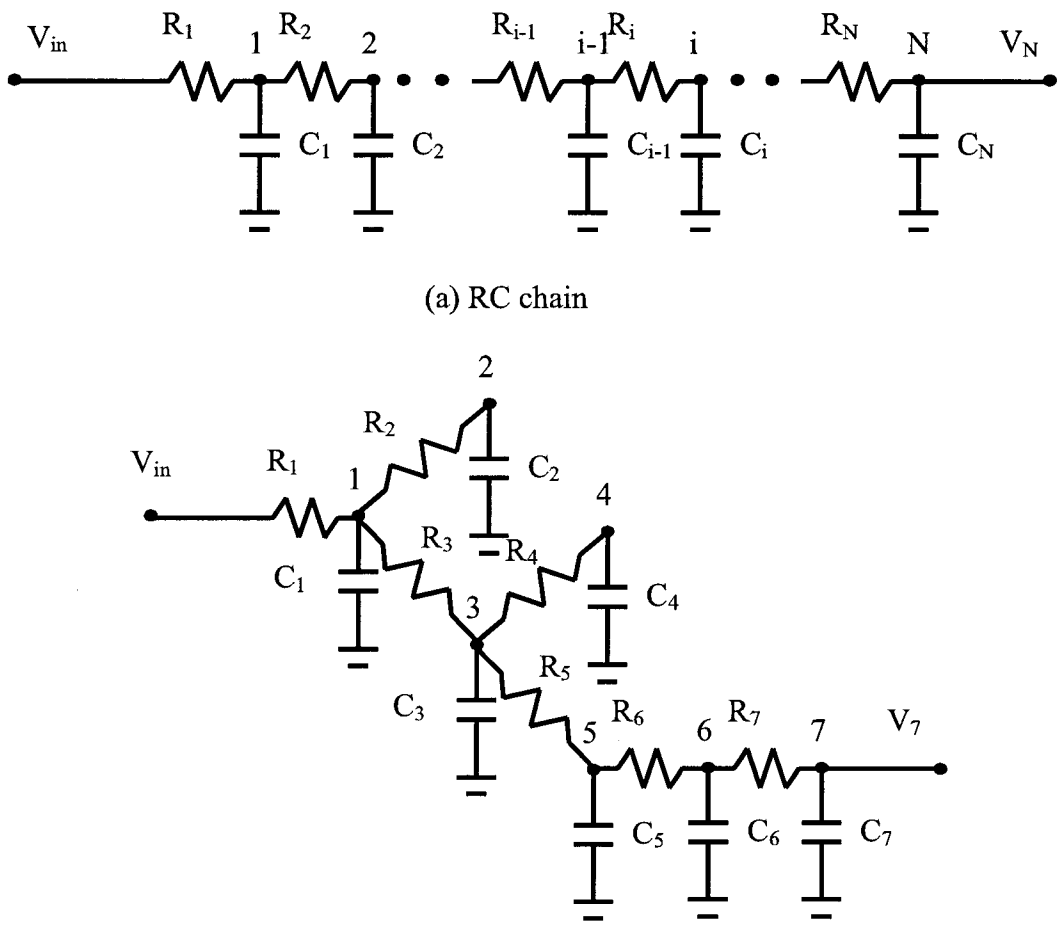


Fig. 2.12 RC chained line and RC tree network

CHAPTER 3 DEEP SUBMICRON MOSFET MODELS

We have reviewed the methodologies of propagation delays for CMOS digital circuits in Chapter 2. Several popular MOSFET models widely used in analytical timing approaches are also examined. These existing MOSFET models are either not accurate enough to model static *I–V* characteristics of a deep submicron MOSFET or too complex to be directly used in the analytical timing model. Most of them do not include the important second-order effects of a deep submicron MOSFET such as charge carrier velocity saturation and channel length modulation.

In order to exploit analytical timing approaches to obtain accurate propagation delay of a deep submicron CMOS inverter, we have to develop a simple and accurate current model for a deep submicron MOSFET. It should include the important second-order effects of a deep submicron MOSFET. In this chapter we first discuss the important second-order effects of a deep submicron MOSFET such as charge carrier velocity saturation and channel length modulation. Then we show that the power index of a deep submicron MOSFET is gradually reduced from two to a value less than two, or even to one in the extreme case when gate-source voltage increases. Based on the investigation, we derive a new deep submicron MOSFET model (DSMM). It is accurate to describe the static *I–V* behaviors of a deep submicron MOSFET since the important second-order MOSFET effects are included in the DSMM. However, it is still complicated to be used in analytical timing models. Then an accurate proper deep submicron MOSFET model (PDSMM) [47], simplified from the DSMM, is proposed. The simple PDSMM has the main features of the important second-order effects so that it is sufficiently accurate and

easy to be used to analyze the timing models of a deep submicron CMOS inverter. The details of analytical timing approaches with PDSMM will be developed in Chapter 4.

3.1 CHANNEL LENGTH MODULATION

The space-charge region at the drain end of a MOSFET channel for saturation mode $(V_{ds} > V_{sat})$ actually varies with drain-source voltage [42] [51] [67] [88]. For long channel MOSFETs, the variation of conducting channel length with regard to drain-source voltage in saturation mode can be ignored comparing with conducting channel length. Shockley MOSFET model was developed for long channel MOSFETs. The conducting channel length is treated as being constant in Shockley MOSFET model. However, for deep submicron technology the variation of conducting channel length as a function of drain-source voltage becomes increasingly obvious and cannot be ignored since the percentage of conducting channel length shorting is significant comparing to deep submicron conducting channel length. Although an exact theory of conducting channel length shortening is complicated, for most design calculations it can be modeled adequately [42] [51] [67] [88] as being linearly proportional to drain-source voltage V_{ds}. Taking channel length modulation into account, Shichman and Hodges [42] [88] proposed the following MOSFET model:

$$I_{ds} = \begin{cases} 0 & V_{gs} \le V_{th} \\ k\left(V_{sat} - \dfrac{V_{ds}}{2}\right)V_{ds}(1+\lambda V_{ds}) & V_{gs} > V_{th}, V_{ds} \le V_{sat} \\ \dfrac{1}{2}kV_{sat}^{\,2}(1+\lambda V_{ds}) & V_{gs} > V_{th}, V_{ds} > V_{sat} \end{cases} \tag{3.1}$$

where k is the MOSFET transconductance parameter, V_{th} the gate-source threshold

voltage, V_{sat} the drain-source saturation voltage. The above MOSFET model is termed *Shichman-Hodges* model. The proportionality constant λ is called channel length modulation coefficient. It is quite small and can be ignored for long channel MOSFETs. But it increases considerably for a deep submicron MOSFET. The term of $(1+\lambda \cdot V_{ds})$ in the equation of linear mode is used to meet the continuity of drain-source current at linear-saturation boundary $(V_{ds} = V_{sat})$. There is no physical justification to do so. The only difference between *Shockley* model and *Shichman-Hodges* model is caused with channel length modulation. *Shichman-Hodges* model becomes *Shockley* modelif channel length modulation parameter is set to zero. The effect of channel length modulation degrades dynamic performance of CMOS logic gates. For example, the transition region of static *I–V* characteristics of CMOS inverter is widened due to channel length modulation, which reduces noise margins of CMOS logic gates. The effect of channel length modulation has to be incorporated in deep submicron MOSFET model [17].

3.2 CHARGE CARRIER VELOCITY SATURATION

When gate-source voltage V_{gs} of an nMOSFET exceeds threshold voltage V_{th} $(V_{gs} > V_{th})$, the p-type surface under gate oxide between source and drain is inverted due to the accumulation of attracted electrons. The inverted surface forms an n-channel as a conduction path between source and drain. The electric field in the channel contains two components: one is the parallel electric field component E_p tangential to the interface between silicon and silicon dioxide and the other is the vertical electric field component E_v perpendicular to it. When drain-source voltage V_{ds} is applied, electrons move from

source to drain along the n-channel under the influence of parallel electric field E_p while they are attracted near the interface by vertical electric field E_v at the same time. Both electric fields E_p and E_v depend on position and applied voltages. The velocity of charge carriers drifting along the channel is influenced by electric fields E_p and E_v. The effect of vertical electric field E_v is complicated and relatively small comparing with the effect of parallel electric field E_p so it is neglected in the following analysis. Hereafter, we simply refer channel electric field to the parallel electric field E_p.

The parallel electric field E_p significantly affects charge carrier velocity of deep submicron MOSFETs [22] [36] [69] [92] [93] [97]. Experiments have shown that charge carrier velocity near the surface tends to saturate at higher electric field. The average drifting velocity v of surface electrons against channel electric field E_p is shown in Fig. 3.1. This behavior is qualitatively similar to drifting velocity of bulk electrons as a function of electric field [69] [92] [93]. It is generic for drifting velocity of charge carriers as a function of electric field and can be modeled by the simplified expression

$$v = \frac{\left|\mu_0 \cdot E_p\right|}{1+\left|E_p/E_c\right|} \tag{3.2}$$

where v represents the charge carrier velocity magnitude along the channel. The parameters μ_0 and E_c are low-field carrier mobility and critical electric field of charge carriers. The low-field carrier mobility μ_0 determines carrier velocity under low electric filed, i.e. $v \approx \mu_0 E_p$ if $E_p << E_c$. Drifting velocity of charge carriers saturates when channel electric field increases. From (3.2), carrier mobility as a function of channel

electric field can be expressed as follows:

$$\mu(E_p) = \frac{\mu_0}{1 + |E_p / E_c|} \tag{3.3}$$

Fig. 3.2 plots the degradation of carrier mobility with channel electric field E_p from (3.3). Obviously charge carrier mobility decreases due to velocity saturation when channel electric field increases. The critical electric field E_c is assumed to be the ratio of critical voltage V_c to its effective length L_{eff} $(E_c = V_c / L_{eff})$. The channel electric field E_p is approximated by drain-source voltage V_{ds} divided by effective length L_{eff} $(E_p = V_{ds} / L_{eff})$. With these relations, (3.2) can be rewritten as

$$v = \frac{|\mu_0 \cdot E_p|}{1 + V_{ds} / V_c} \tag{3.4}$$

where V_c is critical voltage of charge carriers of a MOSFET. The term of V_{ds} / V_c in (3.4), which does not exist in long channel low-field MOSFETs, is caused by the effect of charge carrier velocity saturation. The critical voltage V_c is defined as the product of critical electric field E_c and effective channel length L_{eff} $(V_c = E_c L_{eff})$. For a long channel MOSFET the critical voltage V_c proportional to L_{eff} is much larger than drain-source voltage V_{ds} so that the term of V_{ds} / V_c is small and can be ignored without causing significant error. When effective length L_{eff} becomes smaller, critical voltage V_c proportional to L_{eff} is comparable to or even smaller than drain-source voltage V_{ds}. Therefore, the effect of charge carrier velocity saturation has to be considered in deep submicron MOSFET models.

If channel electric field E_p of (3.2) tends to be infinity, the drifting velocity is termed the charge carrier saturation velocity indicated with v_{sat}. From (3.2), we can obtain the relationship between saturation velocity v_{sat} and the critical electric field E_c

$$v_{sat} = \lim_{E_p \to \infty} \frac{|\mu_0 \cdot E_p|}{1 + |E_p / E_c|} = |\mu_0 \cdot E_c| \tag{3.5}$$

In other words, the critical electric field E_c and critical voltage V_c can be expressed as

$$E_c = \frac{v_{sat}}{\mu_0} \text{ and } V_c = \frac{v_{sat}}{\mu_0} L_{eff} \tag{3.6}$$

Therefore, the critical electric field E_c is equal to charge carrier saturation velocity v_{sat} divided by low-field carrier mobility μ_0, which is illustrated in Fig. 3.1.

3.3 DERIVED MOSFET MODEL WITH EFFECT OF CARRIER VELOCITY SATURATION

By combining the gradual-channel approximation (GCA) [42] [51] [78] used in the derivation of Shockley MOSFET model and the effect of carrier velocity saturation predicted by (3.4), we obtain an expression for drain-source current I_{ds} in linear mode as follows:

$$I_{ds} = \frac{k}{1 + V_{ds}/V_c} \left(V_{gs} - V_{th} - \frac{V_{ds}}{2} \right) \cdot V_{ds} \qquad \left(V_{gs} > V_{th}, V_{ds} \le V_{sat} \right) \tag{3.7}$$

where k is the transconductance parameter, V_{th} the threshold voltage, V_c the critical voltage, V_{sat} the drain-source saturation voltage needed to discuss later. The above expression I_{ds} is identical to Shockley MOSFET model except for the term $1 + V_{ds}/V_c$ in the denominator caused by the effect of charge carrier velocity saturation. If the other

model parameters are the same as in Shockley MOSFET model it is obvious that the drain-source current I_{ds} from (3.7) is less than that from Shockley MOSFET model due to the effect of charge carrier velocity saturation. The saturation voltage V_{sat} with the effect of carrier charge velocity saturation of a MOSFET can be obtained from (3.7)

$$V_{sat} = V_c \cdot \left(\sqrt{1 + 2 \cdot \frac{V_{gs} - V_{th}}{V_c}} - 1 \right) \tag{3.8}$$

The detailed derivation of saturation voltage V_{sat} is shown in Section I.1 of Appendix I. Substituting V_{ds} of (3.7) with saturation voltage V_{sat} of (3.8) yields drain-source saturation current I_{sat}. The detailed derivation of drain-source saturation current I_{sat} is demonstrated in Section I.2 of Appendix I. The resulting expression of drain-source saturation current I_{ds} is simply equal to

$$I_{ds} = \frac{k}{2} V_{sat}^2 \qquad \left(V_{gs} > V_{th}, V_{ds} > V_{sat} \right) \tag{3.9}$$

Except for saturation voltage V_{sat} expressed in (3.8), the expression of drain-source saturation current I_{ds} with regard to saturation voltage V_{sat} is the same as the one of Shockley MOSFET mode in saturation mode. Hence, in the presence of charge carrier velocity saturation, the drain-source current I_{ds} in saturation mode still varies quadratically with saturation voltage V_{sat}. The saturation voltage V_{sat} in (3.8), however, is no longer linearly dependent on gate-source voltage V_{gs} so that the drain-source saturation current I_{sat} is not quadratically dependent on gate-source voltage V_{gs}. Combining the above discussions into one equation yields a derived MOSFET mode with the effect of charge carrier velocity saturation as follows:

$$I_{ds} = \begin{cases} 0 & V_{gs} \le V_{th} \\ \dfrac{k}{1+V_{ds}/V_c} \cdot \left(V_{gs} - V_{th} - \dfrac{V_{ds}}{2} \right) \cdot V_{ds} & V_{gs} > V_{th}, V_{ds} \le V_{sat} \\ \dfrac{1}{2} \cdot k \cdot V_{sat}^2 & V_{gs} > V_{th}, V_{ds} \le V_{sat} \end{cases} \quad (3.10)$$

where k is the transconductance parameter, V_{th} the threshold voltage, V_{sat} the drain-source saturation voltage given in (3.8). For deep submicron technology, the critical voltage V_c proportional to effective channel length is around or even smaller than drain-source voltage V_{ds}. The effect of channel charge carrier velocity saturation has to be considered. For a long channel MOSFET operating voltage is much smaller than the critical voltage V_c proportional to effective channel length. The ignorance of the term V_{ds}/V_c makes (3.7) become the drain-source current expression of Shockley MOSFET model in linear mode. Moreover, the second term in the Taylor series expansion of the square root of (3.8) is much less than unity in this case. The drain-source saturation voltage V_{sat} of (3.8) is reduced to

$$V_{sat} = V_c \cdot \left(\sqrt{1 + 2 \cdot \frac{V_{gs} - V_{th}}{V_c}} - 1 \right) \approx V_{gs} - V_{th} \qquad \left(V_{gs} << V_c \right) \quad (3.11)$$

Thus the drain-source saturation voltage V_{sat} in (3.8) tends to be $\left(V_{gs} - V_{th} \right)$ as predicted for saturation voltage V_{sat} of (2.5) in Shockley MOSFET model. Therefore, this derived MOSFET model is reduced to Shockley MOSFET model when the critical voltage V_c is sufficiently large comparing with operating voltage. In other words, it is an extension of Shockley MOSFET model.

3.4 POWER INDEX OF A DEEP SUBMICRON MOSFET

The static I–V characteristics of a deep submicron MOSFET in saturation mode are completely described with the dependence of drain-source current I_{ds} on gate-source input voltage V_{gs} and drain-source voltage V_{ds}. The relationship between drain-source current I_{ds} and drain-source saturation voltage V_{sat} for a given gate-source voltage V_{gs} is also important for static I–V characteristics. It depicts the boundary between linear mode and saturation mode of a deep submicron MOSFET. In this section let's investigate the relationships of a deep submicron MOSFET in saturation mode.

As we have already known, the relationship between drain-source current I_{ds} and drain-source voltage V_{ds} for a given gate-source voltage V_{gs} in saturation mode is well described with the effect of channel length modulation as in Section 3.1. In the following discussions of this section, we temporarily ignore the effect of channel length modulation.

In saturation mode the dependence of drain-source current I_{ds} on drain-source saturation voltage V_{sat} for a given gate-source voltage V_{gs} is the same for Shockley MOSFET mode in (2.2) and for the derived MOSFET model in (3.9) or (3.10), i.e.,

$$I_{ds} = \frac{k}{2} V_{sat}^2 \qquad \left(V_{gs} > V_{th}, V_{ds} > V_{sat}\right) \quad (3.12)$$

In the other words, the expression of (3.12) is generic to describe the dependence of drain-source current I_{ds} on saturation drain-source voltage V_{sat}. It is sufficiently accurate to depict the drain-source saturation current I_{ds} of not only a long channel MOSFET, but a deep submicron MOSFET also. We will further demonstrate that it is still valid for the chain current of a series-connect MOSFET chain in saturation mode in Chapter 5 as well.

The saturation drain-source voltage V_{sat} in (3.12) depending on gate-source voltage for a long channel MOSFET is however different for a deep submicron MOSFET. For a long channel MOSFET, the drain-source saturation voltage V_{sat} is given by (2.5) of Shockley MOSFET model. For a deep submicron MOSFET, the drain-source saturation voltage V_{sat} is given by (3.8). The relationship of drain-source current I_{ds} and gate-source voltage V_{gs} in saturation mode is obtained by substituting the corresponding drain-source saturation voltage V_{sat} into (3.12). In general, drain-source current I_{ds} in saturation mode has a power-law relationship with gate-source voltage V_{gs}, i.e.,

$$I_{ds} \propto \left(V_{gs} - V_{th}\right)^{\beta} \tag{3.13}$$

The parameter β in (3.13) is defined as power index of a MOSFET. Obviously, Shockley MOSFET model predicts that drain-source current I_{ds} quadratically depends on gate-source input voltage V_{gs} in saturation mode for a long channel MOSFET. Thus, the power index β of a long channel MOSFET is two. The MOSFET power index β, derived from (3.13) is expressed as

$$\beta = \frac{\left(V_{gs} - V_{th}\right)}{I_{ds}} \frac{dI_{ds}}{dV_{gs}} \tag{3.14}$$

Fig. 3.3 shows the dependence of drain-source current I_{ds} on gate-source voltage V_{gs} at $V_{ds} = V_{dd}$ for a deep submicron pMOSFET and a deep submicron nMOSFET, respectively. The power indices β of a deep submicron pMOSFET and a deep submicron nMOSFET by using (3.14) and the derived MOSFET model (3.10) of Section 3.3 are plotted in Fig. 3.4. It shows that the power index β of a deep submicron MOSFET is two

when gate-source voltage is equal to threshold voltage and is gradually and monotonically decreasing as gate-source voltage V_{gs} increases. This can be explained with (3.8), (3.12), and (3.14).

The reduction of power index β is caused by charge carrier velocity saturation of a deep submicron MOSFET so that power index β is also called the velocity saturation index [85]. As we can see from (3.8) and (3.12), the upper limit of power index β of a deep submicron MOSFET is two, which is the same as a long channel MOSFET. The lower limit of power index β of a deep submicron MOSFET is one. We will demonstrate that power index of a series-connected MOSFET chain is less than one in Chapter 5.

3.5 DEEP SUBMICRON MOSFET MODEL (DSMM)

In order to obtain accurate deep submicron MOSFET models, the important second-order effects of a deep submicron MOSFET such as carrier velocity saturation and channel length modulation have to be considered. In Section 3.3, we have achieved a derived MOSFET current model incorporating the effect of velocity saturation of charge carriers. In the analysis of the derived MOSFET current model, the gradual-channel approximation (GCA) is assumed to be valid [42] [51] [78], as in the derivation of *Shockley* MOSFET model. The derived MOSFET current model, however, does not include the effect of channel length modulation. Taking into account the effect of channel length modulation as in *Shichman-Hodges* model, we extend the derived MOSFET current model to the deep submicron MOSFET model (DSMM). Clearly, the important second-order effects of a deep submicron MOSFET are included in DSMM. The drain-source current I_{ds} in DSMM is expressed as

$$I_{ds}=\begin{cases}0 & V_{gs}\le V_{th}\\ \dfrac{k}{1+V_{ds}/V_c}\left(V_{gs}-V_{th}-\dfrac{V_{ds}}{2}\right)V_{ds}(1+\lambda V_{ds}) & V_{gs}>V_{th},V_{ds}\le V_{sat}\\ \dfrac{1}{2}kV_{sat}{}^2(1+\lambda V_{ds}) & V_{gs}>V_{th},V_{ds}>V_{sat}\end{cases} \tag{3.15}$$

In (3.15) k is the transconductance parameter, V_{th} the gate-source threshold voltage, V_{sat} the drain-source saturation voltage, V_c the critical voltage, and λ the channel length modulation coefficient. The drain-source saturation voltage V_{sat} in DSMM is given by (3.8), i.e., $V_{sat}=V_c\cdot\left(\sqrt{1+2\cdot(V_{gs}-V_{th})/V_c}-1\right)$. Both the additional term $(1+V_{ds}/V_c)$ in the linear mode of (3.15) and the drain-source saturation voltage V_{sat} given by (3.8) deviated from (2.5) are caused by the effect of charge carrier velocity saturation.

It is easily verified that the DSMM model becomes Shichman-Hodges model [42] [88] if operating voltage V_{dd} is much less than critical voltage V_c that is the case of a long channel MOSFET. The DSMM model in (3.15) and (3.8) is the extension of Shockley model and Shichman-Hodges model with strong physical meanings. The DSMM model substantially improves the accuracy to describe deep submicron MOSFET *I–V* behaviors since the second-order effects of deep submicron MOSFETs are considered in DSMM. It is suitable to model both long and short channel MOSFET. Take half-micron technology as an example, assume $L_{eff}=0.5\cdot10^{-4}(cm)$, $v_{sat}\approx10^7(cm/\sec)$, and $\mu_0\approx500(cm^2/V\cdot\sec)$ [92]. From (3.6), we have critical voltage $V_c=\dfrac{v_{sat}}{\mu_0}L_{eff}\approx1(V)$. Comparing the operating voltage range (1.8~3.3volts) of a half-micron technology, the effect of the critical voltage V_c cannot be ignored. In other words,

either Shockley MOSFET model or Shichman-Hodges MOSFET model is inaccurate to describe a half-micron MOSFET.

Fig. 3.5 shows the static I–V characteristics of a deep submicron MOSFET from deep submicron MOSFET model (DSMM). As a comparison, the static I–V characteristics from SPICE BSIM3 model are also plotted. From Fig. 3.5, we can see that the DSMM model accurately describes deep submicron MOSFET behaviors due to incorporating important second-order MOSFET effects. The DSMM, however, will cause the differential equation of analytical timing approaches to become Abel's differential equation, which cannot be solved analytically [1]. The proper deep submicron MOSFET model (PDSMM) is presented to simplify the DSMM with accuracy in the next section.

3.6 PROPER DEEP SUBMICRON MOSFET MODEL (PDSMM)

The DSMM model given in (3.15) and (3.8) describes the static I–V characteristics of a deep submicron MOSFET. It is easy to derive Shockley model [90] or Shichman Hodges model [88] if we ignore some second order effects. It is the extension of Shockley model or Shichman-Hodges model and is easily used in numerical approaches of VLSI circuit analysis. However the differential equation to describe the transient behaviors of a CMOS inverter in timing analysis cannot be solved analytically due to the term $\left(1+V_{ds}/V_c\right)$ of (3.15). In order to solve this issue, a proper deep submicron MOSFET model (PDSMM) is proposed below. The PDSMM in linear mode takes the similar formula of Shichman-Hodges MOSFET model. The drain-source current I_{ds} in saturation mode and the drain-source saturation voltage V_{sat} in PDSMM are kept the same as those in DSMM. With this modification, the PDSMM model is expressed as:

$$I_{ds} = \begin{cases} 0 & V_{gs} \le V_t \\ k\left(V_{sat} - \dfrac{V_{ds}}{2}\right)V_{ds}(1+\lambda V_{sat}) & V_{gs} > V_t, V_{ds} \le V_{sat} \\ \dfrac{1}{2}kV_{sat}{}^2(1+\lambda V_{ds}) & V_{gs} > V_t, V_{ds} > V_{sat} \end{cases} \tag{3.16}$$

In (3.16) k is the transconductance parameter, V_{th} the gate-source threshold voltage, V_{sat} the drain-source saturation voltage, and λ the MOSFET channel length modulation coefficient. The drain-source saturation voltage V_{sat} of PDSMM, is the same as DSMM given by (3.8), i.e., $V_{sat} = V_c \cdot \left(\sqrt{1+2\cdot(V_{gs}-V_{th})/V_c} - 1\right)$. The term of $(1+\lambda \cdot V_{sat})$ in the equation of linear mode is used to meet the continuity of drain-source current at linear-saturation boundary $(V_{ds} = V_{sat})$. Note we use $(1+\lambda \cdot V_{sat})$, instead of $(1+\lambda \cdot V_{ds})$, in the equation of linear mode. If the drain current I_{ds} depends cubically on drain-source voltage, the analytical solution of differential equation in analytical propagation delay approaches is almost impossible. The critical voltage V_c exists in the expression of drain-source saturation voltage of (3.8). Fig. 3.6 plots the static *I–V* characteristics of a deep submicron nMOSFET from PDSMM and from SPICE BSIM3 model. It shows that PDSMM is very accurate to reproduce the static *I–V* characteristics of a deep submicron MOSFET.

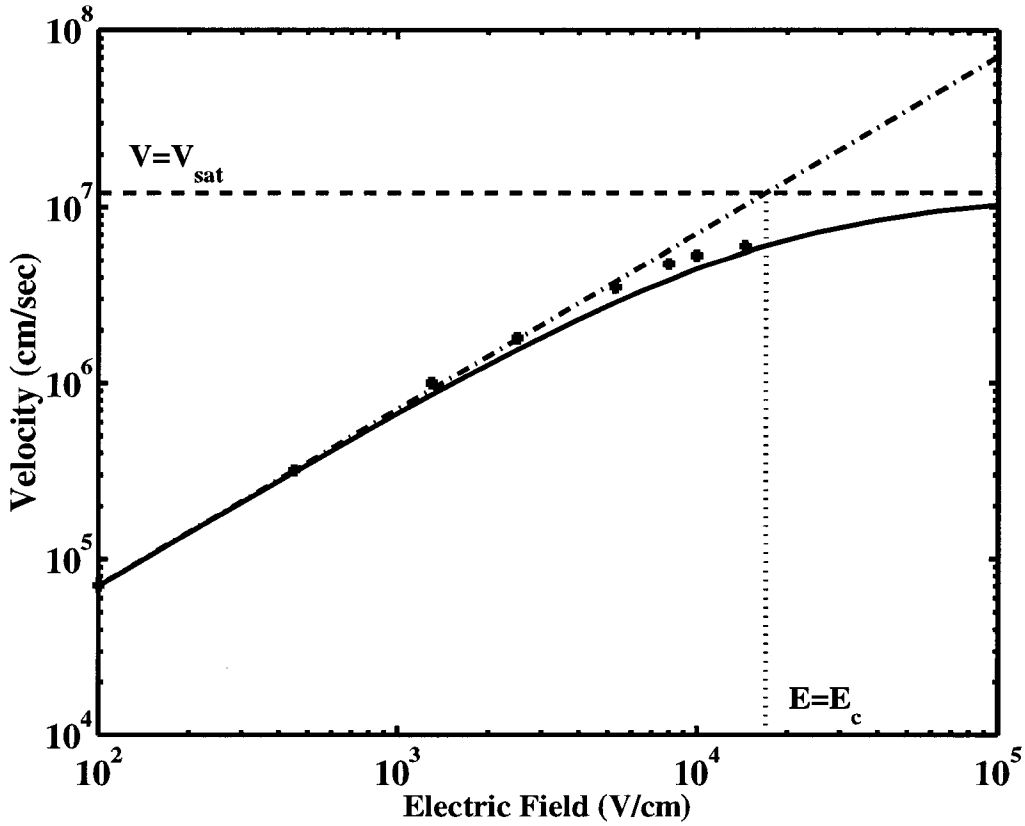


Fig. 3.1 Drifting velocity of surface electrons as a function of channel electric field [22]. Drifting velocity saturates when electric field increases. The critical electric field E_c is obtained by dividing saturation drifting velocity V_{sat} with low-field mobility μ_0.

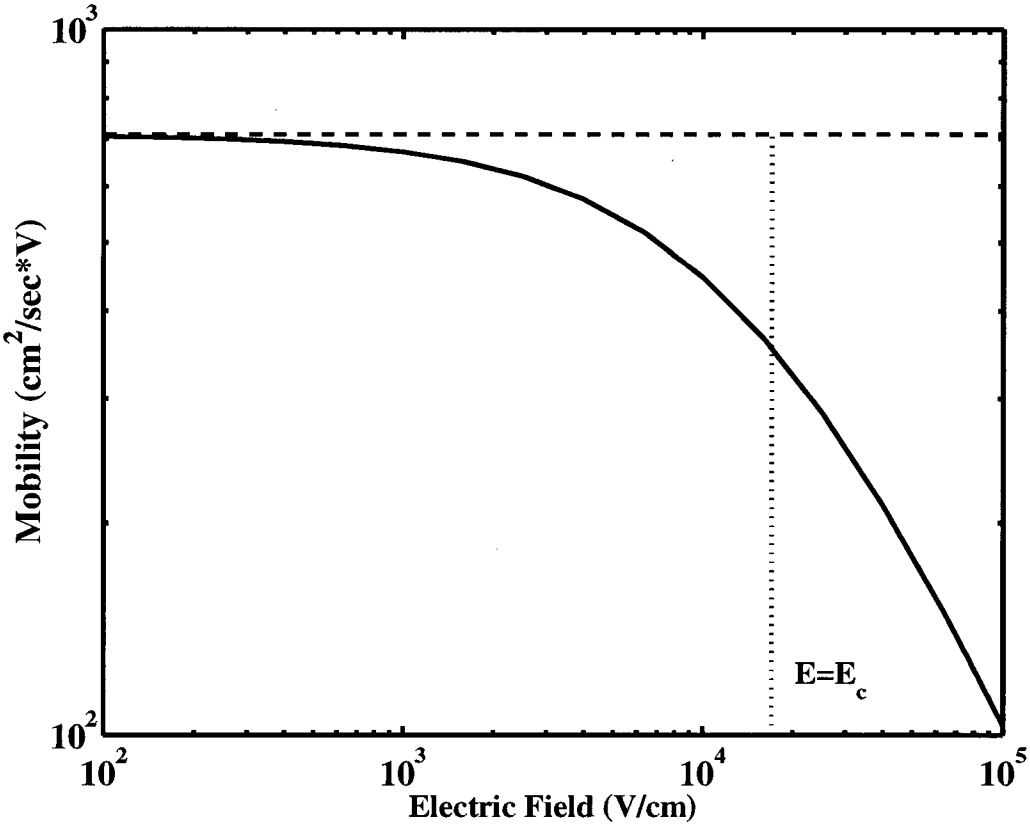


Fig. 3.2 Mobility of surface electrons as a function of channel electric field derived from (3.3) for data of Fig. 3.1. The mobility of surface electrons decreases as channel electric field increases or drifting velocity of surface electrons saturates.

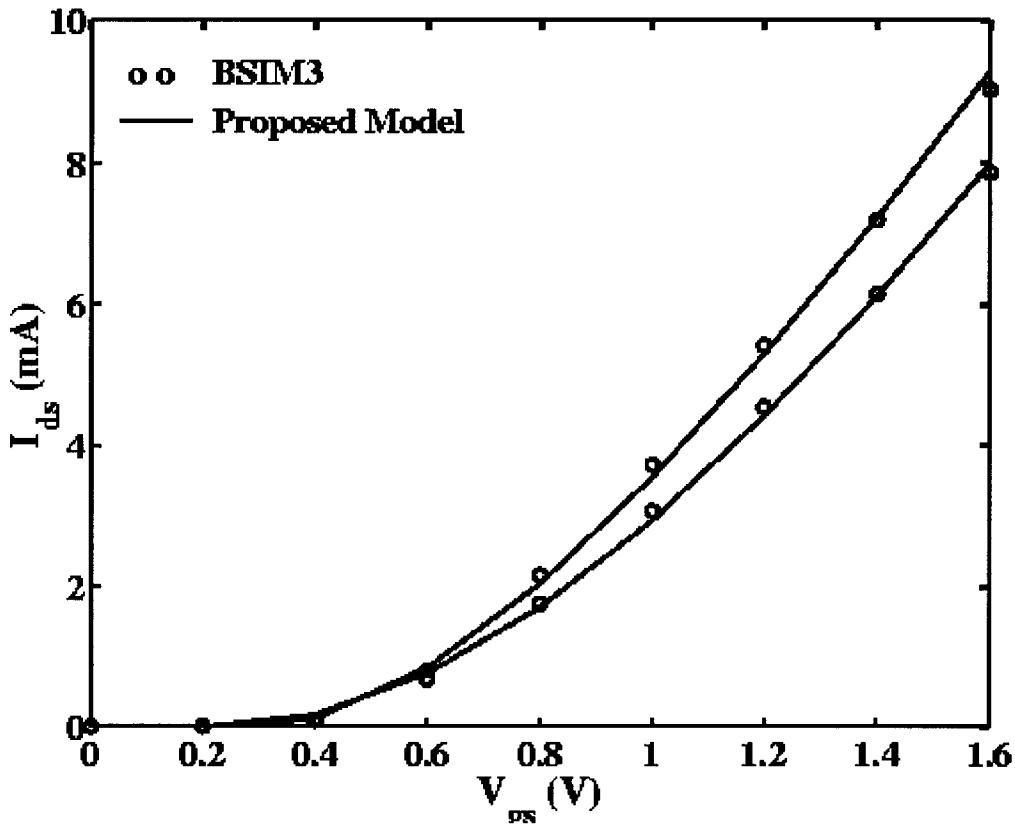


Fig. 3.3 Saturation drain-source currents as a function of gate-source voltage at $V_{ds} = V_{dd}$ from SPICE BSIM3 model and from the proposed models (DSMM and PDSMM) for a deep submicron nMOSFET (upper line) and pMOSFET (lower line), respectively. It shows that the proposed models (DSMM and PDSMM) can accurately reproduce I–V characteristics of a deep submicron MOSFET in saturation region.

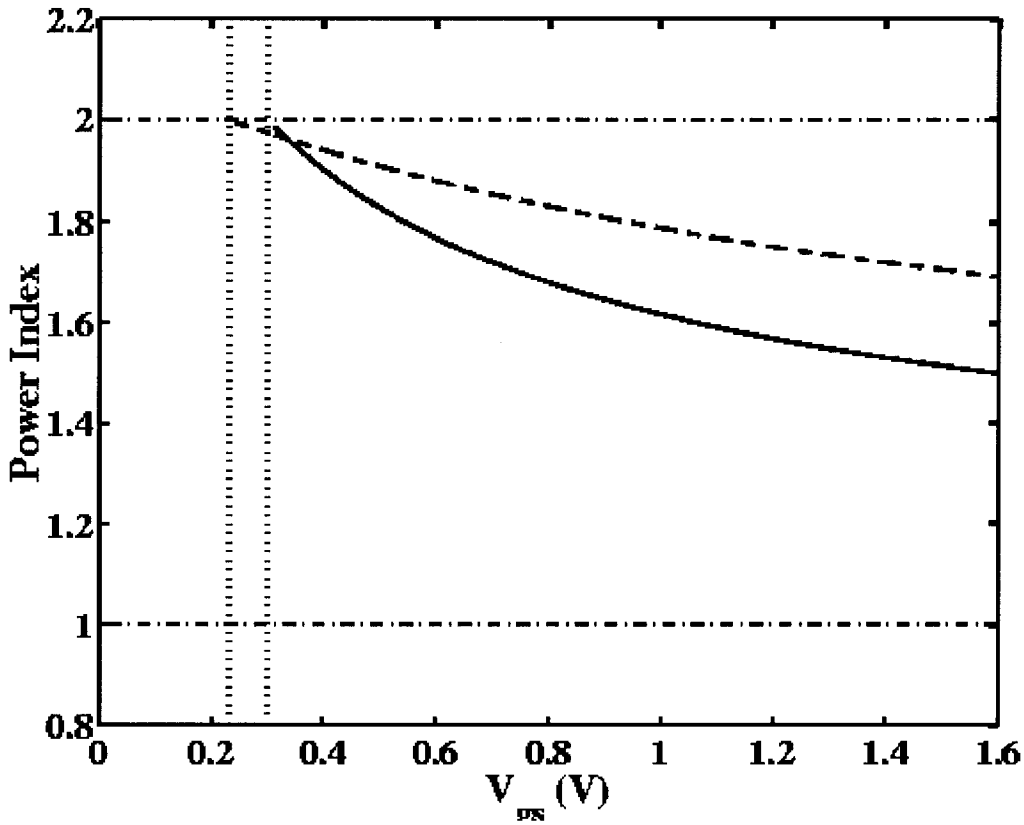


Fig. 3.4 Power indices as a function of gate-source voltage for a deep submicron pMOSFET (dashed line) and nMOSFET (solid line), respectively. The vertical dotted lines indicate the corresponding threshold voltages. The power index β of a deep submicron MOSFET is two when gate-source voltage is equal to threshold voltage and gradually decreases as gate-source voltage increases. The lower limit of power index β of a deep submicron MOSFET is one.

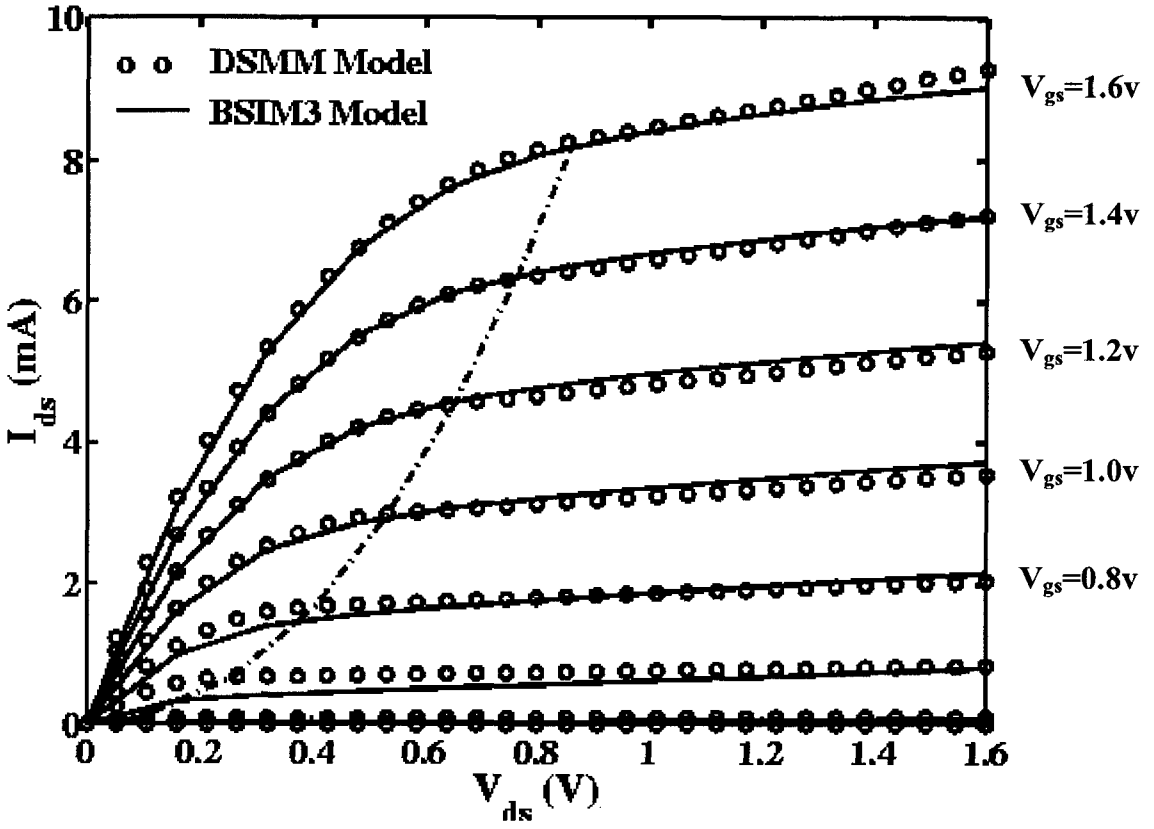


Fig. 3.5 Static I–V characteristics are plotted from deep submicron MOSFET model (DSMM) and SPICE BSIM3 model for a deep submicron MOSFET. The static I–V characteristics from DSMM are well matched with those from SPICE BSIM3 model.

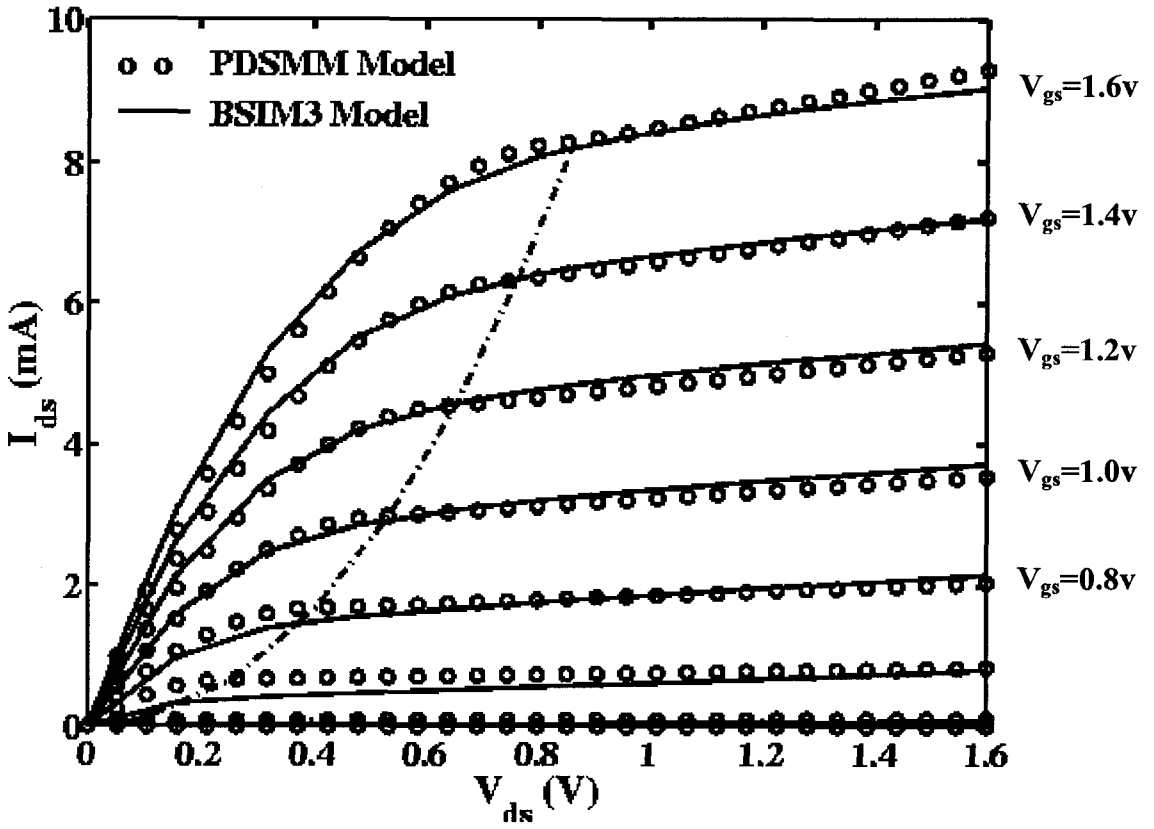


Fig. 3.6 Static I–V characteristics are plotted from proper deep submicron MOSFET model (PDSMM) and SPICE BSIM3 model for a deep submicron MOSFET. The static I–V characteristics from PDSMM are well matched with those from SPICE BSIM3 model.

CHAPTER 4 TIMING MODELS OF A DEEP SUBMICRON CMOS INVERTER

The propagation delay of a deep submicron CMOS inverter is extracted from the analytical expressions of transient output voltage. We start from the differential equation, which controls the transient properties of output voltage of a deep submicron CMOS inverter. The differential equation based on the proper deep submicron MOSFET model (PSDMM) is analytically solved for transient output voltage in three ways according to different combinations of input pattern and circuit model to be used. Input voltage is a step function or a ramp function. A CMOS inverter for a rising input is represented by a dominating nMOSFET with or without a trivial pMOSFET.

As the first method, the step response of a CMOS inverter is derived with the input voltage of a step function. The analytical expressions of step-input propagation delay and output transition time depending on external load capacitance and MOSFET model parameters are obtained from this method. This simple method is accurate for a CMOS inverter with very fast input. However, the effect of input transition time is ignored in the analysis. It yields a significant error for a ramp input in practice.

The second approach of propagation delay is that input is a rising ramp function and only the dominating nMOSFET without a trivial pMOSFET is approximated to model a CMOS inverter. The ramp response of the nMOSFET-only circuit representing a CMOS inverter is analytically developed with input transition time, external load capacitance, and MOSFET model parameters. Ignoring a trivial pMOSFET leads to two phenomena. The first observable fact is when input voltage rises output voltage overshoot

exaggerates due to no inverse discharging current from output node to power rail, especially for small load capacitance. The other phenomenon is the nMOSFET discharging current without the compensated short-circuit current makes the output voltage drop unusually fast, especially for large input transition time. The second approach of propagation delay incorporating input transition time is more complex than the first method with a step input. It is accurate in much broader range of input transition time and external load capacitance, especially for a very fast input or for a comparable external load capacitance. It may cause a significant error with large input transition time and small external load capacitance due to the ignorance of a trivial pMOSFET in the circuit model of a CMOS inverter.

The third method of propagation delay is the comprehensive propagation delay model for a CMOS inverter. The dominating nMOSFET, trivial pMOSFET, input transition time and external load capacitance of a CMOS inverter are taken into account in the analysis of transient output voltage expressions. Their impacts on propagation delay of a CMOS inverter are investigated. For a rising input the trivial pMOSFET provides the inverse discharging current to power rail and the short-circuit current to compensate the nMOSFET discharging current. The output voltage response of the comprehensive propagation delay is very accurate comparing with SPICE BSIM3 simulations, even in the extreme operation condition such as large input transition time and small external load capacitance. From the discussion we can see the tradeoffs between accuracy and simplicity with these three methods. Generally speaking, the computation speeds of propagation delay by using these timing models are more than two orders of magnitude faster than with SPICE [21] [31] [70] [71] [89]. More importantly,

the analytical delay models can be easily incorporated into VLSI CMOS design, optimization, and verification flows.

In the following derivation of the analytical timing model and analysis of a deep submicron CMOS inverter, we concentrate on rising input and falling output. All the arguments made are symmetrically valid for falling input and rising output as well. The difference is that for the case of falling input and rising output, the load capacitance is charged through the dominating pMOSFET while the nMOSFET becomes trivial.

4.1 NORMALIZED VARIABLES

For simplicity, we sometimes use normalized voltage variables and normalized time variables instead of their original variables. Voltage variables are normalized with respect to operating voltage V_{dd}. Normalized voltage variables are expressed with starting letter u instead of original voltage variables starting with V, i.e., we have $u_i = V_i / V_{dd}$, $u_o = V_o / V_{dd}$, $u_{tn} = V_{tn} / V_{dd}$, $u_{tp} = V_{tp} / V_{dd}$, $u_{sn} = V_{sn} / V_{dd}$, $u_{sp} = V_{sp} / V_{dd}$, $u_{cn} = V_{cn} / V_{dd}$ and $u_{cp} = V_{cp} / V_{dd}$. Time variables are normalized with respect to input transition time τ. Normalized time variables are indicated as starting letter s. We have the normalized time variable $s = t / \tau$. Normalized voltage variables during input transition may be equal to the corresponding normalized time variables, for example, $u_{tn} = s_{tn}$.

4.2 CMOS INVERTER

4.2.1 Differential Equation

In analytical timing approaches, the transient expressions of output voltage of a deep submicron CMOS inverter are obtained by analytically solving the differential

equation, which describes its transient response. The propagation delay is then extracted from the transient expressions of output voltage. A generic CMOS inverter shown in Fig. 4.1 is used to build the differential equation. Note that the effects of input-output coupling capacitance C_m and output load capacitance C_{load} are included in Fig. 4.1. The differential equation to describe the transition behaviors of CMOS inverter shown in Fig. 4.1 is obtained from the application of the Kirchoff's current law (KCL) at output node:

$$-I_{C_{load}} + I_{C_m} + I_p - I_n = 0 \tag{4.1}$$

Substituting capacitance characteristics into the above equation yields

$$\frac{dV_o}{dt} = k_c \cdot \frac{dV_i}{dt} + \frac{1}{C_t}\left(I_p - I_n\right) \tag{4.2}$$

where C_t is the total capacitance of output node, which is

$$C_t = C_{load} + C_m \tag{4.3}$$

and k_c the input-output coupling coefficient. It is the ratio of the input-output coupling capacitance C_m to the total capacitance C_t of output node, i.e.,

$$k_c = C_m / C_t \tag{4.4}$$

If input voltage is a ramp function with input transition time τ the differential equation of (4.2) with the normalized time variable s becomes

$$\frac{dV_o}{ds} = k_c V_{dd}\left[u(s) - u(s-1)\right] + \frac{\tau}{C_t}\left(I_p - I_n\right) \tag{4.5}$$

Before we discuss analytical solutions of the differential equations of (4.2) or (4.5), the capacitances and the MOSFET current models I_p and I_n are briefly examined in the next subsections.

4.2.2 Capacitances

The input-output coupling capacitance C_m of a CMOS inverter consists of gate-drain coupling capacitances C_{mn} and C_{mp} for nMOSFET and pMOSFET, respectively,

$$C_m = C_{mn} + C_{mp} \tag{4.6}$$

where gate-drain coupling capacitances C_{mn} and C_{mp} are characterized with SPICE. For simplicity, gate-drain coupling capacitances C_{mn} and C_{mp} per unit channel width are assumed to be voltage-independent and determined by manufacturing technology. In practice, their values depend on both input voltage and output voltage [2] [51] [59] [104].

The total capacitance C_t of output node is the summation of output load capacitance C_{load} and input-output coupling capacitance C_m as in (4.3). It is also the summation of total parasitic drain diffusion capacitance C_d and external load capacitance C_o, i.e.,

$$C_t = C_d + C_o \tag{4.7}$$

where parasitic drain diffusion capacitance C_d is the summation of parasitic drain diffusion capacitances C_{dn} and C_{dp} for nMOSFET and pMOSFET, respectively,

$$C_d = C_{dn} + C_{dp} \tag{4.8}$$

In (4.8), parasitic drain diffusion capacitances C_{dn} and C_{dp} are characterized with SPICE. For convenience, the parasitic drain diffusion capacitances C_{dn} and C_{dp} per unit channel width are assumed to be voltage-independent and determined by the process technology. In practice, their values depend on both input voltage and output voltage [2] [51] [59]

[104]. The external load capacitance C_o in (4.7) is a primary variable in the analytical timing approaches. In fact, it is compound of wire capacitances at output node and fanout input capacitances of next stage.

4.2.3 MOSFET Current Models

In order to analytically solve the differential equations (4.2) or (4.5), it is necessary to carefully choose the MOSFET current models to be used. These MOSFET models should be as simple and accurate as necessary. Fig. 4.2 plots the typical discharging traces of both pMOSFET and nMOSFET for rising ramp inputs. The static *I–V* characteristics of both nMOSFET and pMOSFET are also shown in the figure. For rising ramp inputs, the nMOSFET experiences cutoff mode, saturation mode, and linear mode. The nMOSFET drain-source current I_{ds} dominates the output transition or discharging process. Therefore, the nMOSFET should be modeled as accurately as possible. As we have discussed in Chapter 3, the proper deep submicron MOSFET model (PDSMM) can accurately depict nMOSFET static *I–V* behaviors. The proposed accurate PDSMM is used in the following analyses of output expressions and propagation delays. The nMOSFET discharging current I_n of a CMOS inverter is equal to nMOSFET drain-source current I_{ds}. The input voltage V_i and output voltage V_o are equal to nMOSFET gate-source voltage V_{gs} and drain-source voltage V_{ds}, respectively. With new variables of I_n, V_i, and V_o, the nMOSFET drain-source current of PDSMM from Chapter 3 yields

$$I_n = \begin{cases} 0 & V_i \le V_{tn} \\ k_n\left(V_{sn} - \frac{V_o}{2}\right)V_o(1+\lambda_n V_{sn}) & V_i > V_{tn}, V_o \le V_{sn} \\ \frac{1}{2}k_n {V_{sn}}^2(1+\lambda_n V_o) & V_i > V_{tn}, V_o > V_{sn} \end{cases} \quad (4.9)$$

In (4.9), k_n is the nMOSFET drivability coefficient, λ_n is the nMOSFET channel length modulation coefficient, and V_{tn} is the nMOSFET threshold voltage. The nMOSFET drain-source saturation voltage V_{sn} from PDSMM as in (3.8) is expressed as

$$V_{sn} = V_{cn} \cdot \left(\sqrt{1+2\cdot(V_i - V_{tn})/V_{cn}} - 1\right) \quad (4.10)$$

where V_{cn} is the nMOSFET critical voltage.

As we can see from Fig. 4.2, the trivial pMOSFET of a CMOS inverter does not dominate the discharging process. It still affects the output transition for a rising input. In order to develop the input rising propagation delay of a deep submicron CMOS inverter accurately and thoroughly, the trivial pMOSFET cannot be ignored. There are two major behaviors in the pMOSFET operation trajectory traces as shown in Fig. 4.2. The pMOSFET operation traces experience either the small pMOSFET source-drain voltage $V_{sd}\left(=V_{op}=V_{dd}-V_o\right)$ or the small pMOSFET source-gate voltage $V_{sg}\left(=V_{ip}=V_{dd}-V_i\right)$. In either case, the effect of channel length modulation is small so that it is ignored without making a comparable error. Furthermore, when the applied input voltage V_i is small, the output voltage V_o is close to operating voltage. The pMOSFET source-drain voltage is small. The linearized Shockley MOSFET model is accurate enough to illustrate the pMOSFET characteristics in this linear mode. When applied input voltage V_i is large, the pMOSFET enters into the saturation mode. The pMOSFET source-gain voltage is small

and the square-law equation of Shockley MOSFET model is accurate enough to reproduce the pMOSFET characteristics in the saturation mode. Therefore, the simplified and linearized MOSFET model from Shockley MOSFET model for a rising input accurately enough describes the pMOSFET static I–V characteristics under both operation conditions, i.e.,

$$I_p = \begin{cases} 0 & V_{ip} \le V_{tp} \\ k_p V_{sp} V_{op} & V_{ip} > V_{tp}, V_{op} \le V_{sp}/2 \\ \frac{1}{2} k_p V_{sp}^{\ 2} & V_{ip} > V_{tp}, V_{op} > V_{sp}/2 \end{cases} \tag{4.11}$$

where the pMOSFET saturation voltage V_{sp} is given by:

$$V_{sp} = V_{ip} - V_{tp} \tag{4.12}$$

The intersection point of the two linearized lines corresponding to linear mode and saturation mode of (4.11) is determined by half of the pMOSFET saturation voltage V_{sp}. Thus, in the following propagation delay analysis of a deep submicron CMOS inverter for a rising input, the proposed accurate PDSMM given by (4.9) and (4.10) is used to describe the nMOSFET static *I–V* characteristics. The pMOSFET is modeled by the simplified equations of (4.11) and (4.12).

4.3 TIMING ANALYSIS OF A CMOS INVERTER FOR A STEP INPUT

Propagation delay is the time taken for a logic transition to pass from input to output. The propagation delay of a CMOS inverter for a step input is usually described as the speed of the CMOS inverter. We investigate the output voltage transition and propagation delay of a CMOS inverter for a step input in this section. First, the step

response of output voltage $V_o(t)$ is analytically calculated. Then the propagation delay time t_{dHL} for a rising step input is determined as the time needed for the output voltage to reduce to half of operating voltage by discharging total load capacitance C_t through the dominating nMOSFET transistor.

4.3.1 Operation of a CMOS Inverter for a Step Input

The input voltage $V_i(t)$ to the CMOS inverter is a step function,

$$V_i(t) = V_{dd}u(t) \tag{4.13}$$

where $u(t)$ is a unit step function. The discharging trajectory traces of both nMOSFET and pMOSFET of a CMOS inverter for a rising step input are shown in Fig. 4.3. The characteristics of both nMOSFET and pMOSFET are also plotted. When time is less than zero $(t < 0\_)$, input voltage $V_i(t)$ is zero. The nMOSFET is in cutoff mode and the pMOSFET is in linear mode. Assuming that the inverter is in the steady state before time of zero, the output voltage $V_o(t)$ is equal to operating voltage V_{dd}, i.e., $\left(V_o(t)\big|_{t=0\_} = V_{dd}\right)$. At the time of zero $(t = 0)$, the nMOSFET trajectory trace of the CMOS inverter starts from $(V_{dd},0)$ of the (V_{ds},I_n) plan at the beginning and jumps to its saturation mode of characteristic curve with its gate-source voltage of V_{dd}. The pMOSFET, instead, starts from $(0,0)$ of the (V_{sd},I_p) plan at the beginning and jumps to its cutoff mode. Once time is greater than zero $(t > 0_+)$, the operating voltage V_{dd} is applied to the input of CMOS inverter. The nMOSFET then discharges along the characteristic curve with its gate-source voltage equal to V_{dd} from its saturation mode into its linear mode. The pMOSFET stays in its cutoff mode. The pMOSFET source-drain current does not affect the step

response of a CMOS inverter. Therefore the output voltage response $V_o(t)$ for a rising step input is derived from the following differential equation by substituting (4.13) into (4.2)

$$\frac{dV_o}{dt} = k_c V_{dd}\delta(t) - \frac{1}{C_t}\cdot I_n \tag{4.14}$$

where $\delta(t)$ is Dirac function.

4.3.2 Timing Analysis of a CMOS Inverter for a Step Input

According to the above discussions, we analyze the transient properties of a deep submicron CMOS inverter in three regions for a step input. The propagation delay for a step input is extracted from the transition of output voltage.

Region one $(0_- < t \le 0_+)$**:** Since input voltage V_i is an ideal step function, input voltage V_i or gate-source voltage V_{gs} of nMOSFET immediately goes from ground to its operating voltage V_{dd} at $t = 0$. Integrating (4.14) with respect to t from 0_- to 0_+ with the initial value of output voltage $V_o(t)$ equal to operating voltage V_{dd}, i.e., $\left(V_o(t)\big|_{t=0_-} = V_{dd}\right)$, yields

$$V_o(t)\big|_{t=0_+} = (1 + k_c)\cdot V_{dd} \tag{4.15}$$

Note that the details of nMOSFET current I_n are not important in the region provided that its value is limited. The output voltage $V_o(t)$ jumps from its initial value V_{dd} to higher value $(1 + k_c)\cdot V_{dd}$ at $t = 0_+$. The overshoot of output voltage $V_o(t)$ is caused by the input-output coupling coefficient k_c given by (4.4). Obviously, there is no such overshoot if there is no input-output capacitance coupling.

Region two $(0_+ < t \le t_{snd})$: In this region, input voltage V_i is equal to V_{dd} and output voltage V_o is larger than nMOSFET saturation voltage V_{snd} with gate-source voltage equal to V_{dd}. The pMOSFET is in cutoff mode and the nMOSFET is in saturation mode with characteristics curve of $V_{gs} = V_i = V_{dd}$. Time t_{snd} occurs when the nMOSFET drain-source voltage V_o reaches its saturation voltage V_{snd}. From (4.10), the saturation voltage V_{snd} with the gate-source voltage equal to V_{dd} is expressed as

$$V_{snd} = V_{cn} \cdot \left(\sqrt{1 + 2 \cdot (V_{dd} - V_{tn})/V_{cn}} - 1\right) \tag{4.16}$$

Substituting current expression of (4.9) in saturation mode with $V_{sn} = V_{snd}$ into (4.14) yields

$$\frac{dV_o}{dt} = -\frac{k_n \cdot V_{snd}^2}{2C_t} \cdot (1 + \lambda_n V_o) \tag{4.17}$$

Assume $\lambda_n > 0$. Solving the above equation of (4.17) with initial value of $V_o(t)|_{t=0_+}$ given by (4.15) at $t = 0_+$ results in the following output expression.

$$V_o(t) = \left[\frac{1}{\lambda_n} + (1 + k_c) \cdot V_{dd}\right] e^{-t/\tau_{sat}} - \frac{1}{\lambda_n} \tag{4.18}$$

where the format of the solution is similar as the analytical solution of first-order RC circuit. The time constant τ_{sat} in saturation mode for a step input is

$$\tau_{sat} = \frac{2 \cdot C_t}{k_n \cdot \lambda_n \cdot V_{snd}^2} \tag{4.19}$$

We can rewrite (4.18) as the dependence of time t on output voltage V_o as

$$t = \tau_{sat} \cdot \ln\left[\frac{1 + (1 + k_c) \cdot \lambda_n \cdot V_{dd}}{1 + \lambda_n \cdot V_o}\right] \tag{4.20}$$

If the drain-source saturation voltage V_{snd} of a deep submicron nMOSFET given by (4.16) is less than half of the operating voltage V_{dd} $(V_{snd} < 0.5V_{dd})$, the propagation delay t_{dHL} of output voltage V_o from high to low for a step input is obtained from (4.20).

$$t_{dHL} = \tau_{sat} \cdot \ln\left[\frac{1+(1+k_c)\cdot\lambda_n \cdot V_{dd}}{1+0.5\cdot\lambda_n \cdot V_{dd}}\right] \qquad (V_{snd} < 0.5V_{dd}) \quad (4.21)$$

The time t_{80}, corresponding to $V_o = 80\%V_{dd}$, can be easily obtained

$$t_{80} = \tau_{sat} \cdot \ln\left[\frac{1+(1+k_c)\cdot\lambda_n \cdot V_{dd}}{1+0.8\cdot\lambda_n \cdot V_{dd}}\right] \qquad (V_{snd} < 0.8V_{dd}) \quad (4.22)$$

The time t_{snd} occurs when the output voltage $V_o(t)$ falls to the drain-source saturation voltage V_{snd} given in (4.16), which is

$$t_{snd} = \tau_{sat} \cdot \ln\left[\frac{1+(1+k_c)\cdot\lambda_n \cdot V_{dd}}{1+\lambda_n \cdot V_{snd}}\right] \quad (4.23)$$

Region three $(t \geq t_{snd})$: In this region, input voltage V_i is equal to V_{dd} and output voltage V_o is smaller than nMOSFET saturation voltage V_{snd} with the gate-source voltage equal to V_{dd} $(V_o \leq V_{snd})$. The pMOSFET is in cutoff mode and the nMOSFET is in linear mode with characteristics curve of $V_{gs} = V_i = V_{dd}$. The differential equation of (4.14) becomes

$$\frac{dV_o}{dt} = -\frac{k_n}{C_t}V_o\left(V_{snd} - \frac{V_o}{2}\right)(1+\lambda_n V_{snd}) \quad (4.24)$$

which can be rearranged as

$$\frac{dV_o}{V_o\cdot(V_{snd} - 0.5\cdot V_o)} = -\frac{k_n(1+\lambda_n V_{snd})}{C_t}dt \quad (4.25)$$

Factoring the left hand side of the above equation yields

$$\left(\frac{1}{V_o}+\frac{0.5}{V_{snd}-0.5\cdot V_o}\right)dV_o=-\frac{dt}{\tau_{lin}} \tag{4.26}$$

where τ_{lin} is the time constant of the MOSFET in the linear mode for a step input

$$\tau_{lin}=\frac{C_t}{k_n\cdot V_{snd}\left(1+\lambda_n V_{snd}\right)} \tag{4.27}$$

Integrating both sides of (4.26) with respect to variables t and V_o from $t=t_{snd}$ and the corresponding initial output voltage of $V_o=V_{snd}$ yields

$$t=t_{snd}+\tau_{lin}\cdot\ln\left(\frac{2V_{snd}}{V_o}-1\right) \qquad \left(V_o\le V_{snd}\right) \tag{4.28}$$

Or

$$V_o=2V_{snd}\cdot\frac{1}{1+e^{(t-t_{snd})/\tau_{lin}}} \qquad \left(t\ge t_{snd}\right) \tag{4.29}$$

If the drain-source saturation voltage V_{snd} of a deep submicron nMOSFET given by (4.16) is larger than half of the operating voltage V_{dd} $\left(V_{snd}>0.5V_{dd}\right)$, then the propagation delay t_{dHL} of output voltage V_o from high to low is obtained from (4.28).

$$t_{dHL}=t_{snd}+\tau_{lin}\cdot\ln\left(\frac{4\cdot V_{snd}}{V_{dd}}-1\right) \qquad \left(V_{snd}>0.5V_{dd}\right) \tag{4.30}$$

The time t_{20}, corresponding to $V_o=20\%V_{dd}$, can be easily obtained

$$t_{20}=t_{snd}+\tau_{lin}\cdot\ln\left(\frac{10\cdot V_{snd}}{V_{dd}}-1\right) \qquad \left(0.2V_{dd}<V_{snd}\right) \tag{4.31}$$

Therefore the output falling time t_f is calculated from the time of $V_o=80\%\cdot V_{dd}$ to the time of $V_o=20\%\cdot V_{dd}$ divided by 0.6, which is

$$t_f = \frac{1}{0.6}(t_{20} - t_{80}) = \frac{1}{0.6}[(t_{20} - t_{sat}) + (t_{sat} - t_{80})] \tag{4.32}$$

where $(t_{20} - t_{snd})$ and $(t_{snd} - t_{80})$ are obtained by (4.31), (4.22), and (4.23), respectively

$$t_{20} - t_{snd} = \tau_{lin} \cdot \ln\left(\frac{10 \cdot V_{snd}}{V_{dd}} - 1\right) \tag{4.33}$$

and

$$t_{snd} - t_{80} = \tau_{sat} \ln\left(\frac{1 + 0.8 \cdot \lambda_n \cdot V_{dd}}{1 + \lambda_n \cdot V_{snd}}\right) \tag{4.34}$$

Therefore the falling time t_f of (4.32) becomes

$$t_f = \frac{1}{0.6}\left[\tau_{lin} \cdot \ln\left(\frac{10 \cdot V_{snd}}{V_{dd}} - 1\right) + \tau_{sat} \ln\left(\frac{1 + 0.8 \cdot \lambda_n \cdot V_{dd}}{1 + \lambda_n \cdot V_{snd}}\right)\right] \tag{4.35}$$

4.3.3 Results and Discussions

In the previous subsection, we have derived the analytical expressions of output voltage waveforms and propagation delays of a CMOS inverter for a step input. In this subsection, we take a specific CMOS inverter for a step input as an example. The output voltage waveforms and propagation delays of a CMOS inverter as function of external load capacitance for a step input are directly obtained from those closed-form expressions derived in the previous subsection. The manufacture technology is 0.24μm. The pMOSFET width of the CMOS inverter is 43.52μm and the nMOSFET width is 21.76μm. The static *I–V* characteristics of the pMOSFET and the nMOSFET of the CMOS inverter are characterized with SPICE. The PDSMM parameters of the pMOSFET and nMOSFET of the CMOS inverter are obtained by fitting the static *I–V* characteristics from SPICE BSIM3. They are listed in Table 4.1. The static *I–V* characteristics of the

nMOSFET from PDSMM with the parameters listed in Table 4.1 are plotted in Fig. 3.6. As a comparison, the static I–V characteristics of the nMOSFET from SPICE BSIM3 model are also shown in Fig. 3.6. Therefore, the static I–V characteristics from PDSMM are well matched with those from SPICE BSIM3 model.

Fig. 4.4 shows output voltage waveforms for a rising step input derived from Subsection 4.3.2 based on PDSMM with solid lines. The external load capacitances of the CMOS inverter are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively. As a comparison, the output voltage waveforms from SPICE BSIM3 are also plotted in dashed lines. The expressions of output voltage waveforms derived from Subsection 4.3.2 are very accurate to reproduce the output transition with a step input.

Fig. 4.5 shows the propagation delay of a CMOS inverter versus external load capacitance for a rising step input. The propagation delay derived in Subsection 4.3.2 based on PDSMM is indicated as the solid line. The dashed line represents the propagation delay obtained from SPICE BSIM3. Obviously, the propagation delay of a CMOS inverter for a step input is linear with the external load capacitance. The propagation delay of a CMOS inverter for a step input is accurately predicted in Subsection 4.3.2 based on PDSMM. In order to see the deviation of propagation delay with input transition time from a step input, Fig. 4.5 also show the propagation delay of a CMOS inverter versus external capacitance for a ramp input of 50ps rising time. Clearly, for the same external load capacitance the propagation delay for a ramp input of 50ps rising time is deviated from the one for a step input. Generally speaking, it is necessary to incorporate input transition time into propagation delay models of a CMOS inverter, which will be investigated in the next two sections.

4.4 TIMING ANALYSIS OF A CMOS INVERTER MODELED WITH A DOMINATING MOSFET FOR A RAMP INPUT

The propagation delay time is the time difference between 50% input transition and 50% output transition. In the previous section, we have derived the timing model of a deep submicron CMOS inverter for a step input where the effect of input transition time is ignored. In fact, there is always a non-zero transition time for actual input voltage. The effect of input transition time on propagation delay of a CMOS inverter has to be incorporated in timing model. In this section, we focus on the output response and timing analysis for a ramp input. Only the dominating nMOSFET without trivial pMOSFET is approximated to model the CMOS inverter. The trivial pMOSFET for a rising ramp input is ignored in the analysis $(I_p = 0)$. The ramp response of the nMOSFET-only circuit representing a CMOS inverter is developed with input transition time, external load capacitance, and PDSMM model parameters. There are two phenomena of output voltages due to ignorance of the trivial pMOSFET. The first observable fact is the voltage overshoot exaggerates when input rises because of no inverse discharging current to power rail, especially for small load capacitance. The other is the nMOSFET discharging current without the compensated short-circuit current makes the output voltage drop abnormally fast, especially for large input transition time. The method of propagation delay developed in this section is accurate in much broader range of input transition time and external load capacitance than the one derived in the previous section. It is accurate not only for a very fast input but for a ramp input with a comparable external load capacitance also.

4.4.1 Operation of a Dominating nMOSFET for a Ramp Input

In this subsection, we investigate the operation of a deep submicron nMOSFET for a ramp input. The input voltage is a ramp function, i.e.,

$$V_i(t) = \frac{V_{dd}}{\tau}[r(t) - r(t-\tau)] \text{ or } V_i(s) = V_{dd}[r(s) - r(s-1)] \tag{4.36}$$

where $r(\cdot)$ is a unit ramp function, τ is the rising time of the input voltage, and s is the normalized time. Fig 4.6 plots typical discharging trajectory traces during output transition of the nMOSFET of a CMOS inverter for ramp inputs. Fig. 4.7 shows the five operation regions denoted with R1, R2, R3, R4, and R5, where discharging trajectory traces of the nMOSFET of a CMOS inverter may experience. The five operation regions are described as follows: R1 is region one where nMOSFET is in cutoff mode; R2 is region two where nMOSFET is in saturation mode; R3 is region three where nMOSFET is in linear mode; R4 is region four where nMOSFET is in saturation mode with input voltage equal to operating voltage V_{dd}; and R5 is region five where nMOSFET is in linear mode with input voltage equal to operating voltage V_{dd}. The interface between R2 and R3 is expressed by (4.10). We will discuss the output voltage expressions according to these five operation regions.

4.4.2 Output Voltages of a Dominating nMOSFET for a Ramp Input

Region one ($0 \le s \le u_m$ and $u_o \ge 1$): The nMOSFET is in cutoff mode $(I_n = 0)$. The input-output coupling current dominates the output transition. The differential equation to describe output voltage V_o is obtained by substituting $(I_n = 0)$ into (4.5)

$$\frac{dV_o}{ds} = k_c V_{dd} \tag{4.37}$$

The initial value of output voltage V_o is V_{dd}. The solution of output voltage $V_o(s)$ is

$$V_o(s) = V_{dd}(1 + k_c s) \tag{4.38}$$

The output voltage $V_o(s)$ linearly rises with the normalized time s or the time t in this region due to input-output capacitance coupling. This is the phenomenon of an overshoot of output voltage V_o, caused by the input-output coupling current. Note that if the trivial pMOSFET is considered, the inverse discharging current through the pMOSFET to power rail leads to the lower voltage overshoot than that given by (4.38). In other words the voltage overshoot expected by (4.38) exaggerates.

Region two ($u_{tn} < s \leq 1$ and $u_o \geq u_{sn}$): In this region, the nMOSFET is in saturation mode. The differential equation to describe the output transition and the current expression I_n of the nMOSFET are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} - \dfrac{\tau}{C_t} I_n \\ I_n = \dfrac{1}{2} k_n V_{sn}^2 (1 + \lambda_n V_o) \\ V_{sn} = V_{cn}\left(\sqrt{1 + 2(s - u_{tn})/u_{cn}} - 1\right) \end{cases} \tag{4.39}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{2i} and $s_{2i}(= u_{tn})$. The current I_n through the nMOSFET dominates the discharging process. As the first step, only the current I_n of the nMOSFET is used in the analysis and the differential equation of (4.39) is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{2C_t} V_{sn}^2 (1 + \lambda_n V_o) \tag{4.40}$$

For simplicity, set the variable V_{dn}

$$V_{dn} = \frac{\tau k_n}{C_t} V_{cn}^2 \tag{4.41}$$

Substituting the drain-source saturation voltage V_{sn} given in (4.39) into (4.40) and using the variable V_{dn} yield

$$\frac{dV_o}{ds} = -\frac{V_{dn}}{2}\left(\sqrt{1+2\frac{s-u_{tn}}{u_{cn}}}-1\right)^2 (1+\lambda_n V_o) \tag{4.42}$$

Assume $\lambda_n > 0$. The solution of the above differential equation is expressed as follows:

$$V_{o2} = \left(\frac{1}{\lambda_n} + V_{2i}\right)\exp[-\lambda_n \cdot f_2(s)] - \frac{1}{\lambda_n} \tag{4.43}$$

where the function $f_2(s)$ is given by

$$f_2(s) = V_{dn}\left\{s - u_{tn} + \frac{(s-u_{tn})^2}{2u_{cn}} - \frac{u_{cn}}{3}\left[\left(1+2\frac{s-u_{tn}}{u_{cn}}\right)^{3/2} - 1\right]\right\} \tag{4.44}$$

If the effect of channel length modulation is ignored, i.e. $(\lambda_n = 0)$, the solution of the differential equation of (4.42) is

$$V_{o2} = V_{2i} - f_2(s) \tag{4.45}$$

The second step is that both the current I_n of the nMOSFET and coupling capacitance current between input node and output node are included in the analysis. Then, the output voltage $V_o(s)$ is

$$V_o(s) = k_c V_{dd}(s - u_{tn}) + V_{o2} \tag{4.46}$$

Region three ($u_{tn} < s \le 1$ and $u_o < u_{sn}$): In this region, the nMOSFET is in linear mode. The differential equation to describe the output transition and the drain-source current I_n of the nMOSFET are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} - \dfrac{\tau}{C_t} I_n \\ I_n = k_n \left(V_{sn} - \dfrac{V_o}{2} \right) V_o \left(1 + \lambda_n V_{sn} \right) \approx k_n \left(V_{sn} - \dfrac{V_o}{2} \right) V_o \\ V_{sn} = V_{cn} \left(\sqrt{1 + 2(s - u_{tn})/u_{cn}} - 1 \right) \approx k_l (s - b_l) \end{cases} \tag{4.47}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{3i} and s_{3i}. In order to derive the closed-form expression of output voltage $V_o(t)$ without the loss of accuracy, the term of $(1 + \lambda_n V_{sn})$ in the current I_n is ignored since the effect of $\lambda_n V_{sn}$ is small in this region. The drain-source saturation voltage V_{sn} of the nMOSFET is approximated by the linear equation with its parameters k_l and b_l as follows:

$$\begin{cases} k_l = V_{cn} \left(\sqrt{1 + 2(1 - u_{tn})/u_{cn}} - 1 \right) / (1 - u_{tn}) \\ b_l = u_{tn} \end{cases} \tag{4.48}$$

The nMOSFET drain-source current I_n dominates the discharging process. We find that only the nMOSFET current I_n is accurate enough for the timing analysis. The differential equation of (4.47) becomes

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{C_t} \left(V_{sn} - \frac{V_o}{2} \right) V_o \tag{4.49}$$

This is a Bernoulli's equation that can be changed into first-order linear differential equation by setting

$$y = \frac{1}{V_o} \tag{4.50}$$

Then the Bernoulli's equation of (4.49) becomes

$$\frac{dy}{ds} = \frac{\tau k_n}{C_t} \left(V_{sn} y - \frac{1}{2} \right) \tag{4.51}$$

The solution of the above first-order linear differential equation is

$$y = \exp\left(\int_{s_{3i}}^{s} \frac{\tau k_n}{C_t} V_{sn} ds\right)\left\{\int_{s_{3i}}^{s} -\frac{\tau k_n}{2C_t} \exp\left(-\int_{s_{3i}}^{s} \frac{\tau k_n}{C_t} V_{sn} ds\right) ds + y_{3i}\right\} \tag{4.52}$$

where y_{3i} is the reciprocal of the initial output voltage V_{3i}, i.e., $y_{3i} = \frac{1}{V_{3i}}$. For simplicity, set the following variables

$$\sigma_l^2 = \frac{C_t}{\tau k_n k_l} \tag{4.53}$$

Substituting the simplified linear equation V_{sn} of (4.47) and the variable σ_l into (4.52) yields

$$y = \exp\left[\frac{(s-b_l)^2 - (s_{3i}-b_l)^2}{2\sigma_l^2}\right] \cdot \left\{\int_{s_{3i}}^{s} -\frac{\tau k_n}{2C_t} \exp\left[-\frac{(s-b_l)^2 - (s_{3i}-b_l)^2}{2\sigma_l^2}\right] ds + y_{3i}\right\} \tag{4.54}$$

which can be simplified to

$$y = \exp\left[\frac{(s-b_l)^2}{2\sigma_l^2}\right] \cdot \left\{\int_{s_{3i}}^{s} -\frac{\tau k_n}{2C_t} \exp\left[-\frac{(s-b_l)^2}{2\sigma_l^2}\right] ds + y_{3i} \exp\left[-\frac{(s_{3i}-b_l)^2}{2\sigma_l^2}\right]\right\} \tag{4.55}$$

Using the expression of (III.15a) in Appendix III gives rise to

$$y = \frac{\tau k_n}{2C_t} P(s, b_l, \sigma_l) + \left(y_{3i} - \frac{\tau k_n}{2C_t} P(s_{3i}, b_l, \sigma_l)\right) \exp\left[\frac{(s + s_{3i} - 2b_l)(s - s_{3i})}{2\sigma_l^2}\right] \tag{4.56}$$

Therefore, the output voltage in this region is

$$V_o(s) = \frac{1}{\frac{\tau k_n}{2C_t} P(s, b_l, \sigma_l) + \left(\frac{1}{V_{3i}} - \frac{\tau k_n}{2C_t} P(s_{3i}, b_l, \sigma_l)\right) \exp\left[\frac{(s + s_{3i} - 2b_l)(s - s_{3i})}{2\sigma_l^2}\right]} \tag{4.57}$$

Region four ($s > 1$ and $V_o \geq V_{snd}$): In this region, input voltage V_i is equal to operation voltage V_{dd}. The output voltage V_o is greater than the drain-source saturation voltage

V_{snd} of the nMOSFET with input voltage $V_i = V_{dd}$. The nMOSFET is in saturation mode with input voltage $V_i = V_{dd}$. Then the differential equation to describe the output transition and the current expression I_n for the nMOSFET are as follows:

$$\begin{cases} \dfrac{dV_o}{dt} = -\dfrac{1}{C_t} I_n \\ I_n = \dfrac{1}{2} k_n V_{snd}^2 (1 + \lambda_n V_o) \\ V_{snd} = V_{cn}\left(\sqrt{1 + 2(1 - u_{tn})/u_{cn}} - 1\right) \end{cases} \tag{4.58}$$

The initial values of output voltage V_o and time t in this region are V_{4i} and t_{4i}. Then the output waveform $V_o(t)$ is

$$V_o(t) = \left(\frac{1}{\lambda_n} + V_{4i}\right) \exp\left(-\frac{t - t_{4i}}{\tau_{sat}}\right) - \frac{1}{\lambda_n} \tag{4.59}$$

where the format of the solution is similar as the analytical expression of first-order RC circuit. The time constant τ_{sat} in this region is the same as one given in (4.19), i.e.,

$$\tau_{sat} = \frac{2C_t}{k_n V_{snd}^2 \lambda_n} \tag{4.60}$$

Equation (4.59) can be expressed as the dependence of time t on output voltage V_o.

$$t = t_{4i} + \tau_{sat} \ln\left(\frac{1 + \lambda_n V_{4i}}{1 + \lambda_n V_o}\right) \tag{4.61}$$

Region five ($s > 1$ and $V_o < V_{snd}$): In this region, input voltage V_i is equal to operation voltage V_{dd} and output voltage V_o is less than the drain-source saturation voltage V_{snd} of the nMOSFET with input voltage $V_i = V_{dd}$. The nMOSFET is in linear mode with input voltage $V_i = V_{dd}$. Then the differential equation to describe the output transition and the

current expression I_n for the nMOSFET are as follows:

$$\begin{cases} \dfrac{dV_o}{dt} = -\dfrac{1}{C_t} I_n \\ I_n = k_n \left(V_{snd} - \dfrac{V_o}{2} \right) V_o \left(1 + \lambda_n V_{snd} \right) \\ V_{snd} = V_{cn} \left(\sqrt{1 + 2(1 - u_{tn})/u_{cn}} - 1 \right) \end{cases} \tag{4.62}$$

For generality, the initial values of output voltage $V_o(t)$ and time t in this region are V_{5i} and t_{5i}. Then the output waveform $V_o(t)$ is

$$V_o(t) = \frac{2V_{snd}}{1 + (2V_{snd}/V_{5i} - 1)\exp[(t - t_{5i})/\tau_{lin}]} \tag{4.63}$$

where the time constant τ_{lin} in this region is the same as the one given in (4.27), i.e.,

$$\tau_{lin} = \frac{C_t}{k_n V_{snd} (1 + \lambda_n V_{snd})} \tag{4.64}$$

Equation (4.63) can be expressed as the dependence of time t on output voltage V_o.

$$t = t_{5i} + \tau_{lin} \ln \left(\frac{V_{5i}}{V_o} \frac{2V_{snd} - V_o}{2V_{snd} - V_{5i}} \right) \tag{4.65}$$

4.4.3 Results and Discussions

In the previous subsection, the CMOS inverter for a rising ramp input is modeled with a dominating nMOSFET. The trivial pMOSFET for a rising ramp input is ignored. Describing the nMOSFET with PDSMM, we have derived the analytical expressions of output voltages of a CMOS inverter for a rising ramp input. In this subsection, we apply these closed-form expressions to a specific CMOS inverter, which is the same as the one in Section 4.3.3. The PDSMM parameters of the nMOSFET are listed in Table 4.1. The resulting output voltage and propagation delay are discussed. The consequences of the

ignorance of the trivial pMOSFET are explored.

Fig. 4.8 shows the output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from the voltage expressions based on PDSMM in Subsection 4.4.2 (solid lines). Only a dominating nMOSFET is used to denote the CMOS inverter. The input transition time is 1ps. The external load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively. The calculated output waveforms from PDSMM and the output responses from SPICE BSIM3 agree with each other excellently.

Figs. 4.9-4.12 indicate the output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from PDSMM in Subsection 4.4.2 (solid lines). The nMOSFET-only is used to denote the CMOS inverter. The input transition times are 50ps, 100ps, 200ps, and 400ps, respectively. In each plot, the external load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF accordingly. Comparing with the results from SPICE BSIM3, we obviously see the two main features of output voltages in the nMOSFET-only circuitry: exaggerated output voltage overshoot, especially for small load capacitance and abnormally fast voltage drop for large input transition time and small load capacitance. The former is caused with no inverse discharging current to power rail through the trivial pMOSFET. This effect is significant for small load capacitance since input-to-output coupling coefficient is large. In practice the input-to-output coupling current is partially compensated by the inverse discharging current of the trivial pMOSFET. Thus the voltage overshoot from SPICE BSIM3 is much lower than the one obtained from the nMOSFET-only circuit as shown in Figs. 4.9-4.12. The side effect of significant voltage overshoot for fast input and small load makes the calculated waveforms obviously slower than the corresponding ones from SPICE BSIM3 as shown

in Figs. 4.9-4.10. The latter is due to the lack of short-circuit current. When input transition time increases, the effect of short-circuit current in a real CMOS inverter increase also. Without it due to the lack of the trivial pMOSFET, the uncompensated nMOSFET current discharge the load capacitance faster than a real CMOS circuit. Eventually it causes a significant error as shown in Fig. 4.12, which makes the propagation delay from the methodology unreliable for small load capacitance and large input transition time.

Fig. 4.13 shows propagation delays versus external load capacitances for input transition times of 1ps, 10ps, 100ps, 200ps, and 400ps (bottom to top), respectively. Fig. 4.14 plots propagation delays versus input transition times with external load capacitances of 10fF, 100fF, 200fF, 400fF, and 800fF (bottom to top), respectively. In both figures, the propagation delays of a CMOS inverter obtained from SPICE BSIM3 are indicated with dashed lines and the propagation delays from the nMOSFET-only model based on PDSMM are represented with circles. For small input transition time or large load capacitance, the method proposed in Subsection 4.4.2 is very accurate. However, large input transition time and small load capacitance cause significant errors, which make the expressions in Subsection 4.4.2 unreliable.

Fig. 4.15 shows the correlation of propagation delays from the nMOSFET-only model based on PDSMM with propagation delays from SPICE BSIM3. The two dashed lines are ±5% of propagation delays from SPICE BSIM3. Most of them are in good agreement. However, the significant errors are found with large input transition time and small load capacitance due to no short-circuit current. We will develop a complete timing model with both the dominating nMOSFET and the trivial pMOSFET in the next section.

4.5 COMPREHENSIVE PROPAGATION DELAY ANALYSIS OF A CMOS INVERTER

In Section 4.3, we have derived the timing model of a CMOS inverter for a step input where the effect of input transition time is ignored. The transient properties of a CMOS inverter for a ramp input with the dominating nMOSFET-only model are analyzed in Section 4.4. In this section, we analyze comprehensive propagation delay of a deep submicron CMOS inverter with a ramp input where both the dominating nMOSFET and the trivial pMOSFET are included. We first examine transition traces of both pMOSFET and nMOSFET shown in Fig. 4.16. Eight possible operation regions are obtained according to distinct operation modes of pMOSFET and nMOSFET for each region. However, the analytical solutions of output voltages, especially for region two, are not accurate due to the simplifications of property equations. Thus we extend the eight regions into eleven regions shown in Fig. 4.17. The differential equation controlling the transition of a CMOS inverter is analytically solved for each operation region.

4.5.1 Operation of a CMOS Inverter for a Ramp Input

The typical transition trajectory traces of both pMOSFET and nMOSFET of a CMOS inverter for ramp inputs are shown in Fig. 4.2. The eight operation regions of a CMOS inverter for ramp inputs are plotted in Fig. 4.16. The input ramp and the transient output voltages for different external load capacitances are also plotted in Fig. 4.16. The eight operation regions cover the complete combinations of different operation modes of nMOSFET and pMOSFET for a ramp input such as linear, saturation, and cutoff. However, the analytical solutions of output voltages, especially for region two, are not accurate due to the simplifications of property equations. Thus we extend the eight

regions into eleven regions shown in Fig. 4.17. The eleven operation regions are detailed as: R1 where pMOSFET is in linear mode and nMOSFET is in cutoff mode; R2 where pMOSFET is in linear mode and nMOSFET is in saturation mode; R3 where both pMOSFET and nMOSFET are in saturation mode; R4 where pMOSFET is in saturation mode and nMOSFET is in linear mode; R5 where pMOSFET is in linear mode and nMOSFET is in saturation mode; R6 where both pMOSFET and nMOSFET are in saturation mode; R7 where pMOSFET is in saturation mode and nMOSFET is in linear mode; R8 where pMOSFET is in cutoff mode and nMOSFET is in saturation mode; R9 where pMOSFET is in cutoff mode and nMOSFET is in linear mode; R10 where pMOSFET is in cutoff mode and nMOSFET is in saturation mode with input voltage equal to operating voltage V_{dd}; and R11 where pMOSFET is in cutoff mode and nMOSFET is in linear mode with input voltage equal to operating voltage V_{dd}. The line between R2 and R3 or R5 and R6 is described by (4.12). The interface between R3 and R4 or R6 and R7 or R8 and R9 is expressed by (4.10). The solution of output voltage V_o is analyzed according to the eleven respective operation regions.

4.5.2 Output Voltages of a CMOS Inverter for a Ramp Input

Region one ($0 \le s \le u_m$ and $u_o \ge 1$): In this region, the nMOSFET is in cutoff mode $(I_n = 0)$ and the pMOSFET is in linear mode. The input-output coupling current I_{C_m}, which may cause an overshoot of output voltage V_o, has the main impact on output voltage V_o in this region. The differential equation to describe the output transition and the pMOSFET current expression are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t} I_p \\ I_p = k_p V_{sp} V_{op} \\ V_{sp} = V_{ip} - V_{tp} = V_{dd}\left(1 - u_{tp} - s\right) \end{cases} \tag{4.66}$$

The initial value of output voltage V_o at initial time of zero in this region is V_{dd} $\left(V_{1i} = V_o(s)\big|_{s=0} = V_{dd}\right)$. Substituting the nMOSFET current expression I_p of (4.66) into the differential equation of (4.66) yields

$$\frac{dV_{op}}{ds} = -\frac{\tau k_p}{C_t} V_{sp} V_{op} - k_c V_{dd} \tag{4.67}$$

where the variable V_o of output voltage to be solved in (4.66) is changed to the variable $V_{op}\left(= V_{dd} - V_o\right)$ with its initial value of zero $\left(V_{op}\big|_{s=0} = 0\right)$. The solution of the first-order linear differential equation of (4.67) is analytically expressed as

$$V_{op} = \exp\left(\int_0^s -\frac{\tau k_p}{C_t} V_{sp}\, ds\right) \cdot \left[\int_0^s (-k_c V_{dd}) \exp\left(\int_0^s \frac{\tau k_p}{C_t} V_{sp}\, ds\right) ds\right] \tag{4.68}$$

For simplicity, set the following variables:

$$\begin{cases} m_p = 1 - u_{tp} \\ \sigma_p^2 = \dfrac{C_t}{\tau k_p V_{dd}} \end{cases} \tag{4.69}$$

Substituting the pMOSFET saturation voltage expression V_{sp} of (4.66) into (4.68) and using the variables m_p and σ_p^2 of (4.69) yield

$$V_{op} = -k_c V_{dd} \exp\left[\frac{(s - m_p)^2}{2\sigma_p^2}\right] \cdot \int_0^s \left\{\exp\left[-\frac{(s - m_p)^2}{2\sigma_p^2}\right]\right\} ds \tag{4.70}$$

By using the expression of (III.15b) in Appendix III with the condition of $0 \le s \le u_{tn} < m_p$, the above equation becomes

$$V_{op} = -k_c V_{dd} \exp\left[\frac{(s-m_p)^2}{2\sigma_p^2}\right] \left\{R(s,m_p,\sigma_p)\exp\left[-\frac{(s-m_p)^2}{2\sigma_p^2}\right] - R(0,m_p,\sigma_p)\exp\left[-\frac{m_p^2}{2\sigma_p^2}\right]\right\} \tag{4.71}$$

where the function $R(t,m,\sigma)$ is defined in (III.9b) of Appendix III. Therefore, the output voltage $V_o(s)$ in this region is obtained:

$$V_o(s) = V_{dd} + k_c V_{dd}\left\{R(s,m_p,\sigma_p) - R(0,m_p,\sigma_p)\exp\left(-\frac{s(2m_p-s)}{2\sigma_p^2}\right)\right\} \tag{4.72}$$

Since $R(s,m_p,\sigma_p) \ge R(0,m_p,\sigma_p)$ with $0 \le s \le u_{tn} < m_p$, the output voltage $V_o(s)$ clearly rises with the normalized time s or the time t in this region. That is the phenomena of an overshoot of output voltage $V_o(s)$, caused by the input-output coupling current I_{C_m}.

Region two ($u_{tn} < s \le 0.5$ and $(1-u_o) \le u_{sp}/2$): In this region, the nMOSFET is in saturation mode and the pMOSFET is in linear mode. The differential equation to describe the output transition and the current expressions for the pMOSFET and the nMOSFET are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t}\left(I_p - I_n\right) \\ I_p = k_p V_{sp} V_{op} \\ I_n = \dfrac{1}{2} k_n V_{sn}^2 \left(1 + \lambda_n V_o\right) \approx \dfrac{1}{2} k_n V_{sn}^2 \left(1 + \lambda_n V_{dd}\right) \\ V_{sp} = V_{ip} - V_{tp} = V_{dd}\left(1 - u_{tp} - s\right) \\ V_{sn} = V_{cn}\left(\sqrt{1 + 2\dfrac{s - u_{tn}}{u_{cn}}} - 1\right) \approx k_l \left(s - b_l\right) \end{cases} \tag{4.73}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{2i} and s_{2i}. In order to simplify the derivation without losing accuracy, the term of $\left(1 + \lambda_n V_o\right)$ in the current I_n is reduced to $\left(1 + \lambda_n V_{dd}\right)$ since output voltage V_o is close to V_{dd} in this region. The saturation voltage V_{sn} is approximated by the linear equation with the parameters k_l and b_l as follows:

$$\begin{cases} k_l = V_{cn}\left[\sqrt{1 + 2(0.5 - u_{tn})/u_{cn}} - 1\right] / (0.5 - u_{tn}) \\ b_l = u_{tn} \end{cases} \tag{4.74}$$

Substituting the expressions of currents I_p and I_n of (4.73) into the differential equation of (4.73) yields

$$\frac{dV_{op}}{ds} = -\frac{\tau k_p}{C_t} V_{sp} V_{op} - k_c V_{dd} + \frac{\tau k_n \left(1 + \lambda_n V_{dd}\right)}{2C_t} V_{sn}^2 \tag{4.75}$$

where the variable V_o of output voltage to be solved is changed to variable $V_{op}\left(= V_{dd} - V_o\right)$ with initial value of $\left(V_{op}\big|_{s=s_{2i}} = V_{dd} - V_{2i}\right)$. The solution of the first-order linear differential equation is analytically expressed as

$$V_{op} = \exp\left(\int_{s_{2i}}^{s} -\frac{\tau k_p}{C_t} V_{sp}\, ds\right) \cdot \left\{\int_{s_{2i}}^{s}\left[\left(\frac{\tau k_n (1+\lambda_n V_{dd})}{2C_t} V_{sn}^2 - k_c V_{dd}\right)\exp\left(\int_{s_{2i}}^{s}\frac{\tau k_p}{C_t} V_{sp}\, ds\right)\right] ds + V_{op}\Big|_{s=s_{2i}}\right\} \tag{4.76}$$

In addition to the variables m_p and σ_p^2 defined in (4.69), set the variable $V_{\ln}$ as:

$$V_{\ln} = \frac{\tau k_n k_l^2 (1+\lambda_n V_{dd})}{C_t} \tag{4.77}$$

Substituting the pMOSFET saturation voltage expression V_{sp} of (4.73) into (4.76) and using variables m_p and σ_p^2 given in (4.69) and the variable $V_{\ln}$ yield

$$V_{op} = \exp\left[\frac{(s-m_p)^2}{2\sigma_p^2}\right] \cdot \left\{\int_{s_{2i}}^{s}\left[\frac{V_{\ln}(s-b_l)^2}{2} - k_c V_{dd}\right]\exp\left[-\frac{(s-m_p)^2}{2\sigma_p^2}\right] ds + (V_{dd}-V_{2i})\exp\left[-\frac{(s_{2i}-m_p)^2}{2\sigma_p^2}\right]\right\} \tag{4.78}$$

With the expressions of (III.15b), (III.16b), and (III.17b) in Appendix III, the above equation becomes

$$V_{op} = f_2(s) - [f_2(s_{2i}) + V_{2i} - V_{dd}]\exp\left(-\frac{(s-s_{2i})(2m_p - s_{2i} - s)}{2\sigma_p^2}\right) \tag{4.79}$$

where the function $f_2(x)$ is given by:

$$f_2(x) = -\frac{(x+m_p-2b_l)\sigma_p^2 V_{\ln}}{2} + \left[\frac{((m_p-b_l)^2+\sigma_p^2)V_{\ln}}{2} - k_c V_{dd}\right] R(x, m_p, \sigma_p) \tag{4.80}$$

Therefore, the output voltage V_o is

$$V_o(s) = V_{dd} - f_2(s) + [f_2(s_{2i}) + V_{2i} - V_{dd}]\exp\left(-\frac{(s-s_{2i})(2m_p - s_{2i} - s)}{2\sigma_p^2}\right) \tag{4.81}$$

Region three ($u_{tn} < s \le 0.5$, $(1-u_o) > u_{sp}/2$, and $u_o \ge u_{sn}$): In this region, both the pMOSFET and the nMOSFET are in saturation mode. The differential equation to describe the output transition and the current expressions for the pMOSFET and the nMOSFET are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t}\left(I_p - I_n\right) \\ I_p = \dfrac{1}{2} k_p V_{sp}^2 \\ I_n = \dfrac{1}{2} k_n V_{sn}^2 \left(1 + \lambda_n V_o\right) \\ V_{sp} = V_{ip} - V_{tp} = V_{dd}\left(1 - u_{tp} - s\right) \\ V_{sn} = V_{cn}\left(\sqrt{1 + 2\dfrac{s - u_{tn}}{u_{cn}}} - 1\right) \end{cases} \tag{4.82}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{3i} and s_{3i}. The nMOSFET current I_n dominates the discharging process in this region. As the first step, only the nMOSFET is analyzed and the differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{2C_t} V_{sn}^2 \left(1 + \lambda_n V_o\right) \tag{4.83}$$

Substituting the nMOSFET saturation voltage V_{sn} of (4.82) into the above differential equation and using the variable V_{dn} of (4.41) yield

$$\frac{dV_o}{ds} = -\frac{V_{dn}}{2}\left(\sqrt{1 + 2\frac{s - u_{tn}}{u_{cn}}} - 1\right)^2 \left(1 + \lambda_n V_o\right) \tag{4.84}$$

Assume $\lambda_n > 0$. The solution of the above differential equation is expressed as follows:

$$V_{o3} = \left(\frac{1}{\lambda_n} + V_{3i}\right)\exp[-\lambda_n f_3(s)] - \frac{1}{\lambda_n} \tag{4.85}$$

In (4.85), the function $f_3(s)$ is given by

$$f_3(s) = V_{dn} \cdot \left\{ s - s_{3i} + \frac{(s-u_{tn})^2 - (s_{3i}-u_{tn})^2}{2u_{cn}} - \frac{u_{cn}}{3}\left[\left(1+2\frac{s-u_{tn}}{u_{cn}}\right)^{\frac{3}{2}} - \left(1+2\frac{s_{3i}-u_{tn}}{u_{cn}}\right)^{\frac{3}{2}}\right]\right\} \tag{4.86}$$

The second step is that nMOSFET, pMOSFET, and coupling capacitance current between input node and output node are included in the analysis. For simplicity, set the variable V_{qp} as:

$$V_{qp} = \frac{\tau k_p V_{dd}^2}{C_t} \tag{4.87}$$

Then, the output voltage $V_o(t)$ in this region is

$$V_o(s) = k_c V_{dd}(s - s_{3i}) + \frac{V_{qp}}{6}\left[(1-u_{tp}-s_{3i})^3 - (1-u_{tp}-s)^3\right] + V_{o3} \tag{4.88}$$

Region four ($u_{tn} < s \le 0.5$ and $u_o < u_{sn}$): In this region, the nMOSFET is in linear mode and the pMOSFET is in saturation mode. The differential equation to describe the output transition and the current expressions for the pMOSFET and the nMOSFET are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t}\left(I_p - I_n\right) \\ I_p = \dfrac{1}{2}k_p V_{sp}^2 \\ I_n = k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o\left(1+\lambda_n V_{sn}\right) \approx k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o \\ V_{sp} = V_{ip} - V_{tp} = V_{dd}\left(1 - u_{tp} - s\right) \\ V_{sn} = V_{cn}\left(\sqrt{1+2\dfrac{s-u_{tn}}{u_{cn}}} - 1\right) \approx k_l\left(s - b_l\right) \end{cases} \tag{4.89}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{4i} and s_{4i}. In order to simplify the derivation without losing accuracy, the term of $\left(1+\lambda_n V_{sn}\right)$ in the current I_n is ignored since the effect of $\lambda_n V_{sn}$ is small in this region. The saturation voltage V_{sn} of the nMOSFET is approximated by the same linear equation as in Region two with the parameters k_l and b_l given in (4.74). The nMOSFET current dominates the discharging process in this region. As the first step, only nMOSFET is considered in the derivation. The differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{C_t}\left(V_{sn} - \frac{V_o}{2}\right)V_o \tag{4.90}$$

This differential equation of (4.90), a Bernoulli's equation, can be obtained by employing the same procedures in Region three of Section 4.4. For simplicity, set the following variable

$$\sigma_l^2 = \frac{C_t}{\tau k_n k_l} \tag{4.91}$$

Note that parameter k_l in the above variable σ_l^2 is given in (4.74). Therefore, the output

voltage without input-output coupling current and the pMOSFET is

$$V_{o4}=\frac{1}{\frac{\tau k_n}{2C_t}P(s,b_l,\sigma_l)+\left(-\frac{\tau k_n}{2C_t}P(s_{4i},b_l,\sigma_l)+\frac{1}{V_{4i}}\right)\exp\left[\frac{(s+s_{4i}-2b_l)(s-s_{4i})}{2\sigma_l^2}\right]} \tag{4.92}$$

where $P(x,m,\sigma)$ is defined in (III.9a) of Appendix III and the parameter b_l is given in (4.74). The second step is that nMOSFET, pMOSFET, and coupling capacitance current between input and output are involved in the analysis. Then, the solution is expressed as follows:

$$V_o(s)=k_c V_{dd}(s-s_{4i})+\frac{V_{qp}}{6}\left[(1-u_{tp}-s_{4i})^3-(1-u_{tp}-s)^3\right]+V_{o4} \tag{4.93}$$

where V_{qp} is defined in (4.87).

Region five ($0.5<s\le(1-u_{tp})$ and $(1-u_o)\le u_{sp}/2$): In this region, the nMOSFET is in saturation mode and the pMOSFET is in linear mode. The differential equation to describe the output transition and the current expressions for the pMOSFET and the nMOSFET are as follows:

$$\begin{cases}\dfrac{dV_o}{ds}=k_c V_{dd}+\dfrac{\tau}{C_t}(I_p-I_n)\\ I_p=k_p V_{sp}V_{op}\\ I_n=\dfrac{1}{2}k_n V_{sn}^2(1+\lambda_n V_o)\approx\dfrac{1}{2}k_n V_{sn}^2(1+\lambda_n V_{dd})\\ V_{sp}=V_{ip}-V_{tp}=V_{dd}(1-u_{tp}-s)\\ V_{sn}=V_{cn}\left(\sqrt{1+2\dfrac{s-u_{tn}}{u_{cn}}}-1\right)\approx k_m(s-b_m)\end{cases} \tag{4.94}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{5i}

and s_{5i}. In order to simplify the derivation without losing accuracy, the term of $(1+\lambda_n V_o)$ in the current I_n is reduced to $(1+\lambda_n V_{dd})$ since output voltage V_o is close to V_{dd} in this region. The saturation voltage V_{sn} is approximated by the linear equation with the parameters k_m and b_m:

$$\begin{cases} k_m = V_{cn}\dfrac{\left[\sqrt{1+2(1-u_{tp}-u_m)/u_{cn}}-\sqrt{1+2(0.5-u_m)/u_{cn}}\right]}{0.5-u_{tp}} \\ b_m = 0.5 - V_{cn}\left[\sqrt{1+2(0.5-u_m)/u_{cn}}-1\right]/k_m \end{cases} \tag{4.95}$$

In addition to the variables m_p and σ_p^2 defined in (4.69), set the variable V_{mn}:

$$V_{mn} = \frac{\tau k_n k_m^2 (1+\lambda_n V_{dd})}{C_t} \tag{4.96}$$

With these variables of k_m, b_m, m_p, σ_p^2, and V_{mn}, the rest of the derivations are exactly the same as in region two. Thus the expression V_o of output voltage in this region is.

$$V_o(s) = V_{dd} - f_5(s) + [f_5(s_{5i}) + V_{5i} - V_{dd}]\exp\left(-\frac{(s-s_{5i})(2m_p - s_{5i} - s)}{2\sigma_p^2}\right) \tag{4.97}$$

where the function $f_5(x)$ is given by:

$$f_5(x) = -\frac{(x+m_p-2b_m)\sigma_p^2 V_{mn}}{2} + \left[\frac{((m_p-b_m)^2+\sigma_p^2)V_{\ln}}{2} - k_c V_{dd}\right] R(x, m_p, \sigma_p) \tag{4.98}$$

Region six ($0.5 < s \le (1-u_{tp})$, $(1-u_o) > u_{sp}/2$), and $u_o \ge u_{sn}$): In this region, both the pMOSFET and the nMOSFET are in saturation mode. The differential equation to describe the output transition and the current expressions for the pMOSFET and the nMOSFET are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t}\left(I_p - I_n\right) \\ I_p = \dfrac{1}{2} k_p V_{sp}^2 \\ I_n = \dfrac{1}{2} k_n V_{sn}^2 \left(1 + \lambda_n V_o\right) \\ V_{sp} = V_{ip} - V_{tp} = V_{dd}\left(1 - u_{tp} - s\right) \\ V_{sn} = V_{cn}\left(\sqrt{1 + 2\dfrac{s - u_{tn}}{u_{cn}}} - 1\right) \end{cases} \tag{4.99}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{6i} and s_{6i}. The solutions of output voltages in this region are the same as those in region three. The nMOSFET current I_n dominates the discharging process in this region. As the first step, only the nMOSFET is analyzed to give the following expression:

$$V_{o6} = \left(\frac{1}{\lambda_n} + V_{6i}\right)\exp\left[-\lambda_n f_6(s)\right] - \frac{1}{\lambda_n} \tag{4.100}$$

In (4.100), the function $f_6(s)$ is

$$f_6(s) = V_{dn} \cdot \left\{ s - s_{6i} + \frac{(s - u_{tn})^2 - (s_{6i} - u_{tn})^2}{2u_{cn}} - \frac{u_{cn}}{3}\left[\left(1 + 2\frac{s - u_{tn}}{u_{cn}}\right)^{\frac{3}{2}} - \left(1 + 2\frac{s_{6i} - u_{tn}}{u_{cn}}\right)^{\frac{3}{2}}\right] \right\} \tag{4.101}$$

where V_{dn} is defined in (4.41). The second step is that nMOSFET, pMOSFET, and coupling capacitance current between input node and output node are included in the analysis. Then, the output voltage $V_o(t)$ in this region is

$$V_o(s) = k_c V_{dd}(s - s_{6i}) + \frac{V_{qp}}{6}\left[\left(1 - u_{tp} - s_{6i}\right)^3 - \left(1 - u_{tp} - s\right)^3\right] + V_{o6} \tag{4.102}$$

where the variable V_{qp} is defined in (4.87).

Region seven ($0.5 < s \le (1-u_{tp})$ and $u_o < u_{sn}$): In this region, the nMOSFET is in linear mode and the pMOSFET is in saturation mode. The differential equation to describe the output transition and the current expressions for the pMOSFET and the nMOSFET are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t}(I_p - I_n) \\ I_p = \dfrac{1}{2} k_p V_{sp}^2 \\ I_n = k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o(1+\lambda_n V_{sn}) \approx k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o \\ V_{sp} = V_{ip} - V_{tp} = V_{dd}(1-u_{tp}-s) \\ V_{sn} = V_{cn}\left(\sqrt{1+2\dfrac{s-u_{tn}}{u_{cn}}} - 1\right) \approx k_m(s-b_m) \end{cases} \tag{4.103}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{7i} and s_{7i}. For simplicity without losing accuracy, the term of $(1+\lambda_n V_{sn})$ in the current I_n is ignored since the effect of $\lambda_n V_{sn}$ is small in this region. The saturation voltage V_{sn} of the nMOSFET is approximated by the same linear equation as in Region five with the parameters k_m and b_m given in (4.95). The nMOSFET current dominates the discharging process in this region. As the first step, only nMOSFET is considered in the derivation. Set the following variable

$$\sigma_m^2 = \frac{C_t}{\tau k_n k_m} \tag{4.104}$$

Note that parameter k_m in the above variable σ_m^2 is given in (4.95). Therefore, the output

voltage without the input-output coupling current and the pMOSFET is

$$V_{o7}=\frac{1}{\frac{\tau k_n}{2C_t}P(s,b_m,\sigma_m)+\left(-\frac{\tau k_n}{2C_t}P(s_{7i},b_m,\sigma_m)+\frac{1}{V_{7i}}\right)\exp\left[\frac{(s+s_{7i}-2b_m)(s-s_{7i})}{2\sigma_m^2}\right]} \tag{4.105}$$

where $P(x,m,\sigma)$ is defined in (III.9a) of Appendix III and the parameter b_l is given in (4.95). The second step is that nMOSFET, pMOSFET, and coupling capacitance current between input and output are involved in the analysis. Then, the solution is expressed as follows:

$$V_o(s)=k_cV_{dd}(s-s_{7i})+\frac{V_{qp}}{6}\left[(1-u_{tp}-s_{7i})^3-(1-u_{tp}-s)^3\right]+V_{o7} \tag{4.106}$$

where V_{qp} is defined in (4.87).

Region eight ($1-u_{tp}<s\le 1$ and $u_o\ge u_{sn}$): In this region, nMOSFET is in saturation mode and pMOSFET is in cutoff mode $(I_p=0)$. The differential equation to describe the output transition and the nMOSFET current expression are as follows:

$$\begin{cases}\dfrac{dV_o}{ds}=k_cV_{dd}-\dfrac{\tau}{C_t}I_n\\ I_n=\dfrac{1}{2}k_nV_{sn}^2(1+\lambda_nV_o)\\ V_{sn}=V_{cn}\left(\sqrt{1+2\dfrac{s-u_{tn}}{u_{cn}}}-1\right)\end{cases} \tag{4.107}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{8i} and s_{8i}. The nMOSFET current dominates the discharging process in this region. In this case, the derivation of output voltage is similar to that in region three except that the pMOSFET is in cutoff mode $(I_p=0)$. As the first step, only nMOSFET is considered.

The differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{2C_t} V_{sn}^2 \left(1 + \lambda_n V_o\right) \tag{4.108}$$

Substituting the nMOSFET saturation voltage V_{sn} of (4.107) into the above differential equation and using the variable V_{dn} of (4.41) yield

$$\frac{dV_o}{ds} = -\frac{V_{dn}}{2}\left(\sqrt{1+2\frac{s-u_{tn}}{u_{cn}}}-1\right)^2 \left(1+\lambda_n V_o\right) \tag{4.109}$$

Assume $\lambda_n > 0$. The solution of the above differential equation is expressed as follows:

$$V_{o8} = \left(\frac{1}{\lambda_n} + V_{8i}\right)\exp\left[-\lambda_n f_8(s)\right] - \frac{1}{\lambda_n} \tag{4.110}$$

In (4.110), the function $f_8(s)$ is

$$f_8(s) = V_{dn} \cdot \left\{ s - s_{8i} + \frac{(s-u_{tn})^2 - (s_{8i}-u_{tn})^2}{2u_{cn}} - \frac{u_{cn}}{3}\left[\left(1+2\frac{s-u_{tn}}{u_{cn}}\right)^{\frac{3}{2}} - \left(1+2\frac{s_{8i}-u_{tn}}{u_{cn}}\right)^{\frac{3}{2}}\right]\right\} \tag{4.111}$$

where V_{dn} is defined in (4.41). The second step is that nMOSFET and coupling capacitance current between input node and output node are included in the analysis. Then, we have

$$V_o(s) = k_c V_{dd}(s - s_{8i}) + V_{o8} \tag{4.112}$$

Region nine ($1-u_{tp} < s \le 1$ and $u_o < u_{sn}$): In this region, the nMOSFET is in linear mode and the pMOSFET is in cutoff mode $(I_p = 0)$. Then the differential equation to describe the output transition and the nMOSFET current expression are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} - \dfrac{\tau}{C_t} I_n \\ I_n = k_n \left(V_{sn} - \dfrac{V_o}{2} \right) V_o \left(1 + \lambda_n V_{sn} \right) \approx k_n \left(V_{sn} - \dfrac{V_o}{2} \right) V_o \\ V_{sn} = V_{cn} \left(\sqrt{1 + 2\dfrac{s - u_{tn}}{u_{cn}}} - 1 \right) \approx k_h \left(s - b_h \right) \end{cases} \tag{4.113}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{9i} and s_{9i}. In order to simplify the derivation without losing accuracy, the term of $\left(1 + \lambda_n V_{sn}\right)$ in the current I_n is ignored since the effect of $\lambda_n V_{sn}$ is small in this region. The nMOSFET saturation voltage V_{sn} in this region is approximated by the linear equation with the parameters k_h and b_h as follows:

$$\begin{cases} k_h = \dfrac{V_{cn}}{u_{tp}} \left(\sqrt{1 + 2(1 - u_{tn})/u_{cn}} - \sqrt{1 + 2\left(1 - u_{tp} - u_{tn}\right)/u_{cn}} \right) \\ b_h = 1 - u_{tp} \dfrac{\sqrt{1 + 2(1 - u_{tn})/u_{cn}} - 1}{\sqrt{1 + 2(1 - u_{tn})/u_{cn}} - \sqrt{1 + 2\left(1 - u_{tp} - u_{tn}\right)/u_{cn}}} \end{cases} \tag{4.114}$$

The nMOSFET current dominates the discharging process in this region. The simulation results show that only the nMOSFET current is accurate enough to describe the discharging process in this region. The differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{C_t} \left(V_{sn} - \frac{V_o}{2} \right) V_o \tag{4.115}$$

This differential equation is a Bernoulli's equation. The solution of the differential equation can be obtained by employing the procedures in region four. Set the following variable σ_h^2

$$\sigma_h^2 = \frac{C_t}{\tau k_n k_h} \tag{4.116}$$

Note that parameter k_h in the above variable σ_h^2 is given in (4.114). Therefore, the output voltage without the input-output coupling current is

$$V_o(s)=\frac{1}{\frac{\tau k_n}{2C_t}P(s,b_h,\sigma_h)+\left(-\frac{\tau k_n}{2C_t}P(s_{9i},b_h,\sigma_h)+\frac{1}{V_{9i}}\right)\exp\left[\frac{(s+s_{9i}-2b_h)(s-s_{9i})}{2\sigma_h^2}\right]} \quad (4.117)$$

where $P(x,m,\sigma)$ is defined in (III.9a) of Appendix III.

Region ten ($s>1$ or $t>\tau$ and $V_o \geq V_{snd}$): In this region, the pMOSFET is in cutoff mode $(I_p=0)$ and the nMOSFET is in saturation mode with input voltage $V_i=V_{dd}$. The parameter V_{snd} is the nMOSFET drain-source saturation voltage with input voltage $V_i=V_{dd}$. Then the differential equation to describe the output transition and the nMOSFET current expression are as follows:

$$\begin{cases} \dfrac{dV_o}{dt}=-\dfrac{1}{C_t}I_n \\ I_n=\dfrac{1}{2}k_nV_{snd}^2(1+\lambda_nV_o) \\ V_{snd}=V_{cn}\left(\sqrt{1+2\dfrac{V_{dd}-V_{tn}}{V_{cn}}}-1\right) \end{cases} \quad (4.118)$$

The initial values of output voltage $V_o(t)$ and time t in this region are V_{10i} and t_{10i}. Assume $\lambda_n>0$, then the output waveform $V_o(t)$ is

$$V_o(t)=\left(\frac{1}{\lambda_n}+V_{10i}\right)\exp\left[-\frac{t-t_{10i}}{\tau_{sat}}\right]-\frac{1}{\lambda_n} \quad (4.119)$$

or

$$t=t_{10i}+\tau_{sat}\ln\left(\frac{1+\lambda_nV_{10i}}{1+\lambda_nV_o}\right) \quad (4.120)$$

The time constant τ_{sat} in (4.119) and (4.120) is the same as one in (4.19), i.e.,

$$\tau_{sat} = \frac{2C_t}{k_n V_{snd}^2 \lambda_n} \tag{4.121}$$

Region eleven ($s > 1$ or $t > \tau$ and $V_o < V_{snd}$): In this region, the pMOSFET is in the cutoff mode $(I_p = 0)$ and the nMOSFET is in the linear mode with input voltage $V_i = V_{dd}$. The parameter V_{snd} is the nMOSFET drain-source saturation voltage with input voltage $V_i = V_{dd}$. Then the differential equation to describe the output transition and the current expression for the nMOSFET are as follows:

$$\begin{cases} \dfrac{dV_o}{dt} = -\dfrac{1}{C_t} I_n \\ I_n = k_n \left(V_{snd} - \dfrac{V_o}{2} \right) V_o \left(1 + \lambda_n V_{snd} \right) \\ V_{snd} = V_{cn} \left(\sqrt{1 + 2\dfrac{V_{dd} - V_{tn}}{V_{cn}}} - 1 \right) \end{cases} \tag{4.122}$$

For generality, the initial values of output voltage $V_o(t)$ and time t in this region are V_{11i} and t_{11i}. Then the output waveform $V_o(t)$ is

$$V_o(t) = \frac{2V_{snd}}{1 + (2V_{snd}/V_{11i} - 1)\exp[(t - t_{11i})/\tau_{lin}]} \tag{4.123}$$

or

$$t = t_{11i} + \tau_{lin} \ln\left(\frac{V_{11i}}{V_o} \frac{2V_{snd} - V_o}{2V_{snd} - V_{11i}} \right) \tag{4.124}$$

where the time constant τ_{lin} in (4.123) and (4.124) is the same as the one defined in (4.27), i.e.

$$\tau_{lin} = \frac{C_t}{k_n V_{snd}(1+\lambda_n V_{snd})} \tag{4.125}$$

4.5.3 Results and Discussions

In the previous subsection, we have derived the analytical expressions of output voltages of a CMOS inverter for a rising ramp input by solving the differential equations of (4.2) or (4.5). The dominating nMOSFET of a CMOS inverter is modeled with PDSMM while the trivial pMOSFET is described with a simplified Shockley model. Figs 4.18-4.22 show the voltage output waveforms with input rising time of 1ps, 50ps, 100ps, 200ps, and 400ps, respectively. In these figures, solid lines represent the output voltages from the comprehensive timing model based on PDSMM developed in Section 4.5.2. The dashed lines indicate the output voltage from SPICE BSIM3. The input waveforms are plotted as dotted lines. The region dividing lines are shown with dashed-dotted lines. For each figure, external output capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF (from left to right), respectively. The overshoot of output voltage partially compensated with the inverse discharging current through the trivial pMOSFET, which is much smaller than those from the nMOSFET-only model in Section 4.4, matches with the results from SPICE BSIM3 very well. These plots show that the comprehensive timing model of output voltages based on PDSMM are very accurate comparing with SPICE BSIM3.

Fig. 4.23 shows propagation delays versus external load capacitances for input transition times of 1ps, 10ps, 100ps, 200ps, and 400ps (bottom to top), respectively. Fig. 4.24 plots propagation delays versus input transition times with external load capacitances of 10fF, 100fF, 200fF, 400fF, and 800fF (bottom to top), respectively. In both figures, the propagation delays of a CMOS inverter obtained from SPICE BSIM3 are

indicated with dashed lines and the propagation delays from the comprehensive timing model based on PDSMM in Subsection 4.5.2 are represented with circles. Both the dominating nMOSFET and the trivial pMOSFET are included in the comprehensive timing model. For fast input, the propagation delay is proportional to external output capacitance. For large external output capacitance, the propagation delay is proportional to input transition time. Generally speaking, their relationships are non-linear. In the extremely broad range of input transition time and external load capacitance the comprehensive timing model based on PDSMM proposed in Subsection 4.5.2 is very accurate comparing with SPICE BSIM3.

Fig. 4.25 shows the correlation of propagation delays from the comprehensive timing model based on PDSMM in Subsection 4.5.2 with propagation delays from SPICE BSIM3. The two dashed lines are ±5% of propagation delays from SPICE BSIM3. All the data are in an excellent agreement. Therefore, in order to completely and accurately model the propagation delays of a CMOS inverter it is necessary to include not only the dominating nMOSFET but also the trivial pMOSFET.

Table 4.1 The PDSMM Parameters of nMOSFET and pMOSFET of a CMOS Inverter.

	width (μm)	k (mA/V^2)	λ (1/V)	V_c (V)	V_{th} (V)
nMOSFET	21.76	22.6968	0.1969	0.6318	0.3232
pMOSFET	43.52	7.7267	0.2592	3.6777	0.1978

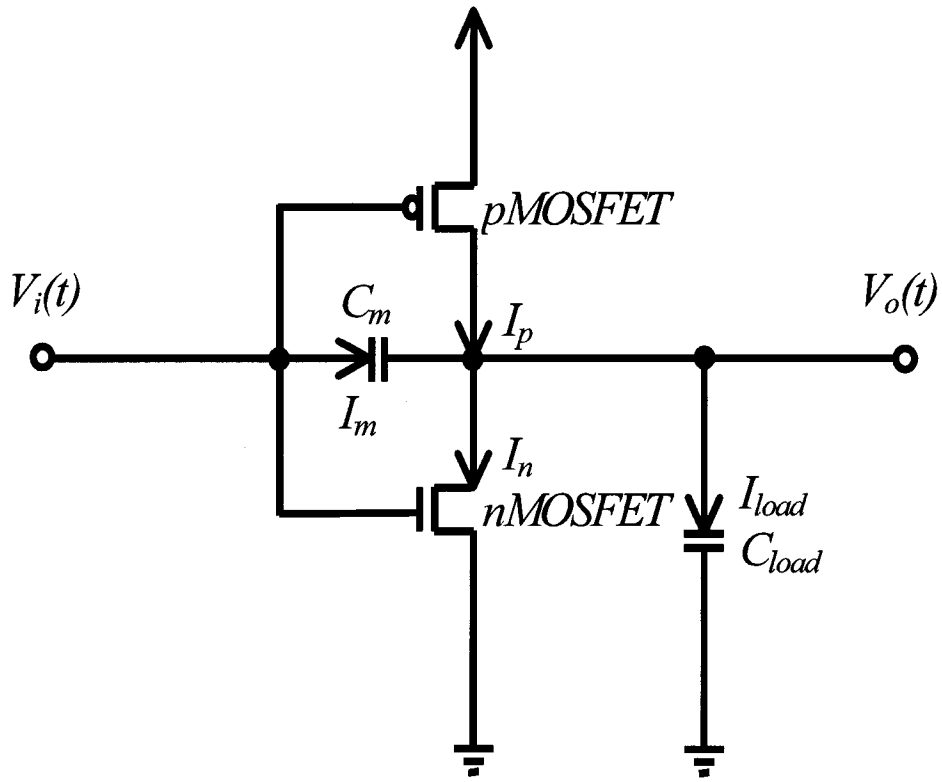


Fig. 4.1 A CMOS Inverter. The input-output coupling capacitance C_m and output load capacitance C_{load} are included.

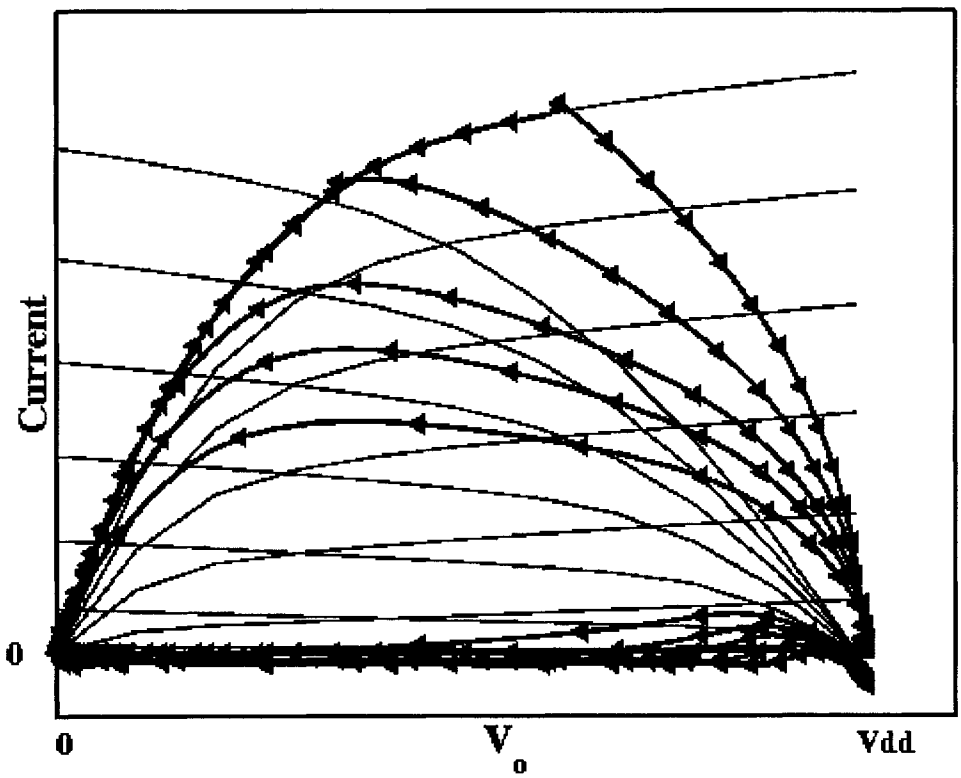


Fig. 4.2 The typical switching trajectory traces of both pMOSFET and nMOSFET of a CMOS inverter for rising ramp inputs. The static I–V characteristics of both pMOSFET and nMOSFET of the CMOS inverter are also plotted. Obviously the nMOSFET drain-source current I_{ds} dominates the output transition or discharging process in case of rising ramp inputs.

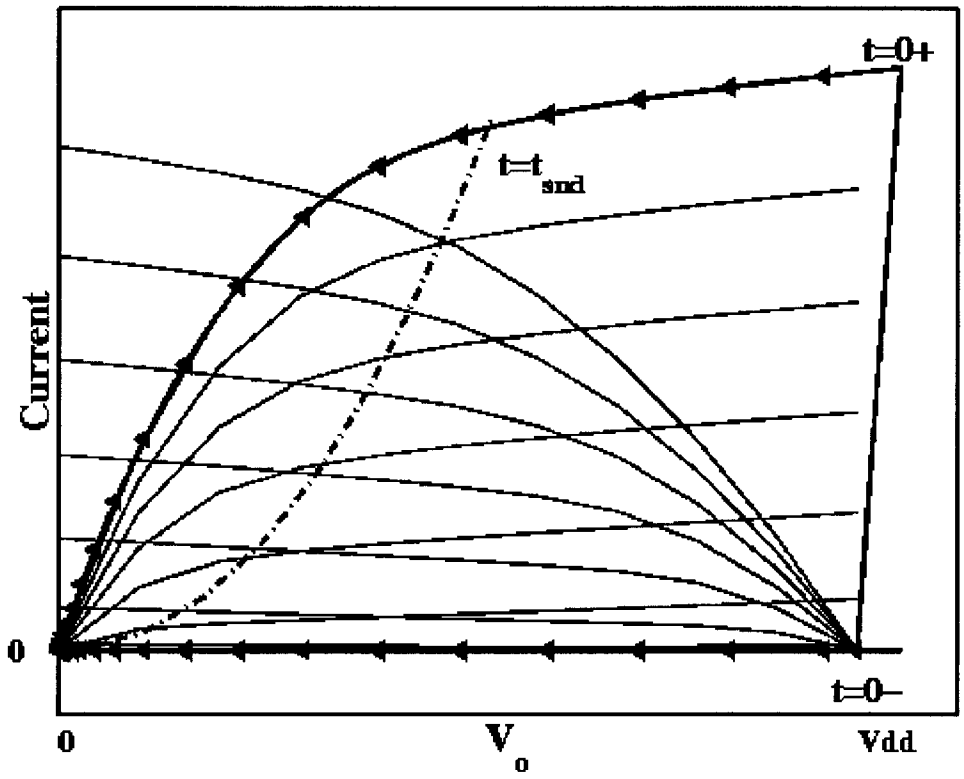


Fig. 4.3 The typical discharging trajectory traces of both pMOSFET and nMOSFET for a step input. In this case, the pMOSFET can be ignored since there is no current passing through the pMOSFET. There is a jump of drain-source current I_{ds} *at* $t = 0$. *The output voltage has an overshoot at* $t = 0$. *When* $t > 0$, *the discharging trajectory trace of nMOSFET goes along the characteristic curve of the nMOSFET with* $V_{gs} = V_{dd}$

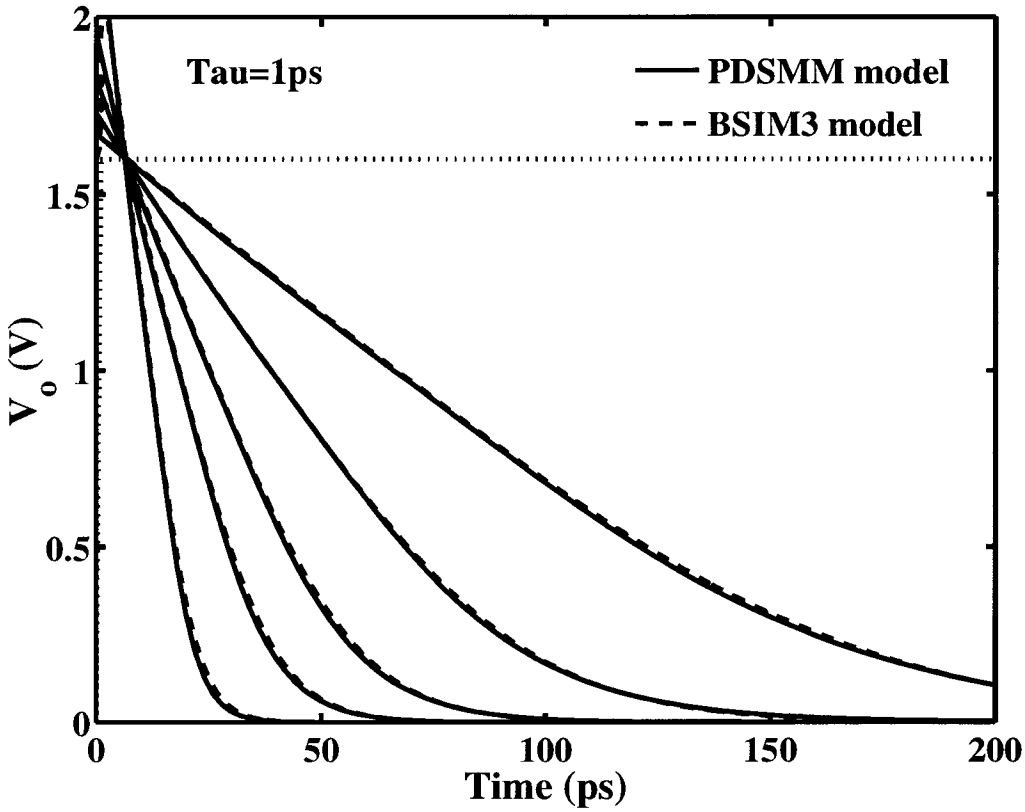


Fig. 4.4 Output voltage waveforms of a CMOS inverter with different external capacitances of 10fF, 100fF, 200fF, 400fF, and 800fF (from left to right) for a rising step input. The transient voltage waveforms of a step input based on PDSMM in Subsection 4.3.2 are shown in solid lines. As a comparison, the output voltage waveforms of a ramp input with input transition time of 1ps from SPICE BSIM3 are also plotted in dashed lines.

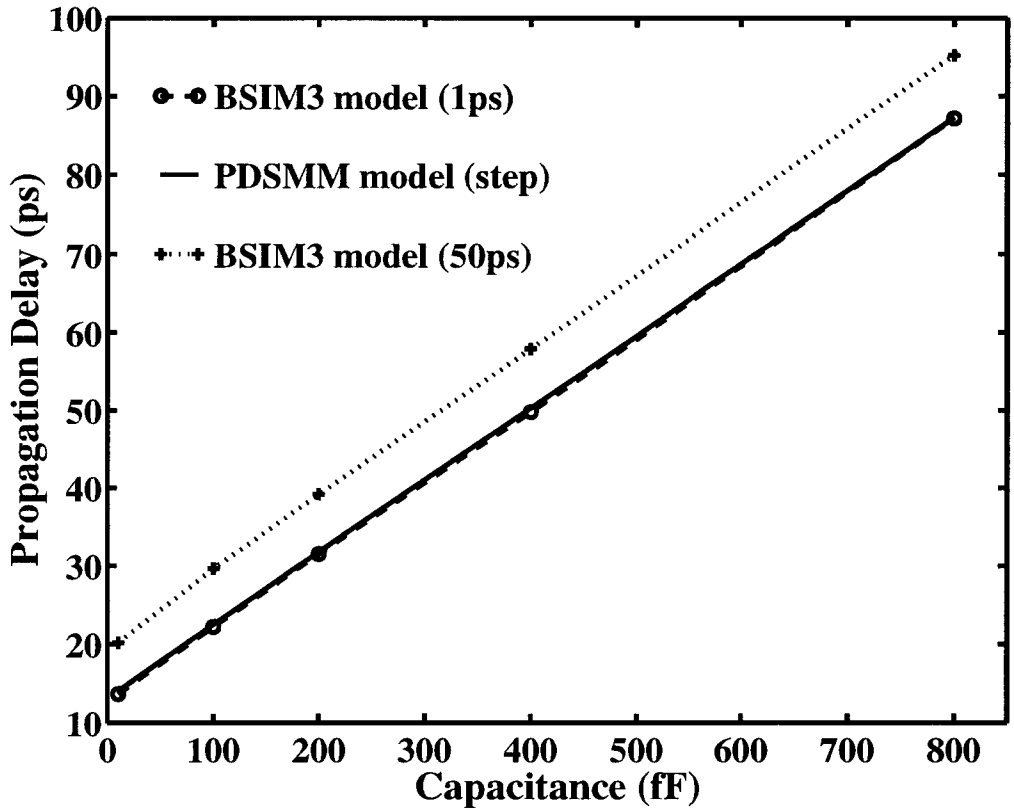


Fig. 4.5 Propagation delay versus external capacitance for a rising step input. The propagation delay derived from PDSMM in Subsection 4.3.2 is indicated as the solid line. The dashed line with circle symbol represents the propagation delay obtained from SPICE BSIM3 for a ramp input of 1ps transition time. As a comparison, the propagation delay of a CMOS inverter with a ramp input of 50ps transition time is shown in dotted line with plus symbol.

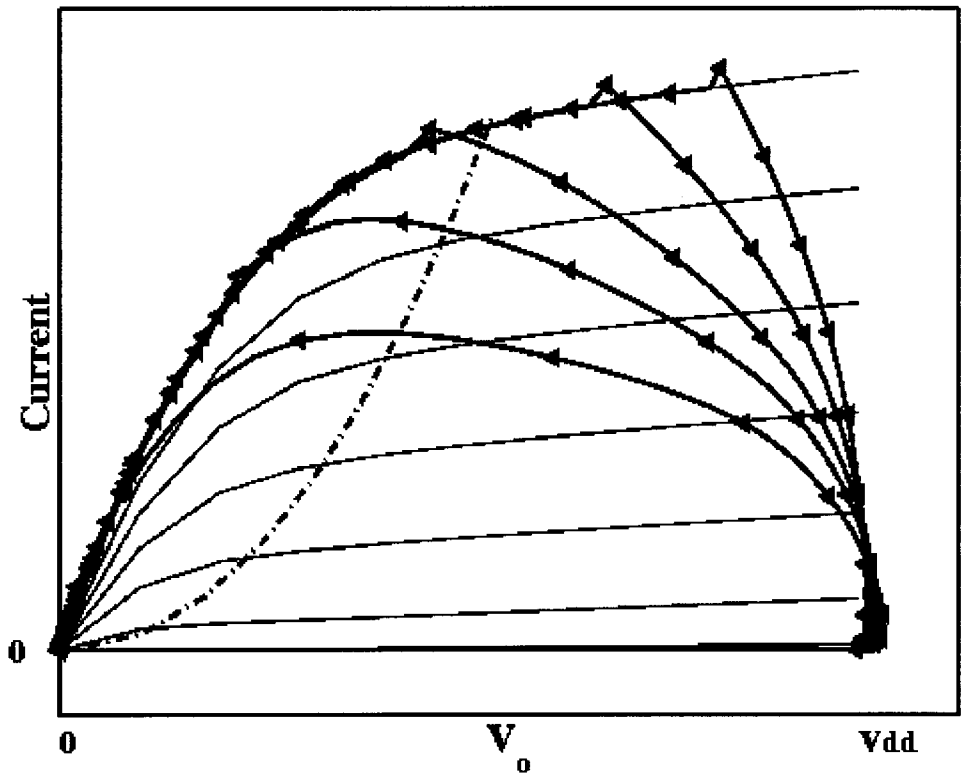


Fig. 4.6 The typical discharging trajectory traces of an nMOSFET for ramp inputs. The static I–V characteristics of only dominating nMOSFET of the CMOS inverter are also plotted.

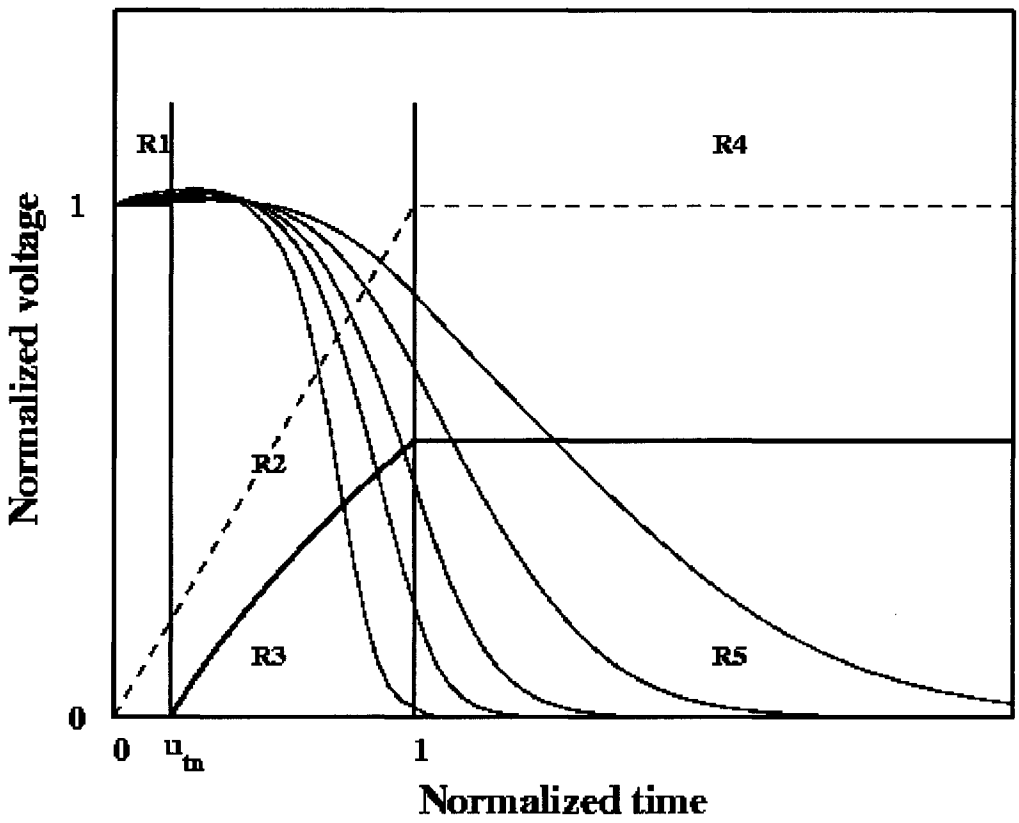


Fig. 4.7 Five operation regions (R1-R5) of the dominating nMOSFET of a CMOS inverter for rising ramp inputs. The five operation regions are introduced in the text. The ramp input is indicated with the dashed line. The typical discharging output waveforms are also plotted.

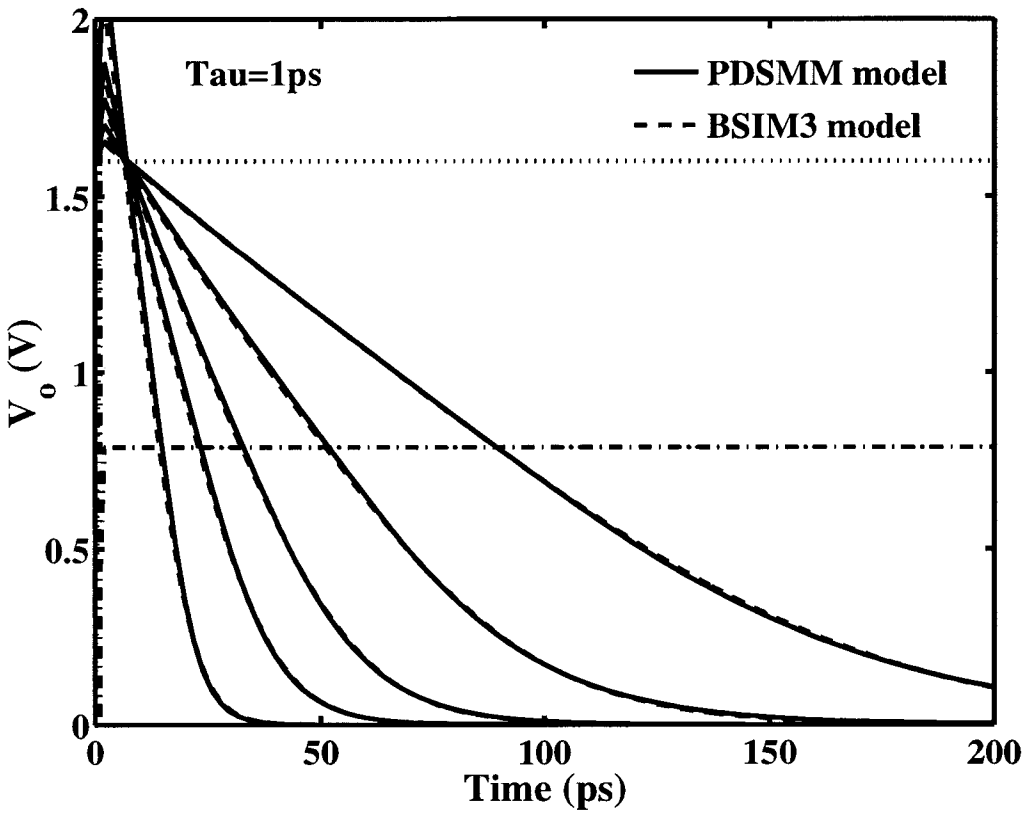


Fig. 4.8 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (solid lines). Only a dominating nMOSFET denotes the CMOS inverter. Input transition time is 1ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). The calculated output waveforms from PDSMM and the output responses from SPICE BSIM3 agree with each other excellently.

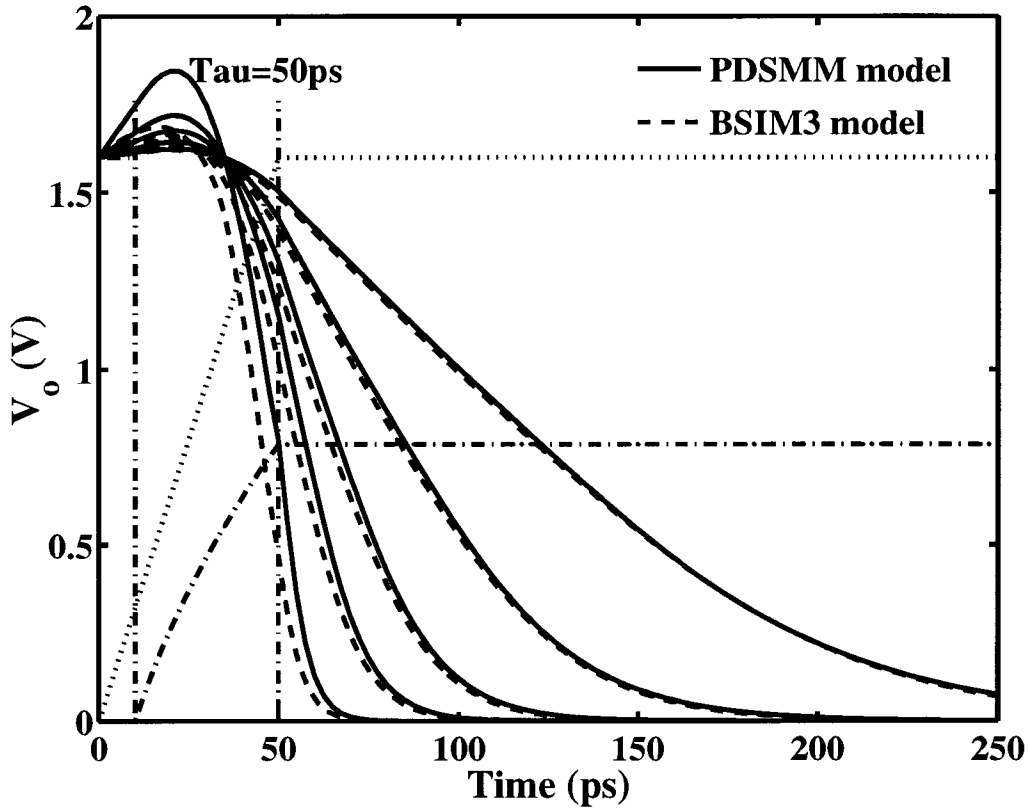


Fig. 4.9 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (solid lines). Input transition time is 50ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). Note that for smaller loads, the calculated waveforms are obviously slower than SPICE BSIM3 results due to exaggerated voltage overshoots caused with input-to-output coupling current. In practice, voltage overshoot is lowered by being partially compensated with inverse discharging current through the trivial pMOSFET of a CMOS inverter to power rail.

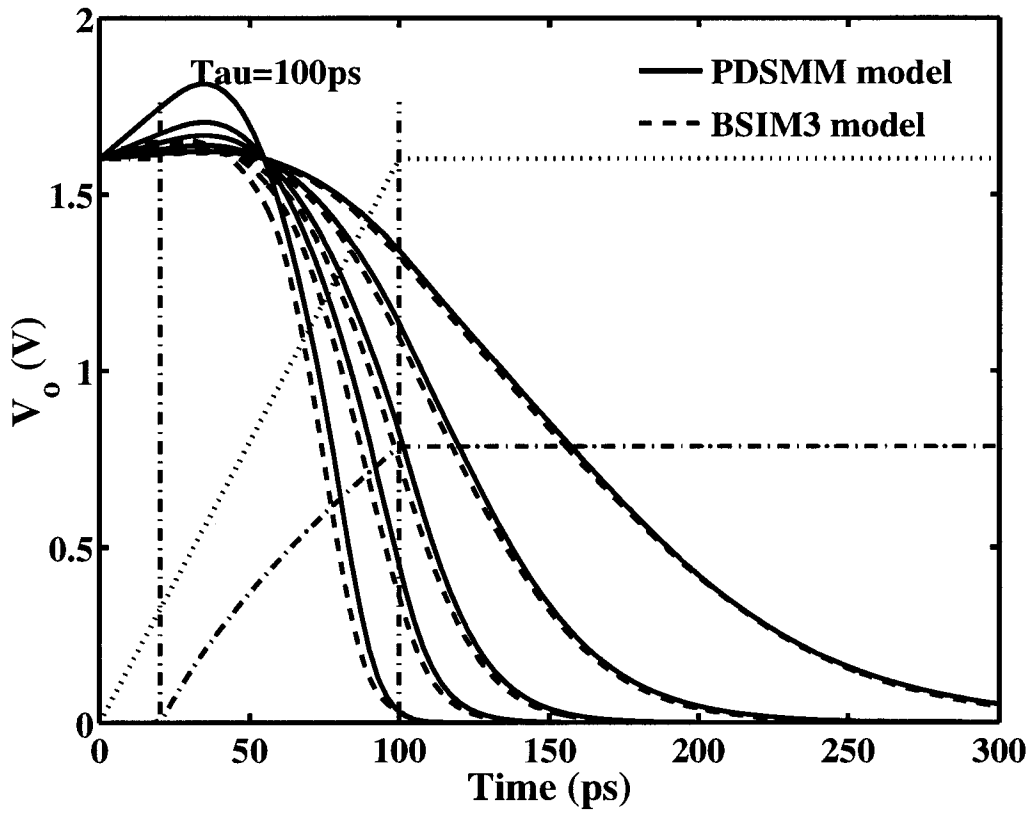


Fig. 4.10 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (solid lines). Input transition time is 100ps. Load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). Voltage overshoots caused with input-to-output coupling capacitance are exaggerated. Comparing with Fig. 4.9, the longer input transition time makes the effect of short-circuit current more significant. The only nMOSFET currents push the calculated waveforms left (or fast). The effect of exaggerated overshoots on propagation delays are partially compensated with no short-circuit currents.

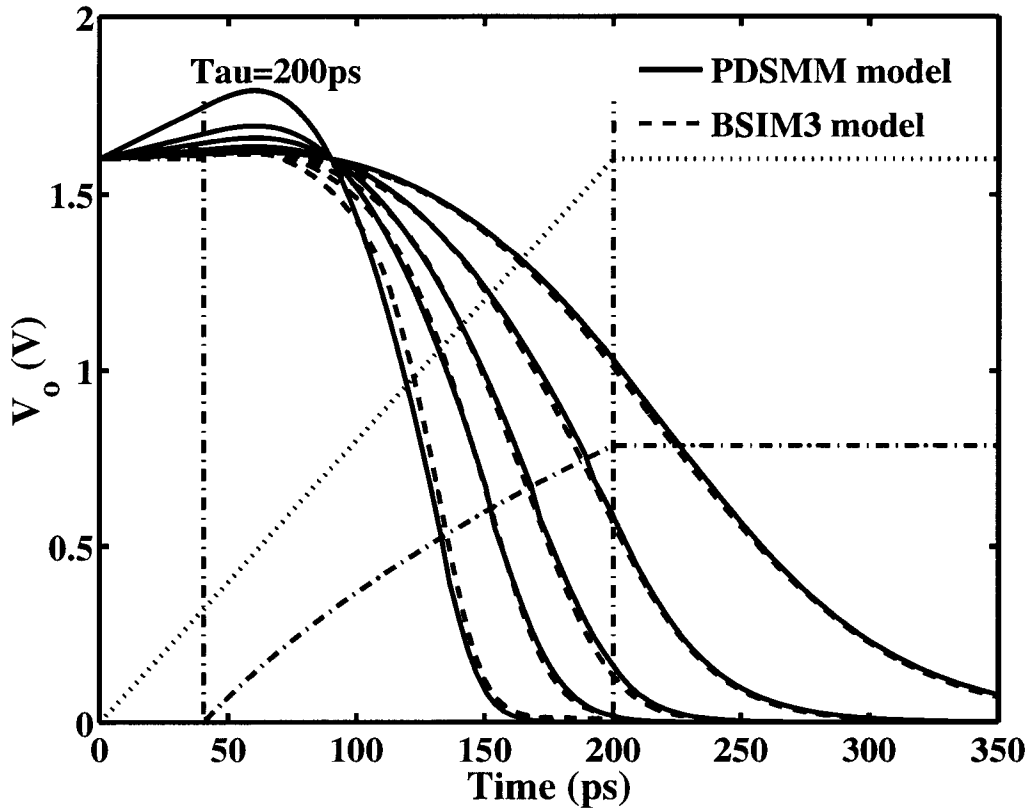


Fig. 4.11 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (solid lines). Input transition time is 200ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). Voltage overshoots caused with input-to-output coupling current are exaggerated. Comparing with Fig. 4.10, the longer input transition time makes the effect of short-circuit current more significant. No short-circuit currents push the calculated waveforms left (or fast). The effect of exaggerated overshoots on propagation delay are almost compensated with no short-circuit currents.

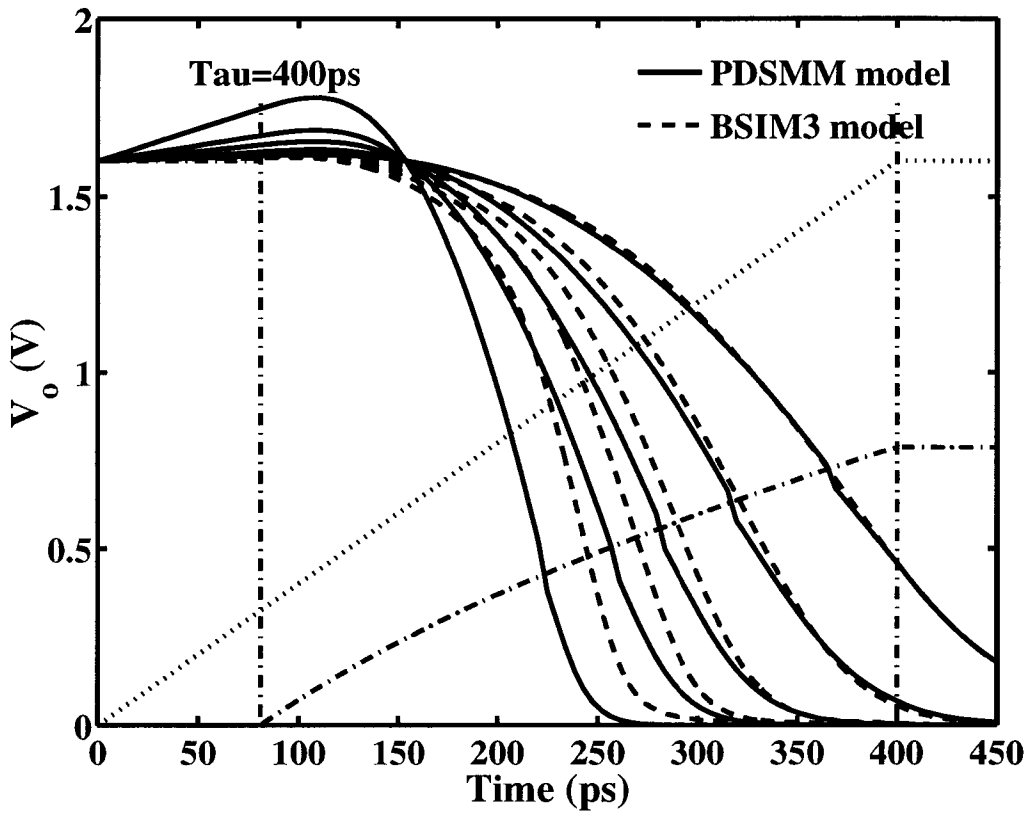


Fig. 4.12 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (solid lines). Input transition time is 400ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). Voltage overshoots are exaggerated. The very slow input transition time makes the effect of no short-circuit current more significant. The uncompensated nMOSFET current discharges the load capacitance much faster. The phenomena make the calculated propagation delays unreliable.

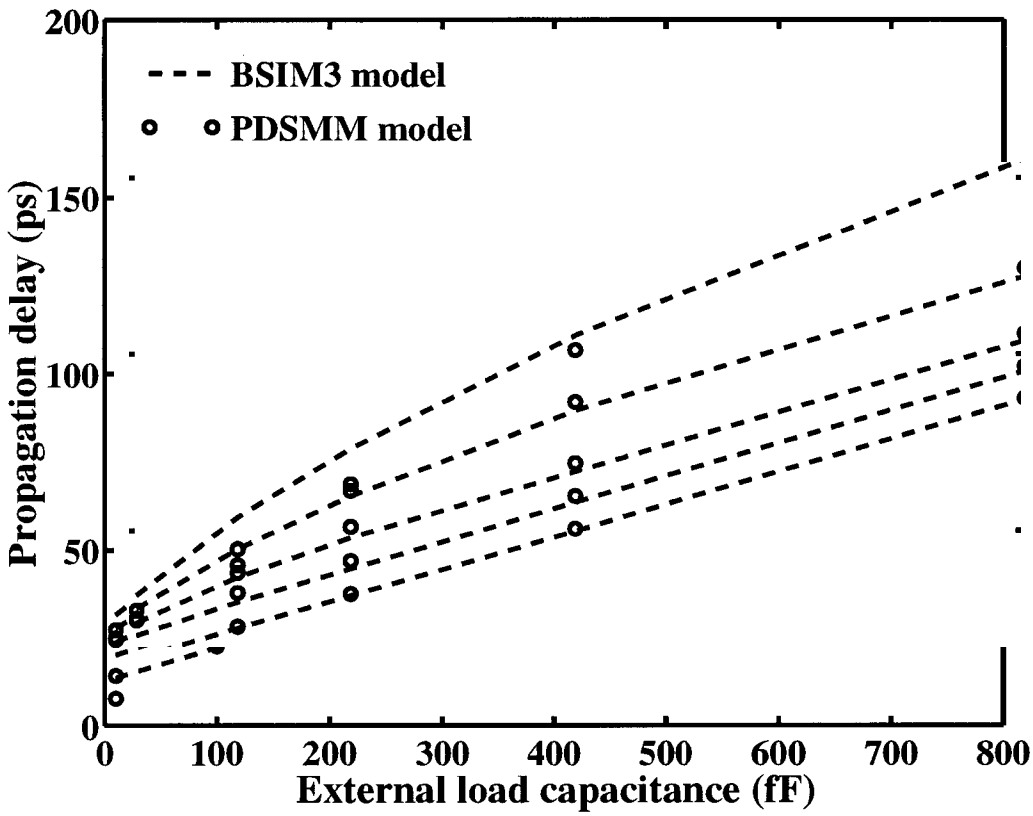


Fig. 4.13 Propagation delays versus external load capacitances for input transition times of 1ps, 10ps, 100ps, 200ps, and 400ps (from bottom to top), respectively. The propagation delays of a CMOS inverter are obtained from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (circles). For small input transition time or large load capacitance, the method proposed in Subsection 4.4.2 is very accurate. However, large input transition time and small load capacitance make a significant error, which cause the expressions in Subsection 4.42 unreliable.

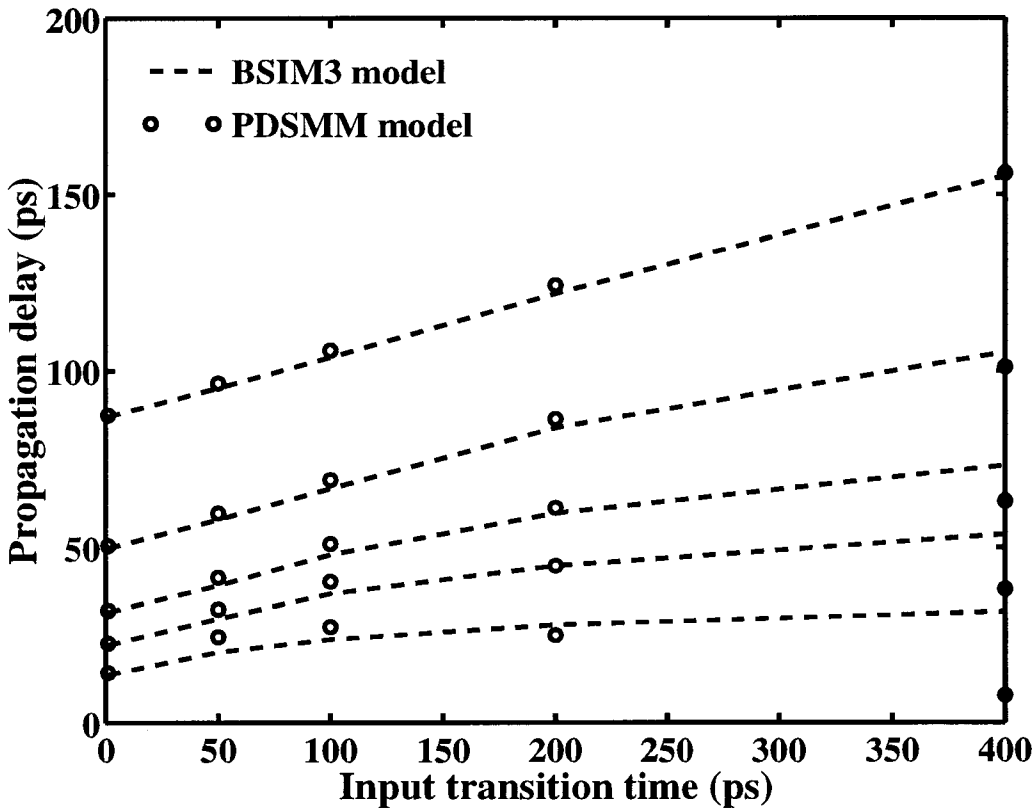


Fig. 4.14 Propagation delays versus input transition times with the external load capacitances of 10fF, 100fF, 200fF, 400fF, and 800fF (from bottom to top), respectively. The propagation delays of a CMOS inverter are obtained from SPICE BSIM3 (dashed lines) and from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 (circles). For small input transition time or large load capacitance, the method proposed in Subsection 4.4.2 is very accurate. However, large input transition time and small load make significant errors, which cause the expressions in Subsection 4.4.2 unreliable.

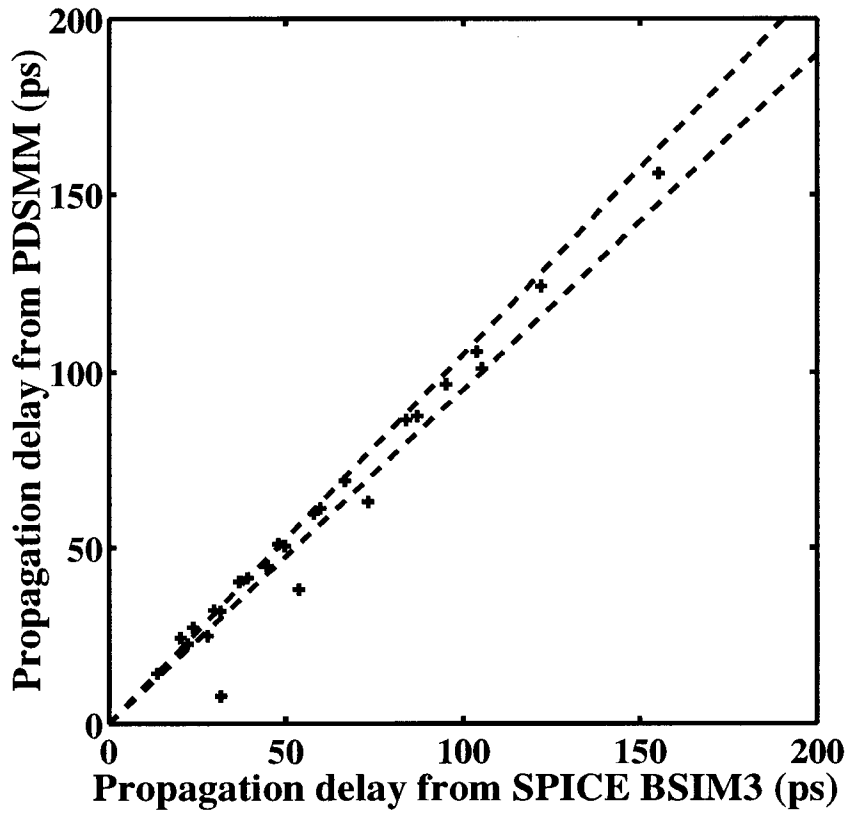


Fig. 4.15 Correlation of propagation delays from the nMOSFET-only timing model based on PDSMM in Subsection 4.4.2 with propagation delays from SPICE BSIM3. Only a dominating nMOSFET denotes the CMOS inverter. The two dashed lines are ±5% of propagation delays deviating from SPICE BSIM3. Significant errors are caused with large input transition time and small load capacitance.

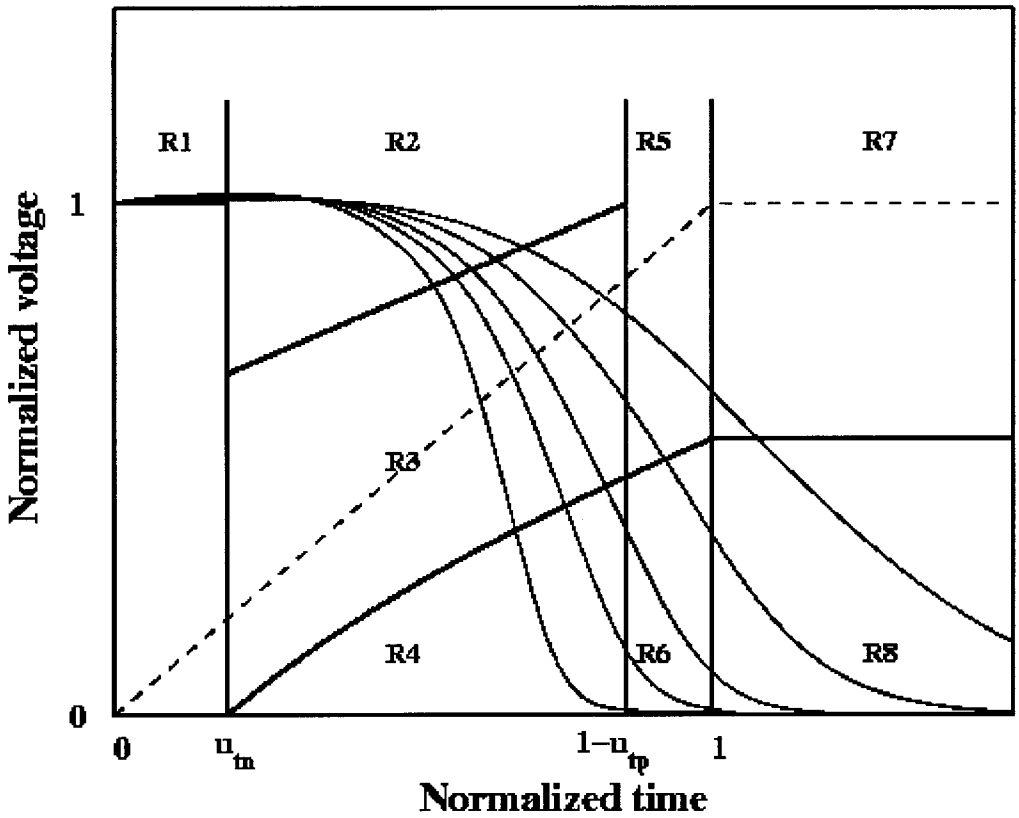


Fig. 4.16 Eight operation regions of a CMOS inverter for a rising ramp input. The input ramp is indicated with dashed line. However, the solution of the differential equation for region two is not accurate enough due to simplification of the property equations. The eleven operation regions shown in Fig 4.17 are used in the comprehensive timing mode.

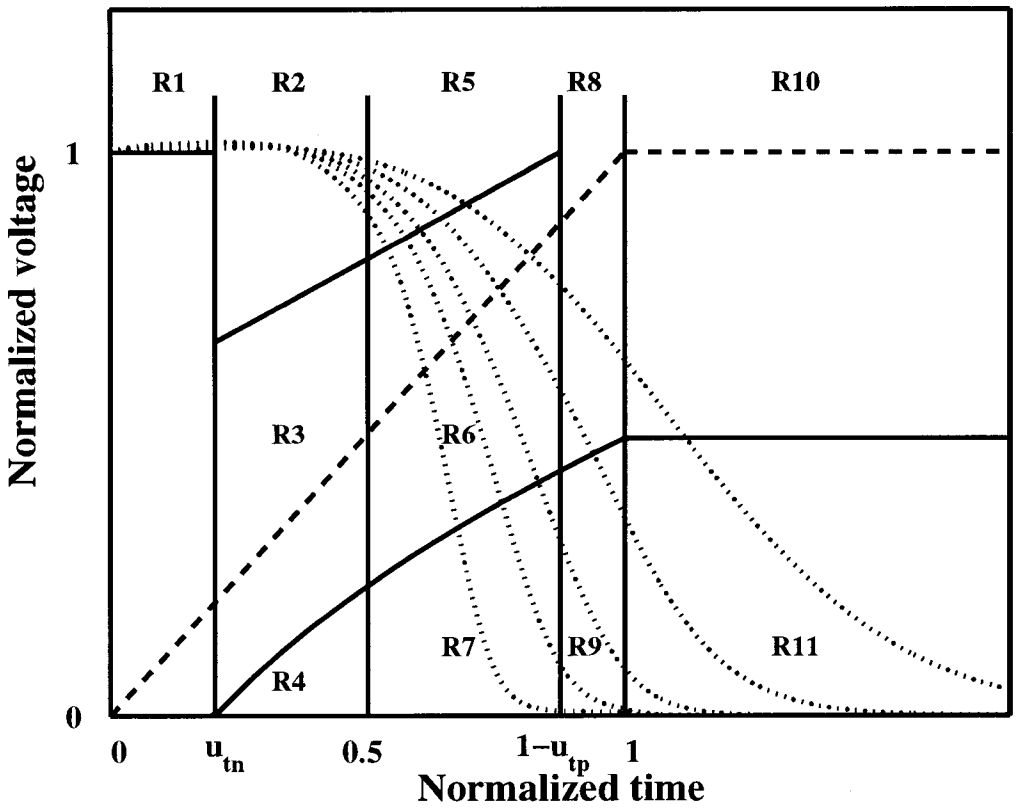


Fig. 4.17 Eleven operation regions of a CMOS inverter with a rising ramp input for the comprehensive timing model. The input ramp is indicated with dashed line. The dotted lines represent the output responses of rising input ramps.

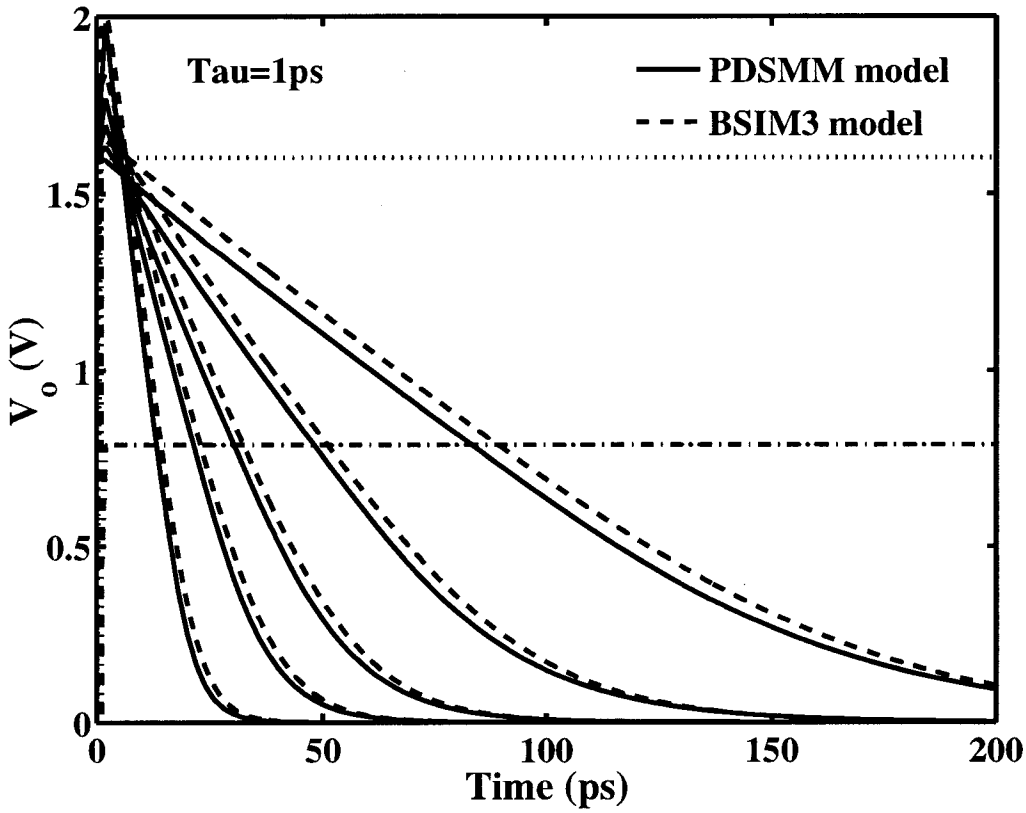


Fig. 4.18 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. Input transition time is 1ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right).

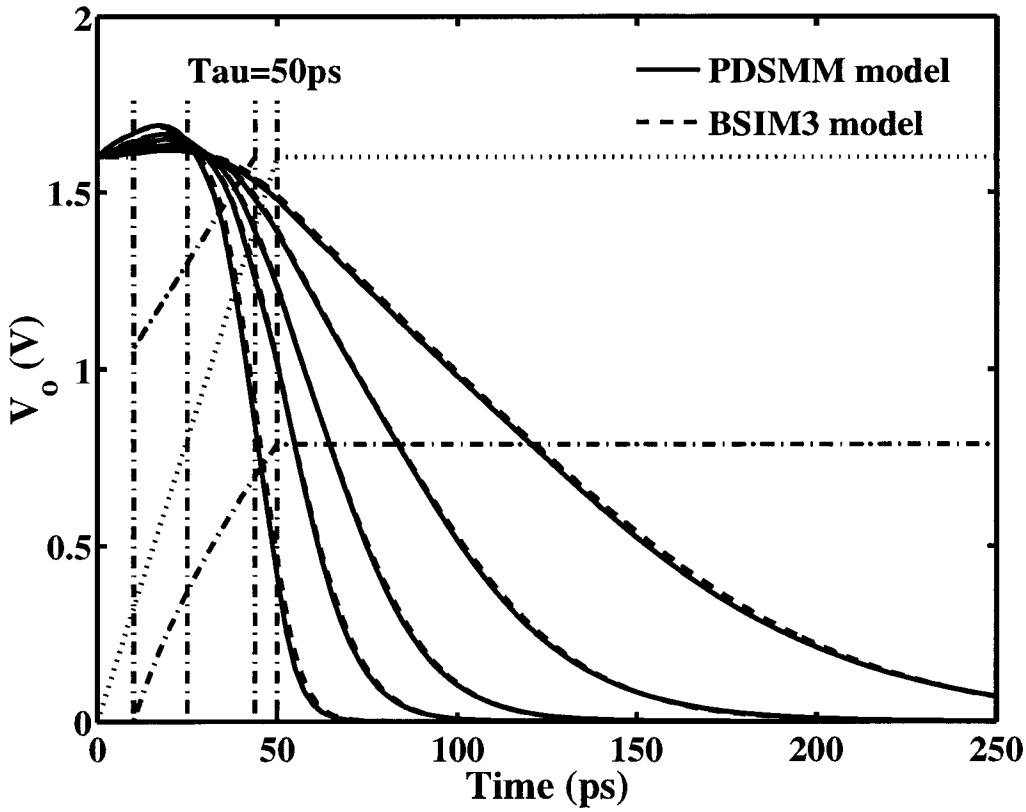


Fig. 4.19 Output voltage waveforms of a CMOS inverter from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. Input transition time is 50ps. External load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). They match with each other excellently.

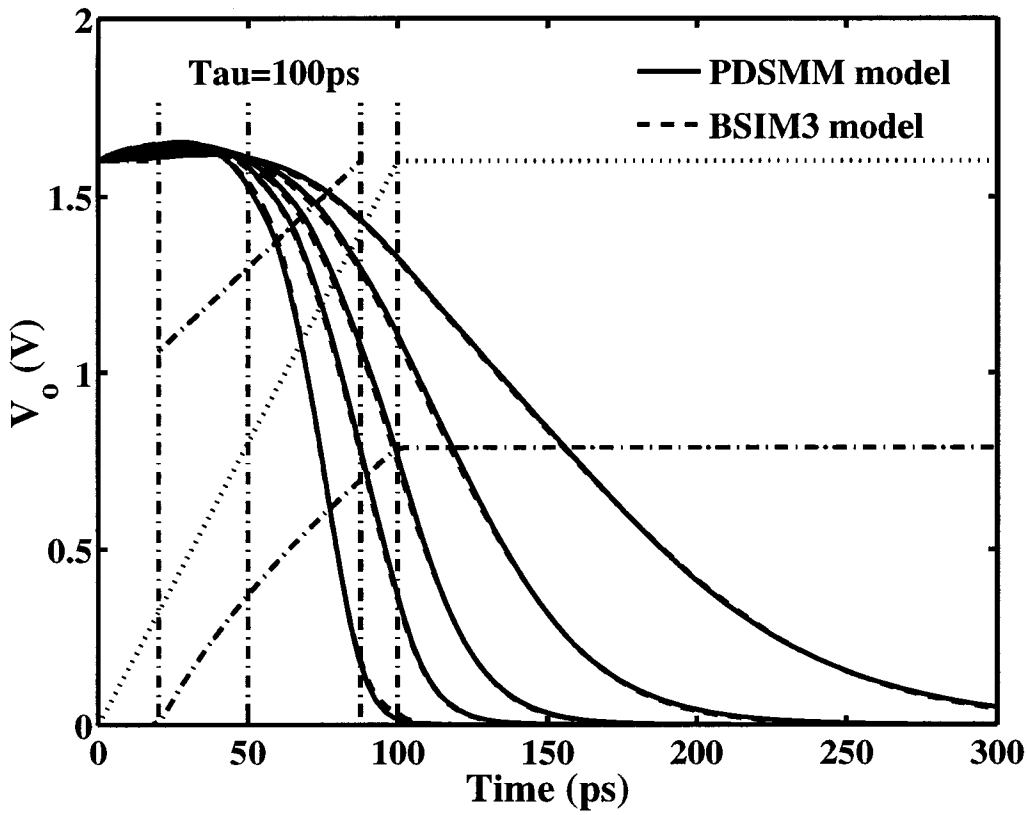


Fig. 4.20 The output voltage waveforms of a CMOS inverter from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. The input transition time is 100ps. The external load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). They agree with each other excellently.

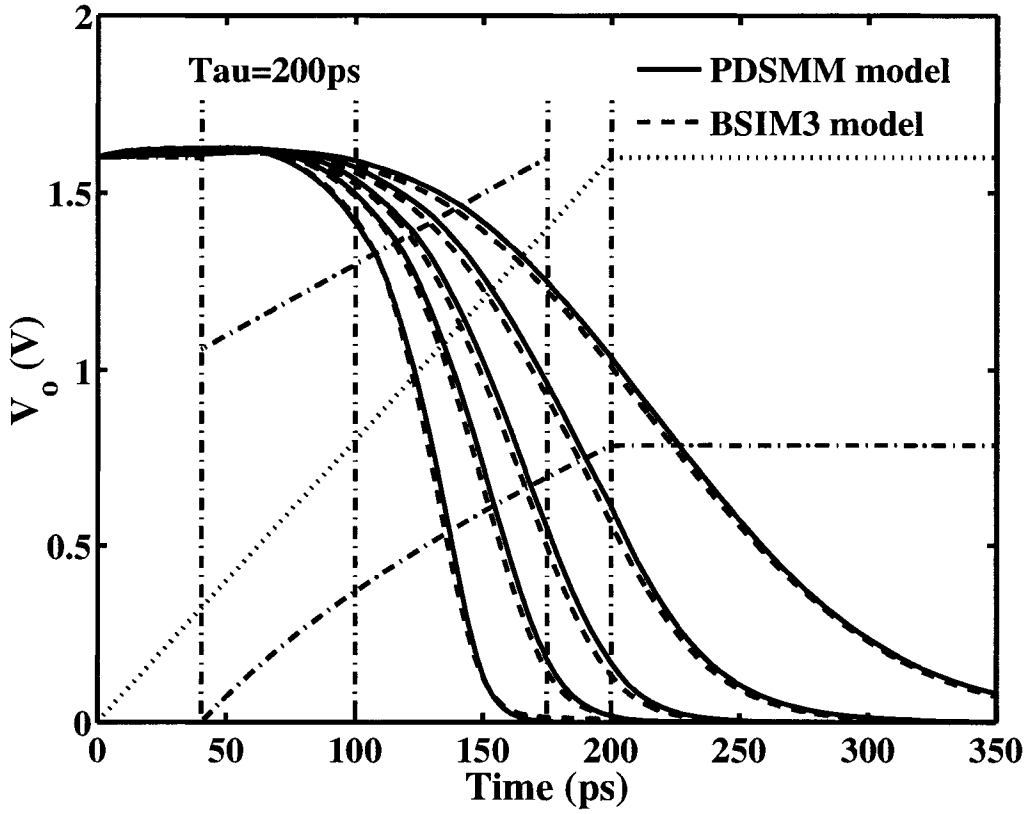


Fig. 4.21 The output voltage waveforms of a CMOS inverter from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. The input transition time is 200ps. The external load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right). They match with each other excellently.

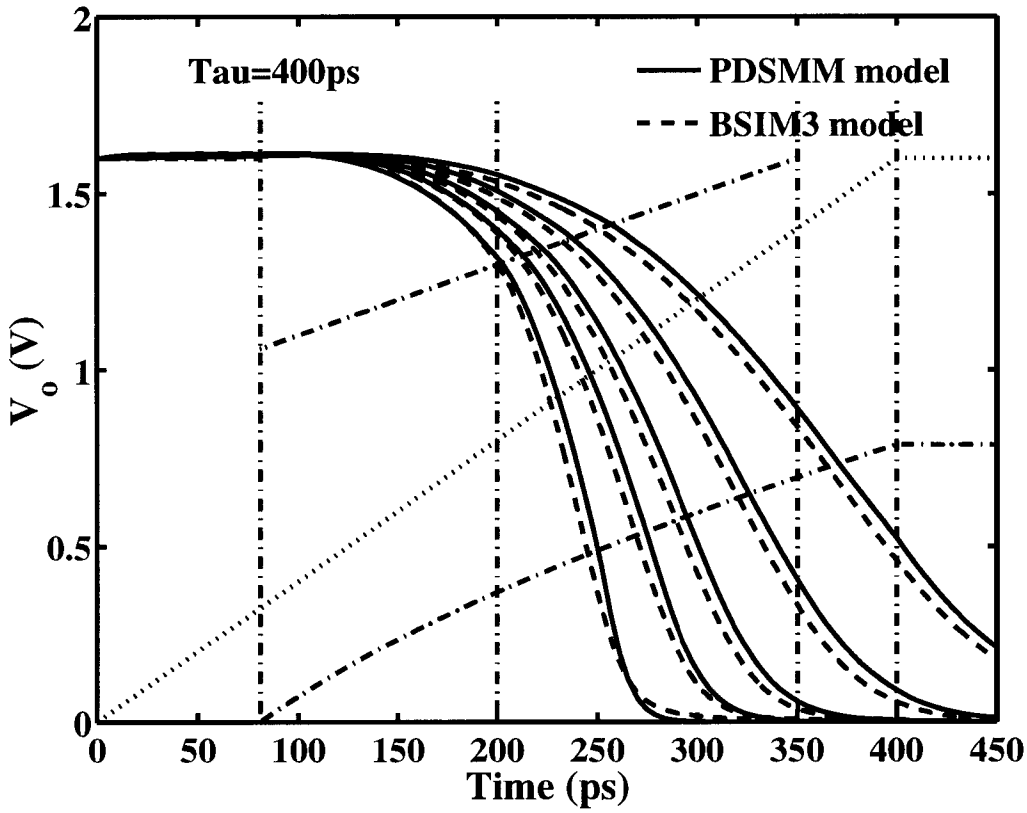


Fig. 4.22 The output voltage waveforms of a CMOS inverter from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. The input transition time is 400ps. The external load capacitances are 10fF, 100fF, 200fF, 400fF, and 800fF, respectively (from left to right).

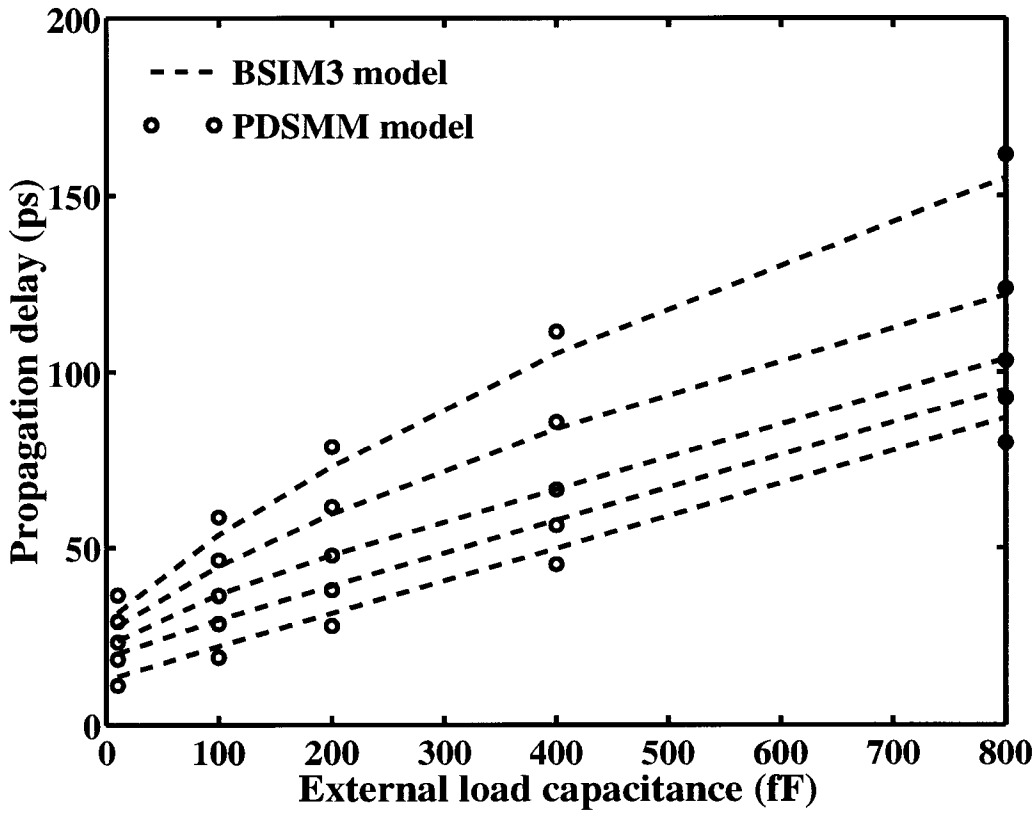


Fig. 4.23 Propagation delay versus external load capacitance. The propagation delays of a CMOS inverter are obtained from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. The input transition times are 1ps, 10ps, 100ps,200ps, and 400ps, respectively (from bottom to top).

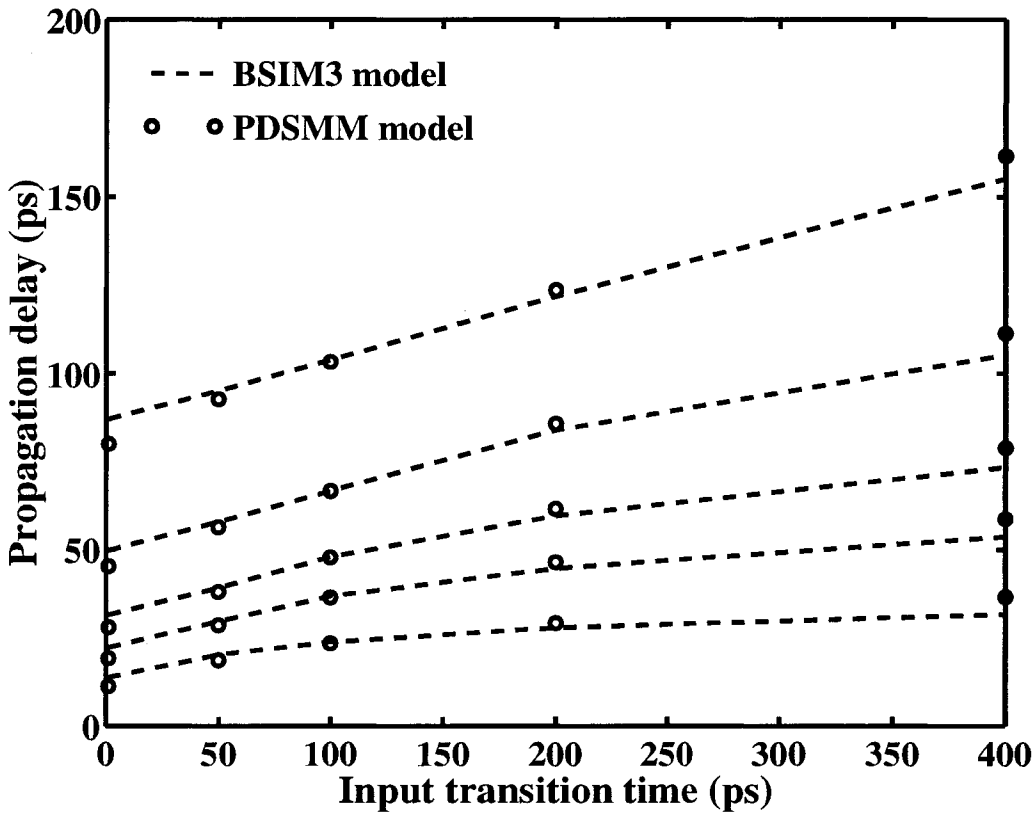


Fig. 4.24 Propagation delay versus external load capacitance. The propagation delays of a CMOS inverter are obtained from SPICE BSIM3 and from the comprehensive timing model based on PDSMM in Subsection 4.5.2. The input transition times are 1ps, 10ps, 100ps, 200ps, and 400ps, respectively (from bottom to top).

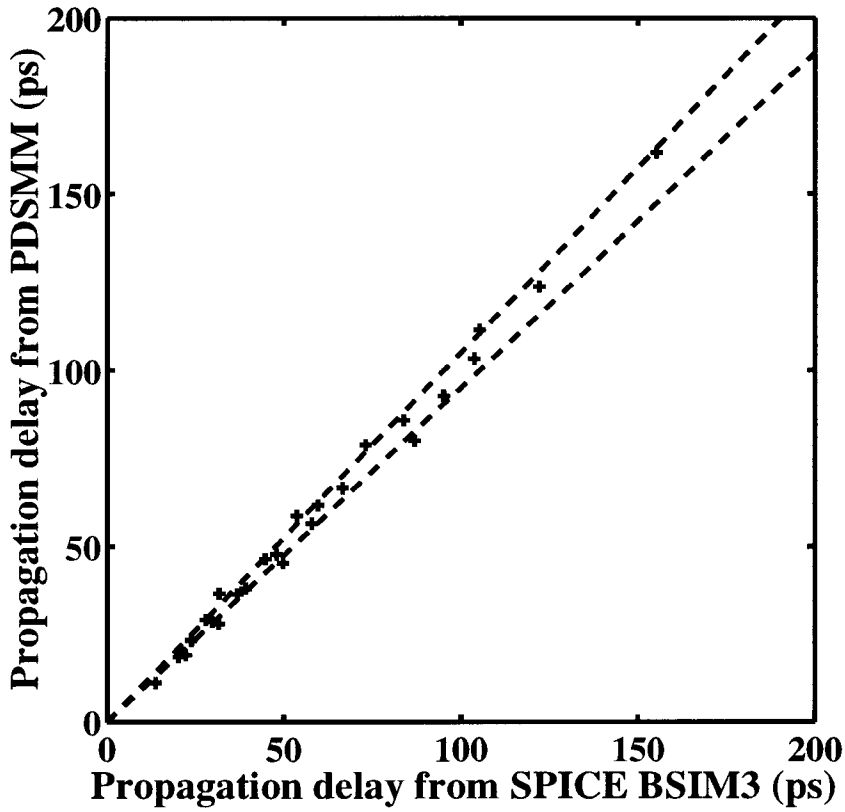


Fig. 4.25 The correlation of propagation delays from the comprehensive timing model based on PDSMM in Subsection 4.5.2 with propagation delays from SPICE BSIM3. The dashed lines indicate ±5% of propagation delays deviating from SPICE BSIM3. All the data match with each other excellently.

CHAPTER 5 A SERIES-CONNECTED MOSFET CHAIN WITH ANEQUAL INPUT PATTERN

Efficient design, optimization and verification of digital integrated circuits require the accurate and fast estimation of propagation delay of a single-stage complex CMOS logic gate such as NAND or NOR gate. The analytical approaches to fast and precisely calculate propagation delay of a deep submicron CMOS inverter are derived in Chapter 4. A single-stage complex CMOS logic gate has been investigated less due to its complexity caused by arbitrary input patterns and multiple intermediate nodes. In order to extend the analytical approaches of propagation delay to a single-stage complex CMOS logic gate, we exclusively examine the characteristic properties of a series-connected MOSFET chain with an equal input pattern in this chapter. A series-connected MOSFET chain is a basic structure in a single-stage complex CMOS logic gate. For simplicity we sometimes refer to a series-connected MOSFET chain as a chain. An equal input pattern means that all the inputs of a chain are either active or inactive. All active inputs are connected to an identical voltage waveform and all inactive inputs are connected to power or ground depending on the type of MOSFETs in a chain. All active inputs determine the sort of a series-connected MOSFET chain with an equal input pattern. The propagation delay of a single-stage complex CMOS logic gate with an equal input pattern will be investigated in the next chapter. The input mapping technique converting an arbitrary input pattern associated with a chain into an equal input pattern will be discussed in Chapter 7.

The behaviors of a series-connected MOSFET chain in a domino CMOS gate were qualitatively described by Shoji [91]. An attempt to study the MOSFET chain was

made, considering a long RC chain and without taking into account any second-order effects. Sakurai and Newton [85] [86] proposed an *nth*-power law MOSFET model to describe the static *I–V* characteristics of a series-connected MOSFET. For some sorts of a series-connected MOSFET chain with an equal input pattern the *nth*-power law MOSFET model is accurate to describe only a portion of their static *I–V* characteristics where input voltage is close to operating voltage. In those cases it is poor to reproduce the other portion of static *I–V* characteristics where input voltage approaches threshold voltage. Generally speaking, the *nth*-power law MOSFET model cannot be accurate to completely depict the whole range of static *I–V* characteristics of all the sorts of a chain with an equal input pattern. They used the *nth*-power law MOSFET model in calculating propagation delay of a logic gate. An input suppressed technique was proposed to illustrate the short-circuit current since a chain without a parallel pMOSFET cluster is used to represent a logic gate. In their derivation the only *nth*-power law MOSFET model was used and a parallel pMOSFET cluster is ignored. They defined a logic transition voltage level as threshold voltage of input suppression. The transient properties of a chain are ignored before input voltage reaches the threshold voltage of input suppression. This method gives rise to three main errors. The first one is caused by the inaccuracy of the *nth*-power law MOSFET model chain with input voltage approaching threshold voltage, especially in a case of large input transition time. The second is the threshold voltage of input suppression cannot truly describe the output transition of a CMOS logic gate for such as a step input. The other one is caused by ignoring of the effect of initial states of parasitic capacitances of intermediate nodes.

Several attempts have been made to collapse a series-connected nMOSFET chain

with an equal input pattern into a MOSFET with two steps, i.e., the chain is reduced into an equivalent circuit consisting of two series-connected MOSFETs and then the two series-connected MOSFETs are converted into a MOSFET [15] [75]. In the first step the MOSFET close to output node of a chain remains unchanged for two series-connected MOSFET equivalent circuit. The other MOSFET in the equivalent circuit is the reduction of the rest of the MOSFETs in the chain. The reduction is based on that all MOSFETs in the chain except the one close to output node operate always in the linear mode during output transition [16]. The MOSFET close to output node initially operates in saturation mode and then enters linear mode. This implies that there are initially no charges in parasitic capacitances of intermediate nodes. In practice, this is not always valid. The arbitrary initial states of parasitic capacitances give rise to distinct operation mode for each MOSFET during transition. As we will see in this chapter, the ignorance of the effect of initial states will result in significant errors.

In this chapter, the naming conventions of the sort of a series-connected MOSFET chain with an equal input pattern are briefly introduced in Section 5.1. The peculiar static I–V characteristics of a series-connected MOSFET chain with an equal input pattern are investigated in Section 5.2. The power index of a series-connected MOSFET chain is examined in Section 5.3. After the investigation of the peculiar static I–V characteristics of a series-connected MOSFET chain with an equal input pattern, a new and accurate generic series-connected MOSFET chain model (SMCM) is proposed in Section 5.4. An input suppression technique is modified to simulate the dependence of propagation delays on initial states of parasitic capacitances of multiple intermediate nodes with an equal input pattern in Section 5.5. Two extreme initial states of a series-connected MOSFET

chain causing fast and slow initial state propagation delays (ISPDs) are introduced in Section 5.6. In Section 5.7, we discuss the step-input activation time of a series-connected MOSFET chain from SPICE simulations. In Section 5.8, we finally analyze the ramp-input activation time of a series-connected MOSFET chain with an equal input pattern. The comprehensive analysis of propagation delay based on the SMCM and the modified input suppressed technique is built in the next chapter.

5.1 NAMING CONVENTIONS

The type of a series-connect MOSFET chain depends on the type of MOSFETs in the chain. We focus on a series-connected nMOSFET chain in this research. The results of a series-connected nMOSFET chain can be applied to a series-connected pMOSFET chain symmetrically. A four-nMOSFET series-connected chain as in a NAND4 logic gate shown in Fig. 5.1 is used to study the static I–V characteristics of a series-connected nMOSFET chain. For simplicity and consistency as shown in Fig. 5.1, the numbering conventions of both nMOSFETs and their corresponding inputs are as follows. The numbering of nMOSFETs in a series-connected chain starts from the nMOSFET connected to ground as number one. The close to output node an nMOSFET in a chain locates, the bigger number it has. An input has the same number as its corresponding MOSFET. Take the four-nMOSFET series-connected chain shown in Fig. 5.1 as an example. The nMOSFET in the series-connected nMOSFET chain, which is connected to ground, is number one. The nMOSFET connected to output node is number four.

The series-connected nMOSFET chain with all its inputs connected either to the identical input pattern or to power is referred to as a series-connected nMOSFET chain

with an equal input pattern. In this chapter we are interested in the sort of a series-connected nMOSFET chain with an equal input pattern. With the numbering conventions of both nMOSFETs and their corresponding inputs, the sort of a series-connected MOSFET chain with an equal input pattern is named as a combination of three main parts. The first one is either letter 'n' or 'p', indicating the type of the series-connected MOSFET chain, namely, the letter 'n' is a series-connected nMOSFET chain and the letter 'p' means a series-connected pMOSFET chain. The second is a single digit denoting the total number of MOSFETs in a chain. The last part has one or more digits representing the active inputs connected to an identical input voltage. The unlisted inputs in the last part are inactive. For example, 'n412' means that it is a four-nMOSFET series-connected chain with an equal input pattern applied to input one and input two. The other two remaining inputs of 'n412' such as input three and input four are connected to power.

5.2 STATIC *I–V* CHARACTERISTICS

Figs. 5.2-5.5 show the static *I–V* characteristics of 'n41', 'n42', 'n43', and 'n44' of a four-nMOSFET series-connected chain with an equal input pattern whose voltage is applied to the individual input from number one to number four, respectively. Fig. 5.6 plots the static *I–V* characteristics of 'n41234' with an equal input pattern applied to all four inputs. As we can see from these figures each sort of a series-connected MOSFET chain with an equal input pattern has its unique static *I–V* characteristics. For example, Fig. 5.2 shows that for saturation mode the square law of chain current versus input voltage is valid when input voltage is small and close to threshold voltage. When input voltage is close to operating voltage the square law of chain current versus input voltage

in saturation mode fails due to the effect of chain current saturation with input voltage.

The unique static I–V characteristics of a series-connected MOSFET chain with an equal input pattern are still composed of three modes – cutoff, linear, and saturation, which are similar with those of a deep submicron MOSFET. When input voltage is smaller than threshold voltage, series-connected MOSFET chain is in cutoff mode. There is no current through the chain. There exists a distinct threshold voltage for each sort of a series-connected MOSFET chain with an equal input pattern. When input voltage exceeds threshold voltage, there is a current through the chain. The static I–V characteristic curve for each input voltage is similar to that of a deep submicron MOSFET except for output saturation voltage. When input voltage exceeds threshold voltage and output voltage is greater than output saturation voltage, the series-connected MOSFET chain enters saturation mode. For a given input voltage the relationship of chain current and output voltage in saturation mode is similar with the effect of channel length modulation of a deep submicron MOSFET. When input voltage exceeds threshold voltage and output voltage is smaller than output saturation voltage, series-connected MOSFET chain is in linear mode. The chain current in linear mode changes dramatically with output voltage, which is the similar with the deep submicron MOSFET current in linear mode.

The common property obtained from Figs 5.2-6 is that the dependence of chain current on output voltage for each input voltage is similar with the dependence of the drain-source current on the drain-source voltage for each gate-source voltage of a deep submicron nMOSFET. In other words, the chain current I_n still shows the square-law dependence on the output saturation voltage V_{sn} for each sort of a series-connected nMOSFET chain with an equal input pattern. The output saturation voltage V_{sn}

depending on input voltage, definitely different from PDSMM or Shockley MOSFET model, needs to be defined for each sort of a series-connected nMOSFET chain with an equal input pattern accordingly. Therefore, the formula of the PDSMM can be extended to accurately depict the peculiar static I–V characteristics of a series-connected nMOSFET chain with an equal input pattern by modifying output saturation voltage V_{sn}, which will be discussed in Section 5.4.

5.3 POWER INDEX

In this section let's investigate the relationship between chain current and input voltage of a series-connected MOSFET chain with an equal input pattern in saturation mode. The power index in the relationship between chain current and input voltage in saturation mode is called the power index of a series-connected MOSFET chain. From Fig. 5.2-6, we can clearly see that the static I–V characteristics of 'n41', 'n42', 'n43', 'n44', and 'n41234' are quite different from each other. In order to investigate the static I–V characteristics, the dependence of chain current I_n on input voltage at $V_o = V_{dd}$ for 'n41', 'n42', 'n43', and 'n44' is plotted in Fig. 5.7. As a comparison, the dependence of drain-source current I_{ds} on gate-source input voltage V_{gs} of a deep submicron nMOSFET with one-quarter of the nMOSFET width in the series-connect chain is also shown in Fig. 5.7. Fig 5.8 illustrates power index as a function of input voltage from (3.14) for 'n41', 'n42', 'n43', 'n44', and 'n11', respectively (from bottom to top).

As we know from Chapter 3 the power index of a long channel MOSFET is two. For a deep submicron MOSFET of 'n11' the power index shown in Fig. 5.8 is close to two at small gate-source voltage V_{gs} and gradually decreases as the gate-source input

voltage V_{gs} increases, which can be predicted by the PDSMM in Chapter 3. The reduction of power index is caused by carrier velocity saturation of a deep submicron MOSFET so that the power index is also called the velocity saturation index. As we know from the PDSMM in Chapter 3, the limiting power index or velocity saturation index of a deep submicron nMOSFET is one. Power index of a deep submicron MOSFET of 'n11' ranges between one and two.

Power indices of a series-connected MOSFET chain with an equal input pattern such as 'n41' are close to two when input voltage is around threshold voltage. They gradually decrease when input voltage increases. Their values are even lower than one at input voltage close to operating voltage. The power index lower than one is introduced with the sort of a series-connected MOSFET chain with an equal input pattern, i.e. the set of active inputs. The power index of a series-connected MOSFET chain with an equal input pattern is an input-dependent index. In addition to the second-order effects of a deep submicron MOSFET such as charge carrier velocity saturation and channel length modulation, the power index of a series-connected MOSFET with an equal input pattern has to be incorporated in the propagation delay model.

5.4 GENERIC SERIES-CONNECTED MOSFET CHAIN MODEL (SMCM)

Obviously all the existing MOSFET models have been developed to describe static *I–V* characteristics of a MOSFET. They cannot accurately or completely reproduce the peculiar static *I–V* characteristics of a series-connected MOSFET chain shown in Figs. 5.2-7. A new generic chain model is needed to accurately describe the static *I–V* characteristics of a series-connected MOSFET chain. This chain model should be

sufficiently simple and accurate to be used in the approaches of propagation delay. As we discussed in Section 5.2, the operation of a series-connected MOSFET mode with an equal input pattern is divided into three modes, i.e., cutoff mode, linear mode, and saturation mode. All the equations of proper deep submicron MOSFET model (PDSMM) for cutoff mode, linear mode, and saturation mode are still valid except for the expression of output saturation voltage V_{sn}. The square-law dependence of chain current I_n on the output saturation voltage V_{sn} at the boundary between linear mode and saturation mode exists. The power index of the chain current I_n on the input voltage V_i at the output saturation for small input voltage V_i is two due to the ignorance of the carrier velocity saturation. The power index of the chain current I_n on the input voltage V_i may reduce to less than one due to the effect of input positioning when the input voltage V_i increases. Therefore, the relationship of the output saturation voltage V_{sn} on the input voltage V_i of a series-connected MOSFET chain deviates from the drain-source saturation voltage V_{sat} of (3.8) in the PDSMM whose minimum power index is one. The generic series-connected MOSFET chain model (SMCM) is proposed to solve the issue. Let us discuss a series-connected nMOSFET chain in details. All the derivations can be applied to a series-connected pMOSFET chain symmetrically. The SMCM takes the same equation format of PDSMM model except output saturation voltage V_{sn}. The chain current I_n of SMCM model is expressed as:

$$I_n = \begin{cases} 0 & V_i \le V_{tn} \\ k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o(1+\lambda_n V_{sn}) & V_i > V_{tn}, V_o \le V_{sn} \\ \dfrac{1}{2}k_n {V_{sn}}^2(1+\lambda_n V_o) & V_i > V_{tn}, V_o > V_{sn} \end{cases} \tag{5.1}$$

where k_n is the transconductance parameter of a series-connected MOSFET chain with an equal input pattern, V_{th} the threshold voltage, and λ_n the channel length modulation coefficient. The output saturation voltage V_{sn} of a series nMOSFET chain in SMCM, however, is quite different from the PDSMM. According to the analysis of static I–V characteristics of a series-connected nMOSFET chain with an equal input pattern in Section 5.3, the output saturation voltage V_{sn} is linear with input voltage V_i when it is small. The output saturation voltage V_{sn} has a power-law relation with input voltage V_i when it is large. For convenience, the joint point between these two pieces is set to $V_{dd}/2$. Thus, the output saturation voltage V_{sn} of a series-connected nMOSFET chain with an equal input pattern in SMCM is

$$V_{sn} = \begin{cases} Q_n(V_i - V_{tn}) & V_{tn} < V_i \le \dfrac{V_{dd}}{2} \\ R_n(V_i - V_{an})^{b_n} & \dfrac{V_{dd}}{2} < V_i \le V_{dd} \end{cases} \tag{5.2}$$

In (5.2), V_{tn} is the threshold voltage of the series-connected nMOSFET chain; Q_n, R_n and V_{an} are the SMCM model parameters. Index b_n is used to describe the effect of power index of the series-connected nMOSFET chain with an equal input pattern. The output saturation voltage V_{sn} of a series-connected nMOSFET chain is continuous when input voltage V_i is approaching $V_{dd}/2$. This gives rise to the following constraint

$$Q_n(V_{dd}/2 - V_{tn}) = R_n(V_{dd}/2 - V_{an})^{b_n} \tag{5.3}$$

Figs. 5.9-13 plot the static *I–V* characteristics of a series-connected nMOSFET chain for 'n41', 'n42', 'n43', 'n44', and 'n41234' calculated from SMCM whose parameters are extracted and listed in Table 5.1. The static *I–V* characteristics obtained from SPICE simulation by using BSIM3 model are also illustrated as a comparison. Figs. 5.9-13 show that proposed SMCM is very accurate to describe the static *I–V* characteristics of a series-connected nMOSFET chain with an equal input pattern.

5.5 MODIFIED INPUT SUPPRESSION TECHNIQUE

Unlike a CMOS inverter, there exist intermediate nodes of a series-connected MOSFET chain in a single-stage complex CMOS logic gate. The effect of initial states of parasitic capacitances of intermediate nodes on propagation delay is complicate and difficult to deal with. The large errors of propagation delays may be caused if the effect of initial states of parasitic capacitances of intermediate nodes is ignored.

One way to model the effect of initial states of parasitic capacitances of intermediate nodes is to make use of input suppression technique. The method to extract the suppressed input from input voltage proposed by Sakurai and Newton in [85] is discussed briefly below. They defined logic threshold voltage of a digital gate as threshold voltage of input suppression. Before input voltage is less than threshold voltage of input suppression, input voltage is suppressed in input suppression technique used in the propagation delay model. Once input voltage is greater than threshold voltage of input suppression, input voltage is bypassed to the suppressed input voltage. Thus there are two portions in the input suppression technique where the input waveform is suppressed in the

first part and is bypassed to the suppressed input voltage in the second part. The input suppression technique makes some corrections in propagation delay with a specific range of input transition time. Obviously, input suppression technique makes no corrections in propagation delay for a step input. Therefore, it does not effectively correct errors in propagation delay of a series-connected MOSFET chain with an equal input pattern.

In this paper, the idea of input suppression technique is modified to describe the effect of initial states of parasitic capacitances of intermediate nodes on propagation delay for single-stage CMOS logic gates with a series-connected MOSFET chain. The new and modified way to extract the suppressed input is exploited to solve the issues mentioned above. We propose the activation time t_a for each sort of a series-connected MOSFET chain with an equal input pattern, instead of the threshold voltage of input suppression, to extract the suppressed input. Before the activation time t_a is reached, the input waveform V_i is suppressed in the suppressed input. Charges stored in parasitic capacitances of intermediate nodes are discharged and adjusted via the series-connected MOSFET chain before the activation time t_a while charges in output load capacitances are not discharged at the same time. After the activation time t_a is reached, the suppressed input is equal to the real input voltage and the charges in output load capacitances start to discharge via the series-connected MOSFET chain. That is why it is called as the activation time t_a of a series-connected MOSFET chain. There are two pieces in the modified input suppression technique depending on whether the activation time t_a is reached or not. In the modified input suppression technique, the input voltage V_i is suppressed before the activation time t_a and is bypassed to the suppressed input after the activation time t_a. Fig. 5.14 shows

the modified input suppression technique from a ramp input. Therefore, the activation time t_a of a series-connected MOSFET chain with an equal input pattern is the key parameter in the modified input suppression technique.

The charges discharged from parasitic capacitances of intermediate nodes before the activation time t_a depends on the sort of a series-connected MOSFET chain with an equal input pattern and initial states of parasitic capacitances. The analysis of the modified input suppression technique is based on that the charges discharged from parasitic capacitances of intermediate nodes are independent of input transition time. In other words, the charges discharged from parasitic capacitances of intermediate nodes before the activation time t_a keep the same for different input transition time provided that the sort of a series-connected MOSFET chain with an equal input pattern is the same. The distinct input transition time gives rise to the corresponding activation times t_a. The activation time t_{as} for the sort of a series-connected MOSFET chain with an equal step input is obtained from SPICE simulation. By using SMCM the activation time t_a for the sort of a series-connected MOSFET chain with an equal ramp input can be derived. We will investigate the activation time t_{as} for an equal step input in Section 5.7 and then extend it to the activation time t_a for an equal ramp input by using SMCM in Section 5.8.

5.6 TWO EXTREME INITIAL STATES – SLOW AND FAST

Before we further discuss the activation time t_a of a series-connected MOSFET chain, let us examine the effect of initial states of parasitic capacitances of intermediate nodes on propagation delay. Fewer researches devoted to it have been conducted due to complexity and uncertainty of initial states. An initial state of parasitic capacitances of

intermediate nodes is caused by the previous input pattern to deactivate the series-connected MOSFET chain. It is impossible to completely describe arbitrary initial states of internal parasitic capacitances in a series-connected MOSFET chain. In this research, propagation delay when we emphasize on initial state is called initial state propagation delay (ISPD). In practical point of view, two possible extreme initial states called the fast initial state (FIS) and slow initial state (SIS) are the most important. FIS and SIS result in fast initial state propagation delay (FISPD) and slow initial state propagation delay (SISPD), accordingly. Logic designers of digital circuits are interested in FIS and SIS. The FISPD is used in the verification and optimization of minimum timing issues associated with hold time violations. The SISPD is employed in the verification and optimization of maximum timing issues associated with setup time violations.

Fig. 5.15 shows an input pattern causing the FISPD of a CMOS NAND4 gate. With this input pattern applied to inputs of a CMOS NAND4 gate, after the fourth nMOSFET close to output node is first turned off, the other three nMOSFETs still discharge the parasitic capacitances of intermediate nodes due to their inputs connected to operating voltage. Then the third nMOSFET, the second nMOSFET, and the first nMOSFET are turned off in such an order that makes the charges stored in the parasitic capacitances of intermediate nodes minimized. The minimum initial charge of parasitic capacitances of intermediate nodes will lead to fast output discharge, i.e., FISPD. Thus, the input pattern shown in Fig. 5.15 gives rise to the fast initial state (FIS). Similarly the input pattern causing the SIS is illustrated in Fig. 5.16.

Let's investigate the extreme initial states (FIS and SIS) of a series-connected MOSFET chain with an equal input pattern. Take 'n413' as an example. We know that it

is a four-nMOSFET series-connected chain with an equal input pattern applied to inputs one and three. The active inputs are the first and the third one. The inactive inputs are the second and fourth one. The inactive inputs turn on far before the equal input transition and are assumed to connect to power rail. The FIS is caused by the input pattern to deactivate the series-connected chain shown in Fig. 5.15. The $MOSFET_4$ turns off first and $MOSFET_1$ turns off last. The parasitic capacitances of effective intermediate nodes of 'n413' between the first and the third nMOSFETs, associated with the first active input and the last active input, are then completely discharged. The minimum initial charges of effective intermediate nodes between the first and the third nMOSFETs are the FIS of 'n413', which leads to fast output discharge, i.e., FISPD. The SIS of 'n413', on the other hand, is caused by the opposite sequence of inputs to deactivate the series-connected MOSFET chain shown in Fig. 5.16. The first nMOSFET turns off first and the fourth one turns off last. The parasitic capacitances of effective intermediate nodes between the first and the third nMOSFETs are fully charged. The maximum initial charges of effective internal intermediate nodes between the first and the third nMOSFETs are the SIS of 'n413', which gives rise to the slow output discharge, i.e., SISPD.

The output waveforms of 'n413' for an equal input pattern corresponding to FISPD and SISPD are plotted in Fig. 5.17. The difference of FIS and SIS are caused with the two extreme states of the effective parasitic capacitances between the first and the last nMOSFETs of the active inputs. They result in the different output waveforms. The dotted ramp line indicates the equal input pattern of 'n413'. As we see from Fig. 5.17, the effect of initial states cannot be ignored in discussing propagation delays. It is necessary to deal with FISPD and SISPD separately.

For input patterns like 'n42', only one active input swings. The other inputs are inactive and connected to power rail. In this case, there are no effective intermediate nodes between the first and last nMOSFETs associated with the first active input and the last active input. Thus there is only one initial state, i.e., SIS and FIS are the same. Both of initial states can be described as follows. The parasitic capacitances of intermediate nodes between the transitioning input and ground are fully discharged. The parasitic capacitances of intermediate nodes between the transitioning input and the output node are fully charged.

5.7 STEP-INPUT ACTIVATION TIME t_{as}

The modified input suppression technique is adopted to depict the effect of charges discharged from parasitic capacitances of intermediate nodes in a series-connected MOSFET chain. The key parameter in the modified input suppression technique is the activation time t_a of the sort of a series-connected MOSFET chain with an equal input pattern having input transition time τ. We first focus on the step input activation time t_{as} in this section. From SPICE characterizations, we find that the step input activation time t_{as} depends on the sort of a series-connected MOSFET chain with an equal input pattern and the initial state of parasitic capacitances of intermediate nodes. Fig. 5.1 is the schematic circuit to measure the step-input activation time t_{as}. The time between the rising moment of a step input voltage and half of the stable value of chain current I_n through the net close to the output of the series-connected MOSFET chain is quantitatively defined as the step-input activation time t_{as}. The step input voltage and the chain current I_n through the net close to the output of the series-connected MOSFET

chain of NAND413 for fast and slow ISDPs from SPICE are plot in Fig. 5.18. Table 5.2 lists the complete step-input activation times t_{as} for all the sorts of nMOSFET chain and pMOSFET chain with 1 to 4 MOSFETs. The fast and slow cases for each sort of a series-connected MOSFET chain with an equal input pattern are caused by the initial states of parasitic capacitances of intermediate nodes. The same values of fast case and slow case for some sorts of a series-connected MOSFET chain with an equal input pattern mean only one active input associated with a single initial state of parasitic capacitances of intermediate nodes.

Assume that the SMCM is suitable to describe the chain current I_n discharged from parasitic capacitances of intermediate nodes of a series-connected MOSFET chain before step-input activation time t_{as}. The charge Q_{as} discharged from parasitic capacitances of intermediate nodes before the step-input activation time t_{as} of a series-connected MOSFET chain for an equal step input is as follows:

$$Q_{as} = \int_0^{t_{as}} I_n\big|_{V_i(t)=V_{dd}u(t)}\, dt \tag{5.4}$$

Substituting the chain current I_n from the SMCM into the above equation yields

$$Q_{as} = \frac{1}{2} k_n R_n^2 \left(V_{dd} - V_{an}\right)^{2b_n} \left(1 + \lambda_n V_{dd}\right) t_{as} \tag{5.5}$$

5.8 RAMP-INPUT ACTIVATION TIME t_a

In order to calculate the activation time t_a of a series-connected MOSFET chain for an equal input pattern, we assume that the initial charge of parasitic capacitances of intermediate nodes depends only on the initial state of the series-connected MOSFET

chain. It is independent of input transition time. Therefore, the value of charge Q_a discharged from parasitic capacitances of intermediate nodes before the activation time t_a of a series-connected MOSFET chain for a ramp input is same as the value of charge Q_{as} discharged from parasitic capacitances of intermediate nodes before the step-input activation time t_{as} of the series-connected MOSFET chain for a step input. We have

$$Q_a = \int_0^{t_a} I_n dt = Q_{as} \tag{5.6}$$

Substituting (5.5) of charge Q_{as} discharged from parasitic capacitances before the step input activation time t_{as} into the above equation gives rise to

$$Q_a = \int_{t_m}^{t_a} I_n dt = \frac{1}{2} k_n R_n^2 (V_{dd} - V_{an})^{2b_n} (1 + \lambda V_{dd}) t_{as} \tag{5.7}$$

Note that the lower boundary of integral of (5.7) is changed from zero to t_m. Before the time $t_m = u_m \tau = \frac{V_m}{V_{dd}} \tau$, the ramp input voltage is less than the threshold voltage of the SMCM of the series-connected MOSFET chain. The series-connected MOSFET chain is cutoff and there is no chain current I_n through the series-connected MOSFET chain. Since the chain current I_n in SMCM is piecewise, the following analysis on the ramp-input activation time t_a is separated into three cases accordingly.

5.8.1 Large Input Transition Time: $(\tau \geq \tau_u)$

This is the case of the large input transition time τ equal to or exceeding the upper bound τ_u. The ramp-input activation time t_a of modified input suppression technique is equal to or smaller than half of the input transition time τ in this case

$(t_a \le \tau/2)$. In other words, the activation time t_a of modified input suppression technique is equal to half of the input transition time τ_u $(t_a = \tau_u/2)$ when the input transition time τ equal to the upper bound τ_u. Then we have

$$Q_a = \int_{t_{tn}}^{t_a} I_n dt = \int_{t_{tn}}^{t_a'=\tau_u/2} I_n' dt = \frac{1}{2} k_n R_n^2 (V_{dd} - V_{an})^{2b_n} (1 + \lambda V_{dd}) t_{as} \tag{5.8}$$

Substituting the chain current I_n from SMCM into (5.8) yields

$$t_a = u_{tn}\tau + \left(\frac{\tau_u}{\tau}\right)^{1/3} \left(\frac{1}{2} - u_{tn}\right)\tau \tag{5.9}$$

where τ_u is the upper bound of input transition time, one of two special bounds for a series-connected MOSFET chain. The upper bound τ_u of input transition time is expressed as

$$\tau_u = \frac{3}{\left(\frac{1}{2} - u_{tn}\right)} \left(\frac{1 - u_{an}}{\frac{1}{2} - u_{an}}\right)^{2b_n} t_{as} \tag{5.10}$$

Clearly, when the input transition time τ is equal to or larger than the upper bound τ_u of input transition time, the activation time t_a of a series-connected MOSFET chain for a ramp input with input transition time τ from the above equation is equal to or less than half of the input transition time τ $(t_a \le \tau/2)$.

5.8.2 Medium Input Transition Time: $(\tau_l \le \tau < \tau_u)$

This is the case of the intermediate input transition time τ between the lower bound τ_l and upper bound τ_u. The activation time t_a of modified input suppression technique is between the half of the input transition time and the input transition time

$(\tau/2 \le t_a < \tau)$. We rewrite the equation due to the piecewise property of SMCM as follows:

$$\int_{t_{tn}}^{\tau/2} I_n dt + \int_{\tau/2}^{t_a} I_n dt = \frac{1}{2} k_n R_n^2 (V_{dd} - V_{an})^{2b_n} (1 + \lambda V_{dd}) t_{as} \tag{5.11}$$

Replacing the piecewise current expressions of SMCM into the above equation yields

$$t_a = u_{an}\tau + \left(\frac{1}{2} - u_{an}\right)\left[1 + \left(\frac{2b_n + 1}{3}\right)\left(\frac{\frac{1}{2} - u_{tn}}{\frac{1}{2} - u_{an}}\right)\left(\frac{\tau_u - \tau}{\tau}\right)\right]^{\frac{1}{2b_n+1}} \tau \tag{5.12}$$

where τ_u is the upper bound of input transition time for the series-connected MOSFET chain defined as before.

The lower bound τ_l of input transition time for the series-connected MOSFET chain is obtained on the condition of $t_a = \tau_l$ and $\tau = \tau_l$. The lower bound τ_l of input transition time is expressed as

$$\tau_l = \frac{1}{\left(\frac{1-u_{an}}{2b_n+1}\right) + \left[\frac{1}{3}\left(\frac{1}{2} - u_{tn}\right) - \left(\frac{\frac{1}{2} - u_{an}}{2b_n+1}\right)\right]\left(\frac{\frac{1}{2} - u_{an}}{1-u_{an}}\right)^{2b_n}} t_{as} \tag{5.13}$$

5.8.3 Small Input Transition Time: $(\tau < \tau_l)$

This is the case of the short input transition time τ below the lower bound τ_l in which the activation time t_a of modified input suppression technique is larger than the input transition time τ $(t_a \ge \tau)$. We rewrite the equation due to the piecewise property of SMCM as follows:

$$\int_{t_{tn}}^{\tau/2} I_n dt + \int_{\tau/2}^{\tau} I_n dt + \int_{\tau}^{t_a} I_n dt = \frac{1}{2} k_n R_n^2 (V_{dd} - V_{an})^{2b_n} (1 + \lambda V_{dd}) t_{as} \tag{5.14}$$

Replacing the piecewise current expressions of SMCM into the above equation yields

$$t_a = t_{as} + \left(1 - \frac{t_{as}}{\tau_l}\right)\tau \tag{5.15}$$

where τ_l is the lower bound of input transition time for the series-connected MOSFET chain defined as before.

5.8.4 Discussions

From the above analysis, the ramp-input activation time t_a of the modified input suppression technique for a series-connected MOSFET chain is obtained. The effect of initial charges stored in parasitic capacitances of intermediate nodes of a series-connected MOSFET chain is modeled by the modified input suppression technique associated with the ramp-input activation time t_a. The charge conservation and SMCM are assumed in the derivation. The suppressed voltage is applied to the active inputs. The activation time of the fast initial state (FIS) is determined with initial charges of parasitic capacitances of intermediate nodes between the output node and the active nMOSFET closest to output node. The activation time of the slow initial state (SIS) is decided with initial charges of parasitic capacitances of intermediate nodes between the output node and the active nMOSFET closest to ground. If there is only one active input, the activation times for the FIS and the SIS are the same. If there is no intermediate node between the output node and the active nMOSFET closest to output node, the activation time of FIS is zero. If there is no intermediate node between the output node and the active nMOSFET closest to ground such as 'n44', the activation times of both FIS and SIS are zero.

Figs. 5.19-22 show the calculated activation times of 'n413', 'n41234', 'n41', and 'n44' for FIS and SIS as a function of input transition time. Given step-input activation times for SIS and FIS from SPICE simulation, the ramp-input activation times are calculated for SIS and FIS. The difference of the activation times of 'n413' between SIS and FIS shown in Fig. 5.19 is caused by initial charges of parasitic capacitances of intermediate nodes between $nMOSFET_1$ and $nMOSFET_3$. The difference between FIS and the threshold line is due to initial charges of parasitic capacitances of intermediate nodes between $nMOSFET_3$ and output node. In Fig.5.20 there are three active intermediate nodes involved in SIS of 'n41234' while there is no active intermediate node for FIS. The difference of that activation times between SIS and FIS is significantly large. The parasitic capacitances of intermediate nodes of 'n41234' in FIS are initially fully discharged and its activation time is much shorter. It is equal to the time when input voltage reaches its threshold voltage. There is only one active input in 'n41' or 'n44', the activation times of slow ISPD and fast ISPD merges into one curve shown in Figs. 5.21-5.22. There are three charged intermediate nodes between the $nMOSFET_1$ and output node for 'n41'. It gives rise to the significantly big activation time for SIS. There are no charged intermediate nodes between the $nMOSFET_4$ and output node for 'n44'. The merged activation time of 'n44' for SIS and FIS is equal to the time when input voltage reaches its threshold voltage.

Table 5.1 The SMCM Model Parameters for All the Sorts of a Four-MOSFET Series-Connected Chain with an Equal Input Pattern Used in Chapter 5

Chain sort	k_n	R_n	V_a	b_n	V_t	λ_n
n41	29.9330	0.7494	0.5871	0.1094	0.2997	0.1047
n412	30.3200	0.8069	0.5669	0.2645	0.3158	0.0891
n4123	31.2990	0.8398	0.5309	0.4404	0.3151	0.0766
n41234	32.5481	0.8202	0.4562	0.6704	0.3252	0.1102
n4124	32.9243	0.7980	0.4887	0.5298	0.3162	0.1234
n413	31.9233	0.8045	0.5416	0.3432	0.3048	0.0784
n4134	33.2255	0.7968	0.4717	0.5732	0.3152	0.1156
n414	33.7873	0.7631	0.4963	0.4369	0.2989	0.1355
n42	31.2512	0.7616	0.5726	0.1768	0.2793	0.0933
n423	32.1125	0.8075	0.5347	0.3694	0.3017	0.0766
n4234	33.2834	0.7967	0.4626	0.5977	0.3147	0.1132
n424	33.8789	0.7663	0.4888	0.4619	0.2973	0.1302
n43	33.1339	0.7633	0.5416	0.2677	0.2620	0.0814
n434	34.1893	0.7667	0.4720	0.5051	0.2957	0.1211
n44	35.1143	0.7233	0.4859	0.3725	0.2474	0.1486

Table 5.2 The Step Input Activation Time t_{as} for a Four-MOSFET Series-Connected Chain with an Equal Input Pattern

nMOSFET chain	Fast case (ps)	Slow case (ps)	pMOSFET chain	Fast case (ps)	Slow case (ps)
n11	0	0	p11	1	1
n21	4	4	p21	9	9
n212	0	3	p212	0	8
n22	0	0	p22	0	0
n31	9	9	p31	23	23
n312	7	10	p312	18	25
n3123	0	10	p3123	0	23
n313	0	9	p313	0	22
n32	6	6	p32	16	16
n323	0	5	p323	0	14
n33	0	0	p33	0	0
n41	17	18	p41	42	42
n412	15	19	p412	37	43
n4123	11	20	p4123	25	45
n41234	0	19	p41234	0	42
n4124	0	18	p4124	0	41
n413	10	20	p413	24	44
n4134	0	18	p4134	0	41
n414	1	18	p414	1	39
n42	14	15	p42	35	36
n423	10	15	p423	23	37
n4234	0	15	p4234	0	34
n424	0	15	p424	0	34
n43	7	7	p43	16	16
n434	0	6	p434	0	13
n44	0	0	p44	0	0

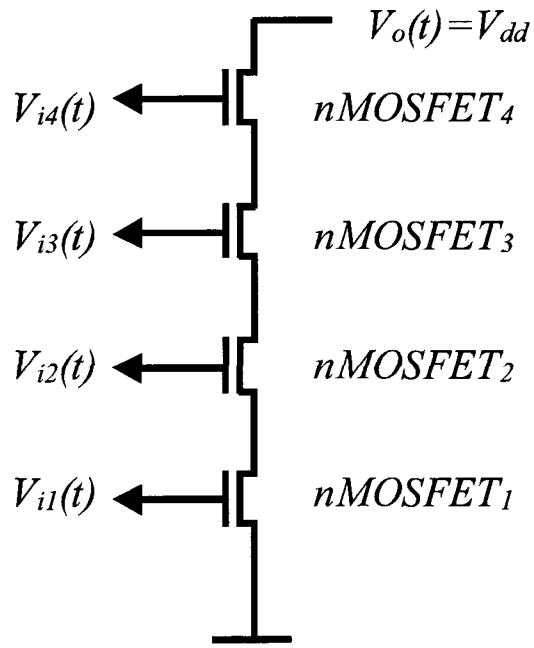


Fig. 5.1 A four-nMOSFET series-connected chain. The numbering of nMOSFET and its corresponding input starts from the nMOSFET connected to ground as number one.

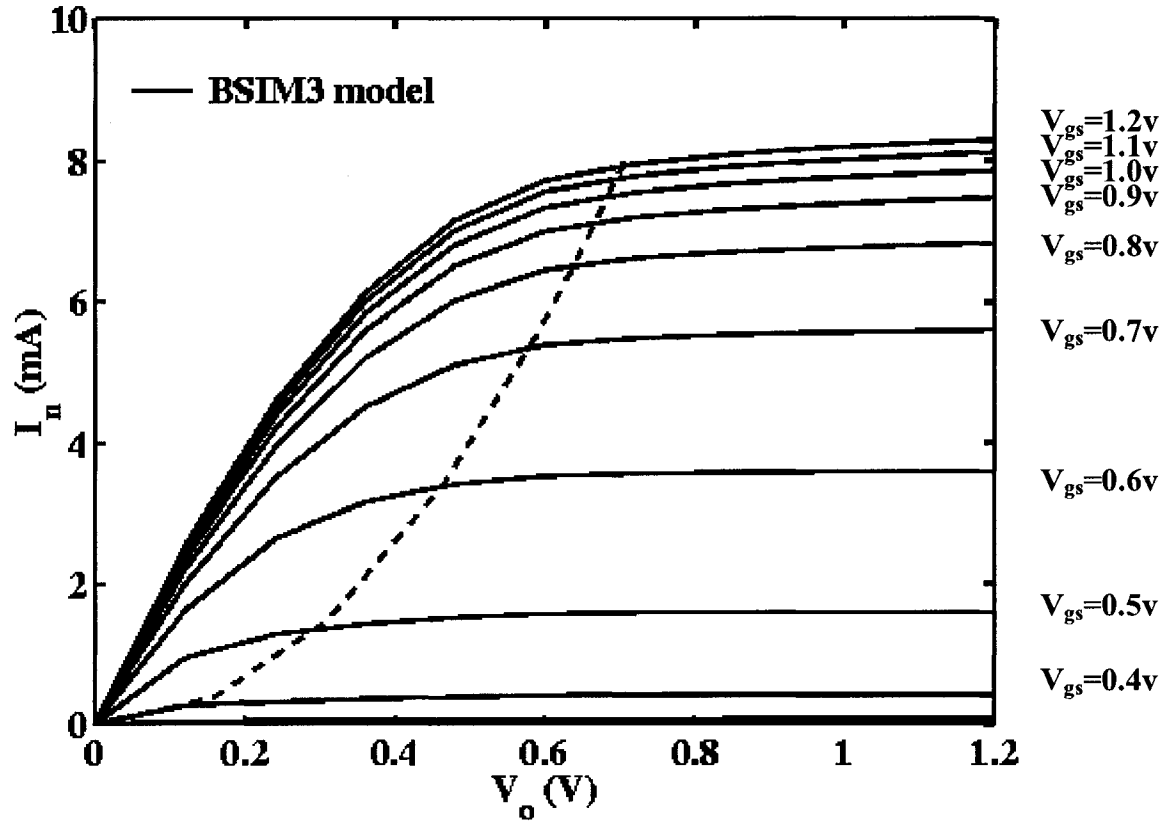


Fig. 5.2 The static I–V characteristics of 'n41' from SPICE. The square-law of chain current versus input voltage is valid when input voltage is small and close to threshold voltage. However, the square-law fails at input voltage close to operating voltage where the chain current strongly saturates as input voltage increases.

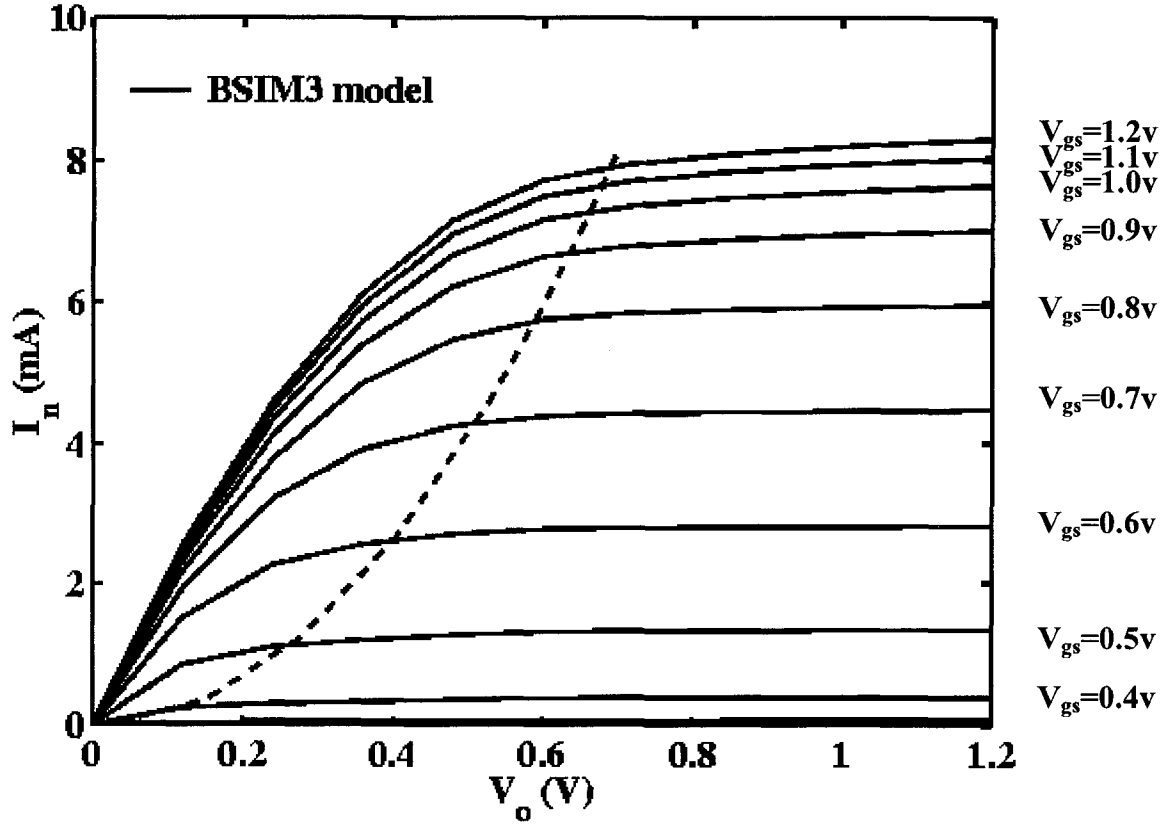


Fig. 5.3 The static I–V characteristics of 'n42' from SPICE. The square-law of chain current versus input voltage is valid when input voltage is small and close to threshold voltage. However, the square-law fails at input voltage close to operating voltage where the effect of chain current saturation is pretty large as input voltage increases.

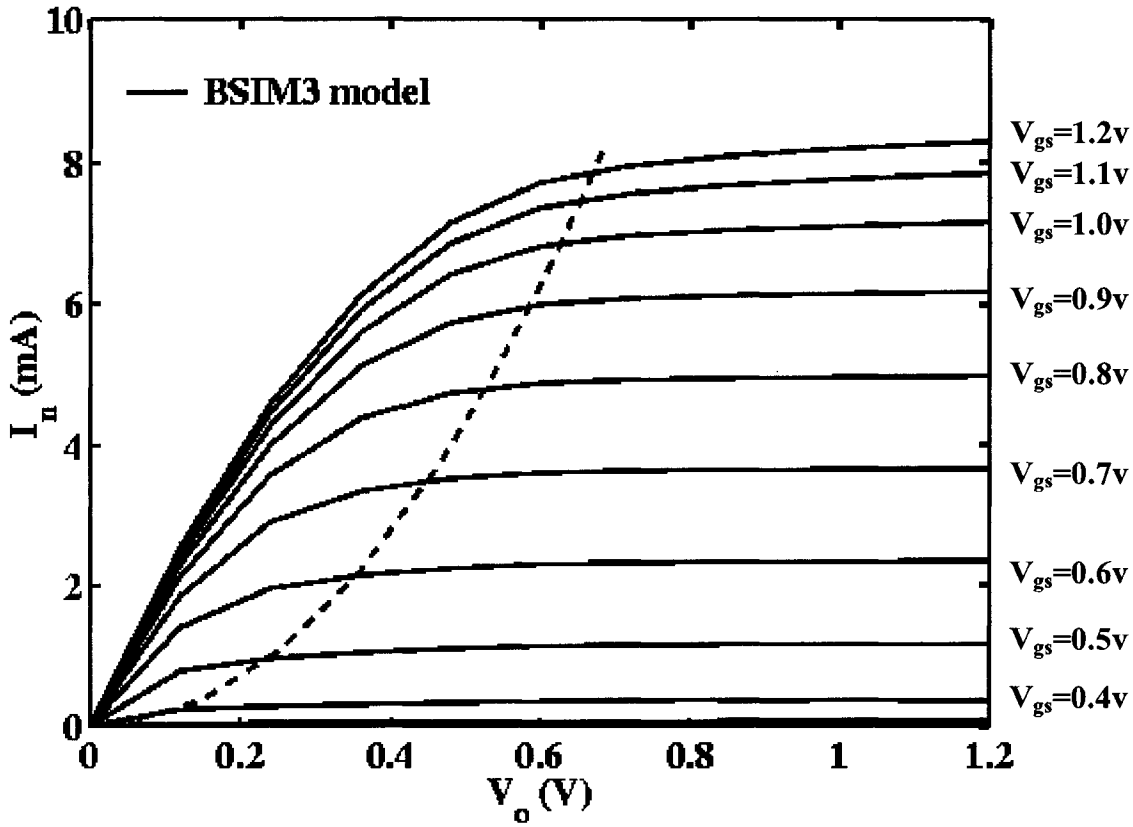


Fig. 5.4 The static I–V characteristics of 'n43' from SPICE. The square-law of chain current versus input voltage is valid when input voltage is small and close to threshold voltage. However, the square-law fails at input voltage close to operating voltage where the effect of chain current saturation is still large as input voltage increases.

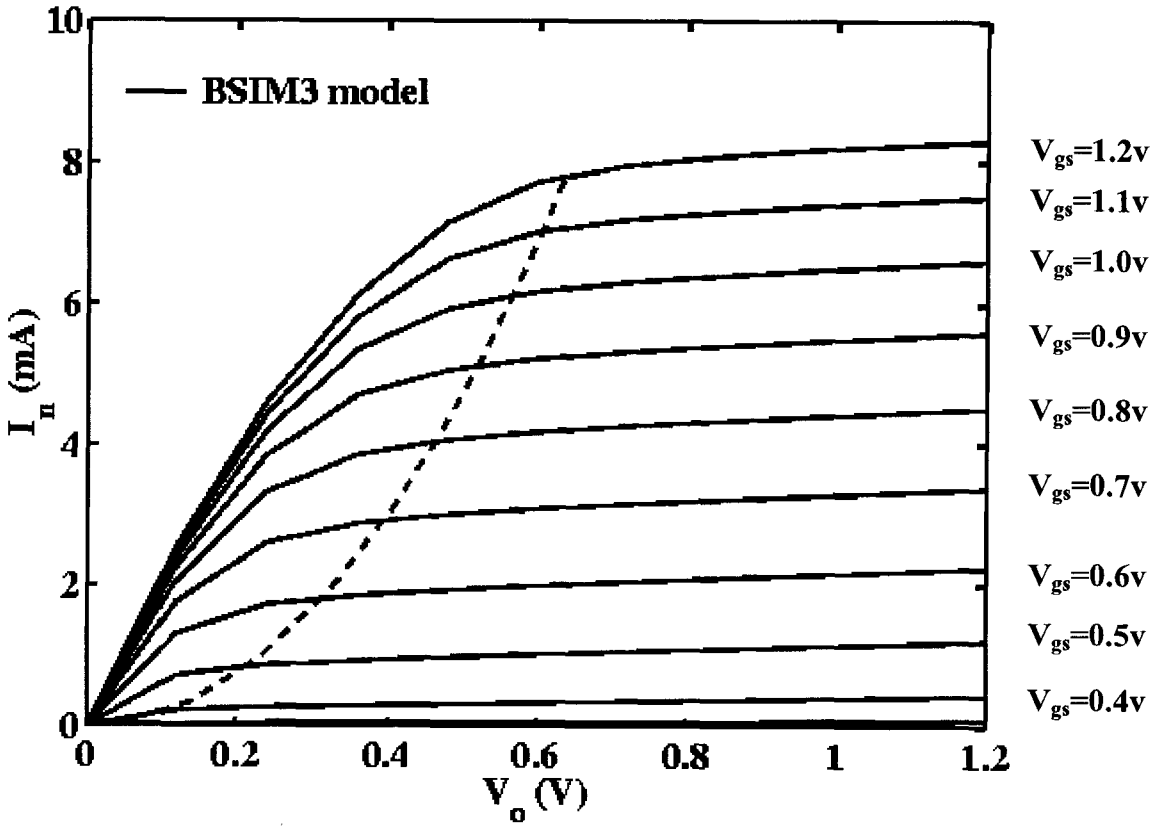


Fig. 5.5 The static I–V characteristics of 'n44' from SPICE. The square-law of chain current versus input voltage is valid when input voltage is small and close to threshold voltage. However, the square-law slightly fails at input voltage close to operating voltage where the effect of chain current saturation is visible as input voltage increases.

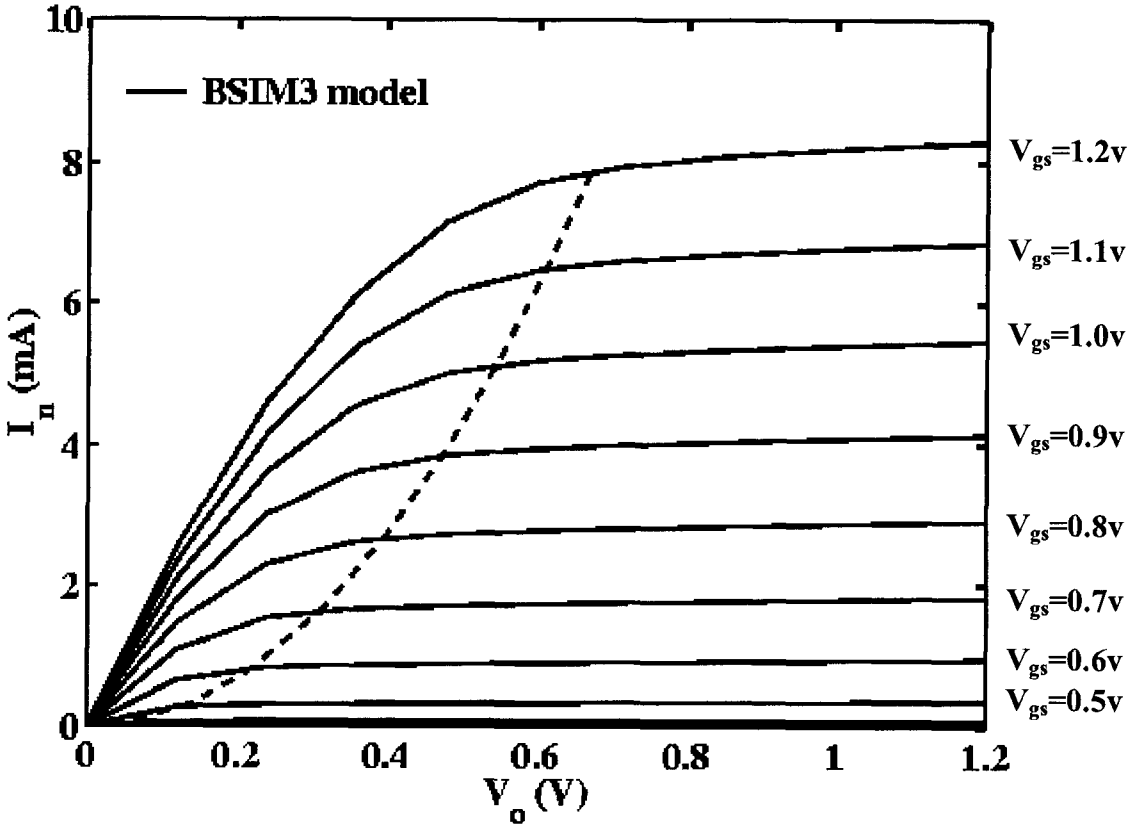


Fig. 5.6 The static I–V characteristics of 'n41234' from SPICE. They are similar with the static I–V characteristics of a deep submicron MOSFET.

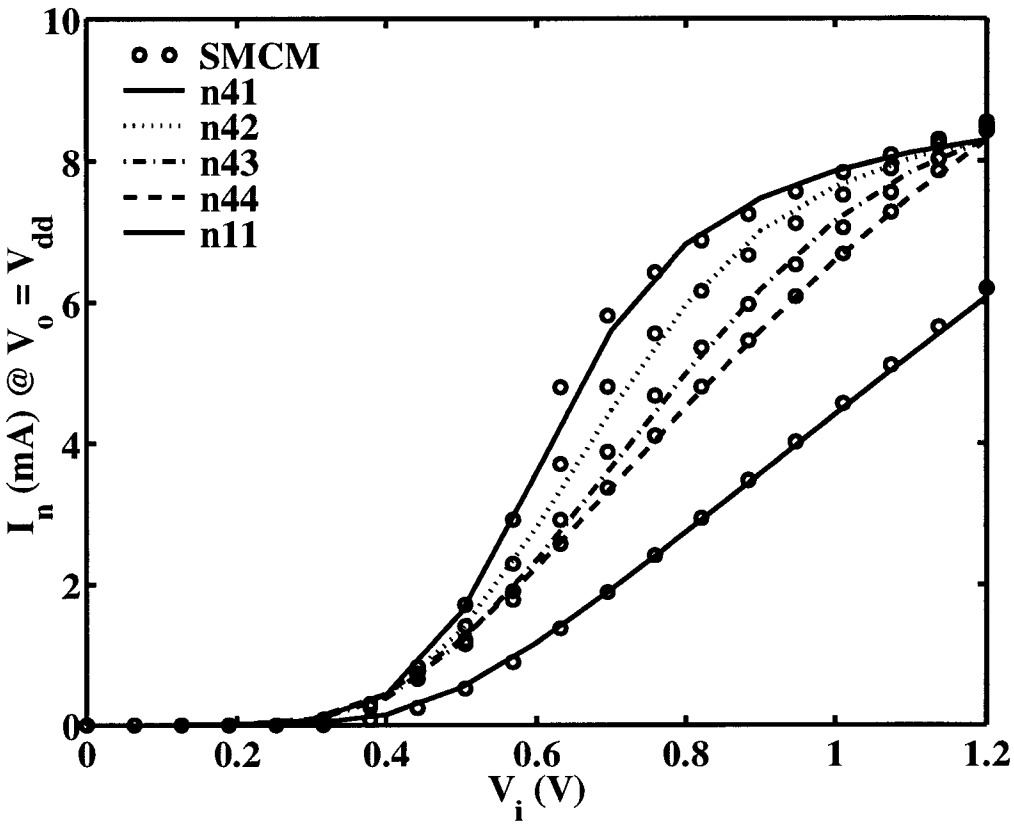


Fig. 5.7 The dependence of chain current I_n on input voltage V_{in} when output voltage V_{out} is equal to V_{dd} for 'n41', 'n42', 'n43', 'n44', and 'n11' (from top to bottom). The deep submicron nMOSFET of 'n11' has the same equivalent width of the series-connected chains of 'n41', 'n42', 'n43', and 'n44'. We can clearly see the effect of chain current saturation in cases like 'n41' and 'n42' as input voltage increases. The current amplitude of an nMOSFET with the same equivalent width of a series-connected MOSFET chain is smaller than that of the chain currents. In other words, the speed of logic NAND gate is faster than an inverter with the same equivalent width.

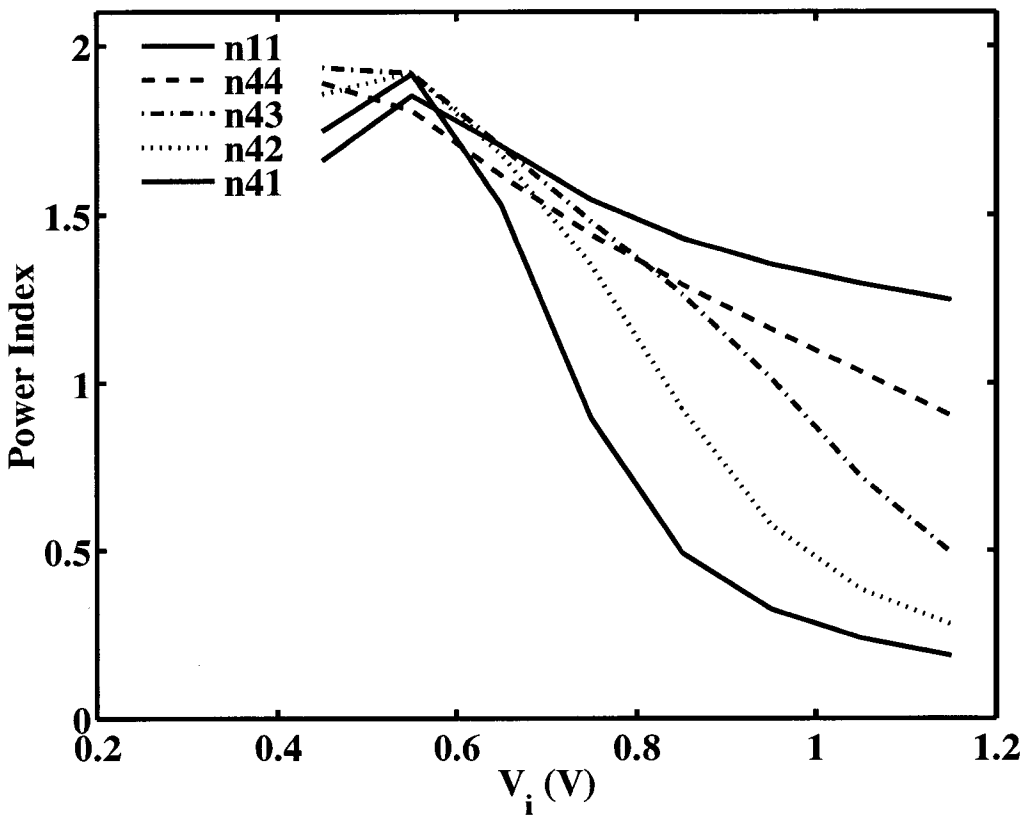


Fig 5.8 Power index as a function of input voltage from (3.14) for 'n41', 'n42', 'n43', 'n44', and 'n11', respectively (from bottom to top). Power index of a deep submicron MOSFET of 'n11' ranges between one and two. Power indices of series-connected MOSFET chains are close to two when input voltage is around threshold voltage. They decrease when input voltage increases. Their values are even lower than one at input voltage close to operating voltage.

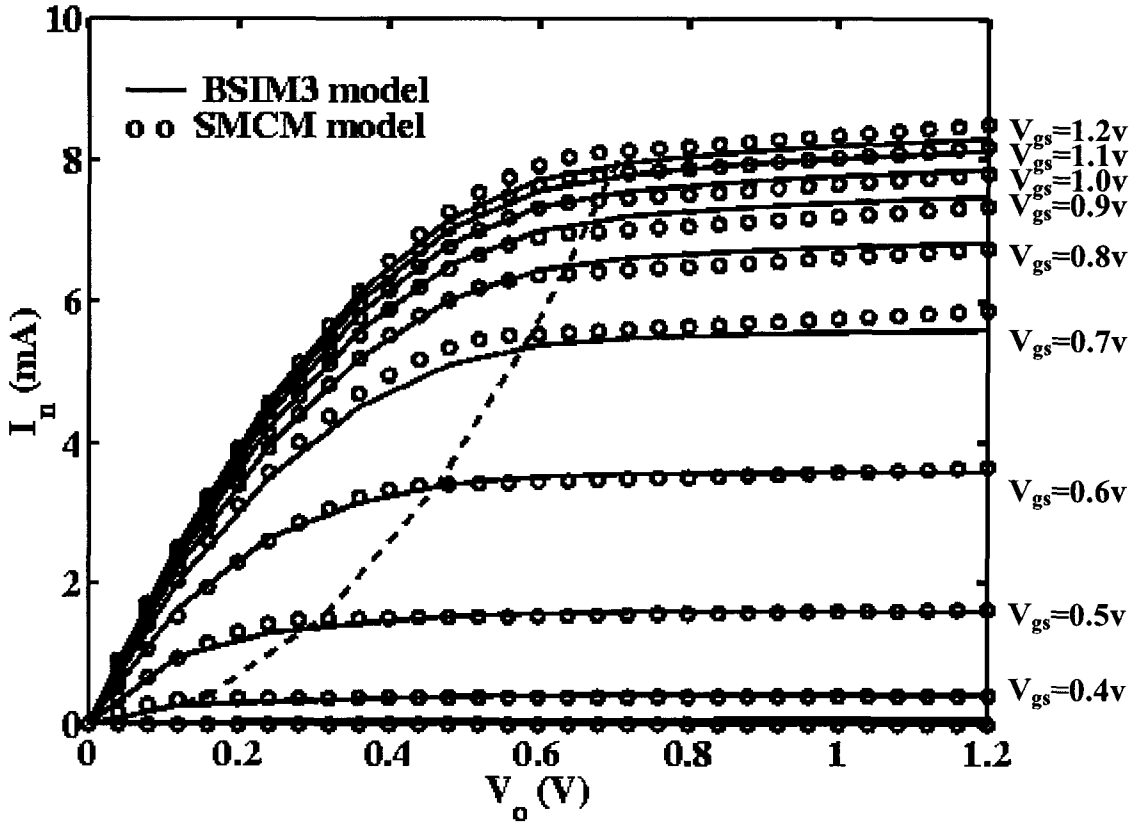


Fig. 5.9 The static I–V characteristics of 'n41' from SPICE and SMCM. From the figure, SMCM can accurately reproduce the peculiar I–V static characteristics of 'n41'. The chain current strongly saturates as input voltage increases.

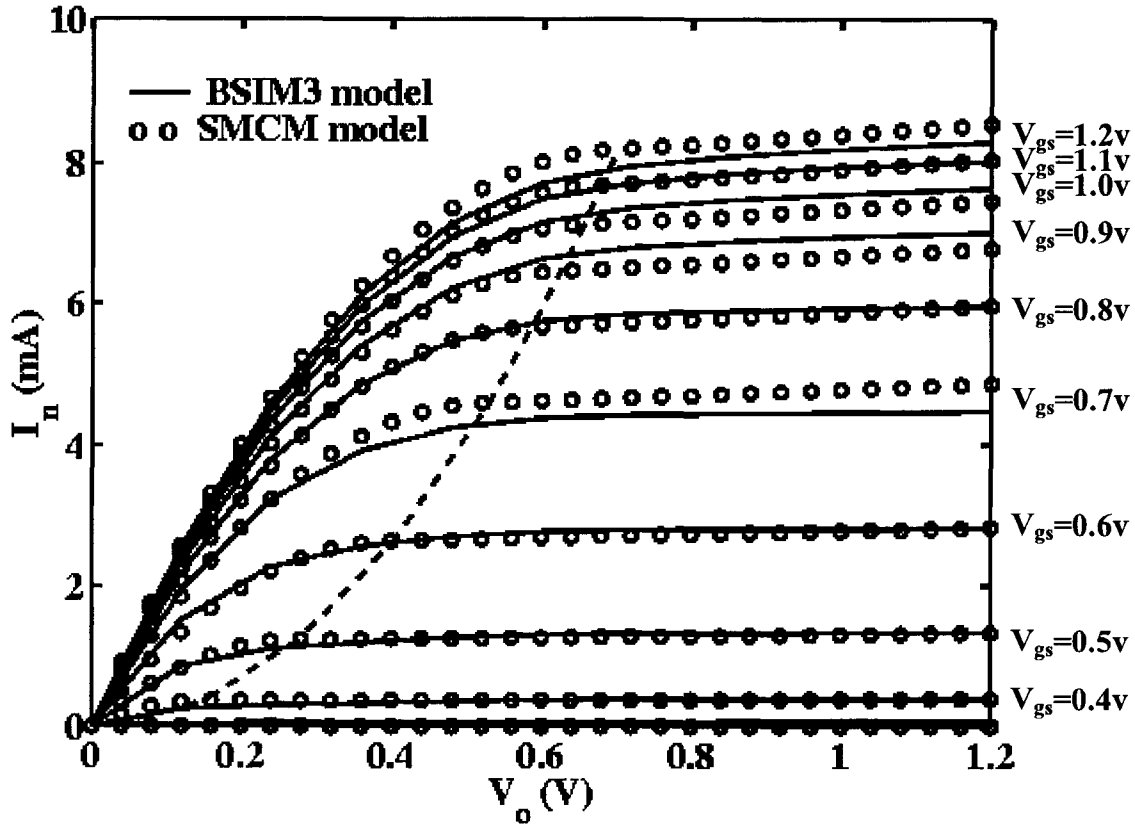


Fig. 5.10 The static I–V characteristics of 'n42' from SPICE and SMCM. From the figure, SMCM can accurately reproduce the peculiar I–V static characteristics of 'n42'. The effect of chain current saturation is pretty large as input voltage increases.

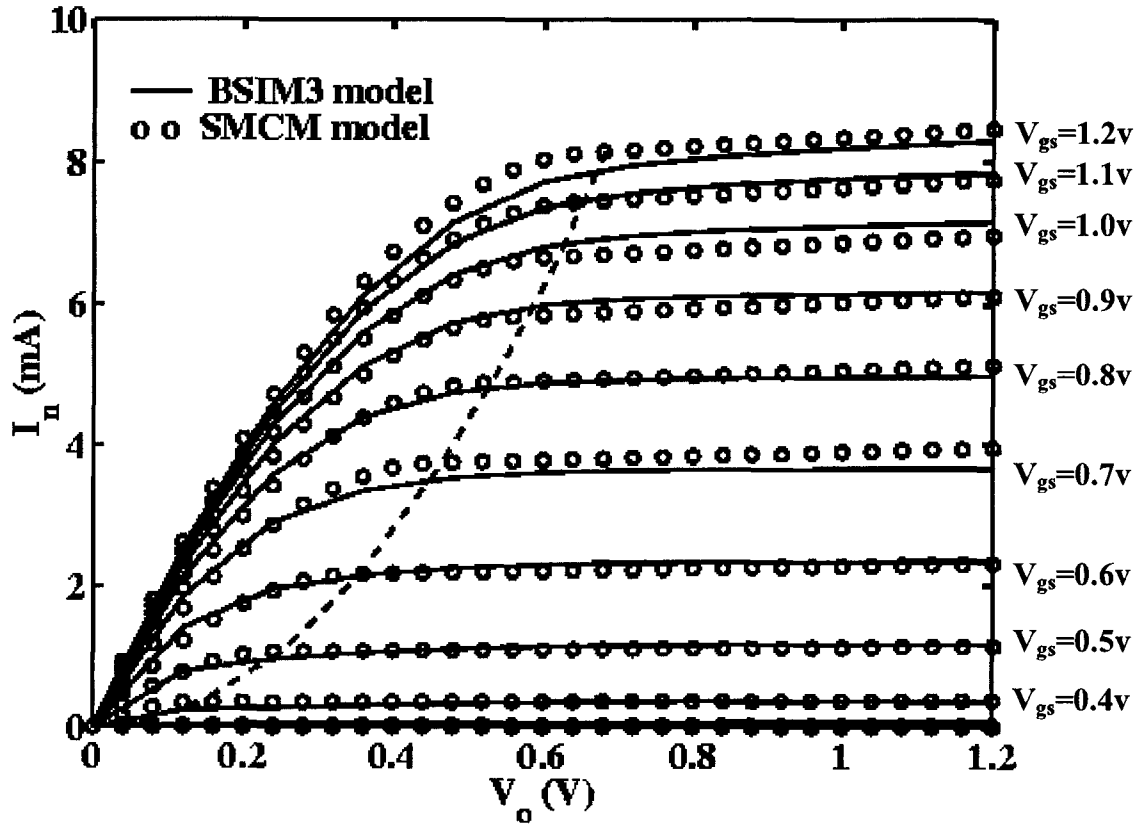


Fig. 5.11 The static I–V characteristics of 'n43' from SPICE and SMCM. From the figure, SMCM can accurately reproduce the peculiar static I–V characteristics of 'n43'. The effect of chain current saturation is still large as input voltage increases.

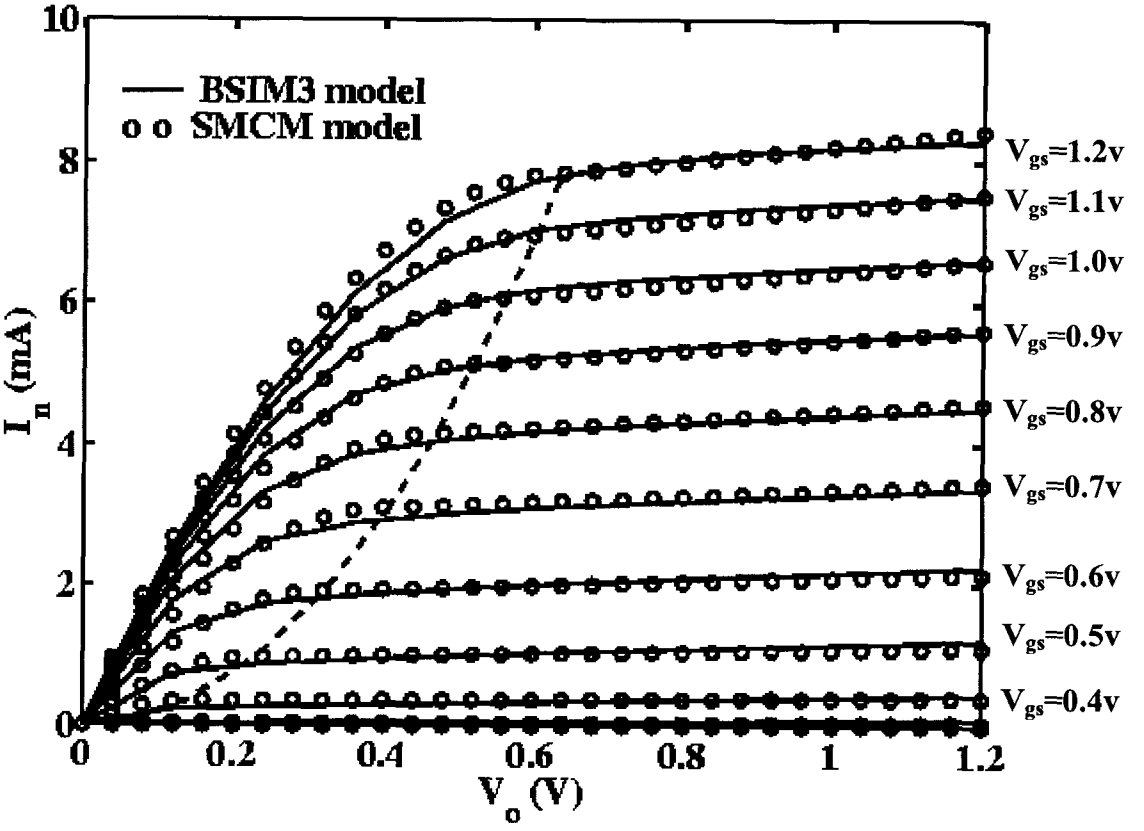


Fig. 5.12 The static I–V characteristics of 'n44' from SPICE and SMCM. From the figure, SMCM can accurately reproduce the peculiar static I–V characteristics of 'n44'. The effect of chain current saturation is visible as input voltage increases.

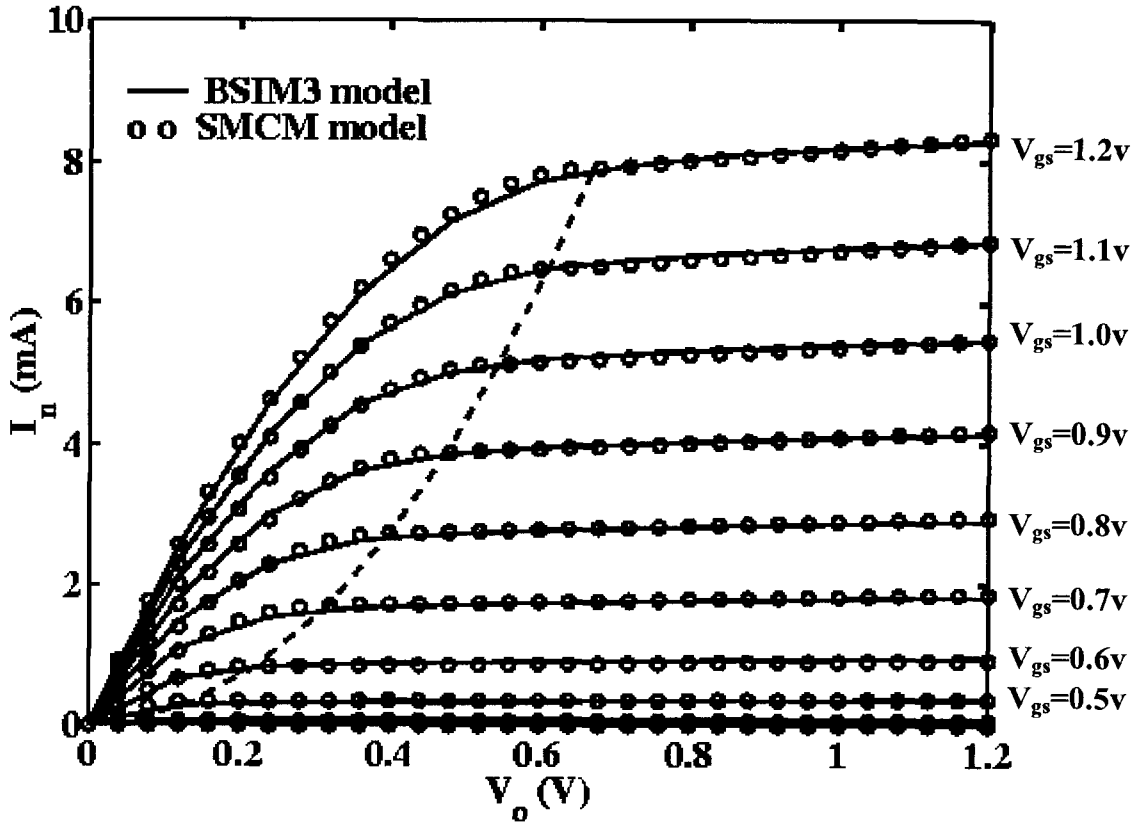


Fig. 5.13 The static I–V characteristics of 'n41234' from SPICE and SMCM. From the figure, SMCM can accurately reproduce the static I–V characteristics of 'n41234'.

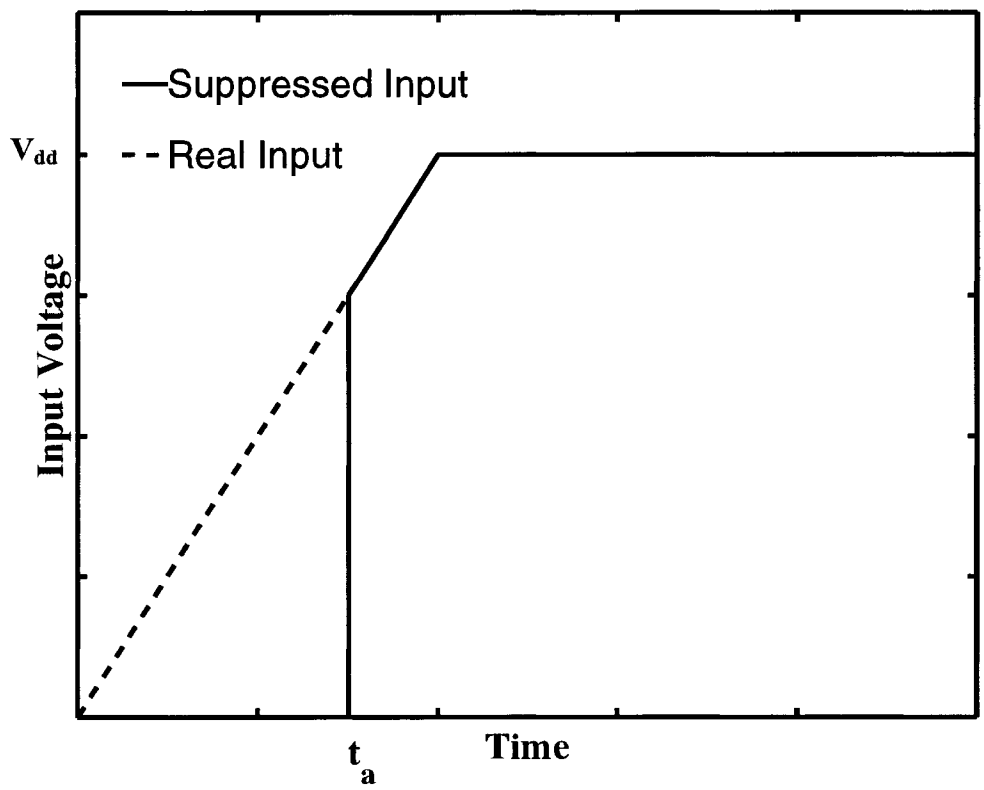


Fig. 5.14 Voltage waveforms of real input and suppressed input of a series-connected MOSFET chain. Before the activation time t_a, the real input is suppressed in the suppressed input. After the activation time t_a, the suppressed input is the same as the real input.

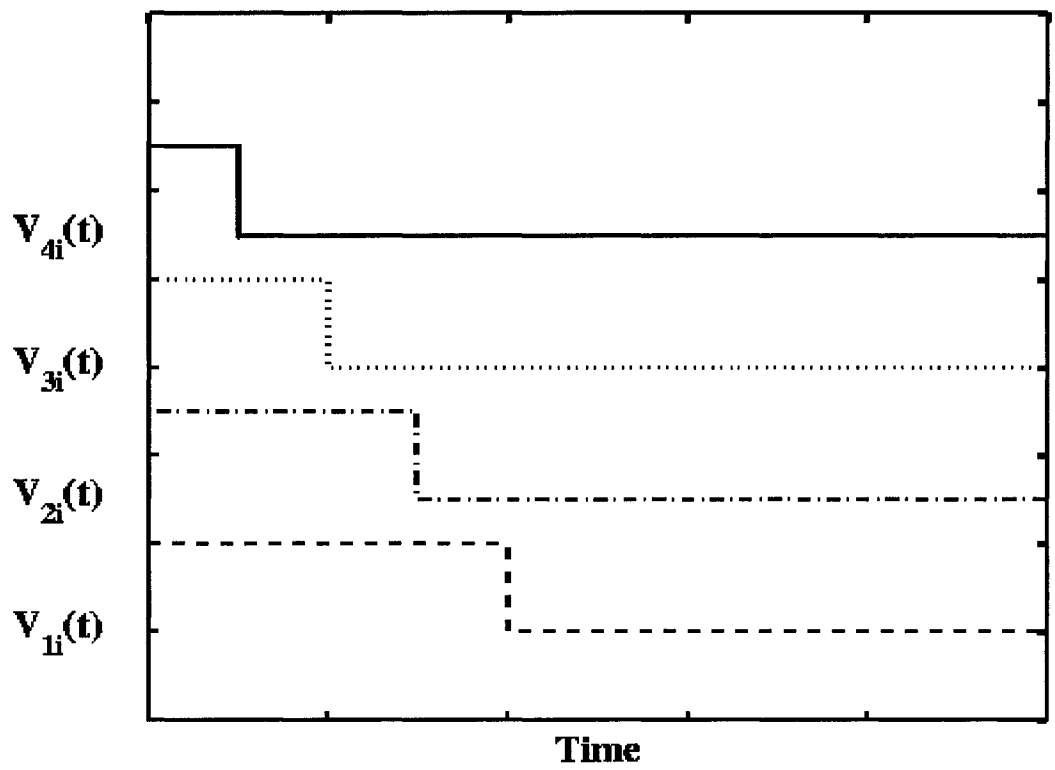


Fig. 5.15 Input pattern to generate fast initial state propagation delay (FISPD) of a series-connected nMOSFET chain. The order of turning nMOSFETs off makes parasitic capacitances of intermediate nodes have the minimized initial charges, which give rise to fast ISPD.

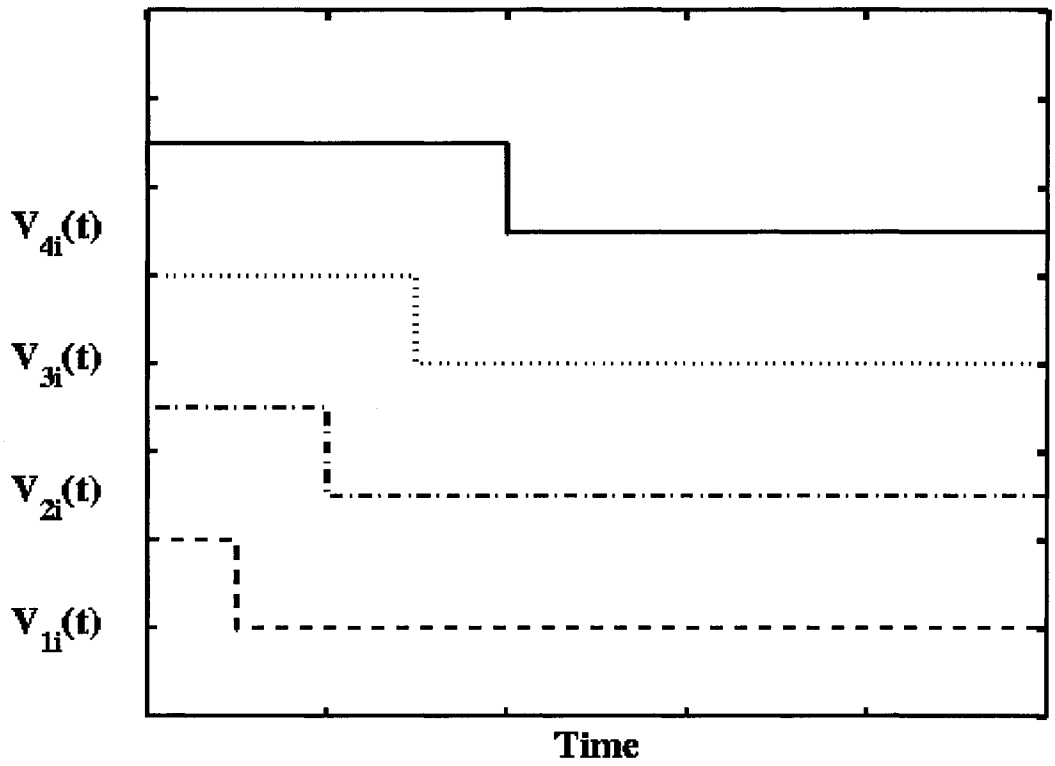


Fig. 5. 16 Input pattern to generate slow initial state propagation delay (SISPD) of a series-connected nMOSFET chain. The order of turning nMOSFETs off makes parasitic capacitances of intermediate nodes have the maximized initial charges, which give rise to slow ISPD.

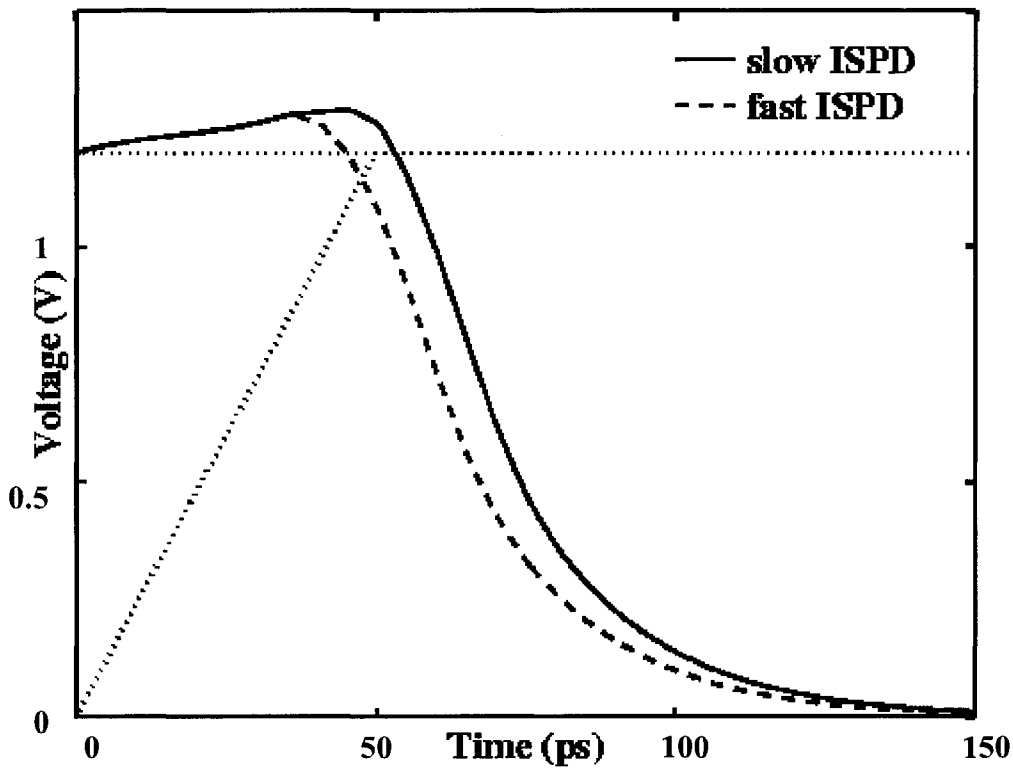


Fig. 5.17 Output voltage waveforms of 'n413' due to fast and slow initial states (FIS and SIS). The effect of initial states of intermediate nodes is clearly shown in this figure. The difference of the waveforms of FISPD and SISPD is caused by initial charges of parasitic capacitances of effective intermediate nodes of a series-connected MOSFET chain.

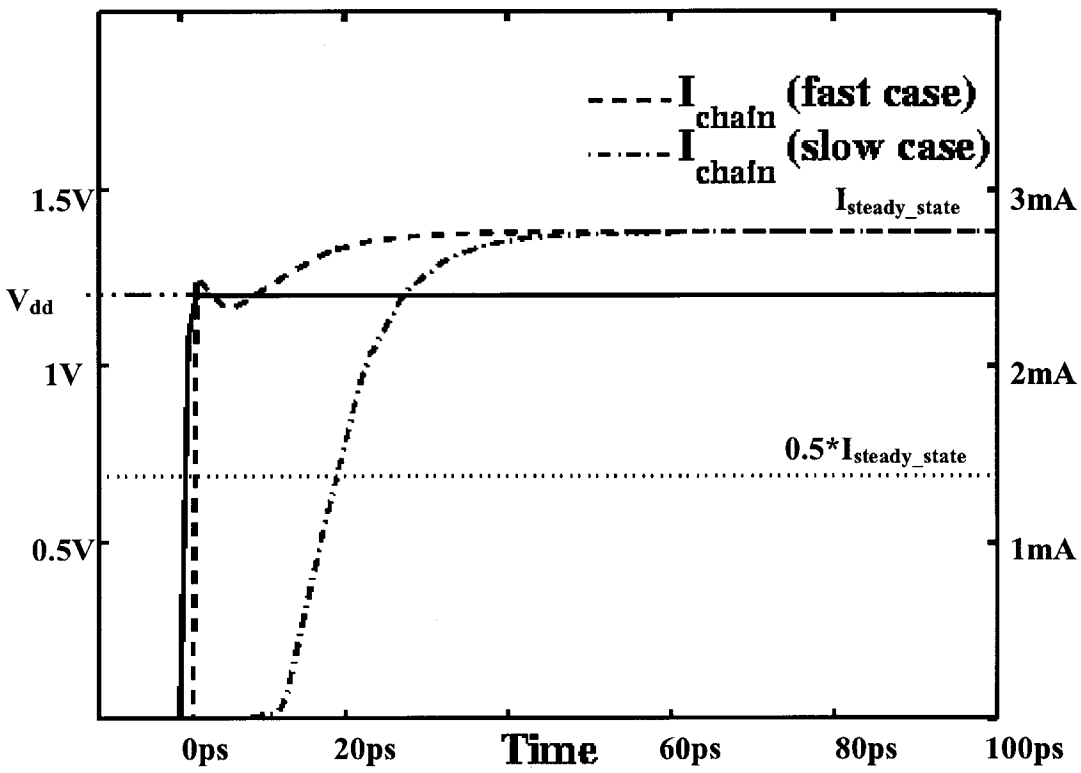


Fig. 5.18 Chain current waveforms of 'n414' for a step input with of fast and slow initial states (FIS and SIS). The step-input activation times for FIS and SIS are measured from the time when input voltage jumps to the time when chain current reaches the half of steady-state current.

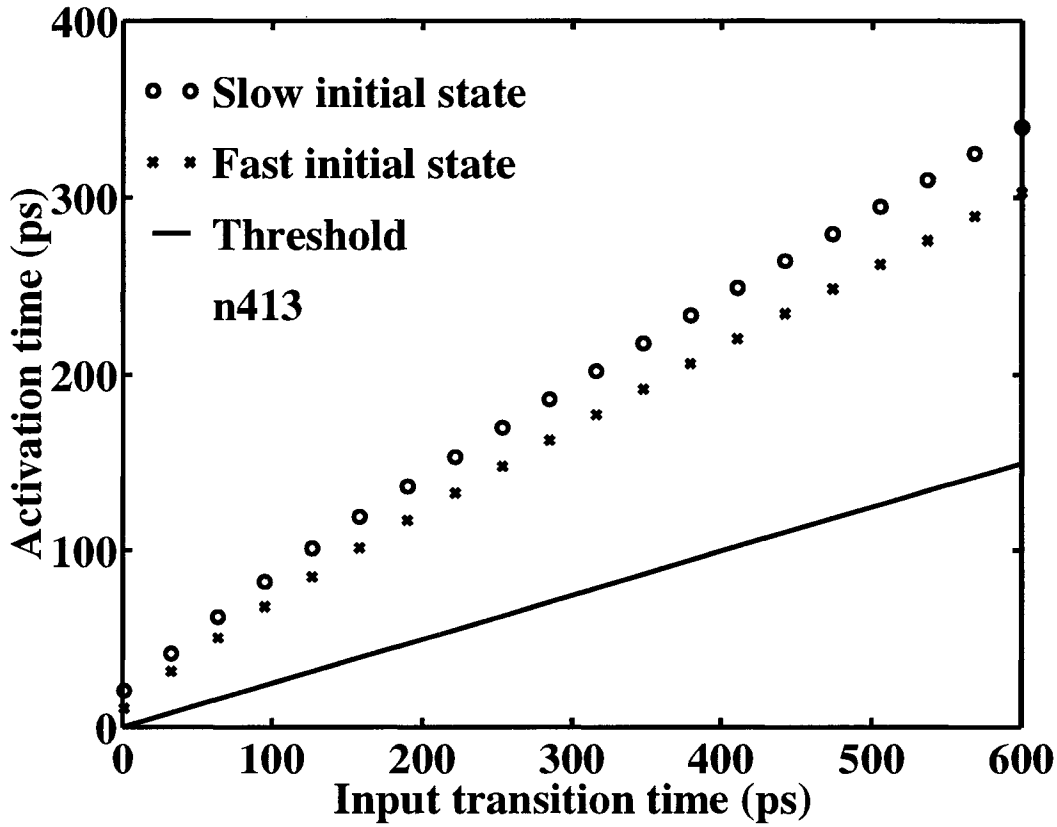


Fig. 5.19 Activation time of 'n413' as a function of input transition time. Given SIS and FIS step-input activation times, the SIS and FIS ramp-input activation times for input transition time are calculated. The charge conservation and SMCM are assumed in the derivation. The difference between SIS and FIS is caused by initial charges of parasitic capacitances of intermediate nodes between nMOSFET$_1$ and nMOSFET$_3$. The difference between FIS and the threshold line is caused by initial charges of parasitic capacitances of intermediate nodes between nMOSFET$_3$ and output node

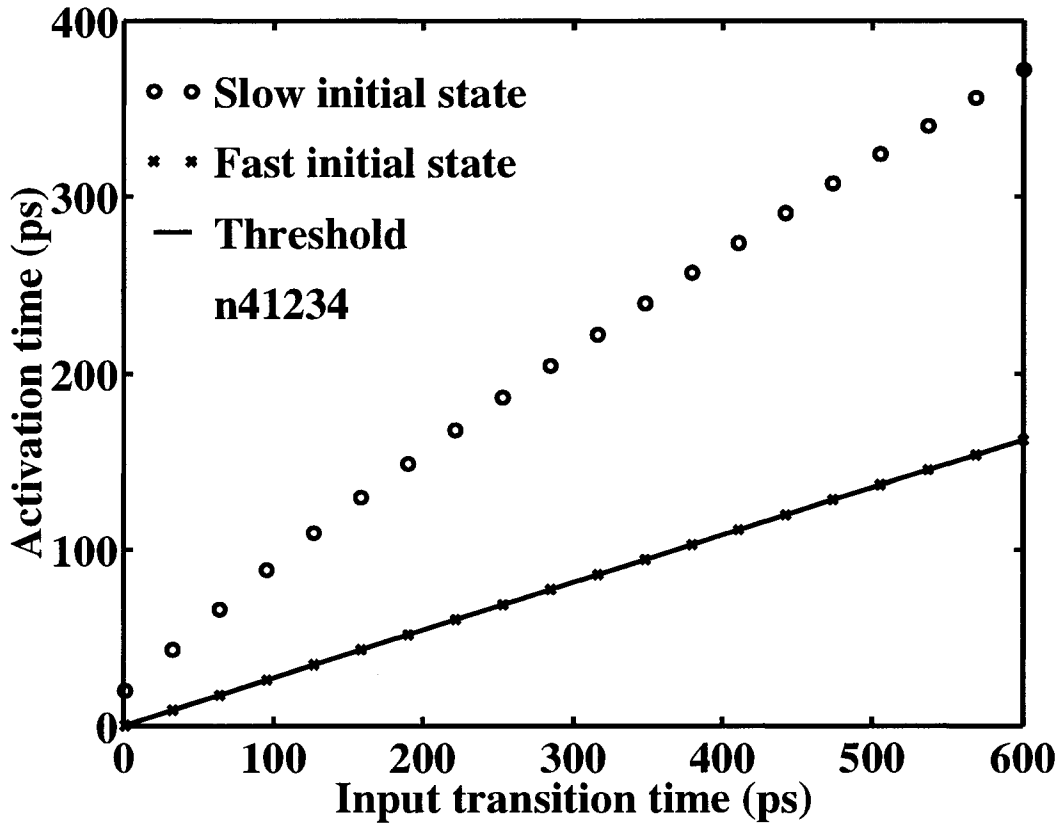


Fig. 5.20 Activation time of 'n41234' as a function of input transition time. The difference between SIS and FIS is caused by initial charges of parasitic capacitances of intermediate nodes between nMOSFET$_1$ and nMOSFET$_4$. There is no difference between FIS and the threshold line due to no intermediate nodes between nMOSFET$_4$ and output node

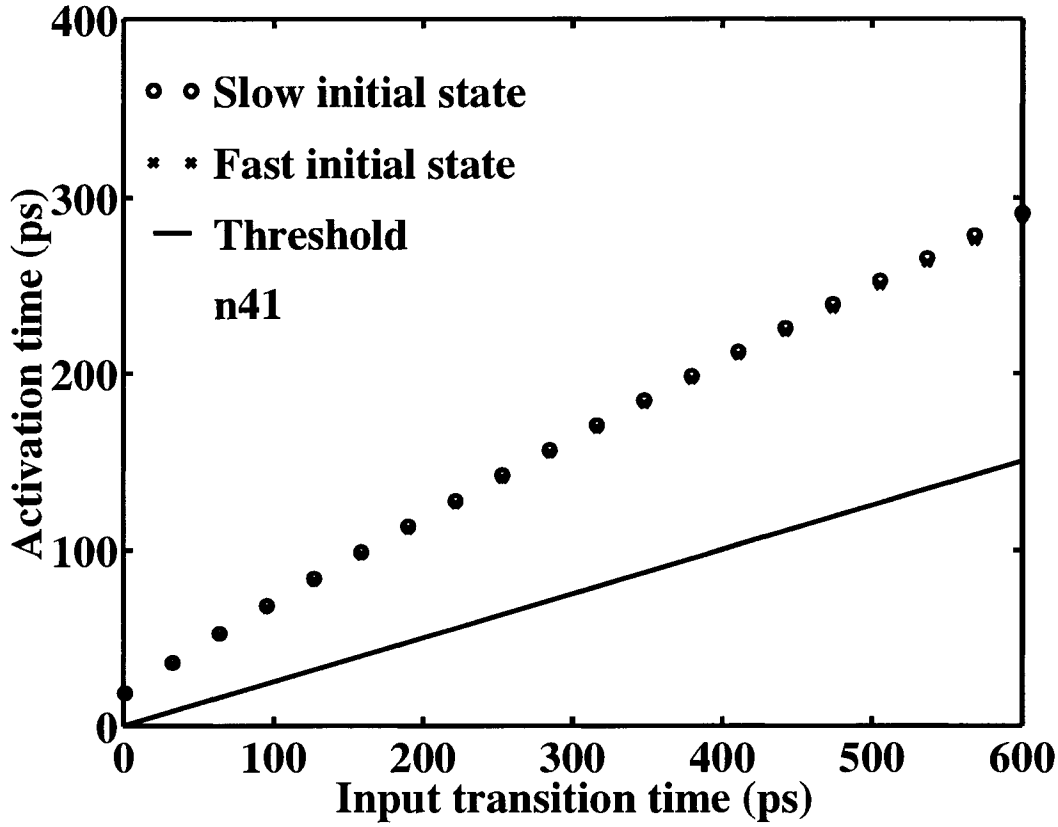


Fig. 5.21 Activation time of 'n41' as a function of input transition time. There is no difference of activation times of SIS and FIS since there is only one active input in n41. There is no effective intermediate node in 'n41'.

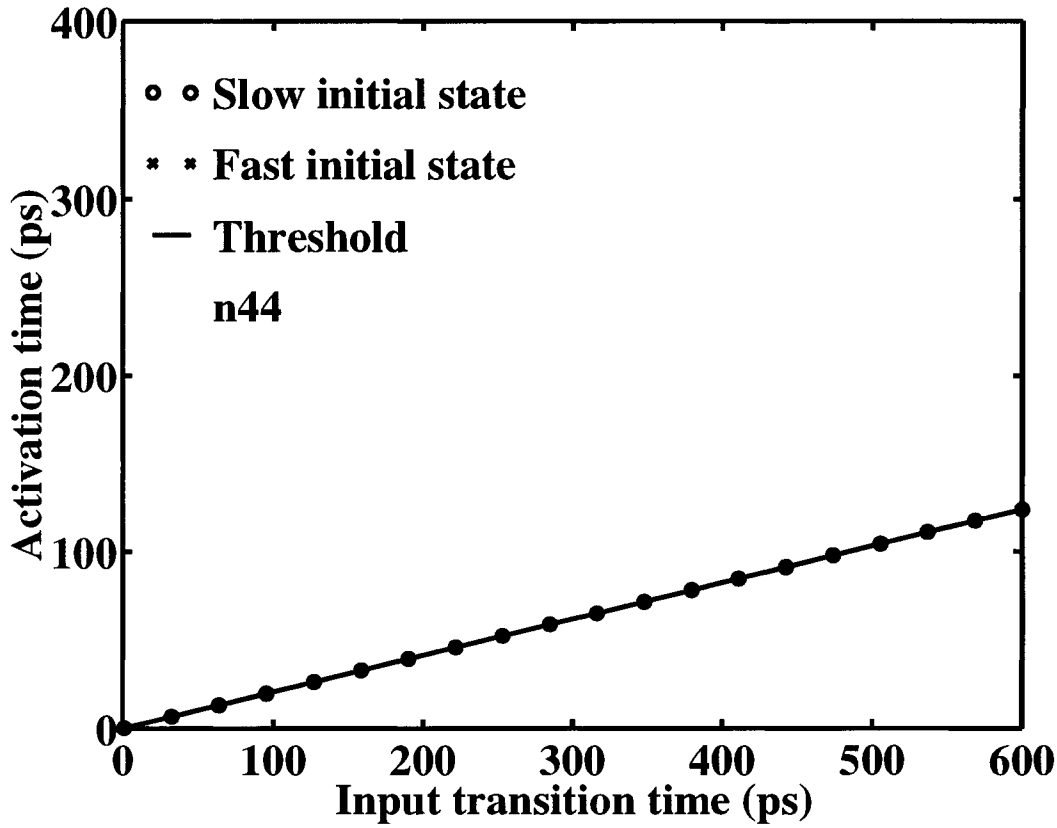


Fig. 5.22 Activation time of 'n44' as a function of input transition time. There is no difference of activation times of SIS and FIS since there is only one active input in n44. There is no effective intermediate node in 'n44'.

CHAPTER 6 TIMING MODELS OF A SINGLE-STAGE CMOS LOGIC GATE WITH AN EQUAL INPUT PATTERN

In this chapter, we focus on the timing models of a single-stage complex CMOS logic gate with an equal input pattern. An equal input pattern means that all inputs are connected to either an identical input waveform or a fixed voltage depending on the type of a CMOS logic gate. Active inputs, which are connected to an identical input waveform, determine the sort of a CMOS logic gate. Inputs connected to a fixed voltage are inactive inputs. For example, inactive inputs of a NAND gate are connected to power.

In order to exploit the methodologies of propagation delay for a deep submicron CMOS inverter developed in Chapter 4, we investigate the peculiar static I–V characteristics of a series-connected MOSFET chain with an equal input pattern in Chapter 5. An accurate series-connected MOSFET chain model (SMCM) is proposed to describe the peculiar static I–V characteristics. The modified input suppression technique associated with activation time is adopted to illustrate the effect of initial states of parasitic capacitances of intermediate nodes on output transition and propagation delay [49]. From the practical point of view we concentrate on the most important initial states, i.e., fast initial state (FIS) and slow initial state (SIS). The step-input activation time for a sort of a series-connected MOSFET chain is extracted from SPICE simulation. Based on charge conservation and SMCM we derive the ramp-input activation time.

In this chapter we extend the methodologies of propagation delay proposed for a CMOS inverter in Chapter 4 to a single-stage complex CMOS logic gate with an equal input pattern. By using the accurate series-connected MOSFET chain model (SMCM) and

the modified input suppression technique, the analytical expressions of transient output voltage of a single-stage CMOS logic gate with an equal input pattern are derived. The propagation delay of a single-stage four-input CMOS NAND4 gate with an equal input pattern is extracted from its analytical transient output expressions. We start from differential equation describing the transient behaviors of output voltage of a single-stage CMOS logic gate. The differential equation is solved for the analytical expressions of output voltage in three ways according to different combinations of input waveform and circuit structure to be used in the analysis. Input voltage can be either a step function or a ramp function. The simulated circuit structure of a single-stage CMOS NAND gate can be modeled as a dominant series-connected MOSFET chain with or without a trivial parallel pMOSFET cluster.

As the first method, the analytical step response of a four-input CMOS NAND4 gate is derived with a step input where all the inputs are connected to either an identical step function or power rail. Propagation delay and output transition time as a function of external load capacitance are obtained. This method is simple and accurate for a single-stage CMOS logic gate with very fast input. However, the effect of input transition time ignored in the derivation may cause a significant error for an equal input pattern with non-zero input transition time. In the second way, the series-connected nMOSFET chain without a trivial parallel pMOSFET cluster is used to model a CMOS NAND4 gate. We concentrate on the effect of input transition time in timing analysis. Without a trivial parallel pMOSFET cluster the voltage overshoots may be exaggerated and the short-circuit current is ignored. The ramp-input response is developed with the consideration of input transition time in addition to external load capacitance. Since the propagation delay

depends on both input transition time and external load capacitance, the complexity of this method is medium. It provides accurate propagation delay in broader range of input transition time and external load capacitance than the first method. The last method is the comprehensive propagation delay model for a four-input CMOS NAND4 gate with an equal input pattern. The trivial pMOSFET cluster is included and the short-circuit current during transition is carefully considered in the method. Since input transition time, and external load capacitance are included in the derivation of transient output voltage expressions, their impacts on propagation delay of a four-input CMOS NAND4 gate are discussed. The tradeoffs between accuracy and simplicity with these three methods are investigated. Generally speaking, the computation speeds of propagation delay by using these analytical timing models are more than two orders of magnitude faster than with SPICE [31] [70] [89]. More importantly, these analytical delay models can be easily incorporated into VLSI CMOS design and verification flows [14] [21].

The approaches of propagation delay developed in this chapter can be applied to other single-stage CMOS logic gates like NAND gates or NOR gates similarly or symmetrically. In the following derivation of the analytical timing model of a series-connected CMOS logic gate, we concentrate on the case of a rising input associated with falling output. All the arguments made are symmetrically valid for a falling input associated with rising output. The difference is that for the case of a falling input, the load capacitance is charged through the parallel pMOSFET cluster, which becomes the major factor on propagation delay. For simplicity, we sometimes use the normalized variables defined in Section 4.1.

6.1 A SINGLE-STAGE CMOSLOGIC GATE WITH AN E QUAL INPUT PATTERN

6.1.1 Differential Equation

In analytical timing approaches, the analytical output voltage expressions of a four-input CMOS NAND4 gate are obtained by solving the differential equation, which describes its transient response. The propagation delay is then extracted from the output voltage expressions. Fig. 6.1 shows the schematic of a four-input CMOS NAND4 gate used to derive the propagation delay. Note that input-output coupling capacitances and output load capacitance C_{load} are included in Fig. 6.1.

The differential equation to describe the transient behaviors of a single-stage CMOS logic gate such as a four-input CMOS NAND4 gate shown in Fig. 6.1 is obtained from the application of the Kirchoff's current law (KCL) at the output node:

$$-I_{C_{load}} + I_{C_m} + I_p - I_n = 0 \tag{6.1}$$

Substituting capacitance characteristics into the above equation yields

$$\frac{dV_o}{dt} = k_c \cdot \frac{dV_i}{dt} + \frac{1}{C_t}\left(I_p - I_n\right) \tag{6.2}$$

where C_t is the total capacitance of the output node, which is

$$C_t = C_{load} + C_m \tag{6.3}$$

and k_c the input-output coupling coefficient. It is the ratio of the effective input-output coupling capacitance C_m to the total capacitance C_t of output node, i.e.,

$$k_c = C_m / C_t \tag{6.4}$$

If input voltage is a ramp function the differential equation of (6.2) with the normalized

time variable s becomes

$$\frac{dV_o}{ds} = k_c V_{dd}[u(s)-u(s-1)]+\frac{\tau}{C_t}(I_p - I_n) \tag{6.5}$$

Before we analytically solve the differential equations of (6.2) or (6.5), the capacitances and the current models of I_p and I_n are briefly discussed in the next subsections.

6.1.2 Capacitances

The input-output coupling capacitance C_m of a single-stage CMOS logic gates is defined as the summation of all the effective input-output coupling capacitances between input and output. Taking a four-input CMOS NAND4 gate for a rising input as an example, it is the summation of the gate-drain capacitances of a parallel switching pMOSFET cluster and the top nMOSFET of the chain connected to the output node if its corresponding input switches. Generally the input-output coupling capacitance C_m of a single-stage CMOS logic gate can be expressed as

$$C_m = n_{mn}C_{mn} + n_{mp}C_{mp} \tag{6.6}$$

where the gate-drain coupling capacitances C_{mn} and C_{mp} are obtained from SPICE characterization. For simplicity, the gate-drain coupling capacitances C_{mn} and C_{mp} per unit channel width are assumed to be voltage-independent and determined by the process technology. In practice they are voltage-dependent [2] [51] [59] [104]. The coefficients of n_{mp} and n_{mn} are dependent of the sort of the series-connected MOSFET chain, i.e., the set of active inputs involved in the coupling current between input node and output node. The coefficients n_{mp} and n_{mn} for INV and NAND2~NAND4 are listed in Table 6.2.

The total capacitance C_t of the output node is the summation of output load

capacitance C_{load} and input-output coupling capacitance C_m as in (6.3). It can be also considered as total parasitic drain diffusion capacitance C_d and external load capacitance C_o, i.e.,

$$C_t = C_d + C_o \tag{6.7}$$

where total parasitic drain capacitance C_d of all the MOSFETs connected to the output node is obtained by

$$C_d = n_{dn}C_{dn} + n_{dp}C_{dp} \tag{6.8}$$

In (6.8) C_{dp} and C_{dn} are the drain capacitances of a pMOSFET and an nMOSFET characterized by SPICE. For simplicity, the parasitic drain diffusion capacitances C_{dn} and C_{dp} per unit channel width are assumed to be voltage-independent and determined by the process technology. In practice they are voltage-dependent [2] [51] [59] [104]. The coefficients n_{dp} and n_{dn} are the numbers of p-tree and n-tree connected to output node. The coefficients n_{dp} and n_{dn} for INV and NAND2~NAND4 are listed in Table 6.2. The external load capacitance C_o in (6.7) is a primary variable in timing models. In fact, it is compound of wire capacitances at output node and input capacitances of next stage.

6.1.3 Current Models of I_p and I_n

In order to analytically solve the differential equations of (6.2) or (6.5), it is necessary to carefully choose the current models I_p and I_n for a trivial parallel pMOSFET cluster and for the series-connected nMOSFET chain, respectively. These current models should be as simple and accurate as necessary. In this chapter, we assume the input voltage applied to a four-input CMOS NAND4 gate is either a rising step

function or a rising ramp function. During output transition the output voltage is discharged from operating voltage to ground. The series-connected nMOSFET chain for a rising input experiences cutoff mode, saturation mode, and linear mode. The series-connected nMOSFET chain current I_n dominates the output transition or discharging process. Therefore, the series-connected nMOSFET chain current I_n should be modeled as accurately as possible. As we discussed in Chapter 5, the generic series-connected MOSFET chain model (SMCM) accurately depict the peculiar static *I–V* behaviors of series-connected nMOSFET chain. The proposed accurate SMCM is used in the following analyses of output expressions and propagation delays. The input voltage V_i and output voltage V_o are equal to nMOSFET gate-to-ground voltage and drain-to-ground voltage, respectively. The current I_n of SMCM from Chapter 5 is expressed as:

$$I_n = \begin{cases} 0 & V_i \le V_{tn} \\ k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o(1+\lambda_n V_{sn}) & V_i > V_{tn}, V_o \le V_{sn} \\ \dfrac{1}{2}k_n {V_{sn}}^2(1+\lambda_n V_o) & V_i > V_{tn}, V_o > V_{sn} \end{cases} \tag{6.9}$$

where k_n is the transconductance parameter of a series-connected MOSFET chain with an equal input pattern, V_{th} the threshold voltage, and λ_n the channel length modulation coefficient. The output saturation voltage V_{sn} of a series-connected nMOSFET chain in SMCM is given by:

$$V_{sn} = \begin{cases} Q_n(V_i - V_{tn}) & V_{tn} < V_i \le \dfrac{V_{dd}}{2} \\ R_n(V_i - V_{an})^{b_n} & \dfrac{V_{dd}}{2} < V_i \le V_{dd} \end{cases} \tag{6.10}$$

In (6.10), Q_n, R_n and V_{an} are the SMCM model parameters. Index b_n is used to describe the effect of power index of the series-connected nMOSFET chain with an equal input pattern. The output saturation voltage V_{sn} of a series-connected nMOSFET chain is continuous when input voltage V_i is approaching $V_{dd}/2$. This gives rise to the following constraint

$$Q_n(V_{dd}/2 - V_{tn}) = R_n(V_{dd}/2 - V_{an})^{b_n} \tag{6.11}$$

In deriving the propagation delay of a deep submicron CMOS NAND4 gate for a rising input accurately and thoroughly, the trivial parallel switching pMOSFET cluster in Fig. 6.1 cannot be ignored though it does not dominate the discharging process. A linearized and simplified model from the series-connected MOSFET chain model (SMCM) is accurate enough to describe the static *I–V* characteristics of a parallel switching pMOSFET cluster since there are two major behaviors in the operation trajectory traces of parallel pMOSFET cluster for a rising input. The operation traces of a parallel pMOSFET cluster experience either the small source-drain voltage $V_{sd}\left(=V_{op}=V_{dd}-V_o\right)$ or the small source-gate voltage $V_{sg}\left(=V_{ip}=V_{dd}-V_i\right)$. In either case, the effect of channel length modulation is small so that it is ignored without making a comparable error. Furthermore, when the applied input voltage V_i is small, the output voltage V_o is close to the operating voltage V_{dd} and the source-drain voltage $V_{sd}\left(=V_{op}=V_{dd}-V_o\right)$ is small. The parallel pMOSFET cluster is in the linear mode where the linearized SMCM is accurate enough to illustrate the pMOSFET characteristic in this mode. When applied input voltage V_i is large, the pMOSFET enters into the saturation mode where the source-gain

voltage $V_{sg}\left(=V_{ip}=V_{dd}-V_i\right)$ is small and the linearized SMCM in the saturation mode is also accurate enough. Therefore, the pMOSFET *I–V* characteristics under both operation conditions are well described by the simplified and linearized SMCM equation:

$$I_p=\begin{cases}0 & V_{ip}\le V_{tp}\\ k_pV_{sp}V_{op} & V_{ip}>V_{tp},V_{op}\le V_{sp}/2\\ \frac{1}{2}k_pV_{sp}{}^2 & V_{ip}>V_{tp},V_{op}>V_{sp}/2\end{cases} \tag{6.12}$$

The parameter k_p is the total drivability coefficient of parallel switching pMOSFET cluster. It can be obtained from $k_p=n_{mp}\cdot k_{p1}$. The coefficient of n_{mp} is the number of active parallel switching pMOSFETs listed in Table 6.2. The parameter k_{p1} is the drivability coefficient of one pMOSFET. In (6.12), the saturation voltage V_{sp} of parallel switching pMOSFET cluster is given by:

$$V_{sp}=Q_p\left(V_{ip}-V_{tp}\right) \tag{6.13}$$

The intersection point of the two linearized lines corresponding to the linear mode and saturation mode is determined by half of the pMOSFET saturation voltage V_{sp} as in (6.12).

Thus, in the following propagation delay analysis for a rising input, the proposed accurate SMCM given by (6.9) and (6.10) is used to describe the series-connected nMOSFET chain *I–V* characteristics. The parallel switching pMOSFET cluster is modeled by using the simplified SMCM equations of (6.12) and (6.13).

6.2 TIMING ANALYSIS OF A SINGLE-STAGE CMOS LOGIC GATE FOR A STEP INPUT

In this section, the propagation delay of a single-stage CMOS logic gate for a step input is derived. The step response of output voltage $V_o(t)$ is first calculated and the propagation delay is determined as the time needed for the output voltage $V_o(t)$ to reduce to half of operating voltage $(V_{dd}/2)$ by discharging the total load capacitance C_t through the series-connected MOSFET chain.

6.2.1 Operation of a Single-Stage CMOS Logic Gate for a Step Input

The input voltage $V_i(t)$ to the single-stage CMOS logic gate is assumed to be an ideal step function,

$$V_i(t) = V_{dd}u(t) \tag{6.14}$$

where $u(t)$ is a unit step function.

When time is less than zero $(t < 0_-)$, input voltage $V_i(t)$ is zero. The series-connected nMOSFET chain is in cutoff mode and the parallel pMOSFET cluster is in linear mode. The initial output voltage $V_o(t)$ is equal to operating voltage V_{dd}, i.e. $\left(V_o(t)\big|_{t=0_-} = V_{dd}\right)$. At the time zero $(t = 0)$, the output voltage $V_o(t)$ of the series-connected nMOSFET chain jumps from V_{dd} to $V_o(t)\big|_{t=0_+}$. During the period between time zero and step-input activation time, i.e., $(0_+ < t \le t_{as})$, the input to the series-connected nMOSFET chain is suppressed according to the modified input suppression technique discussed in Chapter 5. The chain current does not affect the output voltage. Thus output voltage keeps the same for this period. When step-input activation time t_{as} is reached the chain current truns on while output voltage $V_o(t)$ still remains at $V_o(t)\big|_{t=0_+}$ After step-input activation time t_{as}, the output voltage $V_o(t)$ go along the characteristic

curve with its input voltage equal to V_{dd} from its saturation mode into its linear mode.

At the time zero $(t = 0)$, the parallel pMOSFET cluster jumps from linear mode to its cutoff mode since the input jumps to operating voltage. Once the time is greater than zero $(t > 0_+)$, the pMOSFET cluster stays in its cutoff mode. Except for the input-output coupling current the trivial parallel switching pMOSFET cluster current does not affect step response of a single-stage CMOS logic gate. The differential equation to control the output voltage response $V_o(t)$ for a rising step input is obtained by substituting (6.14) into (6.2)

$$\frac{dV_o}{dt} = k_c V_{dd}\delta(t) - \frac{1}{C_t}\cdot I_n \tag{6.15}$$

where $\delta(t)$ is a Dirac function. Therefore the output voltage response $V_o(t)$ for a rising step input is derived for analytically solving the above differential equation.

6.2.2 Timing Analysis of a Single-Stage CMOS Logic Gate for a Step Input

According to the above discussions, we analyze the transient properties of a single-stage CMOS logic gate in four regions for a step input, respectively. The propagation delay for a step input is extracted from the transition of output voltage.

Region one $(0_- < t \le 0_+)$**:** Since input voltage $V_i(t)$ is an ideal step function, input voltage $V_i(t)$ of a single-stage CMOS logic gate like NAND4 immediately goes from ground to operating voltage V_{dd} at $t = 0$. Integrating (6.15) with respect to t from 0- to 0+ yields

$$V_o(t)\Big|_{t=0_+} = (1 + k_c)\cdot V_{dd} \tag{6.16}$$

Note that the details of series-connected nMOSFET chain current I_n are not important

provided that its value is limited. The output voltage $V_o(t)$ jumps from its initial value V_{dd} to higher value $(1+k_c)\cdot V_{dd}$ at $t=0_+$. The overshoot of output voltage $V_o(t)$ is caused with the input-output coupling coefficient k_c given by (6.4). Obviously, there is no such overshoot if there is no input-output capacitance coupling.

Region two $(0_+ < t \le t_{as})$: In this region, the trivial parallel pMOSFET cluster is in cutoff mode and the series-connected nMOSFET is also in cutoff mode due to the suppressed input according to the modified input suppression technique discussed in Chapter 5. The chain current does not affect the output voltage. Thus output voltage keeps the same for this period, i.e.,

$$V_o(t) = V_o(t)\big|_{t=0_+} = (1+k_c)\cdot V_{dd} \tag{6.17}$$

This causes the mesa of output voltage if the step-input activation time is not zero.

Region three $(t_{as} < t \le t_{snd})$: In this region, input voltage V_i is equal to V_{dd} and output voltage V_o is larger than saturation voltage V_{snd} of a series-connected nMOSFET chain $(V_o \ge V_{snd})$. The parallel pMOSFET cluster is in cutoff mode and the series-connected nMOSFET chain is in saturation mode with characteristics curve of $V_{gs} = V_i = V_{dd}$. Time t_{snd} occurs when the output voltage V_o reaches its saturation voltage V_{snd}. From (6.10), the saturation voltage V_{snd} with input voltage equal to V_{dd} is expressed as

$$V_{snd} = R_n(V_{dd} - V_{an})^{b_n} \tag{6.18}$$

Substituting current expression of (6.9) in saturation mode with $V_{sn} = V_{snd}$ into (6.15) yields

$$\frac{dV_o}{dt} = -\frac{k_n \cdot V_{snd}^2}{2C_t} \cdot (1 + \lambda_n V_o) \tag{6.19}$$

Assume $\lambda_n > 0$. Solving the above differential equation of (6.19) with initial value of $V_o(t)\big|_{t=t_{as}} = (1+k_c)\cdot V_{dd}$ given by (6.17) at $t = t_{as}$ results in the following output expression.

$$V_o(t) = \left[\frac{1}{\lambda_n} + (1+k_c)\cdot V_{dd}\right] e^{-(t-t_{as})/\tau_{sat}} - \frac{1}{\lambda_n} \tag{6.20}$$

where the format of the solution is similar as the analytical solution of first-order RC circuit. The time constant τ_{sat} in saturation mode for a step input is

$$\tau_{sat} = \frac{2 \cdot C_t}{k_n \cdot \lambda_n \cdot V_{snd}^2} \tag{6.21}$$

We can rearrange (6.20) as the dependence of time t on output voltage V_o, which gives

$$t = t_{as} + \tau_{sat} \cdot \ln\left[\frac{1 + (1+k_c)\cdot \lambda_n \cdot V_{dd}}{1 + \lambda_n \cdot V_o}\right] \tag{6.22}$$

If the output saturation voltage V_{snd} of a series-connected nMOSFET chain given by (6.18) is less than half of the operating voltage V_{dd} $(V_{snd} < 0.5V_{dd})$, the propagation delay t_{dHL} of output voltage V_o from high to low for step input is obtained from (6.22).

$$t_{dHL} = t_{as} + \tau_{sat} \cdot \ln\left[\frac{1 + (1+k_c)\cdot \lambda_n \cdot V_{dd}}{1 + 0.5 \cdot \lambda_n \cdot V_{dd}}\right] \qquad (V_{snd} < 0.5V_{dd}) \tag{6.23}$$

The time t_{80}, corresponding to $V_o = 80\%V_{dd}$, can be easily obtained

$$t_{80} = t_{as} + \tau_{sat} \cdot \ln\left[\frac{1 + (1+k_c)\cdot \lambda_n \cdot V_{dd}}{1 + 0.8 \cdot \lambda_n \cdot V_{dd}}\right] \qquad (V_{snd} < 0.8V_{dd}) \tag{6.24}$$

The time t_{snd} occurs when the output voltage $V_o(t)$ falls to the drain-source saturation

voltage V_{snd} given in (6.18), which is

$$t_{snd} = t_{as} + \tau_{sat} \cdot \ln\left[\frac{1+(1+k_c)\cdot\lambda_n \cdot V_{dd}}{1+\lambda_n \cdot V_{snd}}\right] \tag{6.25}$$

Region four $(t \geq t_{snd})$: In this region, input voltage V_i is equal to V_{dd} and output voltage V_o is smaller than output saturation voltage V_{snd} of a series-connected nMOSFET chain with input voltage equal to V_{dd}, i.e. $(V_o \leq V_{snd})$. The parallel pMOSFET cluster is in cutoff mode and the series-connected nMOSFET chain is in linear mode with characteristics curve of $V_i = V_{dd}$. The transient differential equation of (6.15) becomes

$$\frac{dV_o}{dt} = -\frac{k_n}{C_t} V_o \left(V_{snd} - \frac{V_o}{2}\right)(1+\lambda_n V_{snd}) \tag{6.26}$$

which can be rearranged as

$$\frac{dV_o}{V_o \cdot (V_{snd} - 0.5 \cdot V_o)} = -\frac{k_n (1+\lambda_n V_{snd})}{C_t} dt \tag{6.27}$$

Factoring the left hand side of the above equation yields

$$\left(\frac{1}{V_o} + \frac{0.5}{V_{snd} - 0.5 \cdot V_o}\right) dV_o = -\frac{dt}{\tau_{lin}} \tag{6.28}$$

where τ_{lin} is the time constant of the series-connected MOSFET chain in the linear mode for a step input is

$$\tau_{lin} = \frac{C_t}{k_n \cdot V_{snd}(1+\lambda_n V_{snd})} \tag{6.29}$$

Integrating (6.28) with respect to variables t and V_o from $t = t_{snd}$ and the corresponding initial output voltage of $V_o = V_{snd}$ yields

$$t = t_{snd} + \tau_{lin} \cdot \ln\left(\frac{2V_{snd}}{V_o} - 1\right) \qquad (V_o \le V_{snd}) \quad (6.30)$$

Or

$$V_o = 2V_{snd} \cdot \frac{1}{1 + e^{(t - t_{snd})/\tau_{lin}}} \qquad (t \ge t_{snd}) \quad (6.31)$$

If the output saturation voltage V_{snd} of a series-connected nMOSFET chain given by (6.18) is larger than half of the operating voltage V_{dd} $(V_{snd} > 0.5V_{dd})$, then the propagation delay t_{dHL} of output voltage V_o from high to low is obtained from (6.30).

$$t_{dHL} = t_{snd} + \tau_{lin} \cdot \ln\left(\frac{4 \cdot V_{snd}}{V_{dd}} - 1\right) \qquad (V_{snd} > 0.5V_{dd}) \quad (6.32)$$

The time t_{20}, corresponding to $V_o = 20\%V_{dd}$, can be easily obtained

$$t_{20} = t_{snd} + \tau_{lin} \cdot \ln\left(\frac{10 \cdot V_{snd}}{V_{dd}} - 1\right) \qquad (0.2V_{dd} < V_{snd}) \quad (6.33)$$

Therefore the falling time t_f is calculated from the time of $V_o = 80\% \cdot V_{dd}$ to the time of $V_o = 20\% \cdot V_{dd}$ divided by 0.6, which is

$$t_f = \frac{1}{0.6}(t_{20} - t_{80}) = \frac{1}{0.6}[(t_{20} - t_{sat}) + (t_{sat} - t_{80})] \quad (6.34)$$

where $(t_{20} - t_{snd})$ and $(t_{snd} - t_{80})$ are obtained from (6.33), (6.24), and (6.25), respectively

$$t_{20} - t_{snd} = \tau_{lin} \cdot \ln\left(\frac{10 \cdot V_{snd}}{V_{dd}} - 1\right) \quad (6.35)$$

and

$$t_{snd} - t_{80} = \tau_{sat} \ln\left(\frac{1 + 0.8 \cdot \lambda_n \cdot V_{dd}}{1 + \lambda_n \cdot V_{snd}}\right) \quad (6.36)$$

Therefore the falling time t_f of (6.34) becomes

$$t_f = \frac{1}{0.6}\left[\tau_{lin}\cdot\ln\left(\frac{10\cdot V_{snd}}{V_{dd}}-1\right)+\tau_{sat}\ln\left(\frac{1+0.8\cdot\lambda_n\cdot V_{dd}}{1+\lambda_n\cdot V_{snd}}\right)\right] \tag{6.37}$$

6.2.3 Results and Discussions

In the previous subsection, we have derived the analytical expressions of output voltages and propagation delays of a complex CMOS logic gate for a step input. In this subsection, we take a specific CMOS NAND4 gate for a step input as an example. The output voltages and propagation delays of the CMOS NAND4 gate as a function of external load capacitance for a step input are directly obtained from those closed-form expressions derived in the previous subsection. The manufacture technology is 0.13μm. The pMOSFET width of the CMOS NAND4 gate is 8μm and the nMOSFET width is 16μm. The static *I–V* characteristics of the pMOSFET and the nMOSFET of the CMOS NAND4 gate are characterized with SPICE. The SMCM parameters of the pMOSFET and nMOSFET of the CMOS NAND4 gate are obtained by fitting the static *I–V* characteristics from SPICE BSIM3. They are listed in Table 6.3.

Fig. 6.2 shows output voltage waveforms of a NAND4234 gate with the active inputs of V_{i2}, V_{i3}, and V_{i4} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. There is a pair of traces for the external load capacitance of 50fF: the left one for the fast initial state (FIS) and the right one for the slow initial state (SIS). We can see an obvious mesa in the voltage waveform for slow initial state (SIS) due to its SIS step-input activation time of 15ps. There is a voltage overshoot, but not a mesa in the fast initial state (FIS) because of its FIS step-input activation time of 0ps. Both phenomena are well

explained by the modified input suppression technique developed in Section 5.5.

Fig. 6.3 plots output voltage waveforms of a NAND41 gate with the active input of V_{i1} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one because both initial states have the same step-input activation time of 18ps.

Fig. 6.4 illustrates output voltage waveforms of a NAND423 gate with the active inputs of V_{i2}, and V_{i3} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are demonstrated in solid lines and in dashed lines. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively (from left to right). There is a pair of traces for each external load capacitance: the left one for fast initial state (FIS) and the right one for slow initial state (SIS). A mesa exists in the voltage waveforms for both slow initial state (SIS) and fast initial state (FIS). Both are accurately explained by the modified input suppression technique developed in Section 5.5 because of the SIS step-input activation time of 15ps and the FIS step-input activation time of 10ps.

Fig. 6.5 displays output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. Note that the left family of traces is for fast initial state (FIS) and the right one for slow

initial state (SIS). For each family the external capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively (from left to right). Note there is a mesa in the voltage waveforms for slow initial state (SIS) due to its SIS step-input activation time of 18ps. No mesa exists in fast initial state (FIS) due to its FIS step-input activation time of 0ps. Both are accurately explained by the modified input suppression technique in Section 5.5.

From the above figures the analytical expressions of output voltages derived from Subsection 6.2.2 are very accurate to reproduce the transient output voltages for a step input, even including the voltage overshoot and voltage mesa. The propagation delay of a CMOS logic gate for a step input is accurately predicted in Subsection 6.2.2 based on SMCM. In practice, a ramp input , instead of a step input, is more realistic. Generally speaking, it is necessary to incorporate input transition time into propagation delay models of a CMOS logic gate, which will be investigated in the next two sections.

6.3 TIMING ANALYSIS OF A SINGLE-STAGE CMOS LOGIC GATE FOR A RAMP INPUT WITHOUT A TRIVIAL PARALLEL MOSFET CLUSTER

The propagation delay time is the time difference between 50% input transition and 50% output transition. In the previous section, we have derived the timing model of a single-stage CMOS logic gate for a step input where the effect of input transition time is ignored. In fact, there is always a non-zero transition time for actual input voltage. The effect of input transition time on propagation delay of a single-stage CMOS logic gate has to be incorporated into timing model. In this section, we focus on the timing analysis of a single-stage CMOS logic gate for a ramp input. For simplicity, the parallel pMOSFET cluster of a single-stage CMOS NAND4 gate is ignored for a rising ramp. The single-

stage CMOS NAND4 gate is modeled with a series-connected nMOSFET chain. Both the reverse discharging current from output node to power during output voltage overshoot and the short-circuit current are neglected in the analysis $(I_p = 0)$.

6.3.1 Operation of a Series Connected nMOSFET Chain for a Ramp Input

In this subsection, we investigate the operation of a series-connected nMOSFET chain for a ramp input. The input voltage in the section is a ramp function, i.e.

$$V_i(t) = \frac{V_{dd}}{\tau}[r(t) - r(t-\tau)] \text{ or } V_i(s) = V_{dd}[r(s) - r(s-1)] \tag{6.38}$$

where $r(\cdot)$ is a unit ramp function, τ is the rising time of the input voltage, and s is the normalized time. Fig. 6.6 shows the seven operation regions of a complex NAND434 gate modeled with a series-connected MOSFET chain n434 with the active inputs of V_{i3} and V_{i4} for a rising ramp input. The typical discharging traces from SPICE BSIM3 are also plotted. Each pair of traces corresponds to the same external load capacitance for fast initial state (FIS) and slow initial state (SIS), respectively. The seven operation regions are denoted with R1, R2, R3, R4, R5, R6, and R7, where discharging trajectory traces of the series-connected nMOSFET chain of a single-stage CMOS NAND4 gate may experience. The seven operation regions of a series-connected MOSFET chain are described in details as follows. R1 is region one where it is in cutoff mode. R2 is region two where it is in saturation mode with input voltage less than half of operation voltage. R3 is region three where it is in linear mode with input voltage less than half of operation voltage. R4 is region four where it is in saturation mode with input voltage larger than half of operation voltage. R5 is region five where it is in linear mode with input voltage larger than half of operation voltage. R6 is region six where it is in saturation mode with

input voltage equal to operating voltage V_{dd}. R7 is region seven where it is in linear mode with input voltage equal to operating voltage V_{dd}. The interfaces between R2 and R3 or between R4 and R5 are expressed by (6.10). We solve for the output voltage expressions according to these seven operation regions.

6.3.2 Output Voltages of a Dominant Series-Connected MOSFET Chain for a Ramp Input without a Trivial Parallel MOSFET Cluster

Region one $(0 \le s \le u_{tn} < 0.5$ and $u_o \ge 1)$: In this region, the series-connected nMOSFET chain is cutoff $(I_n = 0)$. The input-output coupling current I_{C_m}, which may cause an overshoot of output voltage output voltage V_o, has the main impact on the output voltage V_o in this region. The differential equation to describe the output transition is as follows:

$$\frac{dV_o}{ds} = k_c V_{dd} \tag{6.39}$$

The initial value of output voltage V_o at initial time zero in this region is V_{dd} $\left(V_{1i} = V_o(s)\big|_{s=0} = V_{dd}\right)$. Therefore, the output voltage V_o in this region is obtained:

$$V_o(s) = V_{dd}(1 + k_c s) \tag{6.40}$$

The output voltage $V_o(s)$ clearly rises with the normalized time s or the time t in this region. That gives rise to the phenomenon of an overshoot of output voltage V_o, caused by the input-output coupling current I_{C_m}.

Region two $(u_{tn} \le s \le 0.5$ and $u_o > u_{sn})$: In this region, the series nMOSFET chain is in saturation mode. The differential equation to describe the output transition and the current

expressions for the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} - \dfrac{\tau}{C_t} I_n \\ I_n = \dfrac{1}{2} k_n V_{sn}^2 (1 + \lambda_n V_o) \\ V_{sn} = Q_n V_{dd} (s - u_{tn}) \end{cases} \tag{6.41}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{2i} and s_{2i}. The current I_n through the series-connected nMOSFET chain dominates the discharging process in this region. As the first step, only the series-connected nMOSFET chain is analyzed and the differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{2C_t} V_{sn}^2 (1 + \lambda_n V_o) \tag{6.42}$$

For simplicity, set the variable V_{qn}

$$V_{qn} = \frac{\tau k_n Q_n^2 V_{dd}^2}{C_t} \tag{6.43}$$

Substituting the equation V_{sn} of (6.40) into the differential equation and using the variable V_{qn} yield

$$\frac{dV_o}{ds} = -\frac{V_{qn}}{2} (s - u_{tn})^2 (1 + \lambda_n V_o) \tag{6.44}$$

The solution is expressed as follows:

$$V_{o2} = \left(\frac{1}{\lambda_n} + V_{2i} \right) \exp[-f_2(s)] - \frac{1}{\lambda_n} \tag{6.45}$$

In (6.45), V_{2i} is the initial output voltage V_o in region two and the function $f_2(s)$ is given by

$$f_2(s)=\frac{V_{qn}\lambda_n}{6}\left\{(s-u_{tm})^3-(s_{2i}-u_{tm})^3\right\} \tag{6.46}$$

where s_{2i} is the initial normalized time in this region. The second step is that the current of series connected nMOSFET chain and coupling capacitance current I_{C_m} between input net and output net are included in the analysis. Then, the output voltage $V_o(t)$ in this region is

$$V_o = k_c V_{dd}(s-s_{2i})+V_{o2} \tag{6.47}$$

Region three ($u_{tm} \le s \le 0.5$ and $u_o \le u_{sn}$): In this region, the switching series-connected nMOSFET chain is in the linear. The differential equation to describe the output transition and the current expressions for the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} - \dfrac{\tau}{C_t} I_n \\ I_n = k_n\left(V_{sn}-\dfrac{V_o}{2}\right)V_o(1+\lambda_n V_{sn}) \approx k_n\left(V_{sn}-\dfrac{V_o}{2}\right)V_o \\ V_{sn} = Q_n V_{dd}(s-u_{tn}) \end{cases} \tag{6.48}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{3i} and s_{3i}. In order to easily derive the closed-form expression of output voltage $V_o(t)$ without the loss of accuracy, the term of $(1+\lambda_n V_{sn})$ in the current I_n is ignored since the effect of $\lambda_n V_{sn}$ is small in this region. The series-connected nMOSFET chain current I_n dominates the discharging process in this region. As the first step, only series-connected nMOSFET chain is used in the derivation. The differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{C_t}\left(V_{sn} - \frac{V_o}{2}\right)V_o \tag{6.49}$$

This differential equation of (6.49), a Bernoulli's equation, can be obtained by employing the similar procedures in Region three of Section 4.4. For simplicity, set the following variables of m_n and σ_n^2:

$$\begin{cases} m_n = u_{tn} \\ \sigma_n^2 = \dfrac{C_t}{\tau k_n Q_n V_{dd}} \end{cases} \tag{6.50}$$

Therefore, the output voltage without considering the input-output coupling current is

$$V_{o3} = \frac{1}{\dfrac{\tau k_n}{2C_t}P(s, m_n, \sigma_n) + \left(-\dfrac{\tau k_n}{2C_t}P(s_{3i}, m_n, \sigma_n) + \dfrac{1}{V_{3i}}\right)\exp\left[\dfrac{(s + s_{3i} - 2m_n)(s - s_{3i})}{2\sigma_n^2}\right]} \tag{6.51}$$

where $P(x, m, \sigma)$ is defined in (III.9a) of Appendix III. The second step is that current of series-connected nMOSFET chain and coupling capacitance current between input and output are involved in the analysis. Then, the solution is expressed as follows:

$$V_o = k_c V_{dd}(s - s_{3i}) + V_{o3} \tag{6.52}$$

Region four ($0.5 < s \le 1$, and $u_o > u_{sn}$): In this region, the series-connected nMOSFET chain is in saturation mode with the input voltage V_i greater than the half of V_{dd}. The differential equation to describe the output transition and the current expression for the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} - \dfrac{\tau}{C_t} I_n \\ I_n = \dfrac{1}{2} k_n V_{sn}^2 (1 + \lambda_n V_o) \\ V_{sn} = R_n V_{dd}^{b_n} (s - u_{an})^{b_n} \end{cases} \tag{6.53}$$

The current of series-connected nMOSFET chain dominates the discharging process in this region. In this case, the derivation of output is similar to that in Region two except for the different chain current. As the first step, only chain current is considered. The differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{2C_t} V_{sn}^2 (1 + \lambda_n V_o) \tag{6.54}$$

For simplicity, set the variable V_{rn}:

$$V_{rn} = \frac{\tau k_n R_n^2 V_{dd}^{2b_n}}{C_t} \tag{6.55}$$

Substituting the equations of V_{sn} into the differential equation and using the variable V_{rn} yield

$$\frac{dV_o}{ds} = -\frac{V_{rn}}{2} (s - u_{an})^{2b_n} (1 + \lambda_n V_o) \tag{6.56}$$

The solution of the above equation is

$$V_{o4} = \left(\frac{1}{\lambda_n} + V_{4i} \right) \exp[-f_4(s)] - \frac{1}{\lambda_n} \tag{6.57}$$

In (6.57), the function $f_4(s)$ is given by

$$f_4(s) = \frac{V_{rn} \lambda_n}{2(2b_n + 1)} \left\{ (s - u_{an})^{2b_n + 1} - (s_{4i} - u_{an})^{2b_n + 1} \right\} \tag{6.58}$$

The second step is that nMOSFET chain current and coupling capacitance current are

included. Then, we have

$$V_o = k_c V_{dd}(s - s_{4i}) + V_{o4} \tag{6.59}$$

Region five ($0.5 \le s \le 1$ and $u_o \le u_{sn}$): In this region, the series nMOSFET chain is in the linear mode. Then the differential equation to describe the output transition and the current expression for the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} - \dfrac{\tau}{C_t} I_n \\ I_n = k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o\left(1 + \lambda_n V_{sn}\right) \approx k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o \\ V_{sn} = R_n V_{dd}^{b_n}\left(s - u_{an}\right)^{b_n} \approx k_h\left(s - b_h\right) \end{cases} \tag{6.60}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{5i} and s_{5i}. For simplicity, the term of $(1 + \lambda_n V_{sn})$ in the current I_n is ignored since the effect of $\lambda_n V_{sn}$ is small in this region. The saturation voltage V_{sn} in this region is approximated by the linear equation with the parameters k_h and b_h as follows:

$$\begin{cases} k_h = 2R_n V_{dd}^{b_n}\left[(1 - u_{an})^{b_n} - (0.5 - u_{an})^{b_n}\right] \\ b_h = 1 - \dfrac{0.5(1 - u_{an})^{b_n}}{(1 - u_{an})^{b_n} - (0.5 - u_{an})^{b_n}} \end{cases} \tag{6.61}$$

The series-connected nMOSFET chain current dominates the discharging process in this region. As the first step, only nMOSFET is used in the derivation. The differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{C_t}\left(V_{sn} - \frac{V_o}{2}\right)V_o \tag{6.62}$$

This differential equation of (6.62), a Bernoulli's equation, is solved by employing the

similar method in Region three of Section 4.4. For simplicity, set the following variable

$$\sigma_h^2 = \frac{C_t}{\tau k_n k_h} \tag{6.63}$$

Therefore, the output voltage without considering the input-output coupling current is

$$V_{o5} = \frac{1}{\frac{\tau k_n}{2C_t} P(s, b_h, \sigma_h) + \left(\frac{1}{V_{5i}} - \frac{\tau k_n}{2C_t} P(s_{5i}, b_h, \sigma_h) \right) \exp\left[\frac{(s + s_{5i} - 2b_h)(s - s_{5i})}{2\sigma_h^2} \right]} \tag{6.64}$$

where $P(x, m, \sigma)$ is defined in (III.9a) of Appendix III. The second step is that nMOSFET chain current and coupling capacitance current between input and output are involved in the analysis. Then, the solution is expressed as follows:

$$V_o = k_c V_{dd}(s - s_{5i}) + V_{o5} \tag{6.65}$$

In (6.64) and (6.65), V_{5i} and s_{5i} are the initial values of output voltage V_o and the normalized time t in the region, respectively.

Region six ($s \geq 1$ or $t \geq \tau$ and $u_o \geq u_{snd}$): In this region, the series-connected nMOSFET chain is in the saturation mode with input voltage $V_i = V_{dd}$. Then the differential equation to describe the output transition and the current expression for the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{dt} = -\dfrac{1}{C_t} I_n \\ I_n = \dfrac{1}{2} k_n V_{snd}^2 (1 + \lambda_n V_o) \\ V_{snd} = R_n V_{dd}^{b_n} (1 - u_{an})^{b_n} \end{cases} \tag{6.66}$$

The initial values of output voltage $V_o(t)$ and time t in this region are V_{6i} and t_{6i}. Then the output waveform $V_o(t)$ is

$$V_o = \left(\frac{1}{\lambda_n} + V_{6i}\right)\exp\left[-\frac{t - t_{6i}}{\tau_{sat}}\right] - \frac{1}{\lambda_n} \tag{6.67}$$

or

$$t = t_{6i} + \tau_{sat}\ln\left(\frac{1+\lambda_n V_{6i}}{1+\lambda_n V_o}\right) \tag{6.68}$$

where t_{6i} and V_{6i} are initial values of time t and output voltage V_o in the region. The time constant τ_{sat} in this region is the same as in (6.21), i.e.,

$$\tau_{sat} = \frac{2 \cdot C_t}{k_n \cdot \lambda_n \cdot V_{snd}^2} \tag{6.69}$$

In (6.69), the output saturation voltage V_{snd} is given by (6.66).

Region seven ($s \geq 1$ or $t \geq \tau$ and $u_o \leq u_{snd}$): In this region the series-connected nMOSFET chain is in the linear mode with input voltage $V_i = V_{dd}$. Then the differential equation to describe the output transition and the current expression for the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{dt} = -\dfrac{1}{C_t} I_n \\ I_n = k_n\left(V_{snd} - \dfrac{V_o}{2}\right)V_o\left(1+\lambda_n V_{snd}\right) \\ V_{snd} = R_n V_{dd}^{b_n}\left(1-u_{an}\right)^{b_n} \end{cases} \tag{6.70}$$

For generality, the initial values of output voltage $V_o(t)$ and time t in this region are V_{7i} and t_{7i}. Then the output waveform $V_o(t)$ is

$$V_o = \frac{2V_{snd}}{1+\left(2V_{snd}/V_{7i} - 1\right)\exp\left[\left(t - t_{7i}\right)/\tau_{lin}\right]} \tag{6.71}$$

or

$$t = t_{7i} + \tau_{lin} \ln\left(\frac{V_{7i}}{V_o}\frac{2V_{snd} - V_o}{2V_{snd} - V_{7i}}\right) \tag{6.72}$$

where τ_{lin} is the time constant of the series-connected MOSFET chain in the linear mode for a step input is the same as in (6.28), i.e.,

$$\tau_{lin} = \frac{C_t}{k_n \cdot V_{snd}(1 + \lambda_n V_{snd})} \tag{6.73}$$

In (6.73), the output saturation voltage V_{snd} is given by (6.70).

6.3.3 Effect of Initial States of Parasitic Capacitances of Intermediate Nodes

The above analysis of output voltage waveform $V_o(t)$ is based on that there is no effect of initial states of intermediate nodes, i.e., the step-input activation time is zero $(t_{as} = 0)$. From the analysis of the modified input suppression technique the ramp-input activation time t_a given by (5.9) is equal to $u_{tn}\tau$. However, a single-stage CMOS logic gate may have non-zero step-input activation time $(t_{as} \neq 0)$ due to the initial state of parasitic capacitances of intermediate nodes as investigated in Chapter five. For non-zero step-input activation time $(t_{as} \neq 0)$ the modified input suppression technique yields that the ramp-input activation time t_a given by (5.9) is larger than $u_{tn}\tau$. In this case the timing analysis discussed in the previous subsection has to be corrected. For convenience, regions mentioned below refer to the corresponding ones defined in the previous subsection.

Case one $(u_{tn}\tau \leq t_a < \tau)$: Before the ramp-input activation time t_a is reached, i.e., $(0 \leq t < t_a)$, the series-connected nMOSFET chain is not activated $(I_n = 0)$. This is the

same situation as Region one. The differential equation of (6.39) to describe the output voltage $V_o(t)$ caused with the input-to-output coupling current becomes

$$\frac{dV_o}{dt} = k_c \frac{V_{dd}}{\tau} \qquad (0 \le t < t_a) \quad (6.74)$$

The initial value of output voltage V_o at initial time zero in this region is V_{dd}. The solution of the above equation is

$$V_o(t) = V_{dd}(1 + k_c t/\tau) \qquad (0 \le t < t_a) \quad (6.75)$$

After the ramp-input activation time t_a is reached $(t > t_a)$, the series-connected nMOSFET chain is activated and is in saturation mode. The output voltage $V_o(t)$ enters either Region two or Region four depending on the activation time t_a. Then the discussions in the previous subsection are valid. If $(\tau u_{tn} \le t_a < 0.5\tau)$ the series-connected nMOSFET chain enters Region two with the initial time $t_{2i} = t_a$ and initial output voltage $V_{2i} = V_o(t_a)$. If $(0.5\tau \le t_a < \tau)$ the series-connected nMOSFET chain then goes to Region four with the initial time $t_{4i} = t_a$ and initial output voltage $V_{4i} = V_o(t_a)$.

Case two: $(\tau \le t_a)$. Before the activation time t_a $(0 \le t < t_a)$, the series-connected nMOSFET chain is not activated $(I_n = 0)$. When $(0 \le t < \tau)$, the output voltage $V_o(t)$ caused with the input-to output coupling current is given by (6.75). At the time of τ, the output voltage $V_o(t)$ is

$$V_o(t)\big|_{t=\tau} = V_{dd}(1 + k_c) \qquad (6.76)$$

When $(\tau \le t < t_a)$, the series-connected nMOSFET chain is still not activated $(I_n = 0)$ and there is no input-to output coupling current. Therefore, the output voltage $V_o(t)$ stays

constant in this period. We have

$$V_o(t) = V_{dd}(1 + k_c) \qquad (\tau \le t < t_a) \quad (6.77)$$

This clearly explains the occurrence of the mesa of output voltage caused with the initial state of parasitic capacitances of intermediate nodes. After the activation time t_a in this case $(t \ge t_a)$, the series-connected nMOSFET chain is activated. The output voltage $V_o(t)$ enters Region six with the initial time $t_{6i} = t_a$ and initial output voltage $V_{6i} = V_{dd}(1 + k_c)$.

6.3.4 Results and Discussions

In the previous subsection, the CMOS NAND4 gate for a rising ramp input is modeled with a series-connected nMOSFET chain. The trivial parallel pMOSFET cluster for a rising ramp input is ignored. From the series-connected nMOSFET chain modeled with SMCM, we have derived the analytical expressions of output voltages of a CMOS NAND4 gate for a rising ramp input. In this subsection, we apply these closed-form expressions to a specific CMOS NAND4 gate, which is the same as the one in Section 6.2.3. The SMCM parameters of the series-connected nMOSFET chain are listed in Table 6.2. The resulting output voltage and propagation delay are discussed. The consequences of the ignorance of the trivial parallel pMOSFET cluster are explored.

Figs. 6.7-6.9 plot output voltage waveforms of a NAND42 gate with the active input of V_{i2} for external load capacitances of 10fF, 100fF, and 400fF, respectively. The NAND42 gate is modeled with a series-connected MOSFET chain 'n42'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Due to the same step-input activation time of 15ps for slow initial

state (SIS) and fast initial state (FIS) the voltage waveforms each load capacitance merge into one so that we cannot see the difference. The proposed output waveforms from SMCM and the output responses from SPICE BSIM3 agree with each other excellently.

Figs. 6.10-6.12 display output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for external load capacitances of 10fF, 100fF, and 400fF, respectively. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots are exaggerated with no trivial parallel pMOSFET cluster. Of course, the effect of voltage overshoots is reduced with large external load capacitance such as in Fig. 6.12.

Fig. 6.13 plots FIS propagation delays versus external load capacitances for a NAND42 gate. The NAND42 gate is modeled with a series-connected MOSFET chain 'n42'. The propagation delays from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the propagation delays of SIS and FIS for the same load capacitance and input rising time are the same as of FIS due to their same step-input activation times.

Figs. 6.14-6.15 display propagation delays of a NAND4124 gate versus external load capacitances for FIS and SIS, respectively. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The propagation delays from SMCM in

Section 6.3 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). The FIS propagation delays from the proposed method in Section 6.3 are pretty accurate comparing with those from SPICE BSIM3 due to zero step-input activation time of a NAND4123. The SIS propagation delays from the proposed method in Section 6.3 are deviated from SPICE BSIM3 due to non-zero step-input activation time of a NAND4123. The part of error in propagation delays is caused with the exaggerated voltage overshoots.

Figs. 6.16-6.17 show the correlation of proposed propagation delay from SMCM in Section 6.3 with propagation delay from SPICE BSIM3 of fast initial state (FIS) and slow initial state (SIS), respectively. The NAND4 gate is modeled with a series-connected MOSFET chain. The two dashed lines are ±5% of propagation delays from SPICE BSIM3. There are 375 simulation points which cover all the 15 sorts of active inputs. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively. Most of them are in good agreement. The FIS propagation delays from the proposed method in Section 6.3 are pretty accurate comparing with those from SPICE BSIM3. The SIS propagation delays from the proposed method in Section 6.3 are also accurate from SPICE BSIM3. The part of error in propagation delays is caused with the exaggerated voltage overshoots, especially for slow initial state (SIS). We will develop a complete timing model with both the dominating series-connected nMOSFET chain and the trivial parallel pMOSFET cluster in the next section.

6.4 COMPREHENSIVE TIMING ANALYSIS OF A SINGLE-STAGE CMOS LOGIC GATE WITH AN EQUAL INPUT PATTERN

In Section 6.2, we have derived the timing model of a single-stage CMOS logic gate for a step input where the effect of input transition time is ignored. The transient properties of a single-stage CMOS logic gate for a ramp input without short-circuit current are analyzed in Section 6.3. In this section, we analyze comprehensive propagation delay of a single-stage CMOS logic gate where input is a ramp function and short circuit current is considered also. The transition trajectory traces of both series-connected nMOSFET chain and parallel pMOSFET cluster of a single-stage CMOS logic gate are first examined, which gives rise to eleven possible operation regions according to distinct operation modes of series-connected nMOSFET chain and parallel pMOSFET cluster for each region. The differential equation controlling the transition of a single-stage CMOS logic gate is analytically solved for each operation region.

6.4.1 Operation of a Single-Stage CMOS Logic Gate with an Equal Input Pattern

In this subsection, we investigate the operation of a single-stage nMOSFET chain for a ramp input. The input voltage is a ramp function, i.e.,

$$V_i(t) = \frac{V_{dd}}{\tau}[r(t) - r(t-\tau)] \quad \text{or} \quad V_i(s) = V_{dd}[r(s) - r(s-1)] \tag{6.78}$$

where $r(\cdot)$ is a unit ramp function, τ is the rising time of the input voltage, and s is the normalized time. In order to completely analyze the output transition of a single-stage CMOS NAND4 gate with a rising ramp input, Fig. 6.18 shows the eleven operation regions denoted with R1, R2, R3, R4, R5, R6, R7, R8, R9, R10, and R11, where discharging trajectory traces of a single-stage CMOS NAND4 logic gate may experience.

The typical discharging traces from SPICE BSIM3 are also plotted. Each pair of traces corresponds to the same external load capacitance for fast initial state (FIS) and slow initial state (SIS). The eleven separate operation regions correspond to the complete combinations of the different operation modes of the series-connected nMOSFET chain and the parallel pMOSFET cluster due to three modes of linear, saturation, and cutoff. The eleven operation regions are described in details as follows. R1 is region one where series-connected nMOSFET chain is in cutoff mode and parallel switching pMOSFET cluster is in linear mode. R2 is region two where series-connected nMOSFET chain is in saturation mode with input voltage less than half of operating voltage and parallel switching pMOSFET cluster is in linear mode. R3 is region three where series-connected nMOSFET chain is in saturation mode with input voltage less than half of operating voltage and parallel switching pMOSFET cluster is in saturation mode. R4 is region four where series-connected nMOSFET chain is in linear mode with input voltage less than half of operating voltage and parallel switching pMOSFET cluster is in saturation mode. R5 is region five where series-connected nMOSFET chain is in saturation mode with input voltage larger than half of operating voltage and parallel switching pMOSFET cluster is in linear mode. R6 is region six where series-connected nMOSFET chain is in saturation mode with input voltage larger than half of operating voltage and parallel switching pMOSFET cluster is in saturation mode. R7 is region seven where series-connected nMOSFET chain is in linear mode with input voltage larger than half of operating voltage and parallel switching pMOSFET cluster is in saturation mode. R8 is region eight where series-connected nMOSFET chain is in saturation mode with input voltage larger than half of operating voltage and parallel switching pMOSFET cluster is

in cutoff mode. R9 is region nine where series-connected nMOSFET chain is in linear mode with input voltage larger than half of operating voltage and parallel switching pMOSFET cluster is in cutoff mode. R10 is region ten where series-connected nMOSFET chain is in saturation mode with input voltage equal to operating voltage V_{dd} and parallel switching pMOSFET cluster is in cutoff mode. R11 is region eleven where series-connected nMOSFET chain is in linear mode with input voltage equal to operating voltage V_{dd} and parallel switching pMOSFET cluster is in cutoff mode. The interfaces between R3 and R4 or between R6 and R7 or R8 and R9 are expressed by (6.10). The interfaces between R2 and R3 or between R5 and R6 are given by $V_{op} = V_{sp}/2$. We will discuss the output voltage expressions according to these eleven operation regions.

6.4.2 Output Voltages of a Complex CMOS Logic Gate with an Equal Input Pattern

Region one $(0 \le s \le u_{tn} < 0.5$ and $u_o \ge 1)$: In this region, the series-connected nMOSFET chain is cutoff $(I_n = 0)$ and the switching parallel pMOSFET cluster is in linear mode. The input-output coupling current I_{C_m}, which may cause an overshoot of output voltage output voltage V_o, has the main impact on the output voltage V_o in this region. The differential equation to describe the output transition and the current expression for the switching parallel pMOSFET cluster are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t} I_p \\ I_p = k_p V_{sp} V_{op} \\ V_{sp} = Q_p (V_{ip} - V_{tp}) = Q_p V_{dd} (1 - u_{tp} - s) \end{cases} \tag{6.79}$$

The initial value of output voltage V_o at initial time zero in this region is V_{dd}

$\left(V_{1i} = V_o(s)\big|_{s=0} = V_{dd}\right)$. Substituting the current expression of I_p in (6.79) into the differential equation of (6.79) yields

$$\frac{dV_{op}}{ds} = -\frac{\tau k_p}{C_t} V_{sp} V_{op} - k_c V_{dd} \tag{6.80}$$

In (6.80), the variable V_o to be solved in the original differential equation of (6.79) is changed to the variable $V_{op}\left(= V_{dd} - V_o\right)$ with its initial value of zero $\left(V_{op}\big|_{s=0} = 0\right)$. Equation (6.80) is a first-order linear differential equation whose solution is analytically expressed as

$$V_{op} = \exp\left(\int_0^s -\frac{\tau k_p}{C_t} V_{sp}\, ds\right) \cdot \left[\int_0^s \left(-k_c V_{dd}\right) \exp\left(\int_0^s \frac{\tau k_p}{C_t} V_{sp}\, ds\right) ds\right] \tag{6.81}$$

For simplicity, set the following variables:

$$\begin{cases} m_p = 1 - u_{tp} \\ \sigma_p^2 = \dfrac{C_t}{\tau k_p Q_p V_{dd}} \end{cases} \tag{6.82}$$

Substituting the variable V_{sp} in (6.79) into (6.81) and using the variables m_p and σ_p^2 yield

$$V_{op} = -k_c V_{dd} \exp\left[\frac{\left(s - m_p\right)^2}{2\sigma_p^2}\right] \cdot \int_0^s \left\{\exp\left[-\frac{\left(s - m_p\right)^2}{2\sigma_p^2}\right]\right\} ds \tag{6.83}$$

By using the expression (III.15b) in Appendix III with the condition of $0 \le s < m_p$, the above expression becomes

$$V_{op} = -k_c V_{dd} \exp\left[\frac{(s-m_p)^2}{2\sigma_p^2}\right] \left\{R(s,m_p,\sigma_p)\exp\left[-\frac{(s-m_p)^2}{2\sigma_p^2}\right] - R(0,m_p,\sigma_p)\exp\left[-\frac{m_p^2}{2\sigma_p^2}\right]\right\} \tag{6.84}$$

where the function $R(t,m,\sigma)$ is defined in (III.9b) of Appendix III. Therefore, the output voltage V_o in this region is obtained:

$$V_o(s) = V_{dd} + k_c V_{dd}\left\{R(s,m_p,\sigma_p) - R(0,m_p,\sigma_p)\exp\left(-\frac{s(2m_p-s)}{2\sigma_p^2}\right)\right\} \tag{6.85}$$

Since $R(s,m_p,\sigma_p) \geq R(0,m_p,\sigma_p)$ with $0 \leq s \leq u_{tn} < m_p$, the output voltage $V_o(s)$ clearly rises with the normalized time s or the time t in this region. That is the phenomena of an overshoot of output voltage V_o, caused by the input-output coupling current I_{C_m}.

Region two ($u_{tn} \leq s \leq 0.5$ and $(1-u_o) \leq u_{sp}/2$): In this region, the series-connected nMOSFET chain is in saturation mode and the parallel switching pMOSFET cluster is in linear mode. The differential equation to describe the output voltage and the current expressions for the switching parallel pMOSFET cluster and the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t}(I_p - I_n) \\ I_p = k_p V_{sp} V_{op} \\ I_n = \dfrac{1}{2}k_n V_{sn}^2(1+\lambda_n V_o) \approx \dfrac{1}{2}k_n V_{sn}^2(1+\lambda_n V_{dd}) \\ V_{sp} = Q_p(V_{ip} - V_{tp}) = Q_p V_{dd}(1-u_{tp}-s) \\ V_{sn} = Q_n(V_i - V_{tn}) = Q_n V_{dd}(s-u_{tn}) \end{cases} \tag{6.86}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{2i}

and s_{2i}. In order to analytically solve the differential equation, the term of $(1+\lambda_n V_o)$ in the current I_n is replaced with $(1+\lambda_n V_{dd})$ since output voltage V_o is close to V_{dd} in this region. Substituting the current expressions of I_p and I_n in (6.86) into the differential equation in (6.86) yields

$$\frac{dV_{op}}{ds} = -\frac{\tau k_p}{C_t} V_{sp} V_{op} + \frac{\tau k_n (1+\lambda_n V_{dd})}{2C_t} V_{sn}^2 - k_c V_{dd} \tag{6.87}$$

In (6.82), the variable V_o to be solved in the original differential equation is changed to the variable V_{op} defined as $(V_{dd} - V_o)$. The initial value of V_{op} at the normalized initial time of $s = s_{2i}$ in region two is $V_{op}\big|_{s=s_{2i}} (= V_{dd} - V_{2i})$. Equation (6.87) is a first-order linear differential equation whose solution is analytically expressed as

$$\begin{aligned} V_{op} = & \exp\left(\int_{s_{2i}}^{s} -\frac{\tau k_p}{C_t} V_{sp} ds \right) \cdot \\ & \left\{ \int_{s_{2i}}^{s} \left[\left(\frac{\tau k_n (1+\lambda_n V_{dd})}{2C_t} V_{sn}^2 - k_c V_{dd} \right) \exp\left(\int_{s_{2i}}^{s} \frac{\tau k_p}{C_t} V_{sp} ds \right) \right] ds + V_{op}\big|_{s=s_{2i}} \right\} \end{aligned} \tag{6.88}$$

In addition to the variables m_p and σ_p^2 defined in (6.82), set the two more variables V_{qn} and V_{qdn}:

$$\begin{cases} V_{qn} = \dfrac{\tau k_n Q_n^2 V_{dd}^2}{C_t} \\ V_{qdn} = V_{qn}(1+\lambda_n V_{dd}) \end{cases} \tag{6.89}$$

Substituting the saturation voltages of V_{sp} and V_{sn} in (6.86) into (6.88) and using the variables m_p, σ_p^2, and V_{qdn} yield

$$V_{op} = \exp\left[\frac{(s-m_p)^2-(s_{2i}-m_p)^2}{2\sigma_p^2}\right]\cdot \left\{\int_{s_{2i}}^{s}\left[\frac{V_{qdn}}{2}(s-u_{tn})^2-k_cV_{dd}\right]\exp\left[-\frac{(s-m_p)^2-(s_{2i}-m_p)^2}{2\sigma_p^2}\right]ds+(V_{dd}-V_{2i})\right\} \quad (6.90)$$

Simplifying the above equation gives

$$V_{op} = \exp\left[\frac{(s-m_p)^2}{2\sigma_p^2}\right] \left\{\int_{s_{2i}}^{s}\left[\frac{V_{qdn}}{2}(s-u_{tn})^2-k_cV_{dd}\right]\exp\left[-\frac{(s-m_p)^2}{2\sigma_p^2}\right]ds+(V_{dd}-V_{2i})\exp\left[-\frac{(s_{2i}-m_p)^2}{2\sigma_p^2}\right]\right\} \quad (6.91)$$

With the expression of (III.15b) in Appendix III with the condition of $s_{2i} \le s < m_p$, the above equation becomes

$$V_{op} = f_2(s)-[f_2(s_{2i})+V_{2i}-V_{dd}]\exp\left(-\frac{(s-s_{2i})(2m_p-s_{2i}-s)}{2\sigma_p^2}\right) \quad (6.92)$$

In (6.92), the function $f_2(x)$ is defined as:

$$f_2(x) = -\frac{(x+m_p-2u_{tn})\sigma_p^2V_{qdn}}{2}+\left[\frac{((m_p-u_{tn})^2+\sigma_p^2)V_{qdn}}{2}-k_cV_{dd}\right]R(x,m_p,\sigma_p) \quad (6.93)$$

Therefore, the output voltage V_o is

$$V_o(s) = V_{dd}-f_2(s)+[f_2(s_{2i})+V_{2i}-V_{dd}]\exp\left(-\frac{(s-s_{2i})(2m_p-s_{2i}-s)}{2\sigma_p^2}\right) \quad (6.94)$$

Region three ($u_{tn} \le s \le 0.5$, $(1-u_o) > u_{sp}/2$, and $u_o > u_{sn}$): In this region, both the switching parallel pMOSFET cluster and the series-connected nMOSFET chain are in saturation mode. The differential equation to describe the output transition and the current expressions for the switching parallel pMOSFET cluster and the series-connected

nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t}\left(I_p - I_n\right) \\ I_p = \dfrac{1}{2} k_p V_{sp}^2 \\ I_n = \dfrac{1}{2} k_n V_{sn}^2 \left(1 + \lambda_n V_o\right) \\ V_{sp} = Q_p\left(V_{ip} - V_{tp}\right) = Q_p V_{dd}\left(1 - u_{tp} - s\right) \\ V_{sn} = Q_n\left(V_i - V_{tn}\right) = Q_n V_{dd}\left(s - u_{tn}\right) \end{cases} \tag{6.95}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{3i} and s_{3i}. The current I_n through the series-connected nMOSFET chain dominates the discharging process in this region. As the first step, only the series-connected nMOSFET chain is analyzed and the differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{2C_t} V_{sn}^2 \left(1 + \lambda_n V_o\right) \tag{6.96}$$

Substituting the expression of V_{sn} in (6.95) into the differential equation and using the variable V_{qn} of (6.89) yield

$$\frac{dV_o}{ds} = -\frac{V_{qn}}{2}\left(s - u_{tn}\right)^2 \left(1 + \lambda_n V_o\right) \tag{6.97}$$

The solution of output voltage $V_o(t)$ is expressed as follows:

$$V_{o3} = \left(\frac{1}{\lambda_n} + V_{3i}\right) \exp\left[-f_3(s)\right] - \frac{1}{\lambda_n} \tag{6.98}$$

In (6.98), the function $f_3(s)$ is given by

$$f_3(s) = \frac{V_{qn}\lambda_n}{6}\left\{\left(s - u_{tn}\right)^3 - \left(s_{3i} - u_{tn}\right)^3\right\} \tag{6.99}$$

The second step is that currents of series-connected nMOSFET chain, parallel switching pMOSFET cluster, and coupling capacitance between input net and output net are included in the analysis. In order to simplify the expressions in the derivation, set the variable V_{qp} as:

$$V_{qp} = \frac{\tau k_p Q_p^2 V_{dd}^2}{C_t} \tag{6.100}$$

Then, the output voltage $V_o(t)$ in this region is

$$V_o(s) = k_c V_{dd}(s - s_{3i}) + \frac{V_{qp}}{6}\left[(1 - u_{tp} - s_{3i})^3 - (1 - u_{tp} - s)^3\right] + V_{o3} \tag{6.101}$$

Region four ($u_{tn} \le s \le 0.5$ and $u_o \le u_{sn}$): In this region, the switching series-connected nMOSFET chain is in the linear mode and the switching parallel pMOSFET cluster is in saturation mode. The differential equation to describe the output transition and the current expressions for the switching parallel pMOSFET cluster and the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t}(I_p - I_n) \\ I_p = \dfrac{1}{2} k_p V_{sp}^2 \\ I_n = k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o(1 + \lambda_n V_{sn}) \approx k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o \\ V_{sp} = Q_p(V_{ip} - V_{tp}) = Q_p V_{dd}(1 - u_{tp} - s) \\ V_{sn} = Q_n(V_i - V_{tn}) = Q_n V_{dd}(s - u_{tn}) \end{cases} \tag{6.102}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{4i} and s_{4i}. In order to simplify the derivation, the term of $(1 + \lambda_n V_{sn})$ in the current I_n is ignored without significant error since the effect of $\lambda_n V_{sn}$ is small in this region. The

series-connected nMOSFET chain current I_n dominates the discharging process in this region. As the first step, only series-connected nMOSFET chain is used in the derivation. The differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{C_t}\left(V_{sn} - \frac{V_o}{2}\right)V_o \tag{6.103}$$

This differential equation of (6.103), a Bernoulli's equation, can be obtained by employing the similar procedures in Region three of Section 4.4. For simplicity, set the following variables of m_n and σ_n^2:

$$\begin{cases} m_n = u_{tn} \\ \sigma_n^2 = \dfrac{C_t}{\tau k_n Q_n V_{dd}} \end{cases} \tag{6.104}$$

Therefore, the output voltage without the input-output coupling current and the pMOSFET cluster is

$$V_{o4} = \frac{1}{\dfrac{\tau k_n}{2C_t}P(s, m_n, \sigma_n) + \left(\dfrac{1}{V_{4i}} - \dfrac{\tau k_n}{2C_t}P(s_{4i}, m_n, \sigma_n)\right)\exp\left[\dfrac{(s + s_{4i} - 2m_n)(s - s_{4i})}{2\sigma_n^2}\right]} \tag{6.105}$$

where $P(x, m, \sigma)$ is defined in (III.9a) of Appendix III. The second step is that currents of series-connected nMOSFET chain, parallel pMOSFET cluster, and coupling capacitance between input and output are involved in the analysis. Then, the solution is expressed as follows:

$$V_o(s) = k_c V_{dd}(s - s_{4i}) + \frac{V_{qp}}{6}\left[(1 - u_{tp} - s_{4i})^3 - (1 - u_{tp} - s)^3\right] + V_{o4} \tag{6.106}$$

where the variable V_{qp} is given in (6.100).

Region five ($0.5 \le s \le 1 - u_{tp}$ and $(1 - u_o) \le u_{sp}/2$): In this region, the switching series-

connected nMOSFET chain is in saturation mode while switching parallel pMOSFET cluster is in linear mode. The differential equation to describe the output transition and the current expressions for the switching parallel pMOSFET cluster and the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t}\left(I_p - I_n\right) \\ I_p = k_p V_{sp} V_{op} \\ I_n = \dfrac{1}{2} k_n V_{sn}^2 \left(1 + \lambda_n V_o\right) \approx \dfrac{1}{2} k_n V_{sn}^2 \left(1 + \lambda_n V_{dd}\right) \\ V_{sp} = Q_p \left(V_{ip} - V_{tp}\right) = Q_p V_{dd} \left(1 - u_{tp} - s\right) \\ V_{sn} = R_n \left(V_i - V_{an}\right)^{b_n} = R_n V_{dd}^{b_n} \left(s - u_{an}\right)^{b_n} \approx k_m \left(s - b_m\right) \end{cases} \tag{6.107}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{5i} and s_{5i}. In order to simplify the derivation, the term of $\left(1+\lambda_n V_o\right)$ in the current I_n is reduced to $\left(1+\lambda_n V_{dd}\right)$ since output voltage V_o is close to V_{dd} in this region. The saturation voltage V_{sn} is approximated by the linear equation with the parameters k_m and b_m as follows:

$$\begin{cases} k_m = \dfrac{R_n V_{dd}^{b_n} \left[\left(1 - u_{tp} - u_{an}\right)^{b_n} - \left(1/2 - u_{an}\right)^{b_n}\right]}{1/2 - u_{tp}} \\ b_m = \dfrac{1}{2} - \dfrac{\left(1/2 - u_{an}\right)^{b_n} \left(1/2 - u_{tp}\right)}{\left(1 - u_{tp} - u_{an}\right)^{b_n} - \left(1/2 - u_{an}\right)^{b_n}} \end{cases} \tag{6.108}$$

Substituting the current expressions of I_p and I_n in (6.107) into the differential equation in (6.107) yields

$$\frac{dV_{op}}{ds} = -\frac{\tau k_p}{C_t} V_{sp} V_{op} - k_c V_{dd} + \frac{\tau k_n \left(1 + \lambda_n V_{dd}\right)}{2C_t} V_{sn}^2 \tag{6.109}$$

In (6.109), the variable V_o to be solved is changed to variable $V_{op}\left(=V_{dd}-V_o\right)$ with initial value of $\left(V_{op}\big|_{s=s_{5i}}=V_{dd}-V_{5i}\right)$. The solution of the first-order linear differential equation is analytically expressed as

$$V_{op}=\exp\left(\int_{s_{5i}}^{s}-\frac{\tau k_p}{C_t}V_{sp}\,ds\right)\cdot \left\{\int_{s_{5i}}^{s}\left[\left(\frac{\tau k_n\left(1+\lambda_n V_{dd}\right)}{2C_t}V_{sn}^2-k_c V_{dd}\right)\exp\left(\int_{s_{5i}}^{s}\frac{\tau k_p}{C_t}V_{sp}\,ds\right)\right]ds+V_{op}\big|_{s=s_{5i}}\right\} \tag{6.110}$$

In addition to the variables m_p and σ_p^2 defined in (6.82), set the variable $V_{\ln}$:

$$V_{\ln}=\frac{\tau k_n k_m^2\left(1+\lambda_n V_{dd}\right)}{C_t} \tag{6.111}$$

Substituting the equation V_{sp} of (6.107) into (6.110) with the change of variables m_p, σ_p^2, and $V_{\ln}$ yields

$$V_{op}=\exp\left[\frac{\left(s-m_p\right)^2}{2\sigma_p^2}\right]\cdot \left\{\int_{s_{5i}}^{s}\left[\frac{V_{\ln}\left(s-b_m\right)^2}{2}-k_c V_{dd}\right]\exp\left[-\frac{\left(s-m_p\right)^2}{2\sigma_p^2}\right]ds+\left(V_{dd}-V_{5i}\right)\exp\left[-\frac{\left(s_{5i}-m_p\right)^2}{2\sigma_p^2}\right]\right\} \tag{6.112}$$

The above equation is similar to (6.91) in Region two. The similar method is employed to derive the solution of output voltage. With the expressions in Appendix III, the above equation becomes

$$V_{op}=f_5(s)-\left[f_5(s_{5i})+V_{5i}-V_{dd}\right]\exp\left(-\frac{\left(s-s_{5i}\right)\left(2m_p-s_{5i}-s\right)}{2\sigma_p^2}\right) \tag{6.113}$$

where the function $f_5(x)$ is given by:

$$f_5(x)=\left\{-\frac{(x+m_p-2b_m)\sigma_p^2 V_{\ln}}{2}+\left[\frac{((m_p-b_m)^2+\sigma_p^2)V_{\ln}}{2}-k_c V_{dd}\right]R(x,m_p,\sigma_p)\right\} \quad (6.114)$$

Therefore, the output voltage $V_o(s)$ is

$$V_o(s)=V_{dd}-f_5(s)+[f_5(s_{5i})+V_{5i}-V_{dd}]\exp\left(-\frac{(s-s_{5i})(2m_p-s_{5i}-s)}{2\sigma_p^2}\right) \quad (6.115)$$

Region six ($0.5\le s\le 1-u_{tp}$, $(1-u_o)>u_{sp}/2$, and $u_o>u_{sn}$): In this region, both the parallel pMOSFET cluster and the series-connected nMOSFET chain are in saturation mode. The differential equation to describe the output transition and the current expressions for the switching parallel pMOSFET cluster and the series-connected nMOSFET chain are as follows:

$$\begin{cases}\dfrac{dV_o}{ds}=k_c V_{dd}+\dfrac{\tau}{C_t}(I_p-I_n)\\ I_p=\dfrac{1}{2}k_p V_{sp}^2\\ I_n=\dfrac{1}{2}k_n V_{sn}^2(1+\lambda_n V_o)\\ V_{sp}=Q_p(V_{ip}-V_{tp})=Q_p V_{dd}(1-u_{tp}-s)\\ V_{sn}=R_n(V_i-V_{an})^{b_n}=R_n V_{dd}^{b_n}(s-u_{an})^{b_n}\end{cases} \quad (6.116)$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{6i} and s_{6i}. The current I_n of the series-connected nMOSFET chain dominates the discharging process in this region. As the first step, only nMOSFET is considered. The differential equation is reduced to

$$\frac{dV_o}{ds}=-\frac{\tau k_n}{2C_t}V_{sn}^2(1+\lambda_n V_o) \quad (6.117)$$

For simplicity, set the variable V_{rn}:

$$V_{rn} = \frac{\tau k_n R_n^2 V_{dd}^{2b_n}}{C_t} \tag{6.118}$$

Substituting the equations of V_{sn} in (6.116) into the differential equation in (6.116) and using the variable V_{rn} yield

$$\frac{dV_o}{ds} = -\frac{V_{rn}}{2}(s - u_{an})^{2b_n}(1 + \lambda_n V_o) \tag{6.119}$$

The solution of the above differential equation is expressed as follows:

$$V_{o6} = \left(\frac{1}{\lambda_n} + V_{6i}\right)\exp[-f_6(s)] - \frac{1}{\lambda_n} \tag{6.120}$$

In (6.120), the function $f_6(s)$ is given by

$$f_6(s) = \frac{V_{rn}\lambda_n}{2(2b_n + 1)}\left\{(s - u_{an})^{2b_n+1} - (s_{6i} - u_{an})^{2b_n+1}\right\} \tag{6.121}$$

The second step is that currents of series-connected nMOSFET chain, parallel pMOSFET cluster, and coupling capacitance are included in the analysis. Then, we have

$$V_o = k_c V_{dd}(s - s_{6i}) + \frac{V_{qp}}{6}\left[(1 - u_{tp} - s_{6i})^3 - (1 - u_{tp} - s)^3\right] + V_{o6} \tag{6.122}$$

where V_{qp} is defined in (6.100).

Region seven ($0.5 \le s \le 1 - u_{tp}$ and $u_o \le u_{sn}$): In this region, series-connected nMOSFET chain is in linear mode and parallel pMOSFET cluster is in saturation mode. The differential equation to describe the output transition and the current expressions for the switching parallel pMOSFET cluster and the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} + \dfrac{\tau}{C_t}\left(I_p - I_n\right) \\ I_p = \dfrac{1}{2} k_p V_{sp}^2 \\ I_n = k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o\left(1 + \lambda_n V_{sn}\right) \approx k_n\left(V_{sn} - \dfrac{V_o}{2}\right)V_o \\ V_{sp} = Q_p\left(V_{ip} - V_{tp}\right) = Q_p V_{dd}\left(1 - u_{tp} - s\right) \\ V_{sn} = R_n\left(V_i - V_{an}\right)^{b_n} = R_n V_{dd}^{b_n}\left(s - u_{an}\right)^{b_n} \approx k_l\left(s - b_l\right) \end{cases} \tag{6.123}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{7i} and s_{7i}. For simplicity, the term of $(1+\lambda_n V_{sn})$ in the current I_n is ignored since the effect of $\lambda_n V_{sn}$ is small in this region. The saturation voltage V_{sn} of series-connected nMOSFET chain is approximated by the same linear equation as in region five with the parameters k_m and b_m given by (6.108). The series-connected nMOSFET chain current dominates the discharging process in this region. As the first step, only nMOSFET is considered in the derivation. The differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{C_t}\left(V_{sn} - \frac{V_o}{2}\right)V_o \tag{6.124}$$

This differential equation of (6.124), a Bernoulli's equation, is solved by employing the similar method in Region three of Section 4.4. For simplicity, set the following variable

$$\sigma_m^2 = \frac{C_t}{\tau k_n k_m} \tag{6.125}$$

Therefore, the output voltage without the input-output coupling current and the pMOSFET cluster is

$$V_{o7}=\frac{1}{\frac{\tau k_n}{2C_t}P(s,b_m,\sigma_m)+\left(\frac{1}{V_{7i}}-\frac{\tau k_n}{2C_t}P(s_{7i},b_m,\sigma_m)\right)\exp\left[\frac{(s+s_{7i}-2b_l)(s-s_{7i})}{2\sigma_m^2}\right]} \quad (6.126)$$

where $P(x,m,\sigma)$ is defined in (III.9a) of Appendix III. The second step is that currents of series-connected nMOSFET chain, parallel pMOSFET cluster, and coupling capacitance between input and output are involved in the analysis. Then, the solution is expressed as follows:

$$V_o=k_cV_{dd}(s-s_{7i})+\frac{V_{qp}}{6}\left[(1-u_{tp}-s_{7i})^3-(1-u_{tp}-s)^3\right]+V_{o7} \quad (6.127)$$

where V_{qp} is defined in (6.100).

Region eight ($0.5<1-u_{tp}\le s\le 1$, and $u_o>u_{sn}$): In this region, the serial nMOSFET chain is in the saturation mode with the input voltage V_i greater than the half of V_{dd} and all parallel pMOSFETs are in the cutoff mode $(I_p=0)$. The differential equation to describe the output transition and the current expression for the series-connected nMOSFET chain are as follows:

$$\begin{cases}\dfrac{dV_o}{ds}=k_cV_{dd}-\dfrac{\tau}{C_t}I_n\\ I_n=\dfrac{1}{2}k_nV_{sn}^2(1+\lambda_nV_o)\\ V_{sn}=R_n(V_i-V_{an})^{b_n}=R_nV_{dd}^{b_n}(s-u_{an})^{b_n}\end{cases} \quad (6.128)$$

The nMOSFET current dominates the discharging process in this region. In this case, the derivation of output is similar to that in region six except that the pMOSFET cluster is in the cutoff mode. As the first step, only nMOSFET is considered. The differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{2C_t} V_{sn}^2 (1 + \lambda_n V_o) \tag{6.129}$$

Substituting the equations of V_{sn} into the differential equation and using the variable V_{rn} yield

$$\frac{dV_o}{ds} = -\frac{V_{rn}}{2} (s - u_{an})^{2b_n} (1 + \lambda_n V_o) \tag{6.130}$$

The solution of the above equation is

$$V_{o8} = \left(\frac{1}{\lambda_n} + V_{8i} \right) \exp[- f_8(s)] - \frac{1}{\lambda_n} \tag{6.131}$$

In (6.131), the function $f_8(s)$ is given by

$$f_8(s) = \frac{V_{rn} \lambda_n}{2(2b_n + 1)} \left\{ (s - u_{an})^{2b_n + 1} - (s_{8i} - u_{an})^{2b_n + 1} \right\} \tag{6.132}$$

The second step is that nMOSFET and coupling capacitance current are included. Then, we have

$$V_o = k_c V_{dd} (s - s_{8i}) + V_{o8} \tag{6.133}$$

Region nine ($1 - u_{ip} \le s \le 1$ and $u_o \le u_{sn}$): In this region, the series-connected nMOSFET chain is in the linear mode and all the parallel pMOSFETs are in the cutoff mode $(I_p = 0)$. Then the differential equation to describe the output transition and the current expression for the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{ds} = k_c V_{dd} - \dfrac{\tau}{C_t} I_n \\ I_n = k_n \left(V_{sn} - \dfrac{V_o}{2} \right) V_o (1 + \lambda_n V_{sn}) \approx k_n \left(V_{sn} - \dfrac{V_o}{2} \right) V_o \\ V_{sn} = R_n (V_i - V_{an})^{b_n} = R_n V_{dd}^{b_n} (s - u_{an})^{b_n} \approx k_h (s - b_h) \end{cases} \tag{6.134}$$

The initial values of output voltage $V_o(t)$ and normalized time s in this region are V_{9i} and s_{9i}. In order to analytically solve the differential equation, the term of $(1+\lambda_n V_{sn})$ in the current I_n is ignored since the effect of $\lambda_n V_{sn}$ is small in this region. The saturation voltage V_{sn} in this region is approximated by the linear equation with the parameters k_h and b_h as follows:

$$\begin{cases} k_h = \dfrac{R_n V_{dd}^{b_n}\left[(1-u_{an})^{b_n} - (1-u_{tp}-u_{an})^{b_n}\right]}{u_{tp}} \\ b_h = 1 - \dfrac{(1-u_{an})^{b_n} u_{tp}}{(1-u_{an})^{b_n} - (1-u_{tp}-u_{an})^{b_n}} \end{cases} \tag{6.135}$$

The series-connected nMOSFET chain current dominates the discharging process in this region. As the first step, only series-connected nMOSFET chain is considered in the derivation. The differential equation is reduced to

$$\frac{dV_o}{ds} = -\frac{\tau k_n}{C_t}\left(V_{sn} - \frac{V_o}{2}\right)V_o \tag{6.136}$$

This differential equation of (6.136), a Bernoulli's equation, is solved by employing the similar method in Region three of Section 4.4. For simplicity, set the following variable

$$\sigma_h^2 = \frac{C_t}{\tau k_n k_h} \tag{6.137}$$

Therefore, the output voltage without considering the input-output coupling current and the pMOSFET cluster is

$$V_{o9} = \frac{1}{\dfrac{\tau k_n}{2C_t}P(s,b_h,\sigma_h) + \left(-\dfrac{\tau k_n}{2C_t}P(s_{9i},b_h,\sigma_h) + \dfrac{1}{V_{9i}}\right)\exp\left[\dfrac{(s+s_{9i}-2b_h)(s-s_{9i})}{2\sigma_h^2}\right]} \tag{6.138}$$

where $P(x,m,\sigma)$ is defined in (III.9a) of Appendix III. The second step is that

nMOSFET, pMOSFET, and coupling capacitance current between input and output are involved in the analysis. Then, the solution is expressed as follows:

$$V_o = k_c V_{dd}(s - s_{9i}) + V_{o9} \tag{6.139}$$

Region ten ($s \geq 1$ or $t \geq \tau$ and $u_o \geq u_{snd}$): In this region, all the parallel pMOSFETs are in the cutoff mode $(I_p = 0)$ and the series-connected nMOSFET chain is in the saturation mode with input voltage $V_i = V_{dd}$. Then the differential equation to describe the output transition and the current expression for the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{dt} = -\dfrac{1}{C_t} I_n \\ I_n = \dfrac{1}{2} k_n V_{snd}^2 (1 + \lambda_n V_o) \\ V_{snd} = R_n (V_{dd} - V_{an})^{b_n} = R_n V_{dd}^{b_n} (1 - u_{an})^{b_n} \end{cases} \tag{6.140}$$

The initial values of output voltage $V_o(t)$ and time t in this region are V_{10i} and t_{10i}. Then the output waveform $V_o(t)$ is

$$V_o = \left(\frac{1}{\lambda_n} + V_{10i} \right) \exp\left[-\frac{t - t_{10i}}{\tau_{sat}} \right] - \frac{1}{\lambda_n} \tag{6.141}$$

or

$$t = t_{10i} + \tau_{sat} \ln\left(\frac{1 + \lambda_n V_{10i}}{1 + \lambda_n V_o} \right) \tag{6.142}$$

where t_{10i} and V_{10i} are initial values of time t and output voltage V_o in the region. The time constant τ_{sat} in this region is given by

$$\tau_{sat} = \frac{2 \cdot C_t}{k_n \cdot \lambda_n \cdot V_{snd}^2} \tag{6.143}$$

Region eleven ($s \geq 1$ or $t \geq \tau$ and $u_o \leq u_{snd}$): In this region, all the parallel pMOSFETs are in the cutoff mode $(I_p = 0)$ and the series-connected nMOSFET chain is in the linear mode with input voltage $V_i = V_{dd}$. Then the differential equation to describe the output transition and the current expression for the series-connected nMOSFET chain are as follows:

$$\begin{cases} \dfrac{dV_o}{dt} = -\dfrac{1}{C_t} I_n \\ I_n = k_n \left(V_{snd} - \dfrac{V_o}{2} \right) V_o \left(1 + \lambda_n V_{snd}\right) \\ V_{snd} = R_n \left(V_{dd} - V_{an}\right)^{b_n} = R_n V_{dd}^{b_n} \left(1 - u_{an}\right)^{b_n} \end{cases} \tag{6.144}$$

For generality, the initial values of output voltage $V_o(t)$ and time t in this region are V_{11i} and t_{11i}. Then the output waveform $V_o(t)$ is

$$V_o = \frac{2V_{snd}}{1 + \left(2V_{snd} / V_{11i} - 1\right) \exp\left[\left(t - t_{11i}\right) / \tau_{11}\right]} \tag{6.145}$$

or

$$t = t_{11i} + \tau_{lin} \ln\left(\frac{V_{11i}}{V_o} \frac{2V_{snd} - V_o}{2V_{snd} - V_{11i}} \right) \tag{6.146}$$

where the time constant τ_{lin} in this region is given by

$$\tau_{lin} = \frac{C_t}{k_n V_{snd} \left(1 + \lambda_n V_{snd}\right)} \tag{6.147}$$

6.4.3 Effect of Initial States of Parasitic Capacitances of Intermediate Nodes

The above analysis of output voltage waveform $V_o(t)$ is based on that there is no effect of initial states of intermediate nodes, i.e., the step-input activation time of zero $(t_{as} = 0)$. From the analysis of (5.9) in the modified input suppression technique the ramp-

input activation time t_a is equal to $u_{tn}\tau$. However, single-stage CMOS logic gates may have non-zero step-input activation time $(t_{as} \neq 0)$ due to the intermediate node capacitances. The effect of intermediate node capacitances is modeled with the activation time t_a, which is discussed in Chapter five. According to the ramp-input activation time t_a, the timing analysis discussed in the previous subsection has to be corrected. Any region mentioned below refers to the corresponding region defined in the previous subsection.

Case one: $\left(u_{tn}\tau \leq t_a < (1-u_{tp})\tau\right)$. Before the activation time t_a in this case $(0 \leq t < t_a)$, the series-connected nMOSFET chain is not activated $(I_n = 0)$ and the transitioning parallel pMOSFET cluster is in linear mode. This is the same situation as Region one. The equations to describe the output voltage $V_o(t)$ during the period between time of zero and the activation time t_a $(0 \leq t < t_a)$ are:

$$\begin{cases} \dfrac{dV_o}{dt} = k_c \dfrac{V_{dd}}{\tau} + \dfrac{1}{C_t} I_p \\ I_p = k_p V_{sp} V_{op} \\ V_{sp} = Q_p (V_{ip} - V_{tp}) = Q_p V_{dd} \left(1 - u_{tp} - \dfrac{t}{\tau}\right) \end{cases} \tag{6.148}$$

The output voltage $V_o(s)$ derived in the Region one is still valid and is expressed as

$$V_o(s) = V_{dd} + k_c V_{dd} \left\{ R(s, m_p, \sigma_p) - R(0, m_p, \sigma_p) \exp\left(-\frac{s(2m_p - s)}{2\sigma_p{}^2} \right) \right\} \quad (0 \leq t < t_a) \tag{6.149}$$

where all the variables and functions are the same as those in the Region one of the above subsection.

After the activation time t_a in this case $(t > t_a)$, the series-connected nMOSFET

chain is activated and is in saturation mode. The output voltage $V_o(t)$ goes to either Region two or Region five, depending on the activation time t_a. If $(u_m\tau \le t_a < 0.5\tau)$, the output voltage $V_o(t)$ goes to Region two with the initial time $t_{2i} = t_a$ and initial output voltage $V_{2i} = V_o(t_a)$. If $(0.5\tau \le t_a < (1-u_{tp})\tau)$, the output voltage $V_o(t)$ enters Region five with the initial time $t_{5i} = t_a$ and initial output voltage $V_{5i} = V_o(t_a)$.

Case two: $((1-u_{tp})\tau \le t_a < \tau)$. In this case, before the activation time t_a $(0 \le t < t_a)$, the series-connected nMOSFET chain is not activated $(I_n = 0)$. However, the transitioning pMOSFET cluster experiences the linear mode $(0 \le t < (1-u_{tp})\tau)$ and cutoff mode $((1-u_{tp})\tau \le t < t_a)$. Let us discuss these two situations separately.

When $(0 \le s < 1-u_{tp})$, it is the same situation as Region one. Then the output voltage $V_o(s)$ derived in Region one is still valid, which is

$$V_o(s) = V_{dd} + k_c V_{dd} \cdot \left\{ R(s, m_p, \sigma_p) - R(0, m_p, \sigma_p) \exp\left(-\frac{s(2m_p - s)}{2\sigma_p^{\,2}} \right) \right\} \qquad (0 \le s < 1-u_{tp}) \quad (6.150)$$

When $((1-u_{tp})\tau \le t < t_a)$, the series-connected nMOSFET chain is not activated $(I_n = 0)$ and the transitioning pMOSFET cluster is in the cutoff mode $(I_p = 0)$. The equation to describe the output voltage $V_o(t)$ during this period is simplified into:

$$\frac{dV_o}{dt} = k_c \frac{V_{dd}}{\tau} \qquad \left(1-u_{tp} \le s < \frac{t_a}{\tau}\right) \quad (6.151)$$

The solution of the above equation is

$$V_o(s) = V_{a2} + k_c V_{dd}(s - (1-u_{tp})) \qquad \left(1-u_{tp} \le s < \frac{t_a}{\tau}\right) \quad (6.152)$$

In (6.152), V_{a2} is the output voltage $V_o(t)$ at the time of $\tau(1-u_{tp})$, which is

$$V_{a2} = V_{dd} + k_c V_{dd}\left\{R((1-u_{tp}),m_p,\sigma_p) - R(0,m_p,\sigma_p)\exp\left(-\frac{(1-u_{tp})^2}{2\sigma_p^{\,2}}\right)\right\} \qquad (6.153)$$

After the activation time t_a in this case $(t > t_a)$, the series-connected nMOSFET chain is activated. The output voltage $V_o(t)$ enters Region eight discussed in the above subsection with the initial time $t_{8i} = t_a$ and initial output voltage $V_{8i} = V_o(t_a)$ from (6.152).

Case three: $(\tau \le t_a)$. In this case, before the activation time t_a $(0 \le t < t_a)$, the series-connected nMOSFET chain is not activated $(I_n = 0)$. However, the transitioning pMOSFET cluster experiences the linear mode $(0 \le s < 1-u_{tp})$ and cutoff mode $(1-u_{tp} \le s)$. Moreover, the input-to-output coupling current turns to zero when $\left(1 \le s < \frac{t_a}{\tau}\right)$. Let us discuss these three situations separately.

When $(0 \le s < 1-u_{tp})$, it is the same situation as Region one discussed before. Then the output voltage $V_o(t)$ is the same as (6.150)

$$V_o(s) = V_{dd} + k_c V_{dd} \cdot \left\{R(s,m_p,\sigma_p) - R(0,m_p,\sigma_p)\exp\left(-\frac{s(2m_p - s)}{2\sigma_p^{\,2}}\right)\right\} \qquad (0 \le s < 1-u_{tp}) \quad (6.154)$$

When $(1-u_{tp} \le s < 1)$, the series-connected nMOSFET chain is not activated $(I_n = 0)$ and the pMOSFET cluster is in the cutoff mode $(I_p = 0)$. The output voltage $V_o(t)$ during this period similar to (6.152) of case two is:

$$V_o(s) = V_{a2} + k_c V_{dd}(s - (1-u_{tp})) \qquad (1-u_{tp} \le s < 1) \quad (6.155)$$

In (6.155), V_{a2} is defined as in (6.153).

When $\left(1 \le s < \frac{t_a}{\tau}\right)$, the series-connected nMOSFET chain is not activated $(I_n = 0)$ and the pMOSFET cluster is in the cutoff mode $(I_p = 0)$. Moreover, the input-to-output coupling current becomes zero due to constant input. Therefore, the equation to describe the output voltage $V_o(t)$ is

$$\frac{dV_o}{dt} = 0 \text{ or } V_o(t) = V_{a3} \qquad \left(1 \le s < \frac{t_a}{\tau}\right) \quad (6.156)$$

In (6.156), V_{a3} is the output voltage $V_o(t)$ at the time of τ, which is

$$V_{a3} = V_{a2} + k_c V_{dd} u_{tp} \qquad \left(1 \le s < \frac{t_a}{\tau}\right) \quad (6.157)$$

The output in this region $\left(1 \le s < \frac{t_a}{\tau}\right)$ of this case is constant. This is the case for output mesa. After the activation time t_a in this case $(t \ge t_a)$, the series-connected nMOSFET chain is activated. The analysis of output voltage $V_o(t)$ is starting from Region ten of the above subsection with the initial time $t_{10i} = t_a$ and initial output voltage $V_{10i} = V_{a3}$.

6.4.4 Results and Discussions

By using SMCM and the modified input suppression technique and solving the differential equation (6.2) or (6.5) in different regions accordingly, we have developed the complete solution of timing delay of single-stage CMOS logic gates for a rising ramp input. The overshoot of output voltage is obvious in the cases of fast input transition time. These plots show that the proposed SMCM and analysis of output waveform are very accurate. Figs. 6.19-6.21 display output voltage waveforms of a NAND42 gate with the active input of V_{i2} for external load capacitances of 10fF, 100fF, and 400fF, respectively.

The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference.

Figs. 6.22-6.24 plot output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for external load capacitances of 10fF, 100fF, and 400fF, respectively. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots, including a mesa for fast input, are well compensated with a trivial parallel pMOSFET cluster.

Fig. 6.25 shows FIS propagation delays versus external load capacitances for a NAND42 gate. The propagation delays from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the propagation delays of SIS are the same as those of FIS.

Figs. 6.26-6.27 plot propagation delays of FIS versus external load capacitances for a NAND4124 gate for FIS and SIS, respectively. The propagation delays from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the activation time of FIS of a NAND4124 is zero.

Figs. 6.28-6.29 display the correlation of proposed propagation delay from SMCM in Section 6.4 with propagation delay from SPICE BSIM3 for FIS and SIS, respectively. The dotted lines are +/- 5% away from BSIM3 model. There are 375 simulation points which cover all the 15 sorts of active inputs. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively. These figures indicate that the correlation between them is excellent and the developed propagation delay model is accurate to describe the single-stage complex CMOS logic gate.

The calculation speedup of the proposed method with SMCM by using MATLAB over SPICE with BSIM3 model is investigated on the workstation of SUN Ultra-10. The circuit used in the investigation is an inverter ring with the number of inverters equal to 1001, 2001, 3001, 4001, 5001, and 10001, separately. Table 6.1 lists the CPU times of both methods to obtain the periods of inverter rings and the speedup of the proposed method over SPICE with BSIM3 model. Approximately the CPU time of SPICE with BSIM3 is cubically dependent of the number of transistors while the CPU time of the proposed method with SMCM is linearly proportional to the number of transistors. In the other words the proposed method is extremely suitable for large logic circuits.

Table 6.1 The Speedup of the Proposed Method over SPICE with BSIM3 Model

Number of inverters	1001	2001	3001	4001	5001	10001
CPU time of the proposed method (sec)	735	1466	2202	2937	3670	7340
CPU time of SPICE BSIM3 (sec)	1207	5166	12001	21756	33865	136578
Speedup ratio	1.64	3.52	5.45	7.41	9.23	18.61

Table 6.2 The Coefficients for Calculating the Input-to-Output Coupling Capacitances and the Diffusion Capacitances

Sort	nmp	nmn	ndp	ndn
NAND11	1	1	1	1
NAND21	1	0	2	1
NAND212	2	1	2	1
NAND22	1	1	2	1
NAND31	1	0	3	1
NAND312	2	0	3	1
NAND3123	3	1	3	1
NAND313	2	1	3	1
NAND32	1	0	3	1
NAND323	2	1	3	1
NAND33	1	1	3	1
NAND41	1	0	4	1
NAND412	2	0	4	1
NAND4123	3	0	4	1
NAND41234	4	1	4	1
NAND4124	3	1	4	1
NAND413	2	0	4	1
NAND4134	3	1	4	1
NAND414	2	1	4	1
NAND42	1	0	4	1
NAND423	2	0	4	1
NAND4234	3	1	4	1
NAND424	2	1	4	1
NAND43	1	0	4	1
NAND434	2	1	4	1
NAND44	1	1	4	1

Table 6.3 The SMCM Model Parameters Used in Chapter 6

Chain sort	k_n	R_n	V_a	b_n	Q_n	V_t	λ_n
n41	9.9928	0.7401	0.5871	0.1091	1.5415	0.3012	0.1316
n412	10.1023	0.8070	0.5669	0.2645	1.1527	0.3158	0.0890
n4123	10.4288	0.8398	0.5309	0.4404	0.9087	0.3150	0.0767
n41234	10.8437	0.8203	0.4562	0.6704	0.8134	0.3252	0.1101
n4124	10.9679	0.7982	0.4887	0.5298	0.8790	0.3162	0.1233
n413	10.5981	0.8168	0.5417	0.3447	1.0185	0.2989	0.0502
n4134	11.0722	0.7968	0.4717	0.5734	0.8620	0.3152	0.1155
n414	11.2603	0.7630	0.4962	0.4369	0.9416	0.2989	0.1355
n42	10.4140	0.7615	0.5726	0.1767	1.2574	0.2793	0.0935
n423	10.6672	0.8187	0.5343	0.3712	0.9939	0.3002	0.0506
n4234	11.0866	0.7968	0.4627	0.5975	0.8524	0.3146	0.1134
n424	11.2962	0.7650	0.4887	0.4620	0.9185	0.2979	0.1333
n43	11.0375	0.7634	0.5416	0.2677	1.0556	0.2620	0.0815
n434	11.3926	0.7666	0.4720	0.5051	0.8920	0.2957	0.1212
n44	11.6995	0.7234	0.4859	0.3725	0.9142	0.2475	0.1485
p11	4.2654	0.9080	0.4058	0.6602	0.9593	0.2792	0.3382

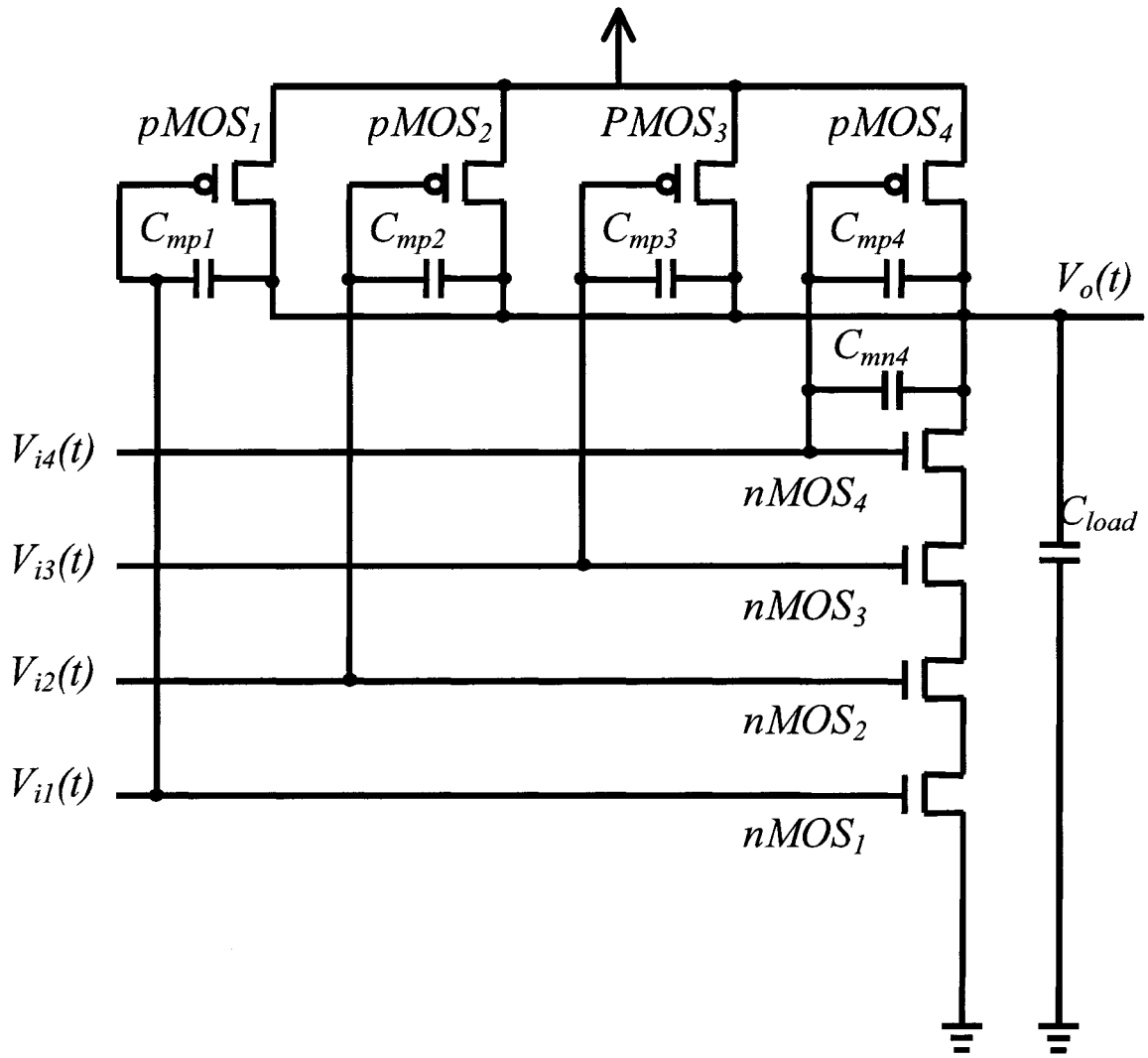


Fig. 6.1 A four-input NAND gate. The numbering of an nMOSFET and its corresponding input starts from the nMOSFET connected to ground as number one. Output load capacitance C_{load} is shown. The possible input-output coupling capacitances of C_{mn4}, C_{mp1}, C_{mp2}, C_{mp3}, and C_{mp4} are also plotted.

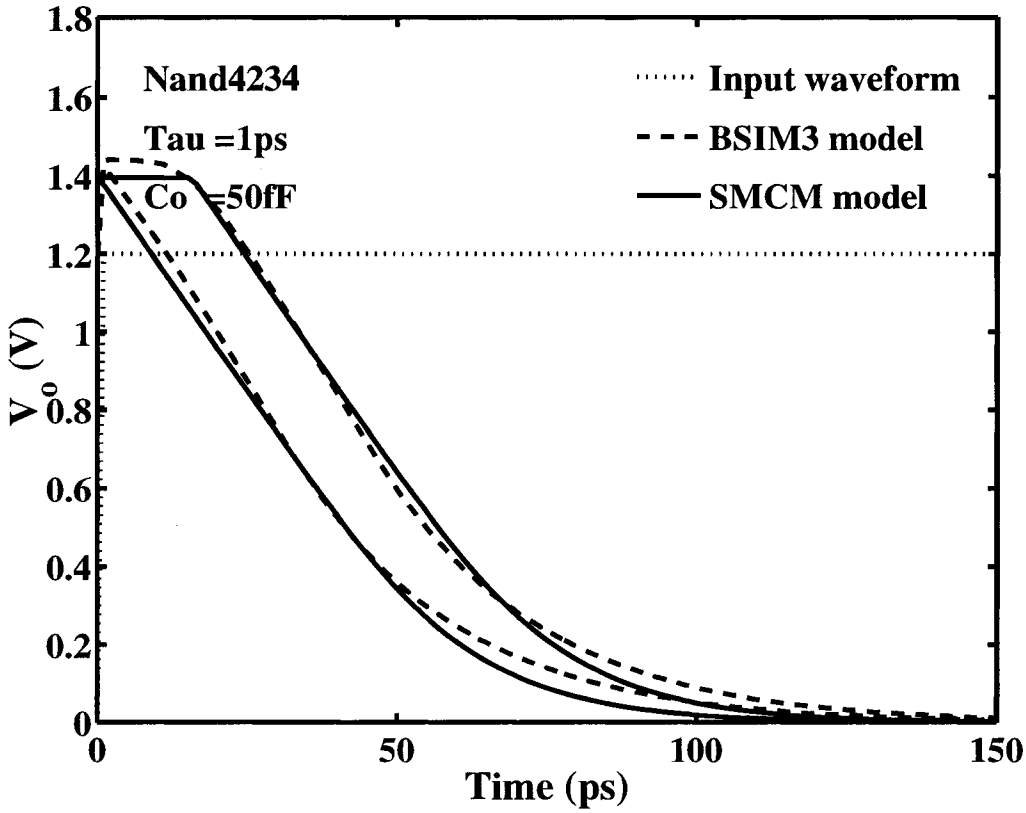


Fig. 6.2 Output voltage waveforms of a NAND4234 gate with the active inputs of V_{i2}, V_{i3}, and V_{i4} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. There is a pair of traces for the external load capacitance of 50fF: the left one for the fast initial state (FIS) and the right one for the slow initial state (SIS). We can see an obvious mesa in the voltage waveform for slow initial state (SIS). There is a voltage overshoot, but not a mesa in the fast initial state (FIS). Both phenomena are well explained by the modified input suppression technique in Section 5.5.

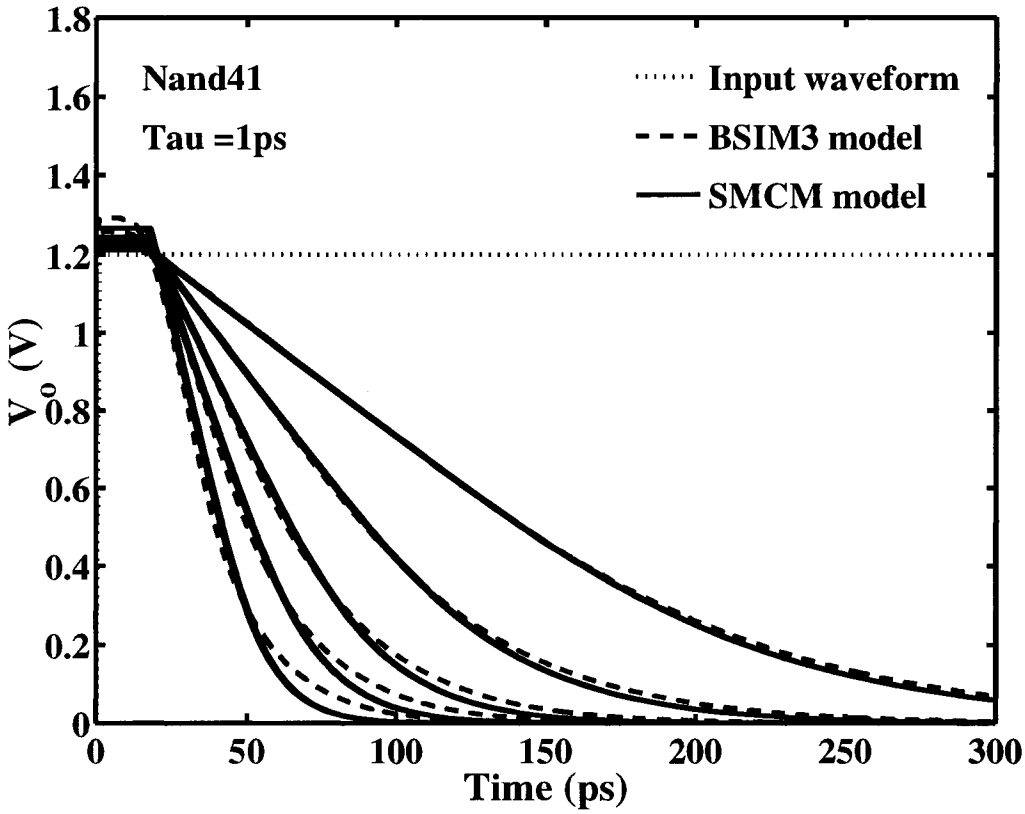


Fig. 6.3 Output voltage waveforms of a NAND41 gate with the active input of V_{i1} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference.

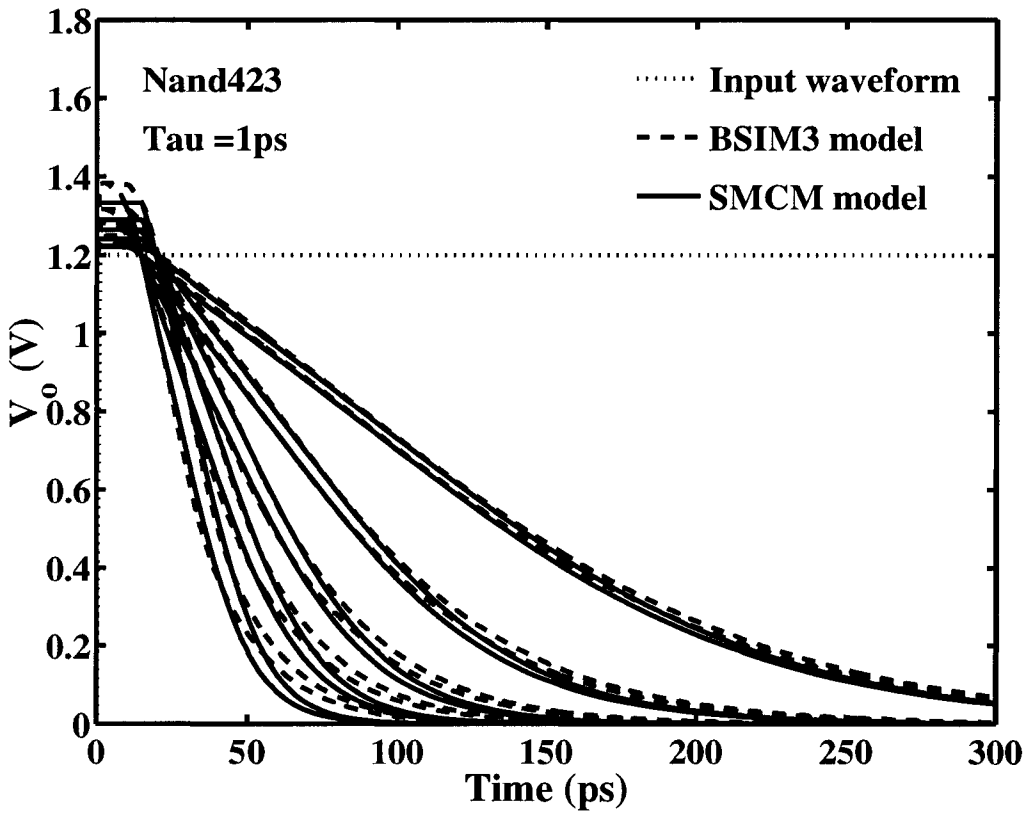


Fig. 6.4 Output voltage waveforms of a NAND423 gate with the active inputs of V_{i2}, and V_{i3} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively (from left to right). There is a pair of traces for each external load capacitance: the left one for fast initial state (FIS) and the right one for slow initial state (SIS). A mesa exists in the voltage waveforms for both slow initial state (SIS) and fast initial state (FIS). Both are accurately explained by the modified input suppression technique in Section 5.5.

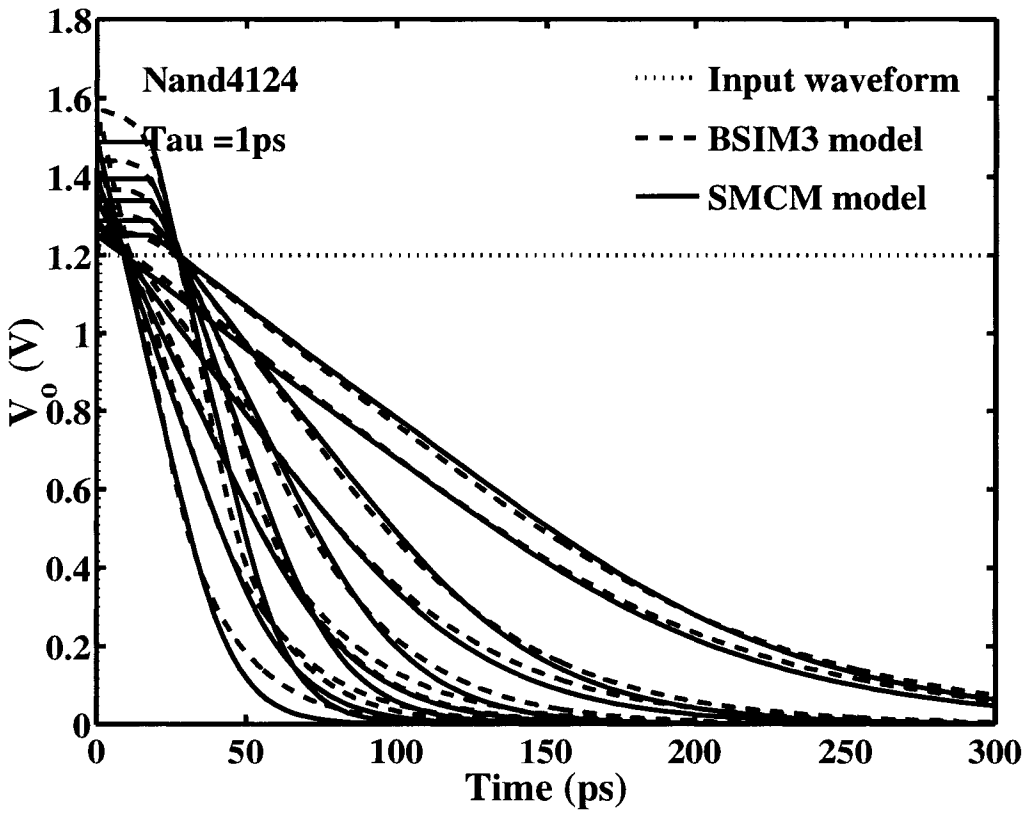


Fig. 6.5 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for a step input. The voltage waveforms from SMCM in Subsection 6.2.2 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. Note that the left family of traces is for fast initial state (FIS) and right one for the slow initial state (SIS). For each family the external capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively (from left to right). Note there is a mesa in the voltage waveforms for slow initial state (SIS). No mesa exists in fast initial state (FIS). Both are explained by the modified input suppression technique in Section 5.5.

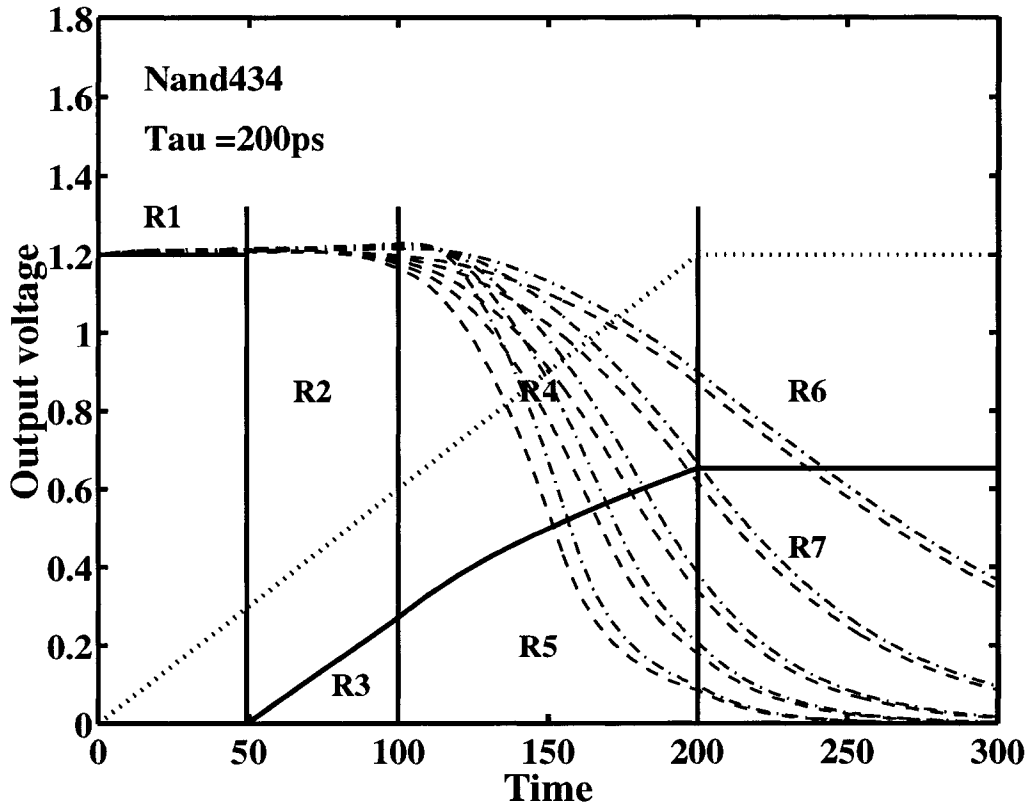


Fig. 6.6 Seven operation regions of a complex NAND434 gate modeled with a series-connected MOSFET chain 'n434' with the active inputs of V_{i3} and V_{i4} for a rising ramp input. The typical discharging traces from SPICE BSIM3 are also plotted. Each pair of traces corresponds to the same external load capacitance for fast initial state (FIS) and slow initial state (SIS).

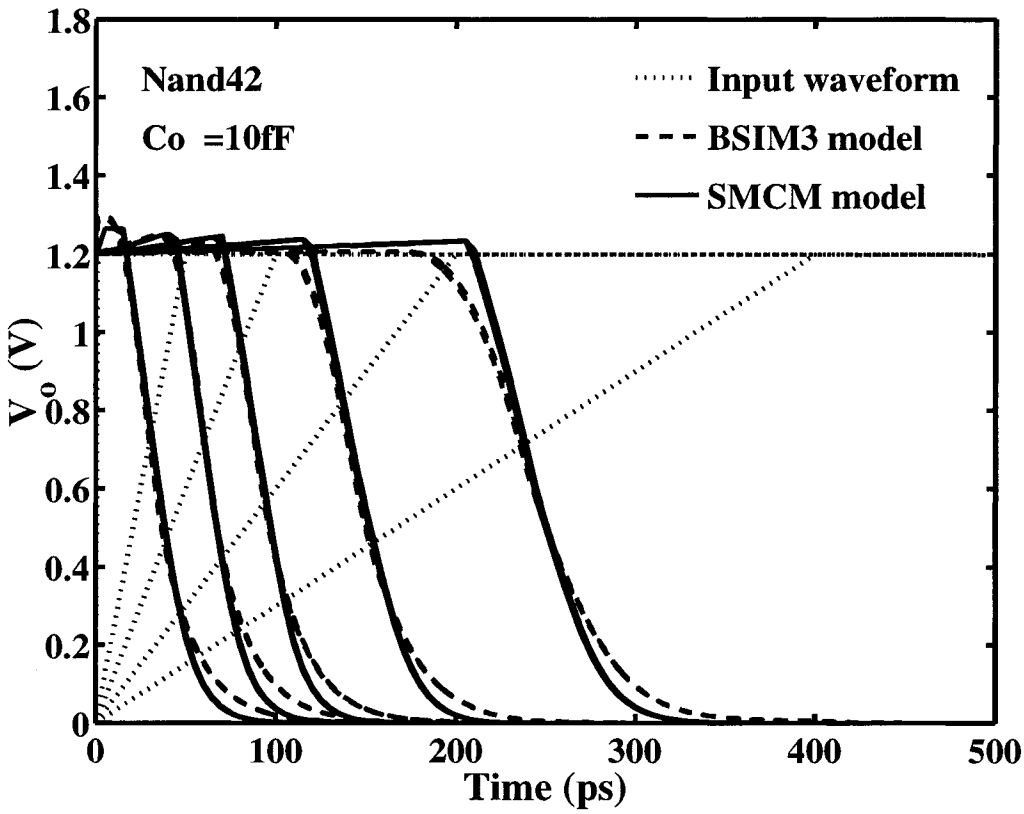


Fig. 6.7 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} for an external load capacitance of 10fF. The NAND42 gate is modeled with a series-connected MOSFET chain 'n42'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference.

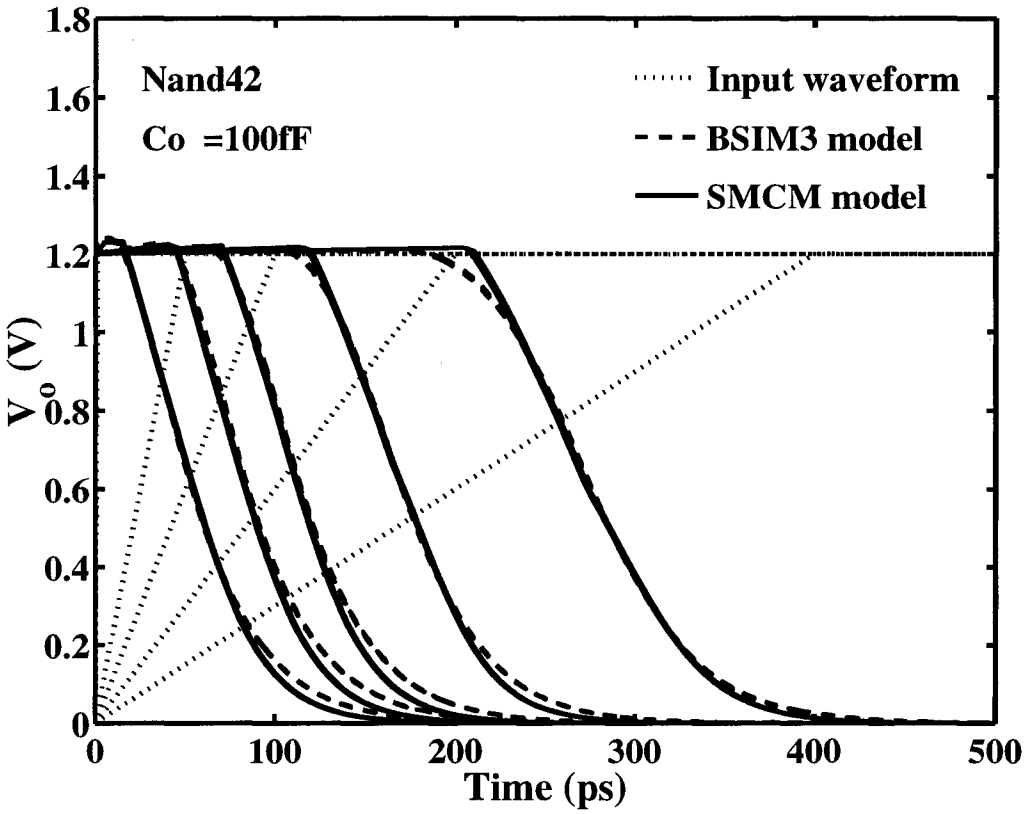


Fig. 6.8 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} for an external load capacitance of 100fF. The NAND42 gate is modeled with a series-connected MOSFET chain 'n42'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference.

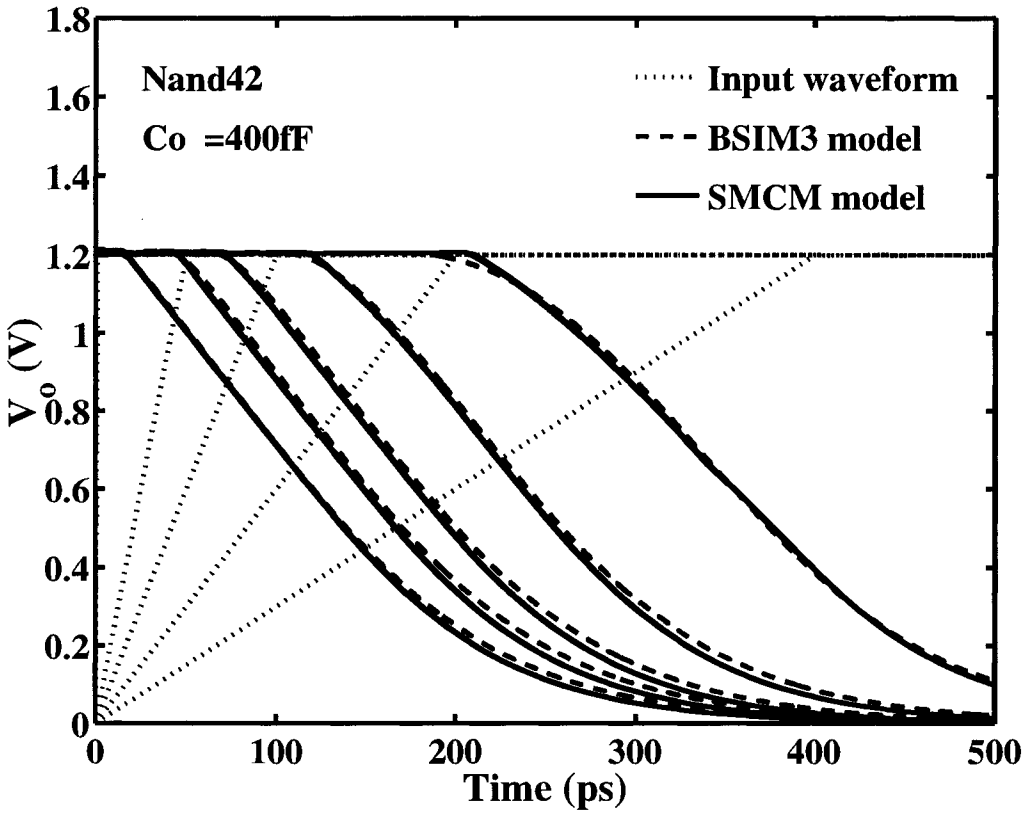


Fig. 6.9 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} for an external load capacitance of 400fF. The NAND42 gate is modeled with a series-connected MOSFET chain 'n42'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference.

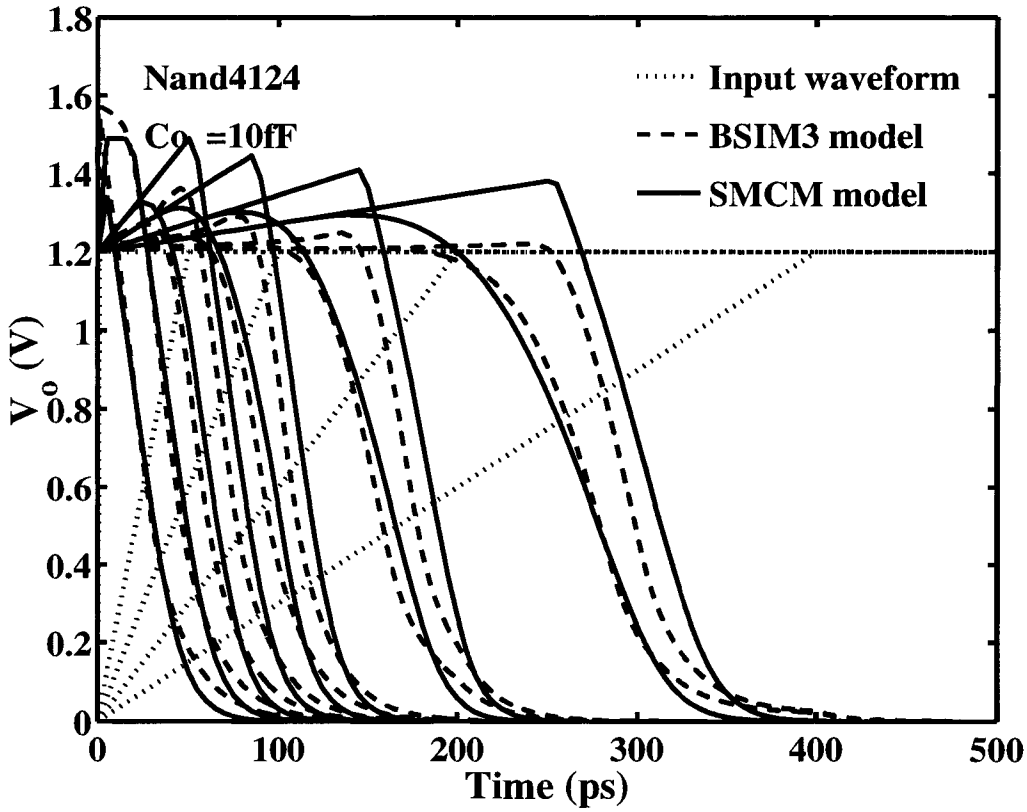


Fig. 6.10 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for an external load capacitance of 10fF. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots are exaggerated with no trivial parallel pMOSFET cluster.

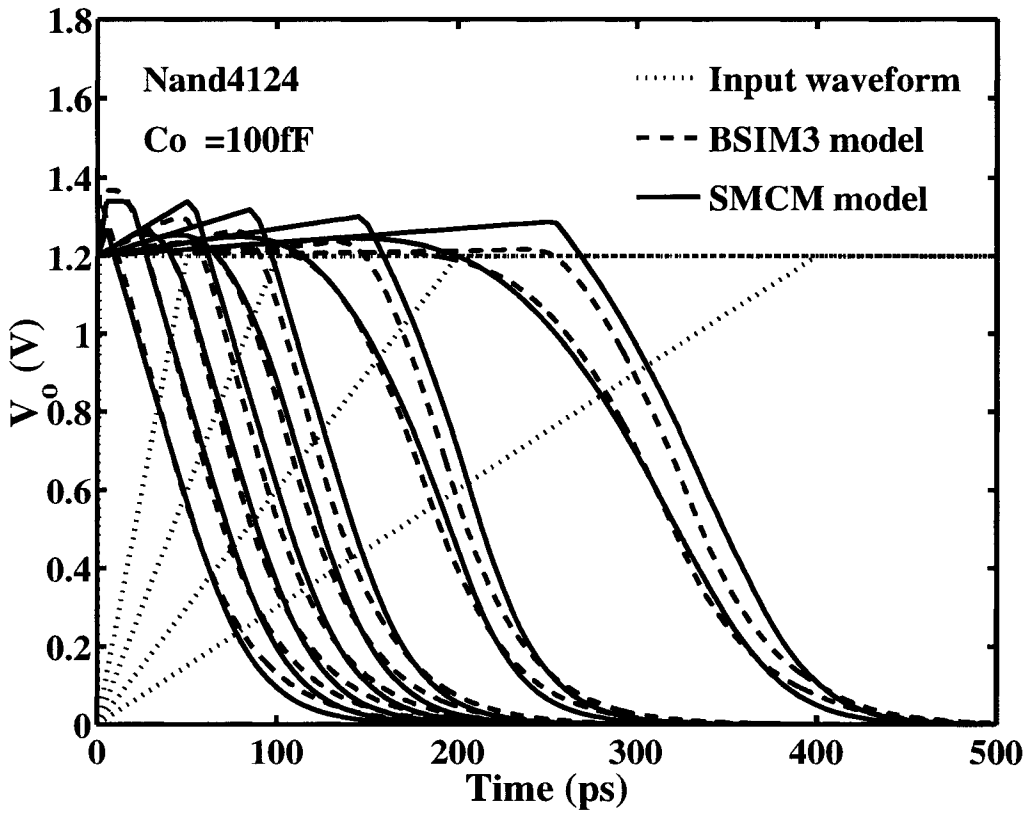


Fig. 6.11 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for an external load capacitance of 100fF. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots are exaggerated with no trivial parallel pMOSFET cluster.

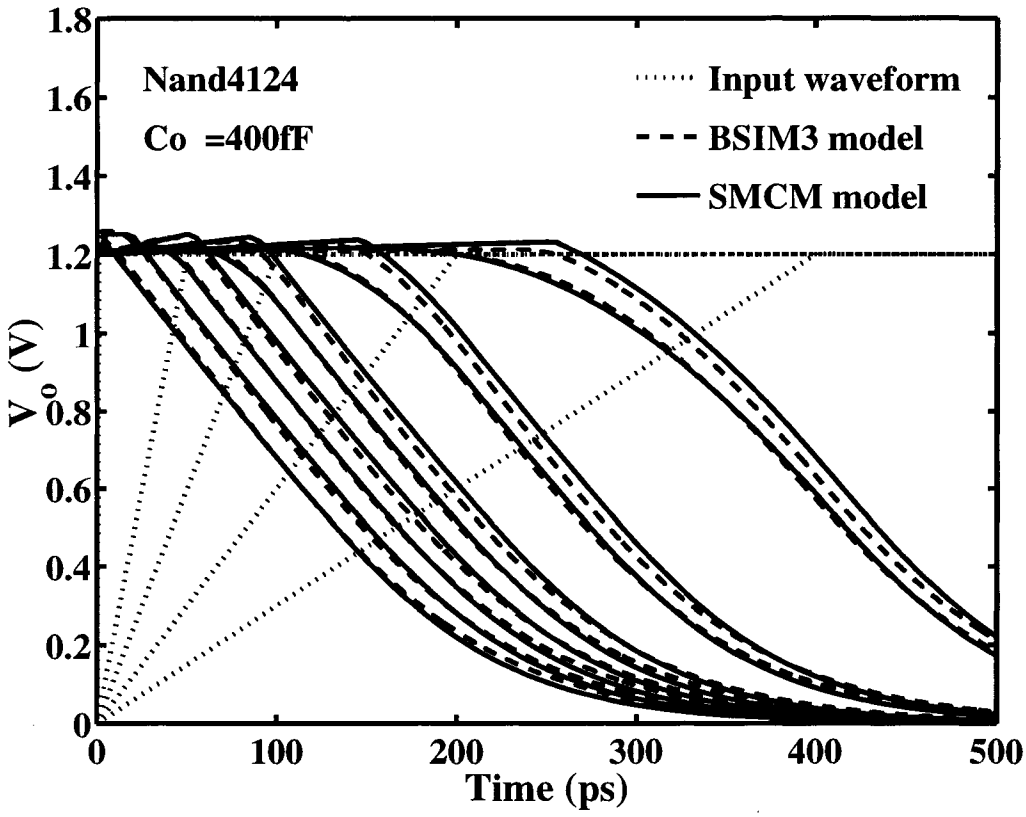


Fig. 6.12 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for an external load capacitance of 400fF. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The voltage waveforms from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The reduction of voltage overshoots is caused with large external load capacitance.

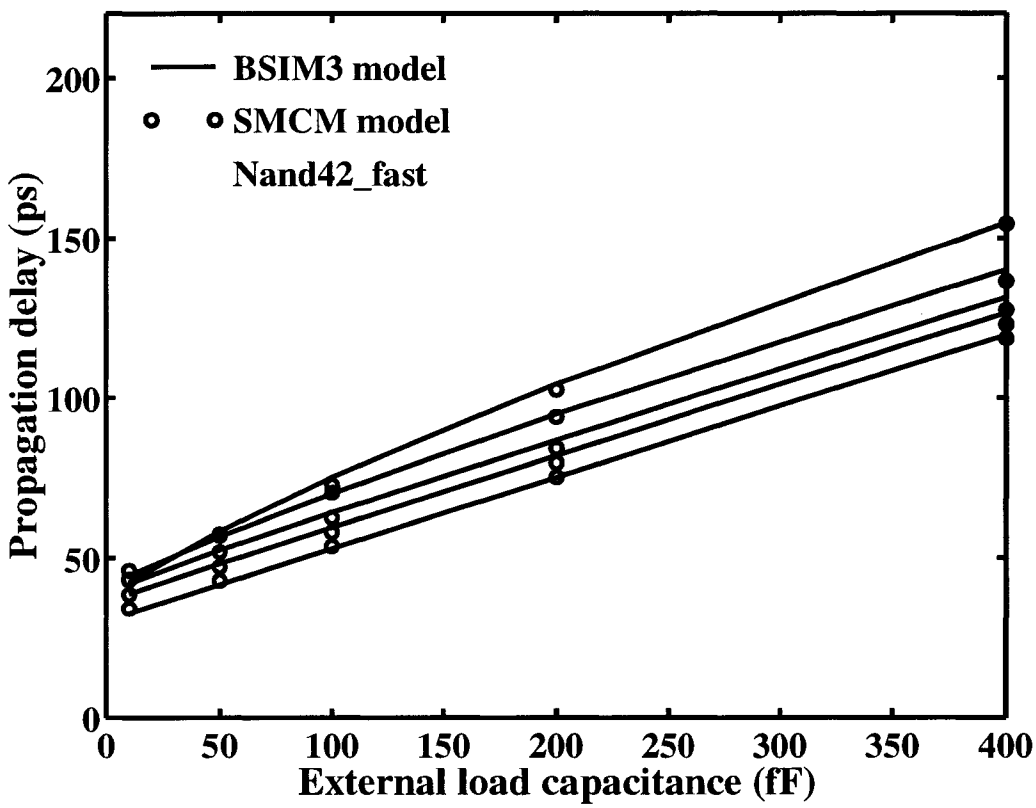


Fig. 6.13 Propagation delays of FIS versus external load capacitances for a NAND42 gate. The NAND42 gate is modeled with a series-connected MOSFET chain 'n42'. The propagation delays from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the propagation delays of SIS are the same.

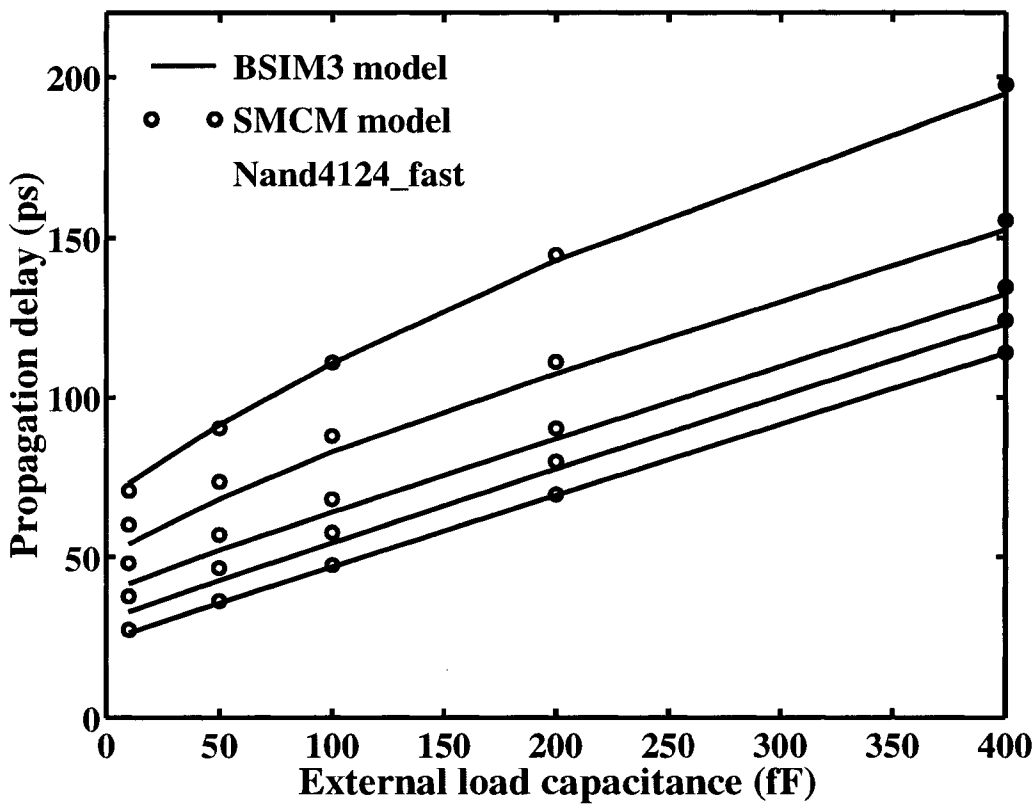


Fig. 6.14 Propagation delays of FIS versus external load capacitances for a NAND4124 gate. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The propagation delays from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the activation time of FIS of a NAND4124 is zero.

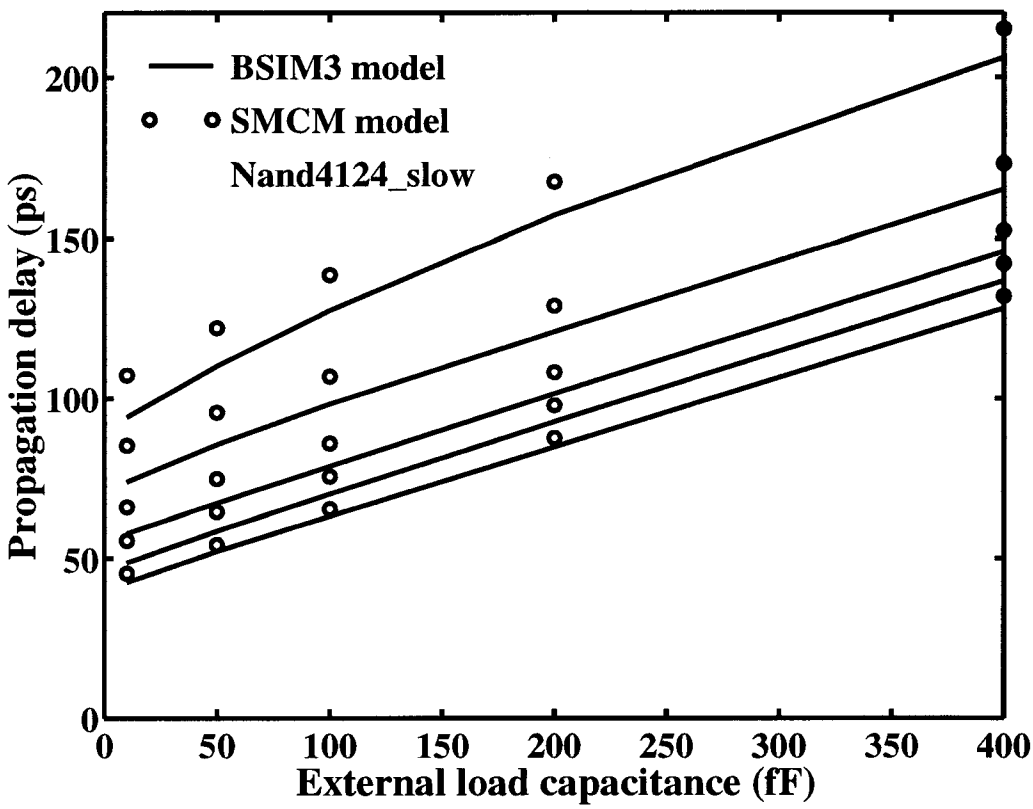


Fig. 6.15 Propagation delays of SIS versus external load capacitances for a NAND4124 gate. The NAND4124 gate is modeled with a series-connected MOSFET chain 'n4124'. The propagation delays from SMCM in Section 6.3 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). The part of error in propagation delays is caused with the exaggerated voltage overshoots.

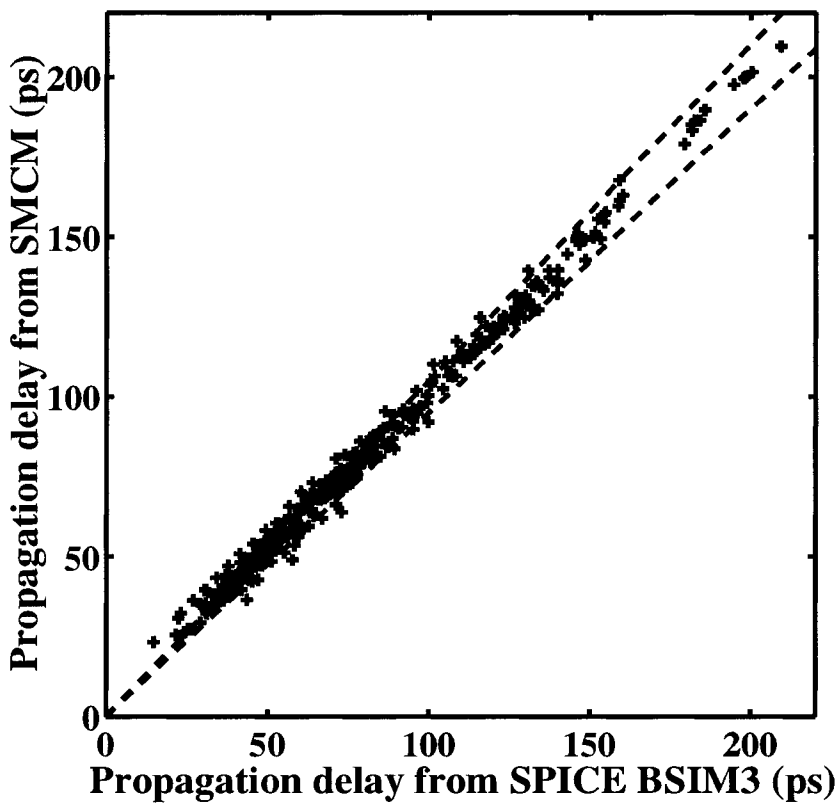


Fig. 6.16 The correlation of derived propagation delay of fast initial state (FIS) from SMCM in Section 6.3 with propagation delay from SPICE BSIM3. The NAND4 gate is modeled with a series-connected MOSFET chain. There are 375 simulation points which cover all the 15 sorts of active inputs. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively.

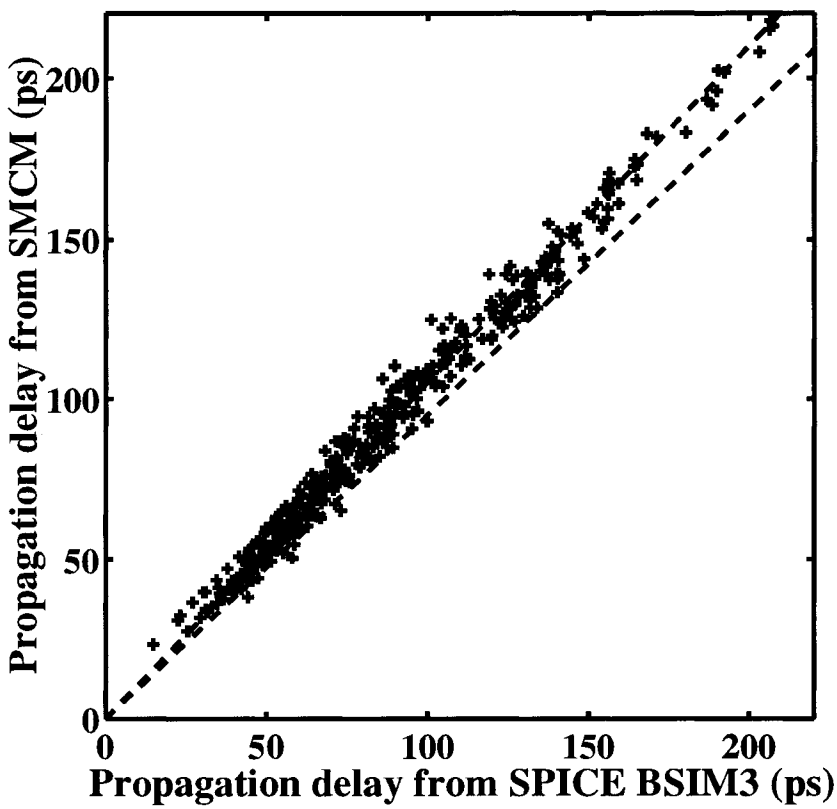


Fig. 6. 17 The correlation of derived propagation delay of slow initial state (SIS) from SMCM in Section 6.3 with propagation delay from SPICE BSIM3. The NAND4 gate is modeled with a series-connected MOSFET chain. There are 375 simulation points which cover all the 15 sorts of active inputs. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively.

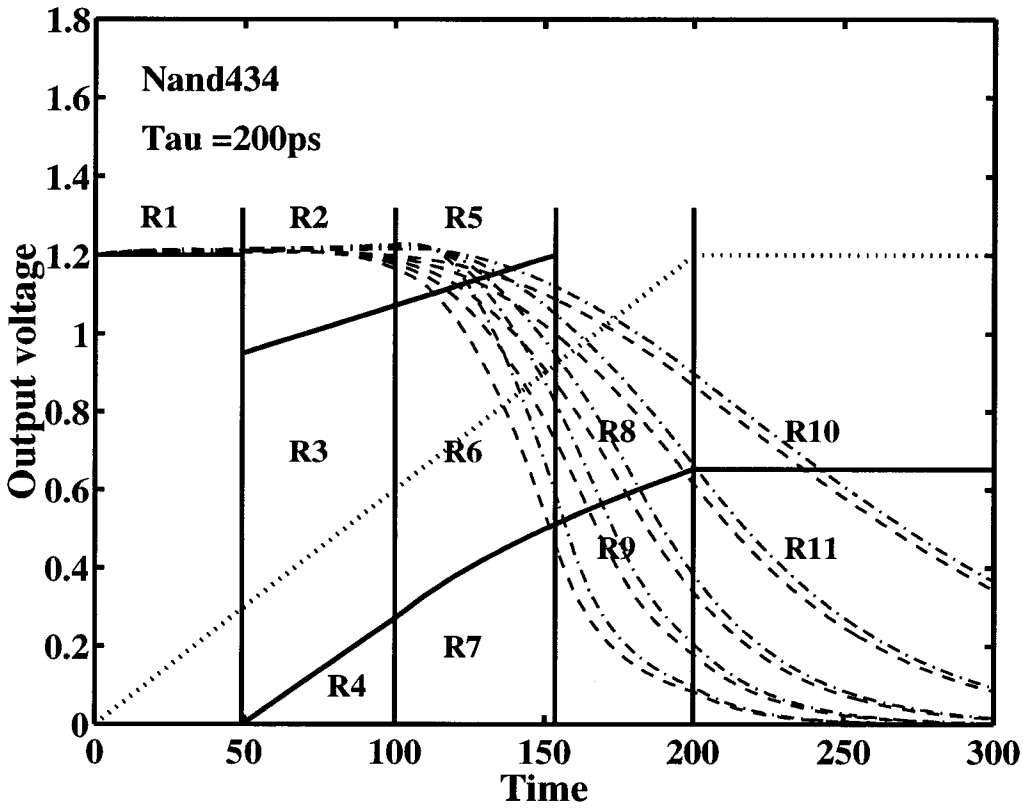


Fig. 6.18 Eleven operation regions of a complex NAND4. The typical discharging traces from SPICE BSIM3 are also plotted. Each pair of traces corresponds to the same external load capacitance for fast initial state (FIS) and slow initial state (SIS).

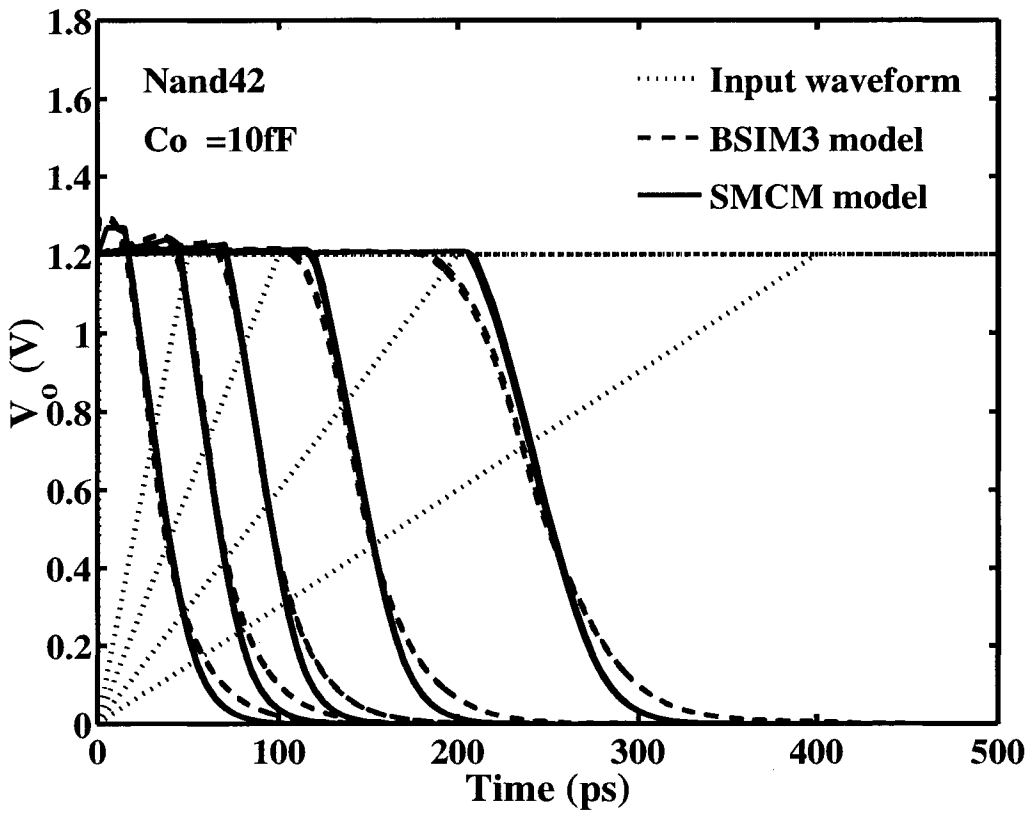


Fig. 6.19 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} for an external load capacitance of 10fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference.

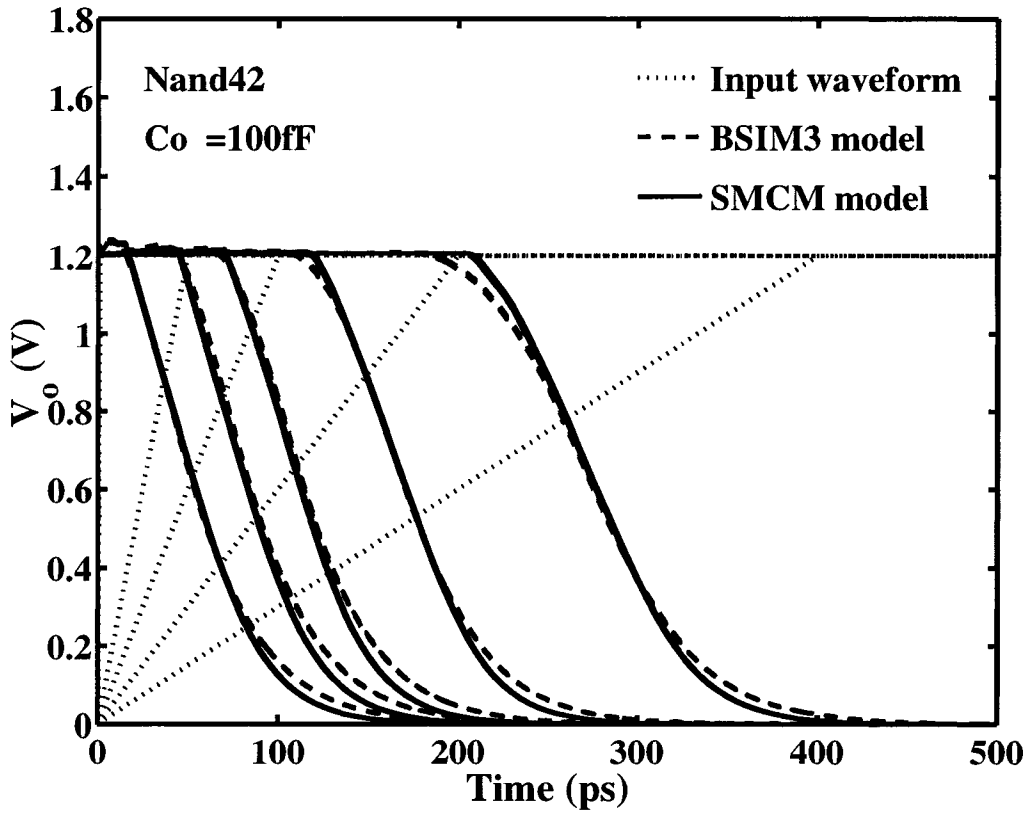


Fig. 6.20 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} for an external load capacitance of 100fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference.

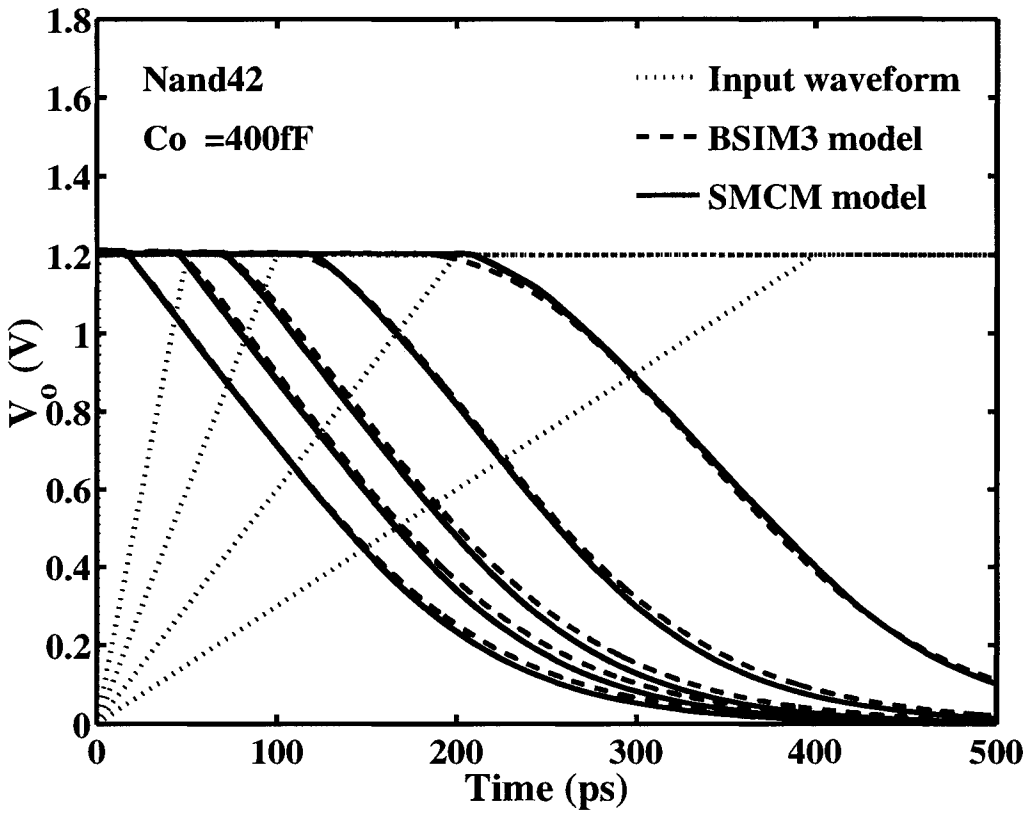


Fig. 6. 21 Output voltage waveforms of a NAND42 gate with the active input of V_{i2} for an external load capacitance of 400fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, (from left to right). Note the voltage waveforms for slow initial state (SIS) and fast initial state (FIS) each load capacitance merge into one so that we cannot see the difference.

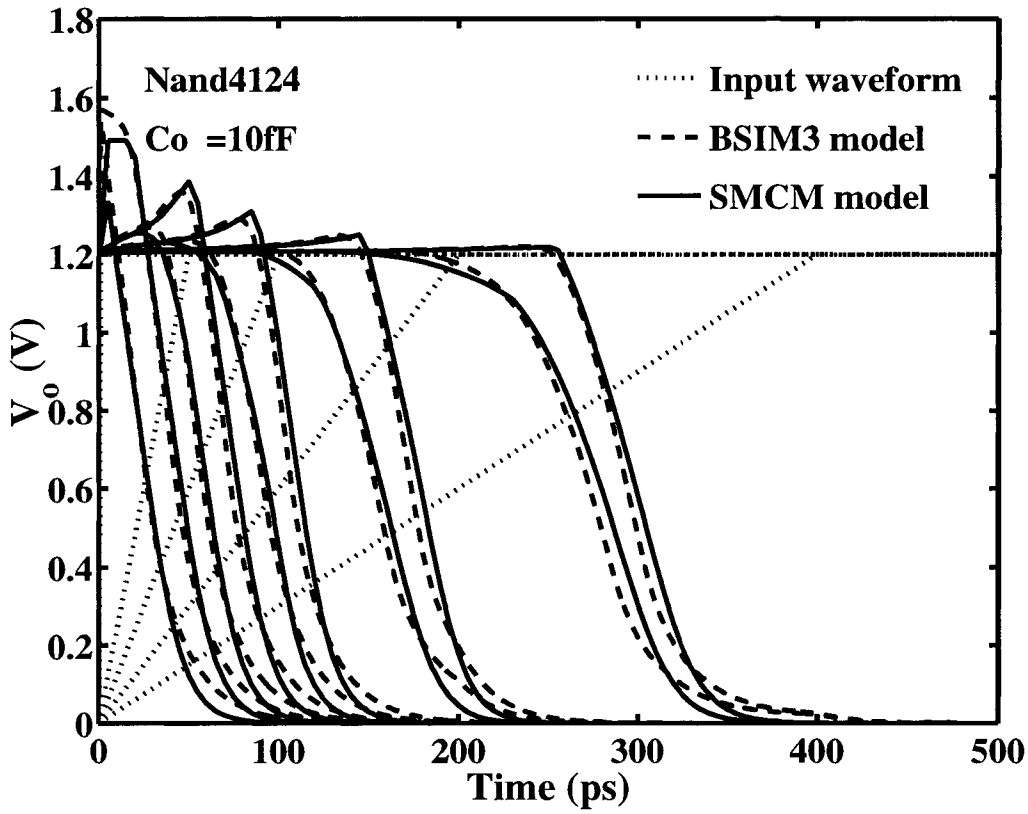


Fig. 6. 22 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for an external load capacitance of 10fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots, including a mesa for fast input, are well compensated with a trivial parallel pMOSFET cluster.

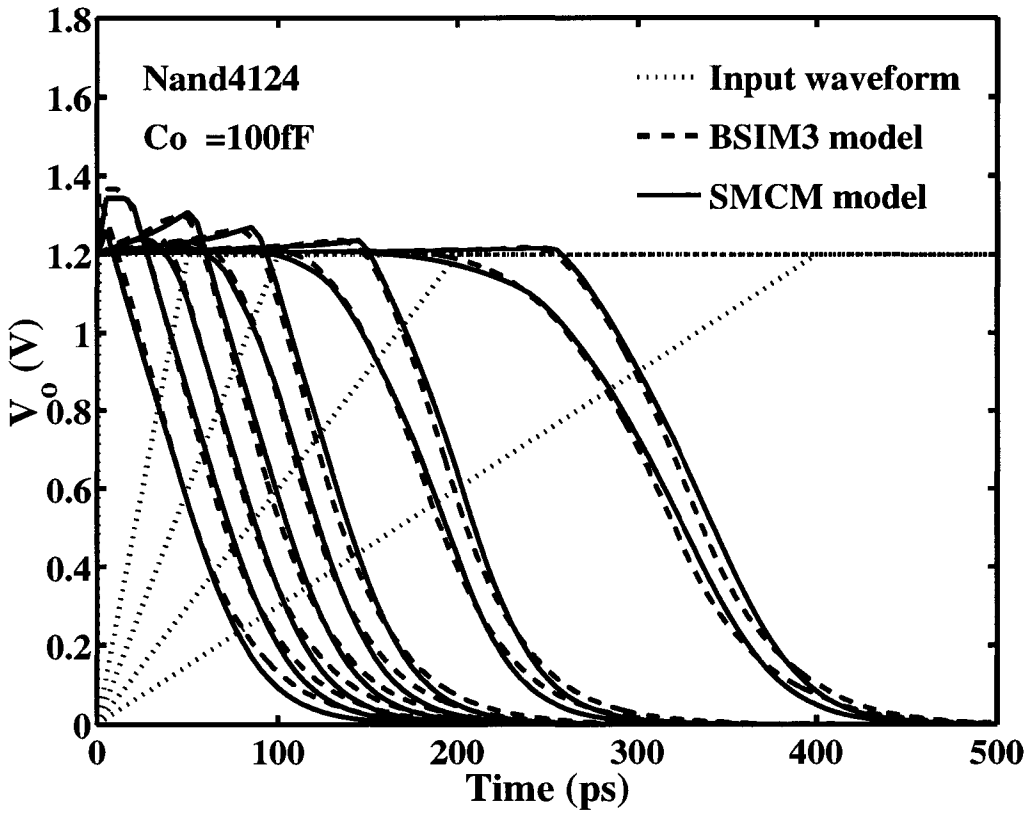


Fig. 6. 23 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for an external load capacitance of 100fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots, including a mesa for fast input, are well compensated with a trivial parallel pMOSFET cluster.

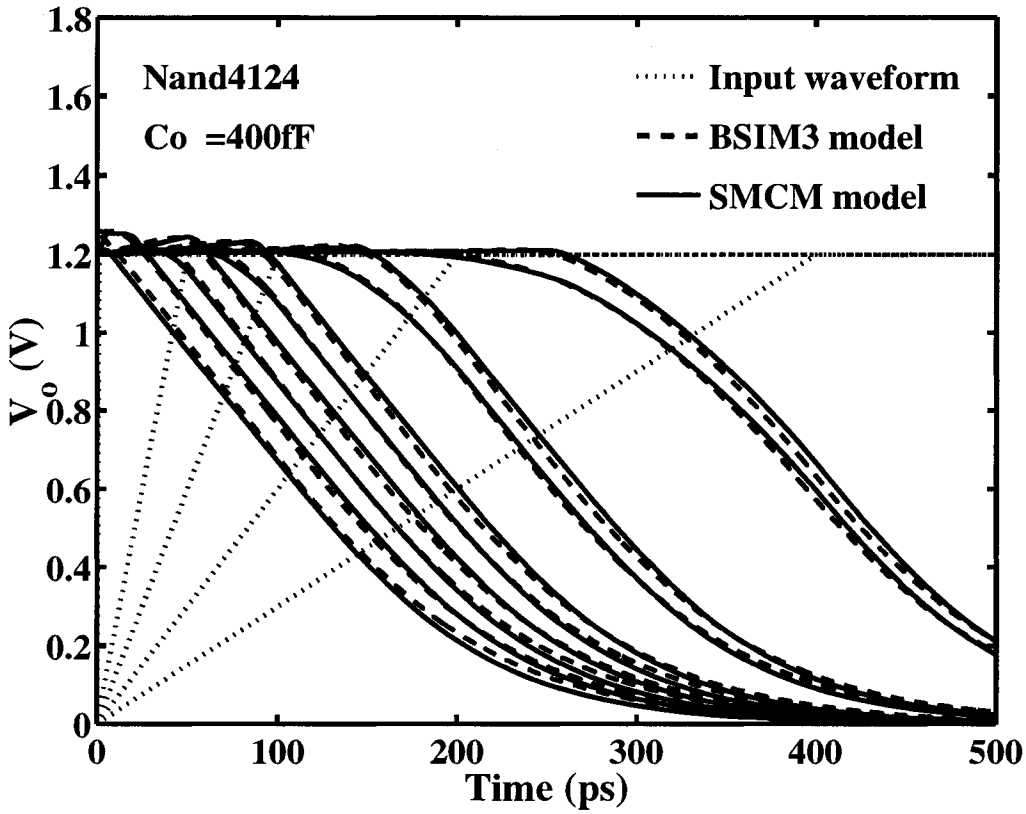


Fig. 6.24 Output voltage waveforms of a NAND4124 gate with the active inputs of V_{i1}, V_{i2}, and V_{i4} for an external load capacitance of 400fF. The voltage waveforms from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in solid lines and in dashed lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from left to right). For each input there is a pair of traces: the left for fast initial state (FIS) and the right for slow initial state (SIS). The voltage overshoots, including a mesa for fast input, are well compensated with a trivial parallel pMOSFET cluster.

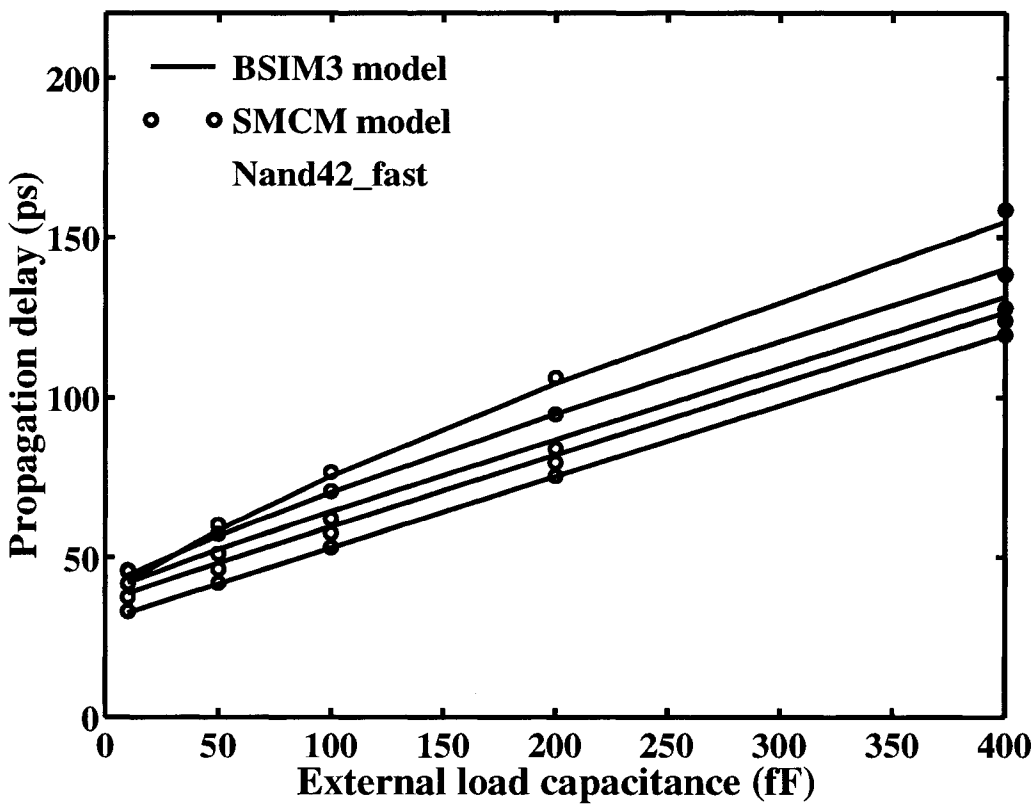


Fig. 6.25 Propagation delays of FIS versus external load capacitances for a NAND42 gate. The propagation delays from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the propagation delays of SIS are the same.

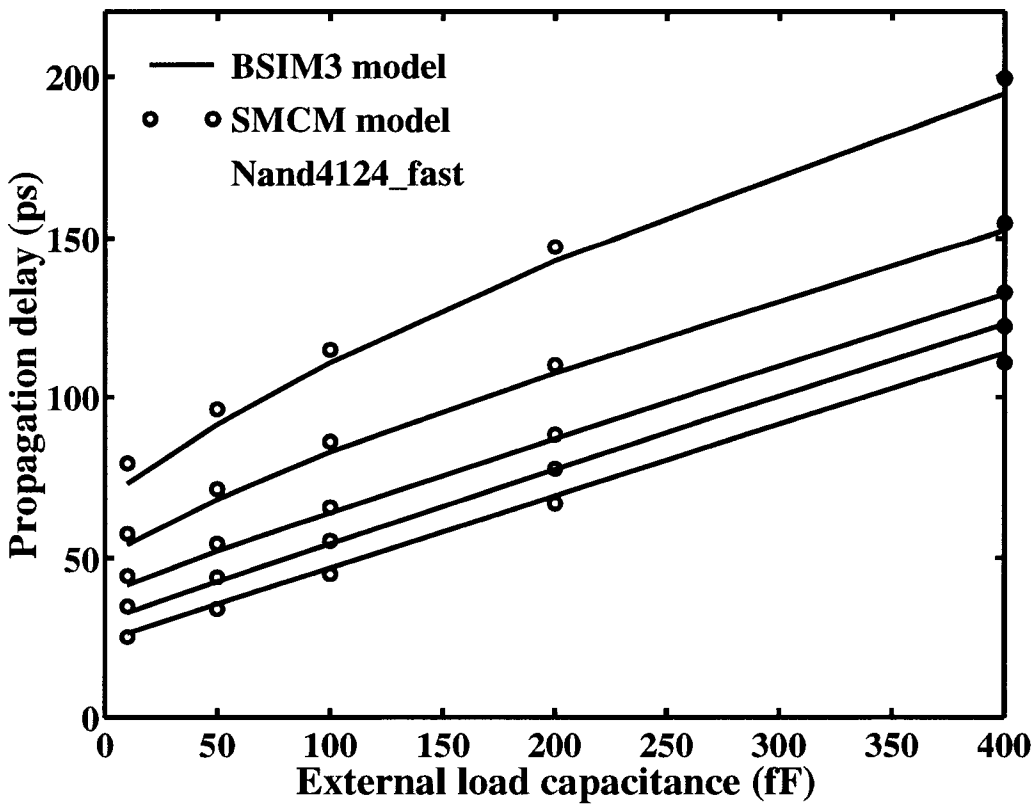


Fig. 6.26 Propagation delays of FIS versus external load capacitances for a NAND4124 gate. The propagation delays from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top). Note the activation time of FIS of a NAND4124 is zero.

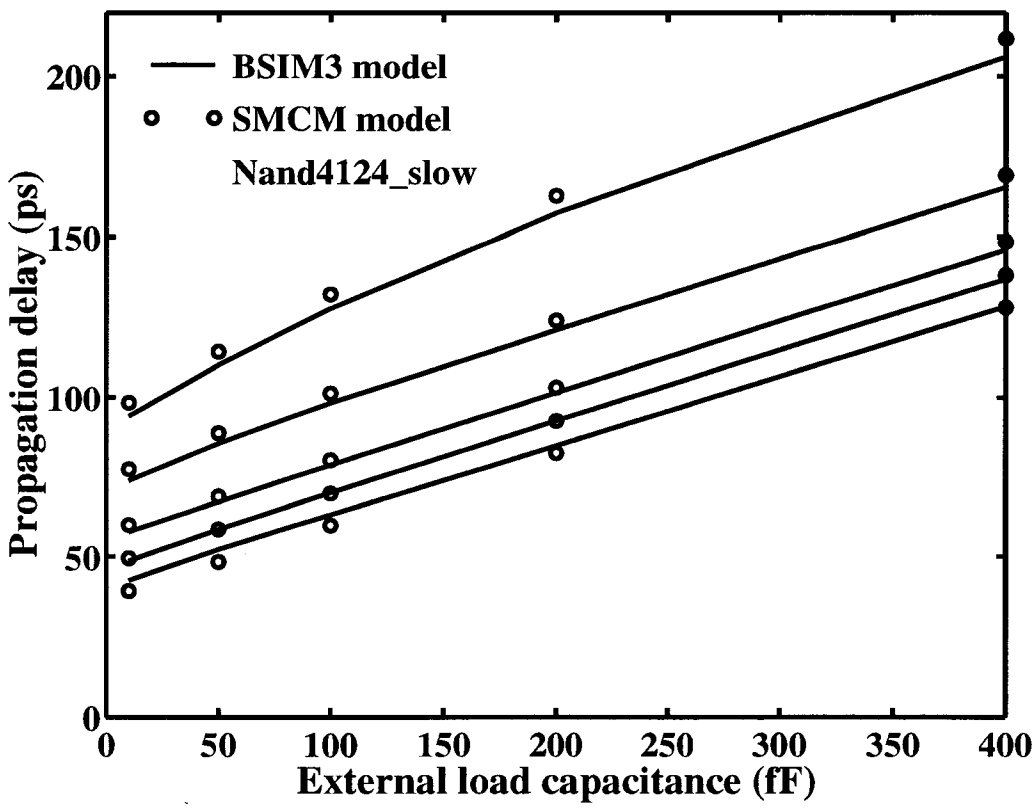


Fig. 6.27 Propagation delays of SIS versus external load capacitances for a NAND4124 gate. The propagation delays from SMCM in Section 6.4 and from SPICE BSIM3 are illustrated in small circuits and in solid lines. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively (from bottom to top).

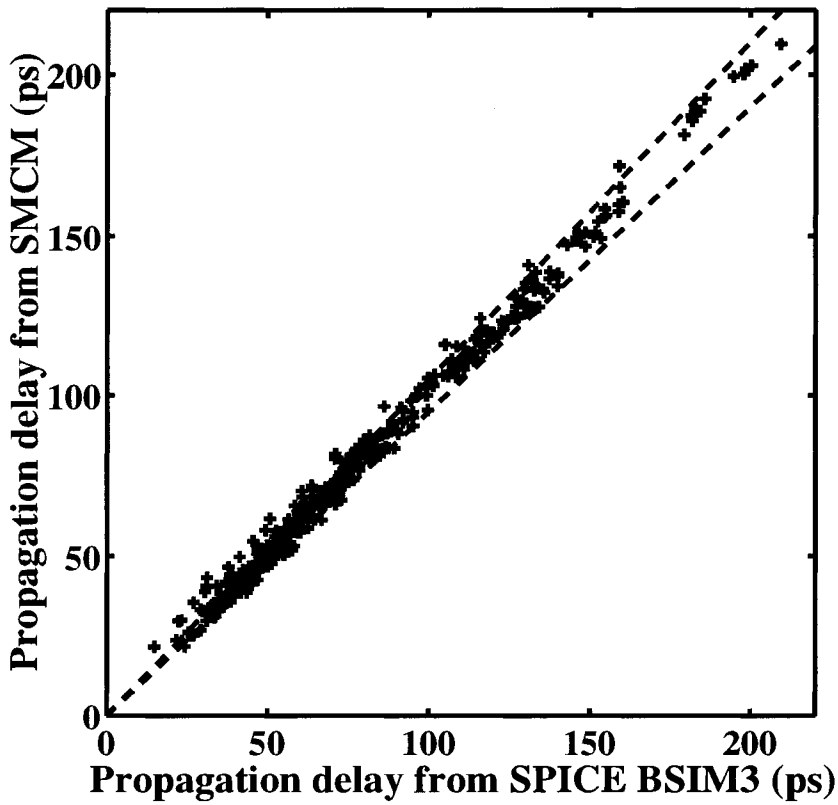


Fig. 6.28 The correlation of proposed propagation delay of fast initial state (FIS) from SMCM in Section 6.4 with propagation delay from SPICE BSIM3. There are 375 simulation points which cover all the 15 sorts of active inputs. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively.

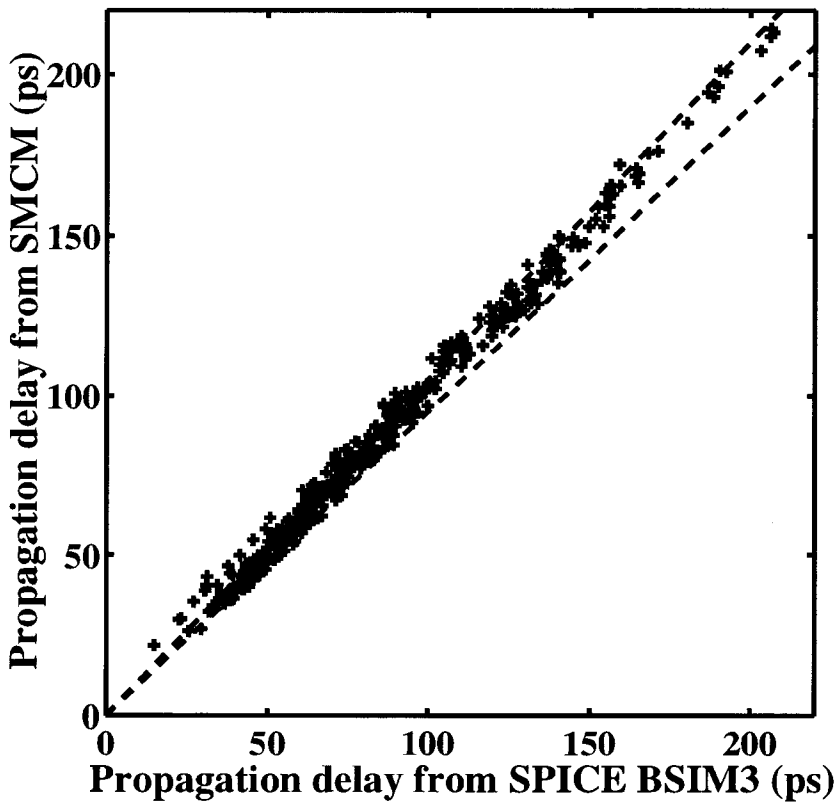


Fig. 6.29 The correlation of proposed propagation delay of slow initial state (SIS) from SMCM in Section 6.4 with propagation delay from SPICE BSIM3. There are 375 simulation points which cover all the 15 sorts of active inputs. The external load capacitances are 10fF, 50fF, 100fF, 200fF, and 400fF, respectively. The input rising times are 1ps, 50ps, 100ps, 200ps, and 400ps, respectively.

CHAPTER 7 TIMING MODEL OF A SINGLE-STAGE CMOS LOGIC GATE WITH AN ARBITRARY INPUT PATTERN

In Chapters 5 and 6, we investigate the characteristic properties of a series-connected MOSFET chain and perform timing analysis on a single-stage complex CMOS logic gate. With these investigations an equal input pattern is assumed. An equal input pattern means that all inputs are connected to either an identical input waveform or a fixed voltage depending on the type of chain or logic gate. Inputs connected to an identical input waveform are active inputs that determine the sort of chain or logic gate. Inputs connected to a fixed voltage are inactive inputs. For a series-connected nMOSFET chain or a CMOS NAND gate inactive inputs are connected to power. In Chapter 5 a generic series-connected MOSFET chain model (SMCM) and a modified input suppression technique are proposed to describe static I–V characteristics and the effect of initial states of parasitic capacitances of intermediate nodes on output transition and propagation delay. In Chapter 6 we derive the accurate closed-form expressions of output voltage of a single-stage complex CMOS logic gate with an equal input pattern. The accurate propagation delay of a single-stage CMOS logic gate with an equal input pattern is developed from the expressions of output voltage. However, in practice, an input voltage pattern to a single-stage CMOS logic gate doesn't have to be an equal input pattern and can be arbitrary. In this chapter, we propose an input mapping technique to extend the accurate timing models of a single-stage CMOS logic gate with an equal input pattern developed in Chapter 6 to any arbitrary input pattern.

Nabavi-Lishi and Rumin [70] presented a semi-empirical method for collapsing a

single-stage CMOS logic gate with an arbitrary input pattern to a single equivalent CMOS inverter with a corresponding input. One of the benefits of the reduction is that accurate timing models of a CMOS inverter can be directly exploited to obtain the propagation delay of a single-stage CMOS logic. However, as we discussed in Chapter 5, the peculiar static I–V characteristics of a series-connected MOSFET chain cannot be accurately reproduced with a single MOSFET. Furthermore, initial states of parasitic capacitances of intermediate nodes and arbitrary input patterns make the operation of a series-connected MOSFET chain much more complex than that of a single MOSFET [9]. The collapsing technique gives rise to significant errors on propagation delay.

Nikolaidis and Chatzigeorgiou [15] [75] proposed another collapsing technique for a series-connected chain. There are four steps involved in it. The first stage is that the arbitrary input pattern is mapped into an equal input pattern, i.e., the sort of a series-connected MOSFET chain and its corresponding equal input pattern transition time are determined. The second step is that the equal input pattern of the sort is empirically converted to an effective input where all inputs of the series-connected MOSFET chain are active. The third one is that the series-connected MOSFET chain gate with all inputs connected to an effective input is reduced to a two-MOSFET equivalent circuit. The one close to output node is the same as the one close to output node in a series-connected MOSFET chain. The other MOSFET is the reduction of the rest of MOSFETs in the chain. The last step is the two-MOSFET equivalent circuit is collapsed into a single equivalent MOSFET. In other words, a single-stage complex CMOS logic gate with an arbitrary input pattern is collapsed into an equivalent inverter with a corresponding input. Assumptions in the collapsing technique such as arbitrary input pattern, initial state of

parasitic capacitances and operation of MOSFETs may cause significant errors on propagation delay.

In this chapter, we propose a new and simple input mapping technique. It efficiently maps an arbitrary input pattern into an equal input pattern. Basically there are two main aspects involved in the input mapping technique. The first one is to find the active inputs from an arbitrary input pattern. The active inputs are associated with an equal input pattern. They determine the sort of a series-connected MOSFET chain with an equal input pattern. The second is to calculate the input transition time of an equal input pattern corresponding to an arbitrary input pattern. Once we have the equal input pattern corresponding to an arbitrary input pattern, the propagation delays of a single-stage CMOS logic gate with an arbitrary input pattern are obtained by exploiting the timing methodologies developed in Chapter 6.

7.1 NAMING CONVENTIONS

In Chapter 5 we have defined the naming conventions of a series-connected MOSFET chain with an equal input pattern. The index of nMOSFET and its corresponding input starts from ground as number one. The active inputs determine the sort of a series-connected MOSFET chain with an equal input pattern. In this chapter we focus on an arbitrary input pattern as well as its individual input. Assume there are totally n inputs. Before we discuss a complex arbitrary n-input pattern, let's define the relevant naming conventions.

For an individual rising ramp of input i, its starting time and input transition time are indicated with t_{si} and T_{ri}, respectively. The moment when its rising ramp reaches half

of operating voltage is defined as its middle time t_{mi} of input i $(t_{mi} = t_{si} + 0.5T_{ri})$. The moment when the rising ramp reaches operating voltage is called its ending time t_{ei} of input i $(t_{ei} = t_{si} + T_{ri})$.

For a set of an arbitrary n-input pattern consisting of its individual rising ramps, suppose that the latest starting time, the latest ending time, and the latest middle time are denoted with t_s, t_e, and t_m as shown in Fig 7.1. They are associated with inputs j, k, and l, respectively, i.e.,

$$\begin{cases} t_s = t_{sj} = \max\limits_{i=1,\ldots,n}(t_{si}) \\ t_e = t_{ek} = \max\limits_{i=1,\ldots,n}(t_{ei}) \\ t_m = t_{ml} = \max\limits_{i=1,\ldots,n}(t_{mi}) \end{cases} \tag{7.1}$$

Evidently the latest starting time t_s, the latest ending time t_e, and the latest middle time t_m are important for an arbitrary input pattern.

7.2 ARBITRARY INPUT PATTERN

It is impossible to deal with every possible arbitrary input pattern directly in analyzing propagation delays of a single-stage CMOS logic gate. We have to classify individual inputs of an arbitrary input pattern into active inputs, partially-active inputs, and inactive inputs. Active inputs determine the sort of a series-connected MOSFET chain with an equal input pattern and dictate input transition time of an equal input pattern. They dominate output voltage transition and propagation delay. Partially-active inputs make some minor corrections of input transition time of an equal input pattern. They slightly affect output voltage waveform and propagation delay. The effect of inactive inputs is insignificant and ignored. With the classified inputs, we can map an

arbitrary input pattern associated with a single-stage CMOS logic gate into an equal input pattern. The sort of the single-stage CMOS logic gate is determined with active inputs. The input transition time of an equivalent equal input pattern is decided mostly with active inputs and slightly with partially-active inputs. Then the methodologies of output voltage transition and propagation delay derived in Chapters 5 and 6 are exploited. Table 7.1 lists the classification of an individual input of an arbitrary input pattern. Fig. 7.1 shows the classification of an arbitrary input pattern.

7.2.1 Inactive Inputs

Obviously the latest starting time t_s of an arbitrary input pattern is the first criterion to determine the effect of input i. If input transition of input i finishes before the latest starting time t_s it is classified as an inactive input. Since the ending time t_{ei} of input i occurs prior to the latest starting time t_s, the effect of input i is assumed to be trivial for output voltage transition of a single-stage complex CMOS logic gate with an arbitrary input pattern. In other words the effect of all inactive inputs on output voltage transition and propagation delay is ignored. All inactive inputs are connected to power or ground depending on the type of logic gate.

7.2.2 Active Inputs

If input i has its value less than $V_{dd}/2$ at the latest starting time t_s and is still transitioning at the latest middle time t_m, it is important for output voltage transition of a complex CMOS logic gate. It is called an active input. The set of all active inputs for an arbitrary input pattern determines the sort of a series-connected MOSFET chain with an equal input pattern. Assume that there are totally m active inputs with the indices of

$m_1, \cdots, m_m$ $(m_1 < \cdots < m_m)$. According to the naming conventions of a series-connected MOSFET chain, the nMOSFET with index m_1 is closest to ground and the nMOSFET with index m_m is closest to output node.

For active input i $(i \in m_1, \cdots, m_m)$, if its starting time t_{si} is equal to the latest starting time t_s its contribution to the input transition time of an equal input pattern is T_{ri}, i.e., $T_{rai} = t_{ei} - t_{si} = t_{ei} - t_s$. If its starting time t_{si} is less than the latest starting time t_s its contribution to the input transition time of an equal input pattern is

$$T_{rai} = \frac{(t_{ei} - t_s)^2}{T_{ri}} \qquad (i \in m_1, \cdots, m_m) \quad (7.2)$$

For an arbitrary input pattern, the average value of contributions to input transition time of all active inputs dictate the input transition time of an equal input pattern.

7.2.3 Partially-Active Inputs

In addition to above-mentioned inactive inputs and active inputs, the remaining inputs are partially-active inputs. For a partially-active input, it has its value larger than $V_{dd}/2$ but less than operating voltage V_{dd} at the latest starting time t_s or its input transition finishes before the latest middle time t_m. This kind of inputs slightly affects output voltage transition but is less important than active inputs. They are not involved in determining the sort of a series-connected MOSFET chain with an equal input pattern, i.e., they are modeled to connect to operating voltage. But they may modify the input transition time of an equal input pattern to some extent. Assume that there are totally q partially-active inputs with the indices of $q_1, \cdots, q_q$ $(q_1 < \cdots < q_q)$.

For fast initial state (FIS), the parasitic capacitances of intermediate nodes

between ground and the nMOSFET of index m_m linked with the top active input are fully discharged before the transition of the top active input. If the partially-active inputs with the index less than m_m switch before the transition of the top active input, there is no redistribution of charges of parasitic capacitances involved. We can ignore their effect on output voltage transition and propagation delay. If the partially-active inputs with the index larger than m_m switch before the transition of the top active input, they may influence the output voltage transition since they may prevent charging the parasitic capacitances between the partially-active inputs and the top active input. By definition of FIS these parasitic capacitances are supposed to fully charge for the sort of the series-connected MOSFET chain. The obstruction of charging these parasitic capacitances due to the transitions of the partially-active inputs may give rise to the non-fully charged state for the sort of the series-connected MOSFET chain. In other words these partially-active inputs may make output voltage transition even faster, which is considered as smaller input transition time of an equal input pattern. Significant effect of these partial inputs on input transition time of an equal input pattern for FIS is taken into account, which is

$$T_{rpf} = -\max_{i\in(q_1,\cdots,q_q)}\left\{\frac{(t_{ei}-t_s)^3}{T_{ri}(t_e-t_s)}u(i-m_m)\right\} \tag{7.3}$$

where $u(\cdot)$ is an ideal step function to guarantee these partially-active inputs position higher that the top active input.

For slow initial state (SIS), the parasitic capacitances of intermediate nodes between output node and the nMOSFET of index m_1 linked with the bottom active input are fully charged before the transition of bottom active input. If the partially-active inputs

with the indices larger than m_1 associated with the bottom active input switch before the transition of bottom active input, there is no redistribution of charges of parasitic capacitances involved. We can ignore their effect on output voltage transition and propagation delay. The partially-active inputs with the indices less than m_1 associated with the bottom active input affect the output voltage transition since it may prevent discharging the parasitic capacitances between the bottom active input and these partially-active inputs. These parasitic capacitances are supposed to fully discharge for the sort of the series-connected MOSFET chain. The obstruction of discharging the parasitic capacitances due to the transitions of these partially-active inputs may give rise to the non-fully charged state for the sort of the series-connected MOSFET chain. In other words these partially-active inputs may make the output voltage transition even slower, which is considered as longer input transition time of an equal input pattern. Significant effect of these partial inputs on input transition time of an equal input pattern for SIS is taken into account, which is

$$T_{rps} = \max_{i\in(q_1,\cdots,q_q)}\left\{\frac{(t_{ei}-t_s)^3}{T_{ri}(t_e-t_s)}u(m_1-i)\right\} \tag{7.4}$$

where $u(\cdot)$ is an ideal step function to guarantee the partially-active inputs position lower that the bottom active input.

7.3 INPUT MAPPING TECHNIQUE

The following proposed empirical algorithm maps an arbitrary input pattern associated with a single-stage CMOS logic gate into an equal input pattern. The single-stage complex CMOS logic gate with an equal input pattern will have the same output

response of the CMOS logic gate applied with the actual arbitrary input pattern. Table 7.1 lists the classification of an individual input of an arbitrary input pattern. Fig. 7.1 plots an arbitrary input pattern mapped into active, partially-active, or inactive inputs. The following five steps of the mapping algorithm are applied for every possible input pattern:

Step 1: Find the latest starting time t_s, the latest ending time t_e, and the latest middle time t_m for an arbitrary input pattern. Assume they are associated with inputs *j*, *k*, and *l*, respectively, i.e. $t_s = t_{sj} = \max_{i=1,\ldots,n}(t_{si})$, $t_e = t_{ek} = \max_{i=1,\ldots,n}(t_{ei})$, and $t_m = t_{ml} = \max_{i=1,\ldots,n}(t_{mi})$. The latest starting time t_s is used as the starting time of the equal input pattern.

Step 2: Determine inactive inputs, m active inputs with the indices of $m_1, \cdots, m_m$ $(m_1 < \cdots < m_m)$ and q partially-active inputs with the indices of $q_1, \cdots, q_q$ $(q_1 < \cdots < q_q)$. The set of all the active inputs decides the sort of a series-connected MOSFET chain with an equal input pattern. The inactive inputs and partially-active inputs are connected to operating voltage.

Step 3: Estimate the effect of all the active inputs on the input transition time of an equal input pattern with the following empirical expression

$$T_{ra} = \frac{1}{m}\sum_{i\in(m_1,\cdots,m_m)} T_{rai} = \frac{1}{m}\sum_{i\in(m_1,\cdots,m_m)} \frac{(t_{ei}-t_s)^2}{T_{ri}} \tag{7.5}$$

Step 4: Estimate the effect of partially-active inputs on the input transition time of an equal input pattern as discussed in Subsection 7.2.3. For fast initial state (FIS), the only paritial inputs higher than the input with index m_m may influence. Their correction of the input transition time of an equal input pattern is given in (7.3). For slow initial state (SIS), the only paritial inputs lower than the input with index m_1 may influence. The

correction of the input transition time of an equal input pattern is given in (7.4).

Step 5: For fast initial state (FIS), combining the effects of active inputs and partially-active inputs the input transition time T_{rf} of an equal input pattern is

$$T_{rf} = T_{ra} + T_{rpf} = \frac{1}{m} \sum_{i \in (m_1, \cdots, m_m)} \frac{(t_{ei} - t_s)^2}{T_{ri}} - \max_{i \in (q_1, \cdots, q_q)} \left\{ \frac{(t_{ei} - t_s)^3}{T_{ri}(t_e - t_s)} u(i - m_m) \right\} \tag{7.6}$$

For slow initial state (SIS), the input transition time T_{rs} of an equal input pattern is

$$T_{rs} = T_{ra} + T_{rps} = \frac{1}{m} \sum_{i \in (m_1, \cdots, m_m)} \frac{(t_{ei} - t_s)^2}{T_{ri}} + \max_{i \in (q_1, \cdots, q_q)} \left\{ \frac{(t_{ei} - t_s)^3}{T_{ri}(t_e - t_s)} u(m_1 - i) \right\} \tag{7.7}$$

The function $u(\cdot)$ is the unit step function in (7.6) and (7.7).

7.4 RESULTS AND DISCUSSIONS

Table 7.2 lists four sets of arbitrary input patterns of a NAND4 gate for FIS. The intermediate variables used in input mapping technique as well as the sort of the NAND4 with an equal input pattern are provided. The external load capacitance is 100fF. The comparison of fast initial state propagation delays from SPICE and from input mapping technique with an equal input pattern shows the excellent accuracy of the proposed algorithm. The corresponding output voltage waveforms from SPICE and from the input mapping technique with an equal input pattern for FIS are plotted in Figs. 7.2-7.5, respectively.

Table 7.3 lists the same four sets of arbitrary input patterns as Table 7.2. The intermediate variables used in input mapping technique for SIS as well as the sort of the NAND4 with an equal input pattern are provided in Table 7.3. Note the input transition times of an equal input pattern for FIS and SIS may be different due to partially-active

inputs. The external load capacitance is 100fF. The comparison of slow initial state propagation delays from SPICE and from input mapping technique with an equal input pattern shows the accuracy of the proposed algorithm. The output voltage waveforms from SPICE and from the input mapping technique with an equal input pattern for SIS are indicated in Figs. 7.6-7.9, respectively.

Tables 7.4-7.5 list 72 sets of input patterns applied to a NAND4 gate for FIS with the external load capacitance of 100fF. The fast initial state propagation delays (ISPDs) from SPICE and from input mapping technique with an equal input pattern are provided. Tables 7.6-7.7 list the same 72 sets of input patterns applied to a NAND4 gate for FIS as in Tables 7.4-7.5. The external load capacitance is 400fF. The fast initial state propagation delays (ISPDs) from SPICE and from the input mapping technique with an equal input pattern are given in the tables. These tables show the good agreement between the fast initial state propagation delays (ISPDs) from SPICE and those from the input mapping technique with an equal input pattern. Fig. 7.10 displays the correlation of fast initial state propagation delays (ISPDs) from SPICE and from the input mapping technique with an equal input pattern. They match with each other excelletly.

Tables 7.8-7.9 list the same 72 sets of input patterns applied to a NAND4 gate for SIS as Tables 7.4-7.5. The external load capacitance of 100fF. The slow initial state propagation delays (ISPD) from SPICE and from the input mapping technique with an equal input pattern are also shown in the tables. Tables 7.10-7.11 list the same sets of 72 input patterns applied to a NAND4 gate for SIS as in Tables 7.8-7.9. The external capacitance is 400fF. The slow initial state propagation delays (ISPDs) from SPICE and from the input mapping technique with an equal input pattern are also listed in the tables.

These tables show the good agreement between the slow initial state propagation delays (ISPDs) from SPICE and those from the input mapping technique with an equal input pattern. Fig. 7.11 displays the correlation of slow initial state propagation delays (ISPDs) from SPICE and from the input mapping technique with an equal input pattern. They match with each other excelletly.

In this chapter we discuss the effect of arbitrary input patterns by classifying them into active inputs, partially-active inputs, and inactive inputs. The active inputs determine the sort of a series-connected MOSFET chain with an equal input pattern. The input transition time of an equal input pattern is decided mainly with the active inputs and slightly with the partially-active inputs. The effect of arbitrary input patterns of CMOS logic gates on propagation delays with two extreme initial states of FIS and SIS is discussed. An efficient and accurate mapping algorithm for a single-stage complex CMOS logic gate with an arbitrary input pattern to a CMOS logic gate with an equal input pattern is proposed. The output voltages and propagation delays developed in Chapters 5 and 6 can be directly exploited for a CMOS gate with an equal input pattern.

Table 7.1 Classification of an Individual Input of an Arbitrary Input Pattern

	$0 \le V_i(t_s) < 0.5V_{dd}$	$0.5V_{dd} \le V_i(t_s) \le V_{dd}$
$t_{ei} \le t_s$	Inactive input	
$t_s < t_{ei} \le t_m$	Partially active input	
$t_m < t_{ei}$	Active input	Partially active input

Table 7.2: Four Examples of Input Mapping Technique for FIS

	Case I	Case II	Cas III	Case IV
t_{s1} (ps)	20	0	20	0
T_{r1} (ps)	160	100	40	100
t_{s2} (ps)	0	40	0	40
T_{r2} (ps)	100	50	100	50
t_{s3} (ps)	40	0	40	0
T_{r3} (ps)	50	70	50	70
t_{s4} (ps)	0	20	0	20
T_{r4} (ps)	70	160	70	40
t_s (ps)	40	40	40	40
t_m (ps)	100	100	65	65
t_e (ps)	180	180	100	100
T_{rf} (ps)	104.6	122.5	36.6	36.6
sort	nand41	nand44	nand423	nand412
t_{df_spice} (ps)*	53.9	61.8	53.4	51.5
t_{df_cal} (ps)*	54.1	63.3	50.9	54.3
Error (%)	0.3	2.4	-4.7	5.5

- Propagation delays count from middle time t_m.

Table 7.3: Four Examples of Input Mapping Technique for SIS

	Case I	Case II	Case III	Case IV
t_{s1} (ps)	20	0	20	0
T_{r1} (ps)	160	100	40	100
t_{s2} (ps)	0	40	0	40
T_{r2} (ps)	100	50	100	50
t_{s3} (ps)	40	0	40	0
T_{r3} (ps)	50	70	50	70
t_{s4} (ps)	0	20	0	20
T_{r4} (ps)	70	160	70	40
t_s (ps)	40	40	40	40
t_m (ps)	100	100	65	65
t_e (ps)	180	180	100	100
T_{rs} (ps)	122.5	140.4	46.3	43.0
sort	nand41	nand44	nand423	nand412
t_{ds_spice} (ps)*	70.2	64.8	61.6	66.5
t_{ds_cal} (ps)*	65.2	74.6	62.3	62.4
Error (%)	-7.1	15.2	1.2	-6.2

- Propagation delays count from middle time t_m.

Table 7.4 Propagation Delays from SPICE and Input Mapping Technique for FIS

t_{s1}	t_{s2}	t_{s3}	t_{s4}	T_{r1}	T_{r2}	T_{r3}	T_{r4}	t_{df_spice}	$t_{df\_collapse}$	Error(%)
20	0	40	0	160	100	50	70	53.9	54.1	0.30
20	0	0	40	160	100	70	50	49.9	54.1	8.34
20	40	0	0	160	50	100	70	55.4	54.1	-2.42
20	40	0	0	160	50	70	100	51.8	54.1	4.37
20	0	0	40	160	70	100	50	50.1	54.1	7.91
20	0	40	0	160	70	50	100	52.0	54.1	3.97
0	20	40	0	100	160	50	70	58.9	55.0	-6.64
0	20	0	40	100	160	70	50	54.1	55.0	1.64
0	40	20	0	100	50	160	70	62.8	62.0	-1.30
0	40	0	20	100	50	70	160	61.8	63.3	2.37
0	0	20	40	100	70	160	50	59.1	52.3	-11.55
0	0	40	20	100	70	50	160	61.6	63.3	2.70
40	20	0	0	50	160	100	70	59.7	56.4	-5.46
40	20	0	0	50	160	70	100	55.6	56.4	1.51
40	0	20	0	50	100	160	70	62.9	62.0	-1.46
40	0	0	20	50	100	70	160	61.9	63.3	2.20
40	0	20	0	50	70	160	100	60.5	53.8	-11.04
40	0	0	20	50	70	100	160	62.1	63.3	1.87
0	20	0	40	70	160	100	50	54.5	55.0	0.90
0	20	40	0	70	160	50	100	56.5	55.0	-2.67
0	0	20	40	70	100	160	50	59.6	52.3	-12.29
0	0	40	20	70	100	50	160	61.9	63.3	2.20
0	40	20	0	70	50	160	100	60.6	53.8	-11.19
0	40	0	20	70	50	100	160	62.0	63.3	2.04
20	0	40	0	40	100	50	70	53.4	50.9	-4.72
20	0	0	40	40	100	70	50	50.4	49.1	-2.59
20	40	0	0	40	50	100	70	53.0	50.9	-4.00
20	40	0	0	40	50	70	100	53.0	49.1	-7.36
20	0	0	40	40	70	100	50	51.3	49.5	-3.49
20	0	40	0	40	70	50	100	52.4	49.5	-5.51
0	20	40	0	100	40	50	70	52.5	50.5	-3.79
0	20	0	40	100	40	70	50	50.3	49.6	-1.49
0	40	20	0	100	50	40	70	52.4	54.3	3.71
0	40	0	20	100	50	70	40	51.5	54.3	5.52
0	0	20	40	100	70	40	50	50.1	49.6	-1.09
0	0	40	20	100	70	50	40	52.3	52.5	0.44

Table 7.5 Propagation Delays from SPICE and Input Mapping Technique for FIS

t_{s1}	t_{s2}	t_{s3}	t_{s4}	T_{r1}	T_{r2}	T_{r3}	T_{r4}	t_{df_spice}	$t_{df\_collapse}$	Error(%)
40	20	0	0	50	40	100	70	54.0	50.5	-6.47
40	20	0	0	50	40	70	100	53.8	49.6	-7.89
40	0	20	0	50	100	40	70	50.9	54.3	6.76
40	0	0	20	50	100	70	40	50.2	54.3	8.25
40	0	20	0	50	70	40	100	53.7	49.6	-7.72
40	0	0	20	50	70	100	40	53.3	52.5	-1.45
0	20	0	40	70	40	100	50	51.2	49.5	-3.30
0	20	40	0	70	40	50	100	52.3	49.5	-5.33
0	0	20	40	70	100	40	50	50.4	49.1	-2.59
0	0	40	20	70	100	50	40	53.1	52.9	-0.34
0	40	20	0	70	50	40	100	53.3	49.1	-7.89
0	40	0	20	70	50	100	40	52.6	52.9	0.61
20	0	40	0	40	50	50	70	48.8	48.1	-1.42
20	0	0	40	40	50	70	50	46.7	50.5	8.06
20	40	0	0	40	50	50	70	48.2	53.4	10.78
20	40	0	0	40	50	70	50	50.4	53.4	5.94
20	0	0	40	40	70	50	50	46.4	50.5	8.76
20	0	40	0	40	70	50	50	50.8	52.5	3.37
0	20	40	0	50	40	50	70	48.9	48.1	-1.62
0	20	0	40	50	40	70	50	46.7	50.5	8.06
0	40	20	0	50	50	40	70	48.8	53.4	9.41
0	40	0	20	50	50	70	40	48.1	53.4	11.01
0	0	20	40	50	70	40	50	46.6	50.5	8.29
0	0	40	20	50	70	50	40	48.9	50.3	2.91
40	20	0	0	50	40	50	70	47.4	55.0	16.03
40	20	0	0	50	40	70	50	49.1	55.0	12.01
40	0	20	0	50	50	40	70	47.8	55.0	15.06
40	0	0	20	50	50	70	40	46.9	55.0	17.27
40	0	20	0	50	70	40	50	48.3	55.0	13.87
40	0	0	20	50	70	50	40	46.3	55.0	18.79
0	20	0	40	70	40	50	50	46.4	50.5	8.76
0	20	40	0	70	40	50	50	50.6	52.5	3.78
0	0	20	40	70	50	40	50	46.3	50.5	9.00
0	0	40	20	70	50	50	40	48.5	50.3	3.75
0	40	20	0	70	50	40	50	50.3	55.6	10.48
0	40	0	20	70	50	50	40	47.6	55.6	16.75

Table 7.6 Propagation Delays from SPICE and Input Mapping Technique for FIS

t_{s1}	t_{s2}	t_{s3}	t_{s4}	T_{r1}	T_{r2}	T_{r3}	T_{r4}	t_{df_spice}	$t_{df\_collapse}$	Error(%)
20	0	40	0	160	100	50	70	122.5	119.5	-2.45
20	0	0	40	160	100	70	50	120.7	119.5	-1.00
20	40	0	0	160	50	100	70	123.1	119.5	-2.93
20	40	0	0	160	50	70	100	121.5	119.5	-1.65
20	0	0	40	160	70	100	50	120.9	119.5	-1.16
20	0	40	0	160	70	50	100	121.8	119.5	-1.89
0	20	40	0	100	160	50	70	127.7	119.7	-6.23
0	20	0	40	100	160	70	50	125.4	119.7	-4.51
0	40	20	0	100	50	160	70	131.8	126.8	-3.76
0	40	0	20	100	50	70	160	131.1	129.5	-1.21
0	0	20	40	100	70	160	50	129.7	117.2	-9.66
0	0	40	20	100	70	50	160	130.8	129.5	-0.98
40	20	0	0	50	160	100	70	127.9	121.2	-5.24
40	20	0	0	50	160	70	100	126.0	121.2	-3.82
40	0	20	0	50	100	160	70	132.0	126.8	-3.91
40	0	0	20	50	100	70	160	131.1	129.5	-1.21
40	0	20	0	50	70	160	100	130.5	118.7	-9.02
40	0	0	20	50	70	100	160	131.4	129.5	-1.44
0	20	0	40	70	160	100	50	125.4	119.7	-4.51
0	20	40	0	70	160	50	100	126.4	119.7	-5.27
0	0	20	40	70	100	160	50	130.1	117.2	-9.94
0	0	40	20	70	100	50	160	131.1	129.5	-1.21
0	40	20	0	70	50	160	100	130.5	118.7	-9.02
0	40	0	20	70	50	100	160	131.2	129.5	-1.29
20	0	40	0	40	100	50	70	121.3	116.4	-4.03
20	0	0	40	40	100	70	50	119.7	115.3	-3.66
20	40	0	0	40	50	100	70	120.4	116.4	-3.31
20	40	0	0	40	50	70	100	121.6	115.3	-5.16
20	0	0	40	40	70	100	50	120.5	116.0	-3.77
20	0	40	0	40	70	50	100	120.8	116.0	-4.01
0	20	40	0	100	40	50	70	120.4	115.9	-3.75
0	20	0	40	100	40	70	50	119.6	116.0	-3.03
0	40	20	0	100	50	40	70	119.9	119.3	-0.54
0	40	0	20	100	50	70	40	119.3	119.3	-0.04
0	0	20	40	100	70	40	50	119.5	116.0	-2.95
0	0	40	20	100	70	50	40	120.3	117.9	-1.97

Table 7.7 Propagation Delays from SPICE and Input Mapping Technique for FIS

t_{s1}	t_{s2}	t_{s3}	t_{s4}	T_{r1}	T_{r2}	T_{r3}	T_{r4}	t_{df_spice}	$t_{df\_collapse}$	Error(%)
40	20	0	0	50	40	100	70	121.7	115.9	-4.77
40	20	0	0	50	40	70	100	122.6	116.0	-5.40
40	0	20	0	50	100	40	70	118.5	119.3	0.64
40	0	0	20	50	100	70	40	117.9	119.3	1.15
40	0	20	0	50	70	40	100	122.5	116.0	-5.32
40	0	0	20	50	70	100	40	121.3	117.9	-2.78
0	20	0	40	70	40	100	50	120.4	116.0	-3.69
0	20	40	0	70	40	50	100	120.6	116.0	-3.85
0	0	20	40	70	100	40	50	119.7	115.3	-3.66
0	0	40	20	70	100	50	40	121.1	118.5	-2.16
0	40	20	0	70	50	40	100	122.0	115.3	-5.47
0	40	0	20	70	50	100	40	120.3	118.5	-1.51
20	0	40	0	40	50	50	70	116.9	113.4	-2.97
20	0	0	40	40	50	70	50	116.0	115.4	-0.48
20	40	0	0	40	50	50	70	116.5	118.4	1.63
20	40	0	0	40	50	70	50	117.3	118.4	0.93
20	0	0	40	40	70	50	50	115.7	115.4	-0.22
20	0	40	0	40	70	50	50	118.3	117.6	-0.56
0	20	40	0	50	40	50	70	117.0	113.4	-3.05
0	20	0	40	50	40	70	50	115.9	115.4	-0.40
0	40	20	0	50	50	40	70	116.6	118.4	1.54
0	40	0	20	50	50	70	40	116.2	118.4	1.89
0	0	20	40	50	70	40	50	115.9	115.4	-0.40
0	0	40	20	50	70	50	40	117.2	115.5	-1.43
40	20	0	0	50	40	50	70	115.6	120.6	4.32
40	20	0	0	50	40	70	50	115.9	120.6	4.05
40	0	20	0	50	50	40	70	116.0	120.6	3.96
40	0	0	20	50	50	70	40	115.4	120.6	4.50
40	0	20	0	50	70	40	50	115.4	120.6	4.50
40	0	0	20	50	70	50	40	114.8	120.6	5.04
0	20	0	40	70	40	50	50	115.7	115.4	-0.22
0	20	40	0	70	40	50	50	118.1	117.6	-0.39
0	0	20	40	70	50	40	50	115.6	115.4	-0.14
0	0	40	20	70	50	50	40	116.7	115.5	-1.01
0	40	20	0	70	50	40	50	117.3	120.6	2.80
0	40	0	20	70	50	50	40	116.2	120.6	3.77

Table 7.8 Propagation Delays from SPICE and Input Mapping Technique for SIS

t_{s1}	t_{s2}	t_{s3}	t_{s4}	T_{r1}	T_{r2}	T_{r3}	T_{r4}	t_{ds_spice}	$t_{ds\_collapse}$	Error(%)
20	0	40	0	160	100	50	70	70.2	65.2	-7.09
20	0	0	40	160	100	70	50	70.9	65.2	-8.00
20	40	0	0	160	50	100	70	70.1	65.2	-6.95
20	40	0	0	160	50	70	100	70.6	65.2	-7.61
20	0	0	40	160	70	100	50	70.7	65.2	-7.74
20	0	40	0	160	70	50	100	70.3	65.2	-7.22
0	20	40	0	100	160	50	70	71.4	76.1	6.58
0	20	0	40	100	160	70	50	72.2	76.1	5.40
0	40	20	0	100	50	160	70	67.8	75.0	10.59
0	40	0	20	100	50	70	160	64.8	74.6	15.17
0	0	20	40	100	70	160	50	67.3	73.5	9.19
0	0	40	20	100	70	50	160	63.6	74.6	17.34
40	20	0	0	50	160	100	70	71.7	77.5	8.13
40	20	0	0	50	160	70	100	72.1	77.5	7.53
40	0	20	0	50	100	160	70	68.1	75.0	10.10
40	0	0	20	50	100	70	160	65.3	74.6	14.28
40	0	20	0	50	70	160	100	68.0	75.0	10.27
40	0	0	20	50	70	100	160	65.0	74.6	14.81
0	20	0	40	70	160	100	50	71.7	68.4	-4.62
0	20	40	0	70	160	50	100	71.4	68.4	-4.22
0	0	20	40	70	100	160	50	67.5	73.5	8.86
0	0	40	20	70	100	50	160	63.6	74.6	17.34
0	40	20	0	70	50	160	100	67.7	75.0	10.76
0	40	0	20	70	50	100	160	64.6	74.6	15.52
20	0	40	0	40	100	50	70	61.6	62.3	1.18
20	0	0	40	40	100	70	50	59.7	66.0	10.59
20	40	0	0	40	50	100	70	66.2	62.3	-5.85
20	40	0	0	40	50	70	100	65.3	66.0	1.10
20	0	0	40	40	70	100	50	57.5	59.6	3.69
20	0	40	0	40	70	50	100	59.6	59.6	0.04
0	20	40	0	100	40	50	70	61.6	64.7	5.04
0	20	0	40	100	40	70	50	58.9	66.5	12.92
0	40	20	0	100	50	40	70	66.4	62.4	-6.03
0	40	0	20	100	50	70	40	66.5	62.4	-6.17
0	0	20	40	100	70	40	50	59.0	66.5	12.73
0	0	40	20	100	70	50	40	61.9	64.7	4.53

Table 7.9 Propagation Delays from SPICE and Input Mapping Technique for SIS

t_{s1}	t_{s2}	t_{s3}	t_{s4}	T_{r1}	T_{r2}	T_{r3}	T_{r4}	t_{ds_spice}	$t_{ds\_collapse}$	Error(%)
40	20	0	0	50	40	100	70	68.1	64.7	-4.98
40	20	0	0	50	40	70	100	66.6	66.5	-0.14
40	0	20	0	50	100	40	70	68.6	62.4	-9.04
40	0	0	20	50	100	70	40	68.6	62.4	-9.04
40	0	20	0	50	70	40	100	66.7	66.5	-0.29
40	0	0	20	50	70	100	40	68.2	64.7	-5.12
0	20	0	40	70	40	100	50	57.1	59.6	4.42
0	20	40	0	70	40	50	100	59.4	59.6	0.37
0	0	20	40	70	100	40	50	59.4	68.1	14.62
0	0	40	20	70	100	50	40	61.7	64.4	4.33
0	40	20	0	70	50	40	100	65.5	68.1	3.95
0	40	0	20	70	50	100	40	66.4	64.4	-3.05
20	0	40	0	40	50	50	70	56.1	55.2	-1.53
20	0	0	40	40	50	70	50	50.7	55.6	9.59
20	40	0	0	40	50	50	70	63.5	61.3	-3.52
20	40	0	0	40	50	70	50	63.1	61.3	-2.91
20	0	0	40	40	70	50	50	51.6	55.6	7.68
20	0	40	0	40	70	50	50	56.6	57.6	1.70
0	20	40	0	50	40	50	70	55.9	55.2	-1.18
0	20	0	40	50	40	70	50	50.1	55.6	10.91
0	40	20	0	50	50	40	70	63.4	59.2	-6.70
0	40	0	20	50	50	70	40	63.5	59.2	-6.85
0	0	20	40	50	70	40	50	49.8	55.6	11.57
0	0	40	20	50	70	50	40	56.4	57.6	2.06
40	20	0	0	50	40	50	70	64.8	60.2	-7.09
40	20	0	0	50	40	70	50	64.5	60.2	-6.66
40	0	20	0	50	50	40	70	64.7	60.2	-6.95
40	0	0	20	50	50	70	40	65.0	60.2	-7.38
40	0	20	0	50	70	40	50	64.5	60.2	-6.66
40	0	0	20	50	70	50	40	65.1	60.2	-7.52
0	20	0	40	70	40	50	50	50.8	55.6	9.38
0	20	40	0	70	40	50	50	56.3	57.6	2.24
0	0	20	40	70	50	40	50	49.9	55.6	11.35
0	0	40	20	70	50	50	40	56.3	57.6	2.24
0	40	20	0	70	50	40	50	63.3	63.4	0.23
0	40	0	20	70	50	50	40	63.9	63.4	-0.71

Table 7.10 Propagation Delays from SPICE and Input Mapping Technique for SIS

t_{s1}	t_{s2}	t_{s3}	t_{s4}	T_{r1}	T_{r2}	T_{r3}	T_{r4}	t_{ds_spice}	$t_{ds\_collapse}$	Error(%)
20	0	40	0	160	100	50	70	137.2	130.4	-4.93
20	0	0	40	160	100	70	50	137.9	130.4	-5.41
20	40	0	0	160	50	100	70	137.4	130.4	-5.07
20	40	0	0	160	50	70	100	137.5	130.4	-5.14
20	0	0	40	160	70	100	50	138.0	130.4	-5.48
20	0	40	0	160	70	50	100	137.5	130.4	-5.14
0	20	40	0	100	160	50	70	141.4	140.7	-0.51
0	20	0	40	100	160	70	50	142.1	140.7	-1.00
0	40	20	0	100	50	160	70	138.9	140.1	0.83
0	40	0	20	100	50	70	160	133.9	141.5	5.64
0	0	20	40	100	70	160	50	138.4	138.5	0.07
0	0	40	20	100	70	50	160	132.6	141.5	6.68
40	20	0	0	50	160	100	70	141.5	142.1	0.45
40	20	0	0	50	160	70	100	141.8	142.1	0.24
40	0	20	0	50	100	160	70	139.0	140.1	0.76
40	0	0	20	50	100	70	160	133.9	141.5	5.64
40	0	20	0	50	70	160	100	138.7	140.1	0.98
40	0	0	20	50	70	100	160	133.8	141.5	5.72
0	20	0	40	70	160	100	50	141.8	133.0	-6.18
0	20	40	0	70	160	50	100	141.3	133.0	-5.85
0	0	20	40	70	100	160	50	138.6	138.5	-0.07
0	0	40	20	70	100	50	160	132.8	141.5	6.52
0	40	20	0	70	50	160	100	138.7	140.1	0.98
0	40	0	20	70	50	100	160	133.7	141.5	5.80
20	0	40	0	40	100	50	70	130.8	127.9	-2.20
20	0	0	40	40	100	70	50	128.0	133.3	4.15
20	40	0	0	40	50	100	70	135.3	127.9	-5.45
20	40	0	0	40	50	70	100	134.3	133.3	-0.73
20	0	0	40	40	70	100	50	126.1	127.1	0.82
20	0	40	0	40	70	50	100	128.9	127.1	-1.37
0	20	40	0	100	40	50	70	130.3	130.1	-0.13
0	20	0	40	100	40	70	50	126.7	133.7	5.52
0	40	20	0	100	50	40	70	135.2	127.3	-5.81
0	40	0	20	100	50	70	40	135.5	127.3	-6.02
0	0	20	40	100	70	40	50	126.6	133.7	5.61
0	0	40	20	100	70	50	40	130.7	130.1	-0.44

Table 7.11 Propagation Delays from SPICE and Input Mapping Technique for SIS

t_{s1}	t_{s2}	t_{s3}	t_{s4}	T_{r1}	T_{r2}	T_{r3}	T_{r4}	t_{ds_spice}	$t_{ds\_collapse}$	Error(%)
40	20	0	0	50	40	100	70	135.8	130.1	-4.18
40	20	0	0	50	40	70	100	133.9	133.7	-0.15
40	0	20	0	50	100	40	70	135.1	127.3	-5.74
40	0	0	20	50	100	70	40	135.3	127.3	-5.88
40	0	20	0	50	70	40	100	133.8	133.7	-0.08
40	0	0	20	50	70	100	40	136.1	130.1	-4.39
0	20	0	40	70	40	100	50	126.1	127.1	0.82
0	20	40	0	70	40	50	100	128.8	127.1	-1.30
0	0	20	40	70	100	40	50	127.9	135.4	5.87
0	0	40	20	70	100	50	40	131.1	130.0	-0.85
0	40	20	0	70	50	40	100	134.4	135.4	0.75
0	40	0	20	70	50	100	40	135.7	130.0	-4.21
20	0	40	0	40	50	50	70	125.6	120.3	-4.23
20	0	0	40	40	50	70	50	119.4	120.7	1.12
20	40	0	0	40	50	50	70	132.4	126.3	-4.62
20	40	0	0	40	50	70	50	132.3	126.3	-4.55
20	0	0	40	40	70	50	50	119.9	120.7	0.70
20	0	40	0	40	70	50	50	126.2	122.6	-2.89
0	20	40	0	50	40	50	70	125.6	120.3	-4.23
0	20	0	40	50	40	70	50	119.3	120.7	1.21
0	40	20	0	50	50	40	70	132.3	124.2	-6.15
0	40	0	20	50	50	70	40	132.6	124.2	-6.36
0	0	20	40	50	70	40	50	118.8	120.7	1.63
0	0	40	20	50	70	50	40	126.2	122.6	-2.89
40	20	0	0	50	40	50	70	131.4	125.8	-4.25
40	20	0	0	50	40	70	50	131.5	125.8	-4.33
40	0	20	0	50	50	40	70	131.6	125.8	-4.40
40	0	0	20	50	50	70	40	132.0	125.8	-4.69
40	0	20	0	50	70	40	50	131.3	125.8	-4.18
40	0	0	20	50	70	50	40	131.7	125.8	-4.47
0	20	0	40	70	40	50	50	119.7	120.7	0.87
0	20	40	0	70	40	50	50	126.1	122.6	-2.81
0	0	20	40	70	50	40	50	118.6	120.7	1.80
0	0	40	20	70	50	50	40	125.9	122.6	-2.66
0	40	20	0	70	50	40	50	132.4	128.5	-2.97
0	40	0	20	70	50	50	40	132.9	128.5	-3.33

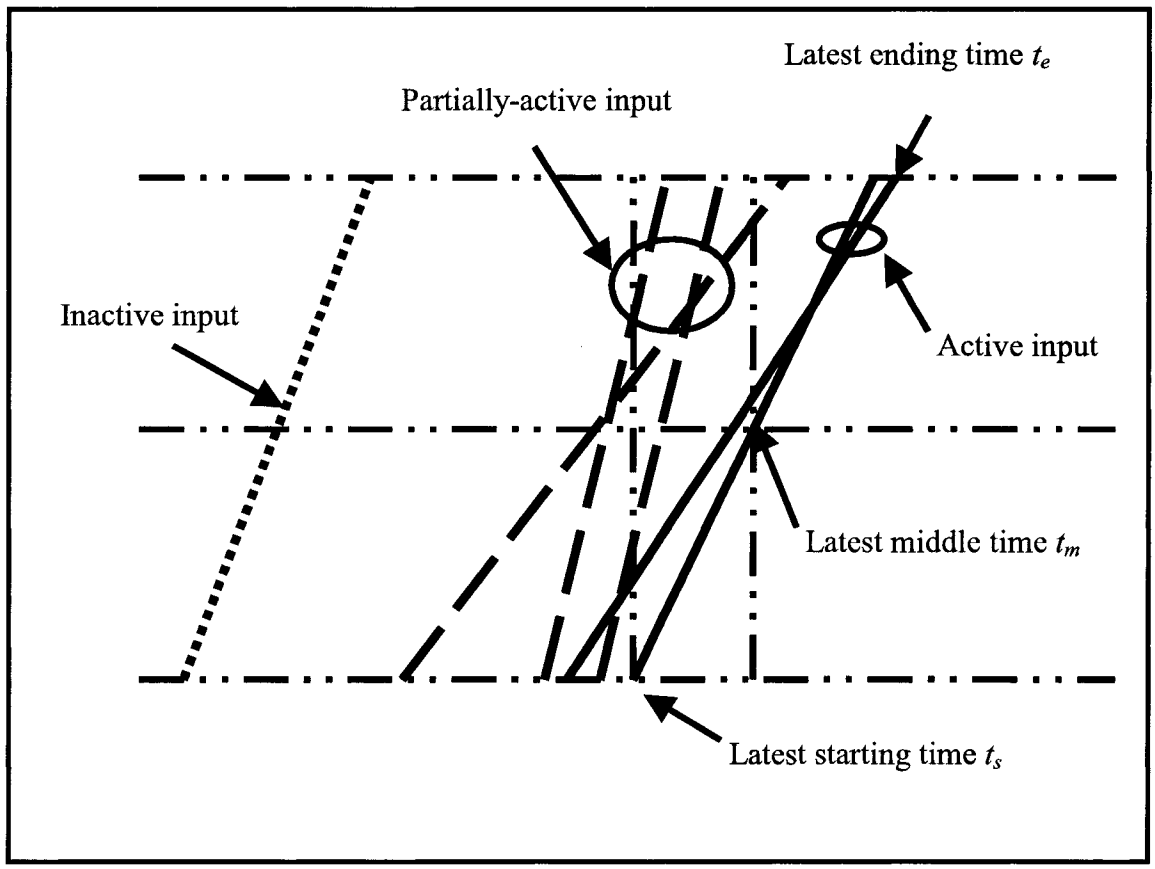


Fig 7.1 Input classification of an arbitrary input pattern in input mapping technique. There are three kinds of inputs – active, partially-active, and inactive.

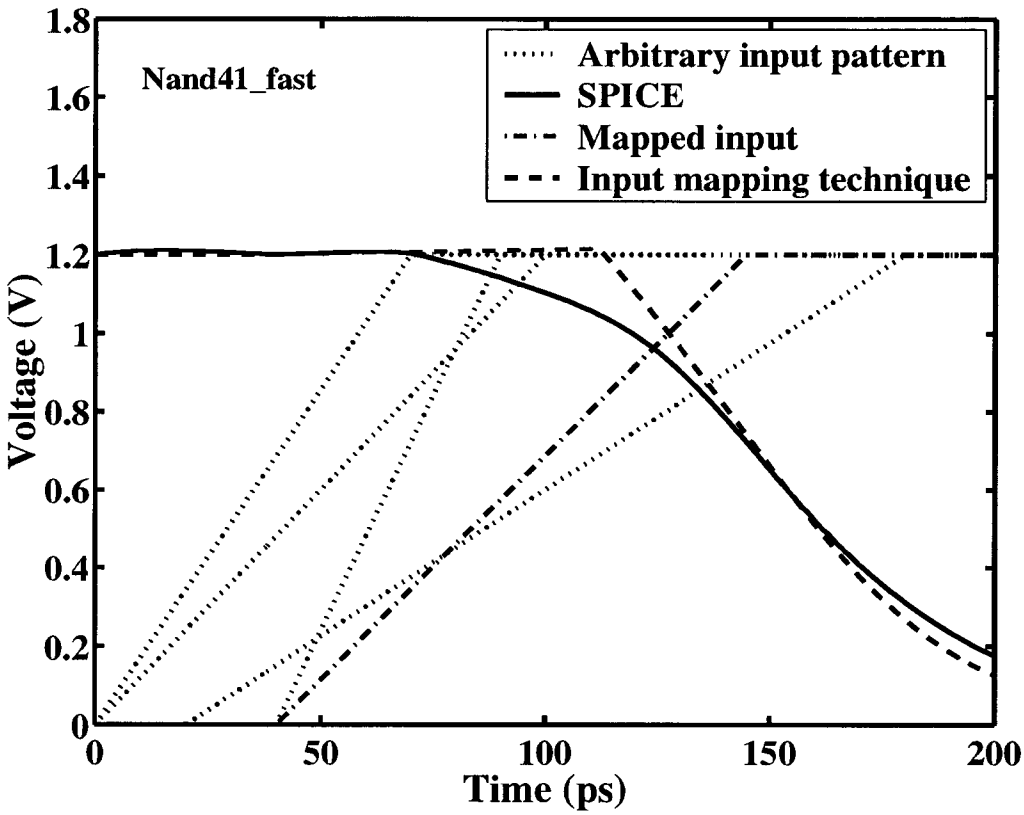


Fig 7.2 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for FIS. The four dotted lines indicate the real input pattern given in case one of Table 7.2. The dashdot line denotes the mapped input.

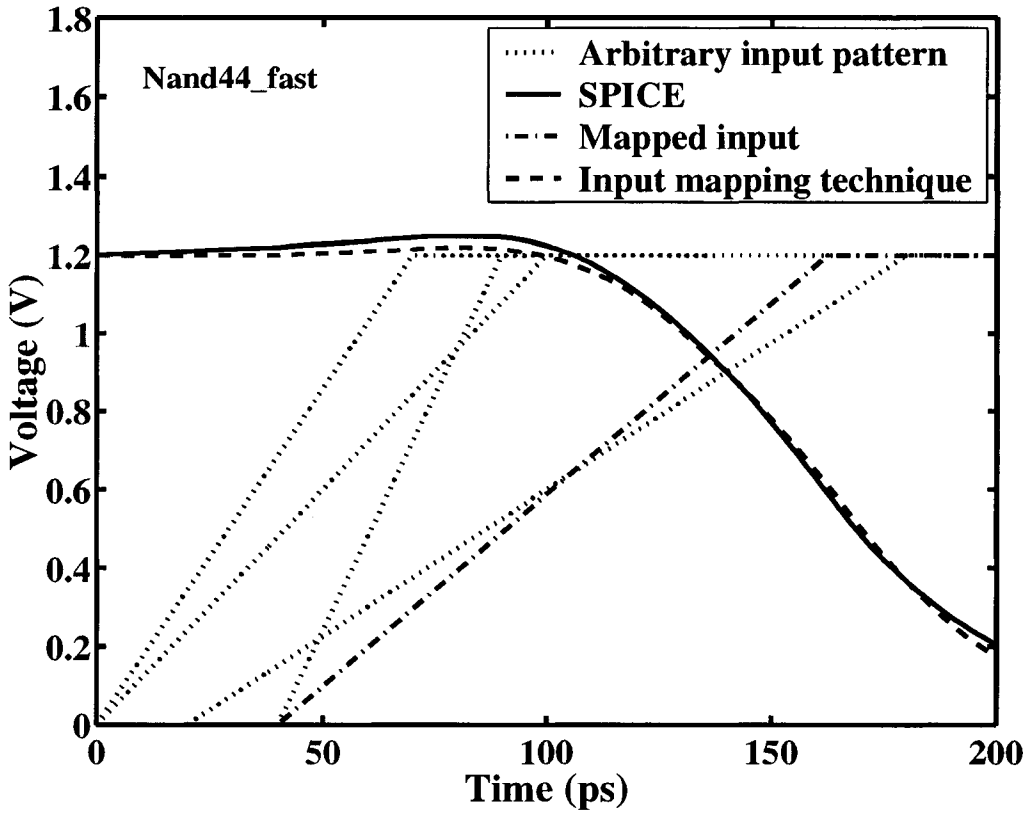


Fig 7.3 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for FIS. The four dotted lines indicate the real input pattern given in case two of Table 7.2. The dashdot line denotes the mapped input.

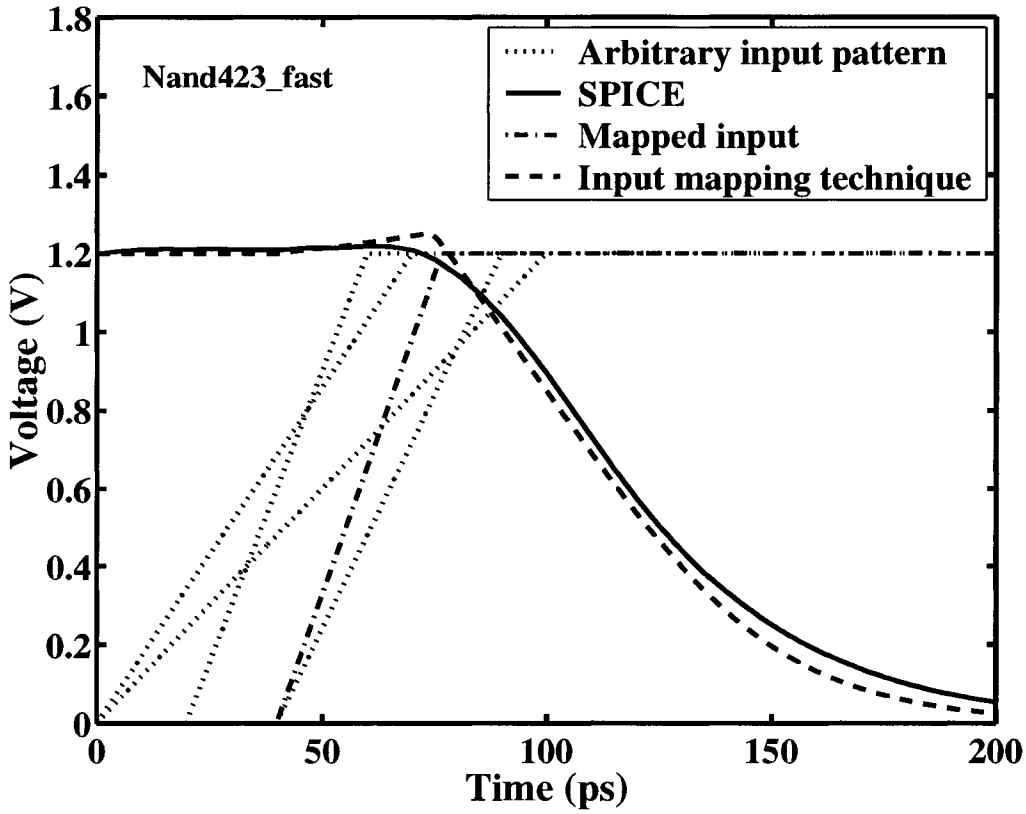


Fig 7.4 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for FIS. The four dotted lines indicate the real input pattern given in case three of Table 7.2. The dashdot line denotes the mapped input.

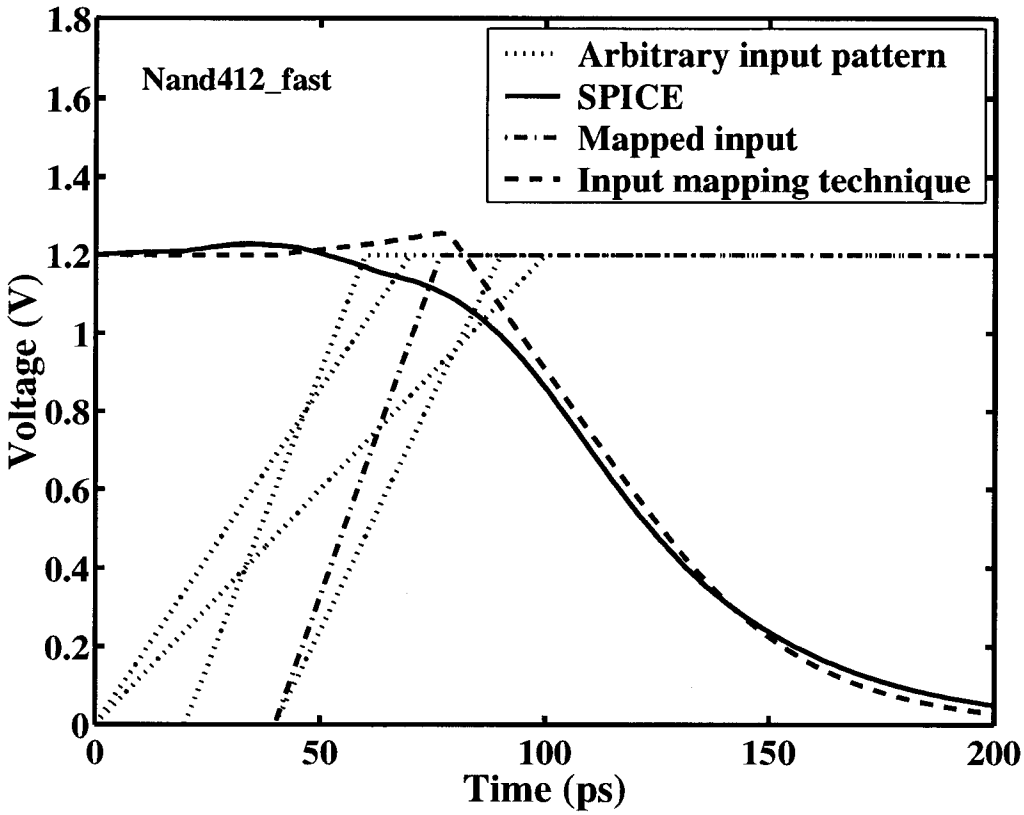


Fig 7.5 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for FIS. The four dotted lines indicate the real input pattern given in case four of Table 7.2. The dashdot line denotes the mapped input.

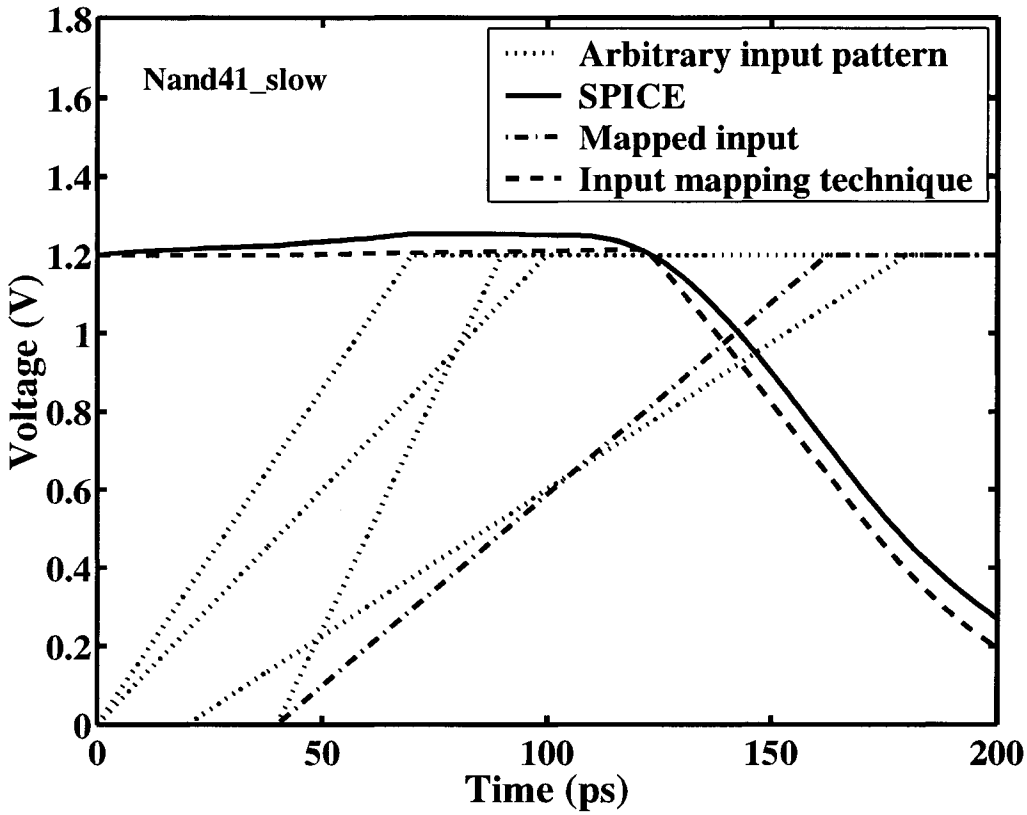


Fig 7.6 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for SIS. The four dotted lines indicate the real input pattern given in case one of Table 7.3. The dashdot line denotes the mapped input.

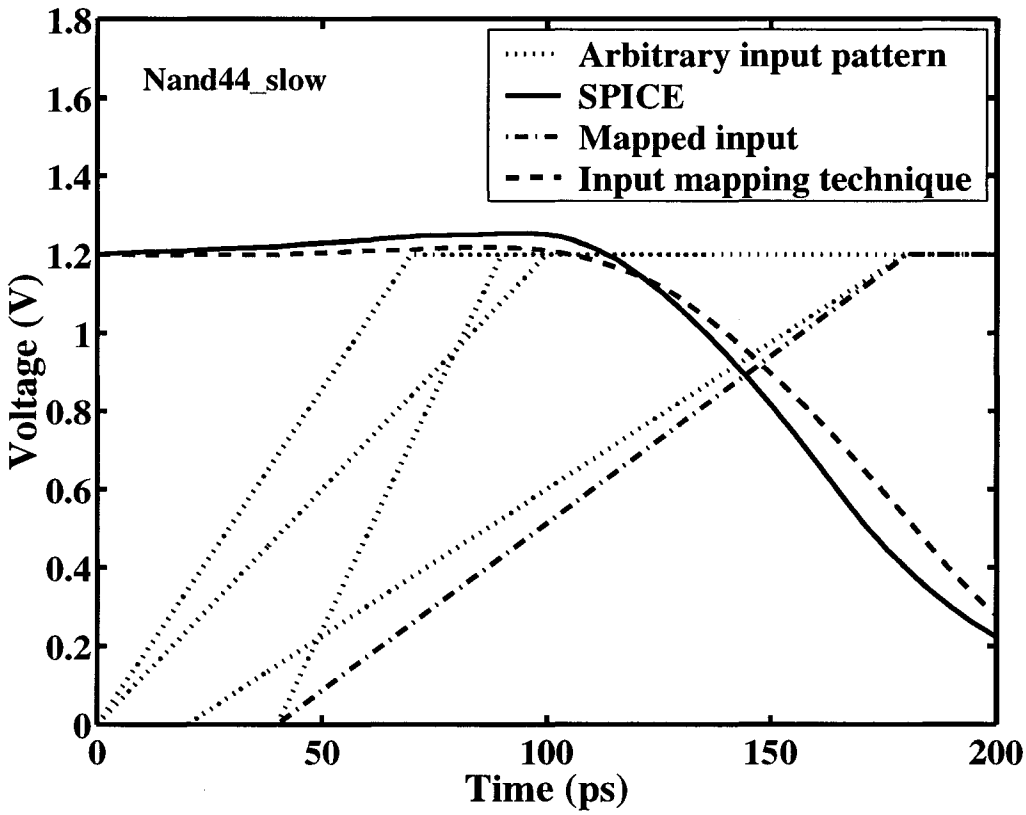


Fig 7.7 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for SIS. The four dotted lines indicate the real input pattern given in case two of Table 7.3. The dashdot line denotes the mapped input.

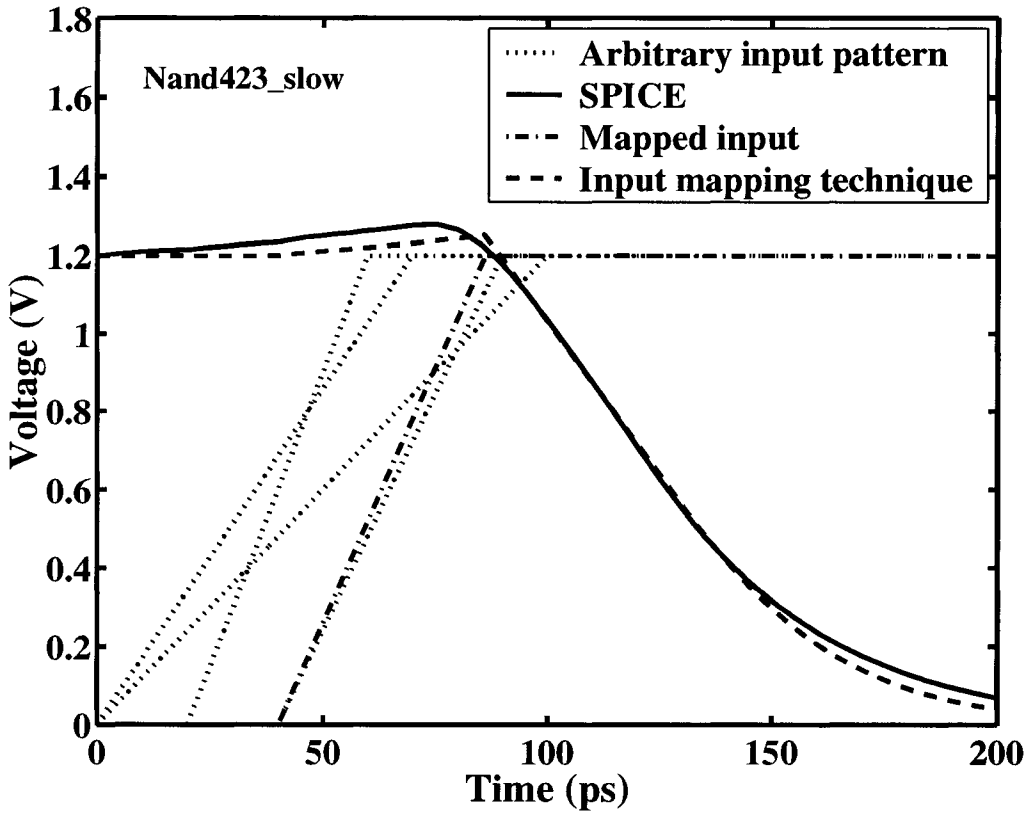


Fig 7.8 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for SIS. The four dotted lines indicate the real input pattern given in case three of Table 7.3. The dashdot line denotes the mapped input.

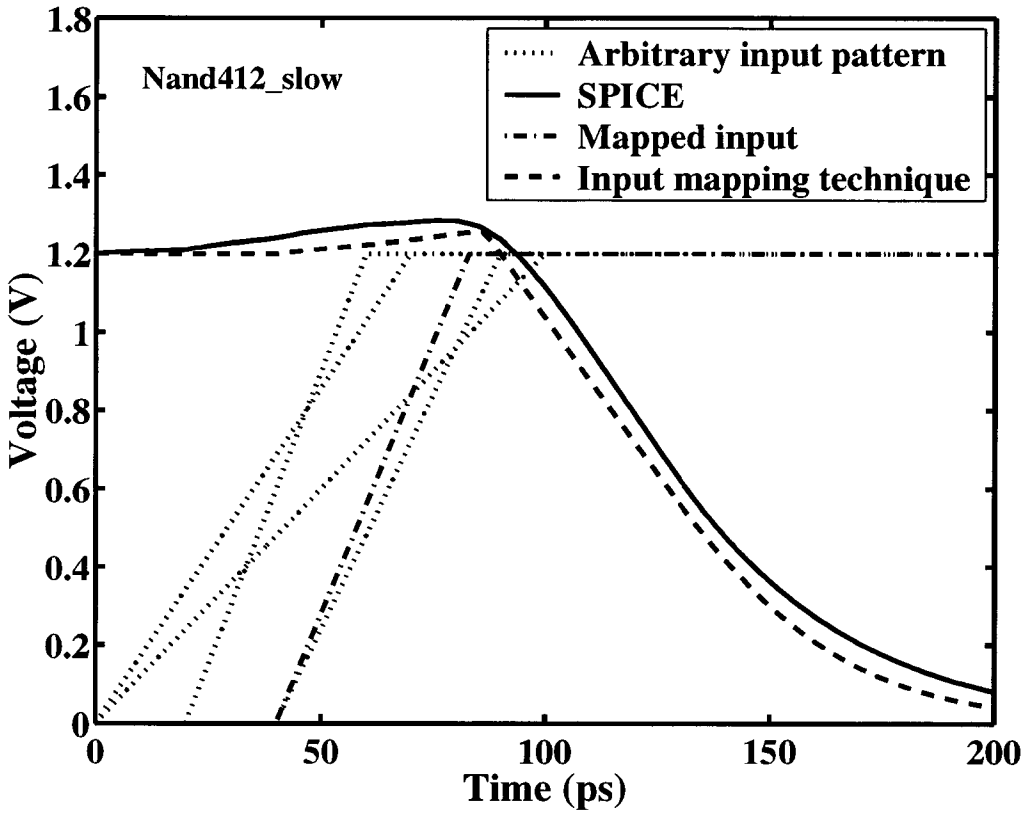


Fig 7.9 The output voltage waveforms of NAND4 from SPICE BSIM3 with solid line and from the input mapping technique with dashed line for SIS. The four dotted lines indicate the real input pattern given in case four of Table 7.3. The dashdot line denotes the mapped input.

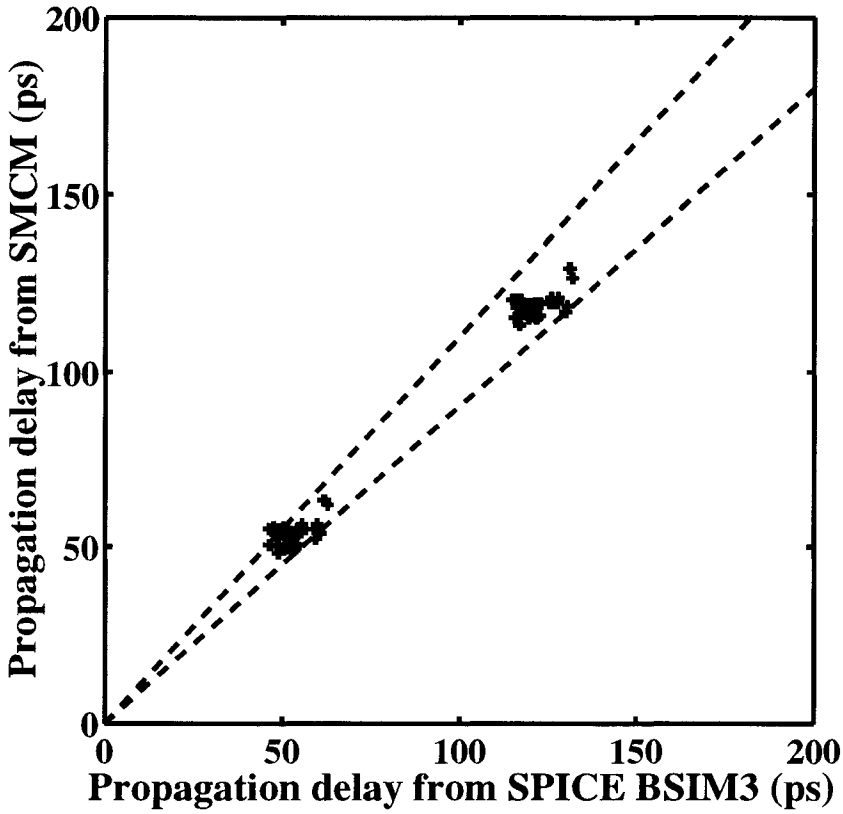


Fig 7.10 The correlation of propagation delays of NAND4 from SPICE BSIM3 and from the input mapping technique for FIS. The real input patterns are given in Tables 7.4-7.7. The external capacitances are 100fF and 400fF, respectively. The dashed lines denote ±10% deviation from SPICE results.

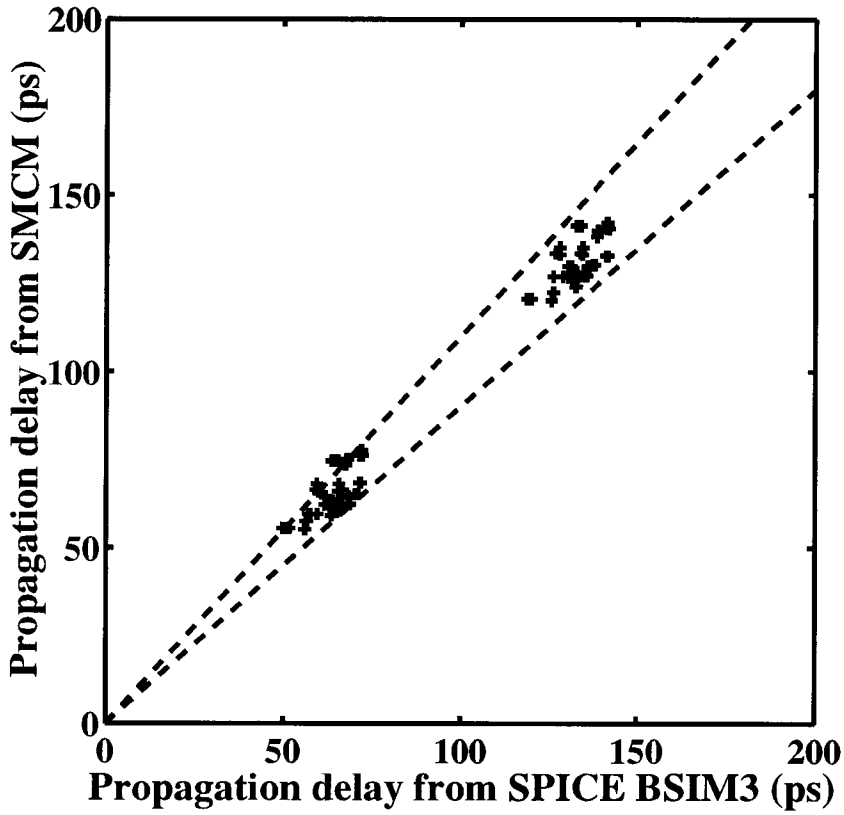


Fig 7.11 The correlation of propagation delays of NAND4 from SPICE BSIM3 and from the input mapping technique for SIS. The real input patterns are given in Tables 7.8-7.11. The external capacitances are 100fF and 400fF, respectively. The dashed lines denote ±10% deviation from SPICE results.

CHAPTER 8 CONCLUSIONS

In this research we have reviewed the methodologies of propagation delays of a CMOS logic gate. The second-order effects of a deep submicron MOSFET such as channel length modulation and charge carrier velocity saturation are introduced. The deep submicron MOSFET model (DSMM) and the proper deep submicron MOSFET model (PDSMM) are proposed to accurately describe the salient features of deep submicron MOSFET I–V characteristics including the important second-order effects. Based on the PDSMM, three analytical approaches of propagation delays are developed from closed-form expressions of output voltage of a deep submicron CMOS inverter. As the first method, the step input is applied. The simple and analytical propagation delay is accurate only for very fast input since the effect of input transition time is ignored. In the second way, the dominating nMOSFET-only circuit is used to model a CMOS inverter. The propagation delay for a ramp input is accurate in a broader range except for large input transition time and small load capacitance. Both the dominating nMOSFET and the trivial pMOSFET are included in the third method. The propagation delays from the third method show excellent agreement with SPICE BSIM3 simulation for a very broad range of input transition time and load capacitance, even for extreme conditions.

The basic structure in a single-stage complex CMOS logic gate such as NAND or NOR gate is a series-connected MOSFET chain. In order to extend the accurate timing approaches of a deep submicron MOSFET inverter to a single-stage complex CMOS logic gate we investigate the characteristic properties of a series-connected MOSFET chain with an equal input pattern. For an equal input pattern, all the inputs are connected

to either an identical input voltage or the corresponding power/ground depending on the type of a series-connected MOSFET chain. A simple and generic series-connected MOSFET chain model (SMCM) is proposed to accurately illustrate the peculiar features of static I–V characteristics of a series-connected MOSFET chain with an equal input pattern.

The intermediate nodes and multiple inputs are unique for a series-connected MOSFET chain, which do not exist in a CMOS inverter. An input suppression technique is modified to describe the effect of initial states of parasitic capacitances of intermediate nodes. From a practical point of view we focus on the two important and extreme cases: the fast initial state (FIS) and slow initial state (SIS). With the SMCM and the modified input suppression technique associated with activation time, output voltages of a single-stage complex CMOS logic gate with an equal input pattern are analytically derived in three analytical approaches. The input in the first approach is a step function. The analytical output voltage and the propagation delay is accurate and suitable only for very fast input. The second way is to model a complex CMOS logic gate with a dominant series-connected MOSFET chain. The trivial parallel MOSFET cluster is ignored. It provides accurate results in broader range than the first approach. The third method is a comprehensive one with taking both a series-connected MOSFET chain and a parallel MOSFET cluster into account. Propagation delays of a single-stage CMOS logic gate with an equal input pattern extracted from output voltages are in excellent agreement with SPICE BSIM3 simulations.

The arbitrary input pattern is approached with an input mapping technique. An efficient and accurate mapping algorithm from an arbitrary input pattern associated with a

single-stage complex CMOS logic gate to an equal input pattern is proposed. Using an input mapping technique, an arbitrary input pattern is mapped to an equal input pattern. Thus the output voltages and propagation delays of a single-stage complex CMOS with an arbitrary input pattern are obtained by exploiting the timing approaches for an equal input pattern.

These analytical approaches provide fast and accurate timing calculation for a deep submicron CMOS logic gate at least two orders of magnitude faster than SPICE, while maintaining reasonable accuracies [31] [70]. The accurate timing methodology can be easily incorporated into deep submicron VLSI digital design, verification, and optimization EDA flows such as timing-driven synthesis, timing closure, floorplan tool, and performance-driven place and route (PAR) [21] [43].

References

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, D. C., 1964.

[2] R. J. Baker, H. W. Li, and D. E. Boyce, *CMOS Circuit Design, Layout, and Simulation*, IEEE Press, 1998.

[3] H. B. Bakoglu, *Circuits, Interconnections, and Packaging for VLSI*, Addison-Wisley, 1990.

[4] A. Bellaouar, S. H. K. Embabi, and M. I. Elmasry, "Scaling of Digital BiCMOS Circuits," *IEEE J. Solid-State Circuits*, vol. 25, pp. 932-941, Aug. 1990.

[5] J. Benkoski and A. J. Strojwas, "A new approach to hierarchical and statistical timing simulations," *IEEE Trans. Computer-Aided Design*, vol. CAD-6, pp. 1039-1052, Nov. 1985.

[6] A. Biran and M. Breiner, *MATLAB for Engineers*, Addison-Wesley, 1995

[7] L. Bisdounism, S. Nikolaidis, and O. Koufopavlou, "Propagation delay and short-circuit power dissipation modeling of the CMOS inverters," *IEEE Trans. Circuits Syst. I*, vol. 45, pp. 259-270, Mar. 1998.

[8] L. Bisdounism, S. Nikolaidis, and O. Koufopavlou, "Analytical transient response and propagation delay evaluation of the CMOS inverter for short-channel devices," *IEEE J. Solid- State Circuits*, vol. 33, pp. 302-306, Feb. 1998.

[9] D. Blaauw, V. Zolotov, and S. Sundareswaran, "Slope Propagation in Static Timing Analysis," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, pp. 1180-1194, Oct. 2002.

[10] P. O. Borjesson and C. E. W. Sundberg, "Simple Approximations of the Error

Function $Q(x)$ for Communications Applications," *IEEE Trans. Communications*, pp. 453-470, Mar. 1979.

[11] K. A. Bowman, B. L. Austin, J. C. Eble, X. Tang, and J. D. Meindl, "A Physical Alpha-Power Law MOSFET Model," *IEEE J. Solid-State Circuits*, vol. 34, pp. 1410-1414, Oct. 1999.

[12] L. M. Brocco, S. P. McCormick, and J. Allen, "Macromodeling CMOS Circuits for Timing Simulation," *IEEE Trans. Computer-Aided Design*, vol. 7, pp. 1237-1249, Dec. 1988.

[13] J. R. Burns, "Switching Response of Complementary-Symmetry MOS Transistor Logic Circuits," *RCA Rev.*, vol. 25, pp. 627-661, Dec. 1964.

[14] B. S. Carlson and S.-J. Lee, "Delay Optimization of Digital CMOS VLSI Circuits by Transistor Reordering," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 1183-1192, Oct. 1995.

[15] A. Chatzigeogiou and S. Nikolaidis, "Collapsing the CMOS transistor chain to an effective single equivalent transistor," *IEE Proc.-Circuits Devices Syst.*, vol. 145, No. 5, pp. 347-353, Oct. 1998.

[16] B. S. Cherkauer and E. G. Friedman, "A Unified Design Methodology for CMOS Tapered Buffers," *IEEE Trans. Very Large Scale Integration (VLSI) Systems*, vol. 3, pp. 99-111, Mar. 1995.

[17] H.-C. Chow and W.-S. Feng, "An Analytical CMOS Inverter Delay Model Including Channel-Length Modulations," *IEEE J. Solid-State Circuits*, vol. 27, pp.1303-1306, Sept. 1992.

[18] H.-C. Chow, W.-S. Feng, and J. B. Kuo, "An improved analytical short-channel

MOSFET model valid in all regions of operation for analog/digital circuit simulation," *IEEE Trans. Computer-Aided Design*, vol. 11, Sept. 1992.

[19] C. Y. Chu and M. A. Horowitz, "Charge-sharing models for switching-level simulation," *IEEE Trans. Computer-Aided Design*, vol. CAD-6, pp. 1053-1061, Nov. 1987.

[20] W. Chuang, S. S. Sapatnekar, and I. N. Hajj, "Timing and Area Optimization for Standard-Cell VLSI Circuit Design," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 308-320, Mar. 1995.

[21] P. Cocchini, G. Piccinini, and M. Zamboni, "A Comprehensive Submicrometer MOST Delay Model and Its Application to CMOS Buffers," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1254-1262, Aug. 1997.

[22] R. Coen and R. S. Muller, "Velocity of Surface Carriers in Inversion Layers on Silicon," *Solid-State Electron.*, vol. 23, p. 35, Jan. 1980.

[23] J. J. Cong and K.-S. Leung, "Optimal Wiresizing Under Elmore Delay Model," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 321-336, Mar. 1995.

[24] W. J. Dally and J. W. Poulton, *Digital Systems Engineering*, Cambridge University Press, 1998.

[25] A.-C. Deng and Y.-C. Shiau, "Generic Linear RC Delay Modeling for Digital CMOS Circuits," *IEEE Trans. Computer-Aided Design*, vol. 9, pp. 367-376, Apr. 1990.

[26] D. Deschacht, C. Dabrin, and D. Auvergne, "Delay Propagation Effect in Transistor Gates," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1184-1189, Aug. 1996.

[27] D. Deschacht, M. Robert, and D. Auvergne, "Explicit Formulation of Delays in CMOS Data Paths," *IEEE J. Solid-State Circuits*, vol. 23, pp. 1257-1264, Oct. 1988.

[28] A. Dharchoudhury and S. M. Kang, "Worst-Case Analysis and Optimization of VLSI Circuit Performances," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 481-492, Apr. 1995.

[29] S. Dutta, S. S. M. Shetti, and S. L. Lusky, "A Comprehensive Delay Model for CMOS Inverters," *IEEE J. Solid-State Circuits*, vol. 30, pp. 864-871, Aug. 1995.

[30] W. C. Elmore, "The Transient Response of Damped Linear Networks with Particular Regard to Wideband Amplifiers," *J. Appl. Phys.*, vol. 19, pp. 55-63, Jan. 1948.

[31] S. H. K. Embabi and R. Damodaran, "Delay Models for CMOS, BiCMOS and BiNMOS Circuits and Their Applications for Timing Simulations," *IEEE Trans. Computer Aided Design*, vol. 13, pp. 1132-1142, Sept. 1994.

[32] Y. Eo and W. R. Eisenstadt, "High-Speed VLSI Interconnection Modeling Based on S-Parameter Measurements," *IEEE Trans. Components, Hybrids, and Manufacture Technology*, vol. 16, pp. 555-562, Aug. 1993.

[33] C.-L. Fang and W.-B. Jone, "Timing Optimization by Gate Resizing and Critical Path Identification," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 201-217, Feb. 1995.

[34] D. Foty, *MOSFET Modeling with SPICE*, Principle and Practice, Prentice Hall, 1997.

[35] R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI Design Techniques for Analog*

and Digital Circuits. New York: McGraw-Hill, 1990.

[36] P. Gray and R. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed., John Wiley and Sons, 1993.

[37] R. X. Gu and M. I. Elmasry, "Power Dissipation Analysis and Optimization of Deep Submicron CMOS Digital Circuits," *IEEE J. Solid-State Circuits*, vol. 31, pp. 707-713, May 1996.

[38] M. Hafed, M. Oulmane, and N. C. Rumin, "Delay and Current Estimation in a CMOS Inverter with an RC Load," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 20, pp. 80-89, Jan. 2001.

[39] M. G. Harbour and J. M. Drake, "Calculation of signal delay in integrated interconnections," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 272-276, Feb. 1989.

[40] N. Hedenstierna and K. O. Jeppson, "CMOS Circuit Speed and Buffer Optimization," *IEEE Trans. Computer-Aided Design*, vol. CAD-6, pp. 270-281, Mar. 1987.

[41] A. Hirata, H. Onodera, and K. Tamaru, "Estimation of Propagation Delay Considering Short-Circuit Current for Static CMOS Gates," *IEEE Trans. Circuits Syst. I*, vol. 45, pp1194-1198, Nov. 1998.

[42] D. A. Hodges and H. G. Jackson, *Analysis and Design of Digital Integrated Circuits*, 2nd ed., McGraw-Hill, 1983.

[43] B. Hoppe, G. Neuendorf, D. Schmitt-Landsiedel, and W. Specks, "Optimization of High-Speed CMOS Logic Circuits with Analytical Models for Signal Delay, Chip Area, and Dynamic Power Dissipation," *IEEE Trans. Computer-Aided Design*, vol. 9, pp. 236-247, Mar. 1990.

[44] H. Iwai, "CMOS Technology—Year 2010 and Beyond," *IEEE J. Solid-State Circuits*, vol. 34, pp. 357-366, Mar. 1999.

[45] K. O. Jeppson, "Modeling the Influence of the Transistor Gain Ratio and the Input-to-Output Coupling Capacitance on the CMOS Inverter Delay," *IEEE J. Solid-State Circuits*, vol. 29, pp. 646-654, June 1994.

[46] M. W. Jetton and A. P. Jayasumana, "Method of Determining Delay in Logic Cell Models," *U.S. Patent 6,028,995*, Feb. 22, 2000.

[47] X. Jiang, A. P. Jayasumana, W. Zhang, and S. Chiao, "A Proper Deep Submicron MOSFET Model (PDSMM) and Its Applications for Delay Modeling of CMOS Inverters," in *Proc. Sixth International Conference on Solid-State and Integrated-Circuit Technology (ICSICT2001)*, Shanghai, China, Vol. 2, 875-878, October, 2001.

[48] X. Jiang and A. P. Jayasumana, "Input Collapse of CMOS Logic Gates with a Series-Connected MOSFET Chain," in *Proc. Seventh International Conference on Solid-State and Integrated-Circuit Technology (ICSICT2004)*, Beijing, China, October, 2004.

[49] X. Jiang and A. P. Jayasumana, "Effect of Initial States of Intermediate Nodes on Propagation Delay of CMOS NAND/NOR Gates," submitted to *International Symposium on Circuits and Systems (ISCAS2005).*

[50] Y.-H. Jun, K. Jun, and S.-B. Park, "An Accurate and Efficient Delay Time Modeling for MOS Logic Circuits Using Polynomial Approximation," *IEEE Trans. Computer-Aided Design*, vol. 8, pp. 1027-1032, Sept. 1989.

[51] S.-M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits: Analysis and*

Design, WCB/McGraw-Hill, 1999.

[52] A. I. Kayssi, K. A. Sakallah, and T. M. Burks, "Analytical Transient Response of CMOS Inverters," *IEEE Trans. Circuits and Syst.-I*, vol. 39, pp. 42-45, Jan. 1992.

[53] P. K. Ko, C. Hu, et al., *BSIM3v3 Manual*, Dept. of EECS, Univ. of California, Berkeley, 1995.

[54] J.-T. Kong and D. Overhauser, "Methods to Improve Digital MOS Macromodel Accuracy," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 868-881, July 1995.

[55] J.-T. Kong and D. Overhauser, *Digital Timing Macromodeling for VLSI Design Verification*, Kluwer Academic Publishers, 1995.

[56] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, 2nd ed., Addison-Wesley Publishing Company, 1994.

[57] J. Lillis, C.-K. Cheng, and T.-T. Y. Lin, "Optimal Wire Sizing and Buffer Insertion for Low Power and a Generalized Delay Model," *IEEE J. Solid-State Circuits*, vol. 31, pp. 437-447, Mar. 1996.

[58] T.-M. Lin and C. A. Mead, "Signal Delay in General RC Networks," *IEEE Trans. Computer Aided Design*, vol. CAD-3, pp. 331-349, Oct. 1984.

[59] J. J. Liou, *Advanced Semiconductor Device Physics and Modeling*, Artech House, 1994.

[60] W. Liu and M.-C. Chang, "Transistor Transient Studies Including Transcapacitive Current and Distributive Gate Resistance for Inverter Circuits," *IEEE Trans. Circuits Syst. I*, vol. 45, pp 416- 422, Apr. 1998.

[61] C. A. Marinov and P. N. Shivakumar, "Delay time bounds for RC ladder networks," *IEEE Trans. Circuits Syst. I*, vol. 43, pp. 613-614, July 1996.

[62] D. Martin and N. C. Rumin, "Delay Prediction from Resistance-Capacitance Models of General MOS Circuits," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 12, pp. 997-1003, July 1993.

[63] G. Massobrio and P. Antognetti, *Semiconductor Device Modeling with SPICE.* New York: McGraw-Hill, 1993.

[64] The Math Works Inc., *The Student Edition of MATLAB*, Version 4 User's Guide, Prentice Hall, 1995.

[65] C. A. Mead and L. A. Conway, *Introduction to VLSI Systems*, Addison-Wesley, 1980.

[66] Meta-Software Staff, *HSPICE User's Manual*, Meta-Software Inc., 1996.

[67] M. Miura-Mattausch, "Analytical MOSFET Model for Quarter Micron Technologies," *IEEE Trans. Computer-Aided Design*, vol. 13, pp. 610-615, May 1994.

[68] G. E. Moore, "Progress in digital integrated circuit," in *IEDM Tech. Dig.*, Dec. 1975, p. 11.

[69] R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, 2nd ed., John Wiley & Sons, 1986.

[70] A. Nabavi-Lishi and N. C. Rumin, "Inverter Models of CMOS Gates for Supply Current and Delay Evaluation," *IEEE Trans. Computer-Aided Design*, vol. 13, pp. 1271-1279, Oct. 1994.

[71] A. Nabavi-Lishi and N. C. Rumin, "Simultaneous delay and maximum current calculation in CMOS gates," *IEE Electron. Lett.*, vol. 28, pp. 682-684, Mar. 1992.

[72] L. W. Nagel, "SPICE2: A Computer Program to Simulation Semiconductor

Circuits," *Rep. No. ERL-M520, Electronic Research Laboratory, Univ. of California, Berkeley*, May 1975.

[73] L. W. Nagel and R. Rohrer, "Computer Analytical of Nonlinear Circuits, Excluding Radiation (CANCER)," *IEEE J. Solid-State Circuits*, vol. SC-6, pp. 166-182, 1971.

[74] N. Nassif, M. P. Desai, and D. H. Hall, "Robust Elmore Delay Models Suitable for Full Chip Timing Verification of a 600MHz CMOS Microprocessor," *Proc. of the 35th Design Automation Conference*, pp. 230-235, June 1998.

[75] S. Nikolaidis and A. Chatzigeogiou, "Modeling the Transistor Chain Operation in CMOS Gates for Short Channel Devices," *IEEE Trans. Circuits Syst. I*, vol. 46, pp. 1191-1202, Oct. 1999.

[76] J. K. Ousterhout, "A Switch-Level Timing Verifier for Digital MOS VLSI," *IEEE Trans. Computer-Aided Design*, vol. CAD-4, pp. 336-349, July 1985.

[77] H.-J. Park and M. Soma, "Analytical Model for Switching Transitions of Submicron CMOS Logics," *IEEE J. Solid-State Circuits*, vol. 32, pp. 880-889, June 1997.

[78] J. M. Rabaey, *Digital Integrated Circuits: A Design Perspective*. Prentice Hall, 1996.

[79] R. Rajaraman and D. F. Wong, "Optimum Clustering for Delay Minimization," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 1490-1495, Dec. 1995.

[80] P. J. Restle, K. A. Jenkins, A. Deutsch, and P. W. Cook, "Measurement and Modeling of On-Chip Transmission Line Effects in a 400 MHz Microprocessor," *IEEE J. Solid-State Circuits*, vol. 33, pp. 662-665, Apr. 1998.

[81] J. L. Rossello and J. Segura, "Simple and accurate propagation delay model for submicron CMOS gates based on charge analysis," *Electronics Letters*, vol. 38, pp 772-774, July 2002.

[82] J. Rubinstein, P. Penfield, Jr., and M. A. Horowitz, "Signal Delay in RC Tree Networks," *IEEE Trans. Computer-Aided Design*, vol. CAD-2, pp. 202-211, July 1983.

[83] C. T. Sah, "Characteristics of the Metal-Oxide-Semiconductor Transistor," *IEEE. Trans. Electron. Devices*, vol. ED-11, pp. 324-345, July 1964.

[84] T. Sakurai and A. R. Newton, "Alpha-Power Law MOSFET Model and its Applications to CMOS Inverter Delay and Other Formulas," *IEEE J. Solid-State Circuits*, vol. 25, pp. 584-594, Apr. 1990.

[85] T. Sakurai and A. R. Newton, "Delay Analysis of Series-Connected MOSFET Circuits," *IEEE J. Solid-State Circuits*, vol. 26, pp. 122-131, Feb. 1991.

[86] T. Sakurai and A. R. Newton, "A Simple MOSFET Model for Circuit Analysis," *IEEE Trans. Electron Devices*, vol. 38, pp. 887-894, Apr. 1991.

[87] E. Seewann, "Switching Speeds of MOS Inverters," *IEEE J. Solid-State Circuits*, vol. SC- 15, pp. 246-252, Apr. 1980.

[88] H. Shichman and D. A. Hodges, "Modeling and Simulation of Insulated-Gate Field-Effect Transistor Switching Circuits," *IEEE J. Solid-State Circuits*, pp. 285-289, Sept. 1968.

[89] Y.-H Shih and S.-M Kang, "Analytical Transient Solution of General MOS Circuit Primitives," *IEEE Trans. Computer-Aided Design*, vol. 11, pp. 719-731, June 1992.

[90] W. Shockley, "A Unipolar Field Effect Transistor," *Proc. IRE*, vol. 40, pp. 1365-1376, Nov. 1952.

[91] M. Shoji, "FET Scaling in Domino CMOS Gates," *IEEE J. Solid-State Circuits*, vol. SC-20, pp. 1067-1071, Oct., 1985.

[92] C. G. Sodini, P. Ko, and J. L. Moll, "The Effect of High Fields on MOS Device and Circuit Performance," *IEEE Trans. Electron Devices*, vol. ED-31, pp. 1386-1393, Oct. 1984.

[93] B. G. Streetman, *Solid State Electronic Devices*, 3rd Ed., Prentice Hall, 1990.

[94] I. Sutherland, B. Sproull, and D. Harris, *Logical Effect: Designing Fast CMOS Circuits,* Morgan Kaufmann Publishers, Inc, 1999.

[95] Synopsys Staff, *TimeMill/PowerMill/ACE Workship*, Synopsys Inc. 1998

[96] S. M. Sze, *VLSI Technology*. New York: McGraw-Hill, Inc, 1983.

[97] K. Takeuchi and M. Fukuma, "Effects of the Velocity Saturated Region on MOSFET Characteristics," *IEEE Trans. Electron Devices*, vol. 41, pp. 1623-1627, Sept. 1994.

[98] T. Toh, P. K. Ko, and R. G. Meyer, "An Engineering Model for Short-Channel MOS Devices," *IEEE J. Solid-State Circuits*, vol. 23, pp. 950-957, Aug. 1988.

[99] T. P. Tsividis, *Operation and Modeling of the MOS Transistor*, McGraw-Hill, 1987.

[100] P. W. Tuinenga, *SPICE: A Guide to Circuit Simulation and Analysis Using PSPICE*, 2nd Ed., Englewood Cliffs: Prentice Hall, 1992.

[101] H. J. M. Veendrick, "Short-Circuit Dissipation of Static CMOS Circuitry and Its Impact on the Design of Buffer Circuits," *IEEE J. Solid-State Circuits*, vol. SC-

19, pp. 468-473, Aug. 1984.

[102] S. R. Vemuru and A. R. Thorbjornsen, "A Model for Delay Evaluation of a CMOS Inverters," in *Proc. IEEE Int. Symp. on Circuits and Systems*, vol. 1, pp. 89-92, May 1990.

[103] S. R. Vemuru and N. Scheinberg, "Short-Circuit Power Dissipation Estimation for CMOS Logic Gates," *IEEE Trans. Circuits Syst. I*, vol. 45, pp. 762-765, Nov. 1994.

[104] N. H. E. Weste and K. Eshraghian, *Principle of CMOS VLSI Design: A System Perspective*, 2nd ed., Addison-Wesley, 1993.

[105] J. L. Wyatt, Jr., "Signal Delay in RC Mesh Networks," *IEEE Trans. Circuits Syst.*, vol. CAS-32, pp. 507-510, May 1985.

[106] C. Xiaohong, H. Chunhai, Q. Jixin, and W. Qidi, "Comments on Delay Time Bounds for RC Ladder Networks," *IEEE Trans. Circuits Syst. I*, vol. 45, pp. 114-115, Jan. 1998.

[107] C. A. Zukowski, "Relaxation Bounds for Linear RC Mesh Networks," *IEEE Trans. Computer-Aided Design*, vol. CAD-5, pp. 395-312, Apr. 1985.

[108] J. M. Zurada and T. Liu, "Equivalent dominant pole approximation of capacitively loaded VLSI interconnection," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 205-207, Feb. 1987.

APPENDIX I DERIVATION OF NEW DSMM MODEL

I.1 DRAIN-SOURCE SATURATION VOLTAGE V_{sat} IN DSMM

Theoretically, the drain-source current I_{ds} in the linear mode maximizes when drain-source saturation voltage V_{sat} is reached. In Section 3.3, we have a generic expression of current I_{ds} in (3.6) for linear mode as

$$I_{ds} = \frac{k}{1+\frac{V_{ds}}{V_c}} \cdot (V_{gs} - V_{th} - \frac{V_{ds}}{2}) \cdot V_{ds} \tag{I.1}$$

The differentiation of I_{ds} with respect to the drain-source voltage V_{ds} is

$$\frac{dI_{ds}}{dV_{ds}} = \frac{k}{\left(1+\frac{V_{ds}}{V_c}\right)^2} \cdot \left\{ (V_{gs} - V_{th} - V_{ds}) \cdot \left(1+\frac{V_{ds}}{V_c}\right) - (V_{gs} - V_{th} - \frac{V_{ds}}{2}) \cdot \frac{V_{ds}}{V_c} \right\} \tag{I.2}$$

The drain-source current I_{ds} in the linear mode maximizes when the differentiation of I_{ds} with respect to V_{ds} is equal to zero and the variable V_{ds} is replaced with drain-source saturation voltage V_{sat}, i.e.

$$(V_{gs} - V_{th} - V_{sat}) \cdot (V_{sat} + V_c) - (V_{gs} - V_{th} - \frac{V_{sat}}{2}) \cdot V_{sat} = 0 \tag{I.3}$$

Simplifying the above equation yields

$$V_{sat}^2 + 2V_c V_{sat} - 2(V_{gs} - V_{th})V_c = 0 \tag{I.4}$$

The drain-source saturation voltage V_{sat} is obtained from the solution of the above equation

$$V_{sat} = V_c \cdot \left(\sqrt{1 + 2 \cdot \frac{V_{gs} - V_{th}}{V_c}} - 1 \right) \tag{I.5}$$

I.2 DRAIN-SOURCE SATURATION CURRENT I_{sat} IN DSMM

In order to obtain drain-source saturation current I_{sat}, we substitute drain-source voltage V_{ds} of (3.6) with drain-source saturation voltage V_{sat} given by (I.5), which is

$$I_{sat} = \left(\frac{k}{1 + \frac{V_{ds}}{V_c}} \cdot (V_{gs} - V_{th} - \frac{V_{ds}}{2}) \cdot V_{ds} \right) \Bigg|_{V_{ds} = V_{sat}} \tag{I.6}$$

Putting drain-source saturation voltage V_{sat} yields

$$I_{sat} = \frac{k}{1 + \frac{V_{sat}}{V_c}} \cdot (V_{gs} - V_{th} - \frac{V_{sat}}{2}) \cdot V_{sat} \tag{I.7}$$

Rearranging the expression of drain-source saturation voltage V_{sat} in (I.5) gives rise to

$$2\frac{V_{gs} - V_{th}}{V_c} = \left(1 + \frac{V_{sat}}{V_c} \right)^2 - 1 \tag{I.8}$$

Substituting the above equation of (I.8) into (I.7) yields

$$I_{sat} = \frac{k}{2\left(1 + \frac{V_{sat}}{V_c} \right)} \cdot \left[V_c \left(1 + \frac{V_{sat}}{V_c} \right)^2 - V_c - V_{sat} \right] \cdot V_{sat} \tag{I.9}$$

Simplifying the above equation gives

$$I_{sat} = \frac{k}{2\left(1 + \frac{V_{sat}}{V_c} \right)} \cdot \left(V_{sat} + \frac{V_{sat}^2}{V_c} \right) \cdot V_{sat} \tag{I.10}$$

The term $\left(1+\frac{V_{sat}}{V_c}\right)$ is cancelled from the above equation. Then the drain-source saturation current I_{sat} becomes

$$I_{sat}=\frac{k}{2}\cdot V_{sat}^2 \tag{I.11}$$

APPENDIX II FIRST-ORDER LINEAR DIFFERENTIAL EQUATION, BERNOULLI EQUATION, AND RACCATI EQUATION

In this appendix, we review the solutions of three linear differential equations with initial values. We first introduce the solution of first-order linear differential equation, then extend it to Bernolli equation. In the end of the appendix, we briefly discuss Raccati equation.

II.1 FIRST-ORDER LINEAR DIFFERENTIAL EQUATION

The first-order linear differential equation is

$$\frac{dy(x)}{dx} = p(x)y + q(x) \tag{II.1}$$

As a first step, we look into the corresponding homogenous equation,

$$\frac{dy(x)}{dx} = p(x)y \tag{II.2}$$

The general solution of the above homogenous equation is

$$y(x) = C \cdot e^{\int p(x)dx} \tag{II.3}$$

Then the general solution is extended to the solution of first-order linear differential equation of (II.1) by the following expression

$$y(x) = C(x) \cdot e^{\int p(x)dx} \tag{II.4}$$

where $C(x)$ is obtained by substituting the above expression into (II.1). It gives rise to

$$C(x) = \int q(x) e^{-\int p(x)dx} dx + B \tag{II.5}$$

where B is a constant. Therefore the solution of first-order linear differential equation is

$$y = e^{\int p(x)dx}\left\{\int q(x)e^{-\int p(x)dx}dx + B\right\} \tag{II.6}$$

II.2 BERNOULLI EQUATION

The format of Bernoulli equation is

$$\frac{dy(x)}{dx} = p(x)y + q(x)y^n \tag{II.7}$$

where $q(x)$ is not zero and n is not equal to zero or one. For $q(x)=0$ or $n=0$ or $n=1$, the differential equation of (II.7) has become the first-order linear differential equation whose solution is obtained in Section II.1. By using the following transformation

$$z = y^{1-n} \tag{II.8}$$

Bernoulli equation is then changed into

$$\frac{dz(x)}{dx} = (1-n)p(x)z + (1-n)q(x) \tag{II.9}$$

which is a first-order linear differential equation. Its solution is well known and expressed analytically in Section II.1. Therefore the solution of Bernoulli equation is

$$y^{1-n} = e^{\int (1-n)p(x)dx}\left\{\int (1-n)q(x)e^{-\int (1-n)p(x)dx}dx + B\right\} \tag{II.10}$$

II.3 RACCATI EQUATION

The format of Raccati equation is

$$\frac{dy(x)}{dx} = p(x)y^2 + q(x)y + r(x) \tag{II.11}$$

where $p(x)$ and $r(x)$ is non-zero. For $p(x)=0$, the above equation is reduced to the first-order linear differential equation that can be solved analytically as in Section II.1. For $r(x)=0$, it belongs to Bernoulli equation whose solution is shown in Section II.2.

In general, Raccati equation cannot be solved analytically. However, if one particular solution $y_p(x)$ is known, then Raccati equation can be solved analytically as follows. Assume $y(x)$ is expressed as the addition of the particular solution $y_p(x)$ and the other unknown function $u(x)$

$$y(x) = y_p(x) + \frac{1}{u(x)} \tag{II.12}$$

Substituting the solution $y(x)$ into Raccati equation of (II.11), the left hand side of (II.11) becomes

$$LHS = \frac{dy_p(x)}{dx} - \frac{1}{u(x)^2}\frac{du(x)}{dx} \tag{II.13}$$

and the right hand side is

$$RHS = p(x)y_p(x)^2 + q(x)y_p(x) + r(x) + \frac{1}{u(x)}\{2p(x)y_p(x) + q(x)\} + \frac{p(x)}{u(x)^2} \tag{II.14}$$

Equating the above two equations of (II.13) and (II.14) yields

$$-\frac{du(x)}{dx} = \{2p(x)y_p(x) + q(x)\}u(x) + p(x) \tag{II.15}$$

which is a first-order linear differential equation for $u(x)$ and can be solved analytically as in Section II.1.

APPENDIX III CUMULATIVE DISTRIBUTION FUNCTION (CDF), *Q-FUNCTION*, AND USEFUL EXPRESSIONS

III.1 CUMULATIVE DISTRIBUTION FUNCTION (CDF) AND *Q-FUNCTION*

The cumulative distribution function (cdf) of a normalized Gaussian random variable is the integral of

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} dx \tag{III.1}$$

It does not have a closed-form expression. In electrical engineering, it is customary to work with the *Q–function* or error function integral [56], which is defined by

$$Q(t) = 1 - \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{t}^{\infty} e^{-x^2/2} dx \tag{III.2}$$

The *Q–function* is simply the probability of the "tail" of the probability density function (pdf) of a Gaussian random variable. From (III.2), we have

$$\Phi(t) = 1 - Q(t) \tag{III.3}$$

The symmetry of the pdf with respect to t implies that

$$Q(-t) = 1 - Q(t) \tag{III.4}$$

Setting $t = 0$ in the above equation yields

$$Q(0) = \frac{1}{2} \tag{III.5}$$

Then, comparing (III.3) and (III.4) gives rise to

$$\Phi(t) = Q(-t) \tag{III.6}$$

Traditionally the integral in *Q–function* has been evaluated by looking up tables that list

$Q(t)$ or by using approximations that require numerical evaluation [1]. Borjesson and Sundberg have found the following expression to give good accuracy of $Q(t)$ over the entire range $0 \le t < \infty$ [10]:

$$Q(t) = \left[\frac{1}{(1-1/\pi)t + 1/\pi\sqrt{t^2+2\pi}}\right] \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-t^2/2} \qquad 0 \le t < \infty \quad \text{(III.7)}$$

Fig III.1 plots $Q(t)$ from the above equation and from numerical approach. It shows that (III.7) is very accurate to describe $Q(t)$. The integral of $\Phi(t)$, cumulative distribution function, over the range of $-\infty < t \le 0$ can be obtained by substituting the above expression into (III.6)

$$\Phi(t) = Q(-t) = \left[\frac{1}{(1-1/\pi)(-t) + 1/\pi\sqrt{t^2+2\pi}}\right] \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-t^2/2} \qquad -\infty < t \le 0 \quad \text{(III.8)}$$

III.2 USEFUL EXPRESSIONS

In order to derive useful expressions from the above expressions of (III.7) and (III.8), the following functions $P(t,m,\sigma)$ and $R(t,m,\sigma)$ for simplicity are set:

$$P(t,m,\sigma) = \frac{\sigma^2}{(1-1/\pi)(t-m) + 1/\pi\sqrt{(t-m)^2 + 2\pi\sigma^2}} \qquad \text{(III.9a)}$$

$$R(t,m,\sigma) = P(2m-t,m,\sigma) = \frac{\sigma^2}{(1-1/\pi)(m-t) + 1/\pi\sqrt{(m-t)^2 + 2\pi\sigma^2}} \qquad \text{(III.9b)}$$

Using the analysis in the previous section and the functions $P(t,m,\sigma)$ and $R(t,m,\sigma)$, the following useful derived expressions are obtained:

$$P(t-m,0,\sigma) = P(t,m,\sigma) \qquad m \le t < +\infty \quad \text{(III.10a)}$$

$$R(t-m,0,\sigma) = R(t,m,\sigma) \qquad -\infty < t \le m \quad \text{(III.10b)}$$

$$P\left(\frac{t-m}{\sigma},0,1\right)=\frac{1}{\sigma}P(t,m,\sigma) \qquad m\le t<+\infty \quad \text{(III.11a)}$$

$$R\left(\frac{t-m}{\sigma},0,1\right)=\frac{1}{\sigma}R(t,m,\sigma) \qquad -\infty<t\le m \quad \text{(III.11b)}$$

$$\int_{t}^{\infty}e^{-x^2/2}dx=P(t,0,1)\cdot e^{-t^2/2} \qquad 0\le t \quad \text{(III.12a)}$$

$$\int_{-\infty}^{t}e^{-x^2/2}dx=R(t,0,1)\cdot e^{-t^2/2} \qquad t\le 0 \quad \text{(III.12b)}$$

$$\int_{t}^{\infty}e^{-\frac{(x-m)^2}{2\sigma^2}}dx=P(t,m,\sigma)\cdot e^{-\frac{(t-m)^2}{2\sigma^2}} \qquad m\le t \quad \text{(III.13a)}$$

$$\int_{-\infty}^{t}e^{-\frac{(x-m)^2}{2\sigma^2}}dx=R(t,m,\sigma)\cdot e^{-\frac{(t-m)^2}{2\sigma^2}} \qquad t\le m \quad \text{(III.13b)}$$

$$\int_{t_1}^{t_2}e^{-x^2/2}dx=P(t_1,0,1)\cdot e^{-t_1^2/2}-P(t_2,0,1)\cdot e^{-t_2^2/2} \qquad 0\le t_1<t_2 \quad \text{(III.14a)}$$

$$\int_{t_2}^{t_1}e^{-x^2/2}dx=R(t_1,0,1)\cdot e^{-t_1^2/2}-R(t_2,0,1)\cdot e^{-t_2^2/2} \qquad t_2\le t_1<0 \quad \text{(III.14b)}$$

$$\int_{t_1}^{t_2}e^{-\frac{(x-m)^2}{2\sigma^2}}dx=P(t_1,m,\sigma)\cdot e^{-\frac{(t_1-m)^2}{2\sigma^2}}-P(t_2,m,\sigma)\cdot e^{-\frac{(t_2-m)^2}{2\sigma^2}} \qquad m\le t_1<t_2 \quad \text{(III.15a)}$$

$$\int_{t_2}^{t_1}e^{-\frac{(x-m)^2}{2\sigma^2}}dx=R(t_1,m,\sigma)\cdot e^{-\frac{(t_1-m)^2}{2\sigma^2}}-R(t_2,m,\sigma)\cdot e^{-\frac{(t_2-m)^2}{2\sigma^2}} \qquad t_2<t_1\le m \quad \text{(III.15b)}$$

$$\int_{t_1}^{t_2}xe^{-\frac{(x-m)^2}{2\sigma^2}}dx=\left[\sigma^2+mP(t_1,m,\sigma)\right]\cdot e^{-\frac{(t_1-m)^2}{2\sigma^2}}-\left[\sigma^2+mP(t_2,m,\sigma)\right]\cdot e^{-\frac{(t_2-m)^2}{2\sigma^2}} \qquad m\le t_1<t_2 \quad \text{(III.16a)}$$

$$\int_{t_2}^{t_1} xe^{-\frac{(x-m)^2}{2\sigma^2}}dx = \left[-\sigma^2 + mR(t_1,m,\sigma)\right]\cdot e^{-\frac{(t_1-m)^2}{2\sigma^2}} - \left[-\sigma^2 + mR(t_2,m,\sigma)\right]\cdot e^{-\frac{(t_2-m)^2}{2\sigma^2}} \qquad t_2 < t_1 \le m \quad \text{(III.16b)}$$

$$\int_{t_1}^{t_2} x^2 e^{-\frac{(x-m)^2}{2\sigma^2}}dx = \left[(t_1+m)\sigma^2 + (m^2+\sigma^2)P(t_1,m,\sigma)\right]e^{-\frac{(t_1-m)^2}{2\sigma^2}} - \left[(t_2+m)\sigma^2 + (m^2+\sigma^2)P(t_2,m,\sigma)\right]e^{-\frac{(t_2-m)^2}{2\sigma^2}} \qquad m \le t_1 < t_2 \quad \text{(III.17a)}$$

$$\int_{t_2}^{t_1} x^2 e^{-\frac{(x-m)^2}{2\sigma^2}}dx = \left[-(t_1+m)\sigma^2 + (m^2+\sigma^2)R(t_1,m,\sigma)\right]e^{-\frac{(t_1-m)^2}{2\sigma^2}} - \left[-(t_2+m)\sigma^2 + (m^2+\sigma^2)R(t_2,m,\sigma)\right]e^{-\frac{(t_2-m)^2}{2\sigma^2}} \qquad t_2 < t_1 \le m \quad \text{(III.17b)}$$

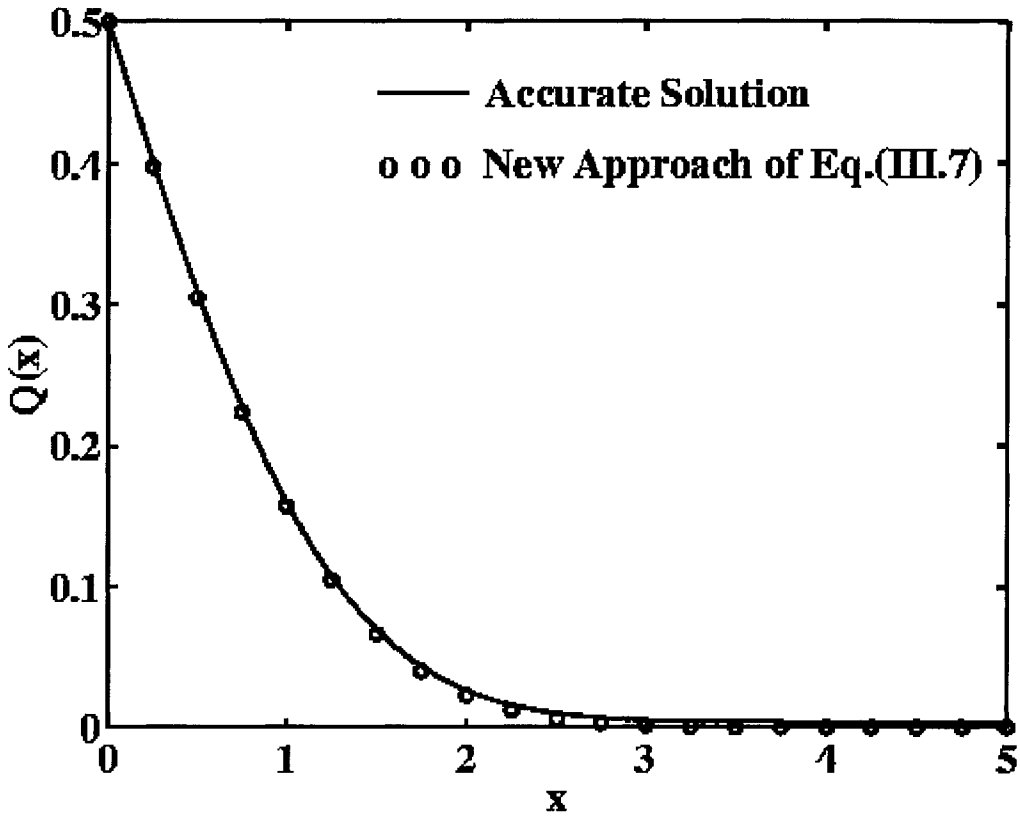


Fig. III.1 Q-function from new approach and numerical approach. It shows that the new approach is very accurate to describe Q-function.