



M.Sc. Thesis

Standard Cell Behavior Analysis and Waveform Set Model for Statistical Static Timing Analysis

Ashish Nigam

Abstract

As we are moving toward nanometre technology, the variability in the circuit parameters and operating environment (*Process, Voltage and Temperature (PVT)*) are increasing, causing uncertainty in the circuit performance. *Statistical Static Timing Analysis (SSTA)* is a category of methodologies to analyse the variations in delay due to PVT variations. This thesis work is a part of the MODERN project, which is involved in developing a new SSTA methodology.

In this thesis, the variation of the delay in 45nm standard cells is analysed. In industry practice, the Monte Carlo method is often used to estimate the statistical moments. This method needs a large number of simulation iterations and these simulations are parameter distribution dependent. A fast statistical moment estimation method is proposed in this work. The proposed methodology is at least 100× faster than the Monte Carlo method and simulations are independent of the parameter distribution.

In the SSTA methodology of the MODERN project, the signal waveforms with their variations are preserved at each pin of the standard cell. The concept of a “set of waveforms” as a representation of a variable electrical signal is also developed in this thesis work. Possible methods to represent the set of waveforms and their integration with the timing analysis methodology are analysed. The pseudo circuit based representation turns out to be the most compact model. A methodology for the analysis of the accuracy and efficiency of the pseudo circuit model is proposed.

Standard Cell Behavior Analysis and Waveform Set Model for Statistical Static Timing Analysis

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

MICROELECTRONICS

by

Ashish Nigam
born in Tulsipur, India

This work was performed in:

Circuits and Systems Group
Department of Microelectronics & Computer Engineering
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology



Delft University of Technology

Copyright © 2010 Circuits and Systems Group
All rights reserved.

DELFT UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF
MICROELECTRONICS & COMPUTER ENGINEERING

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled “**Standard Cell Behavior Analysis and Waveform Set Model for Statistical Static Timing Analysis**” by **Ashish Nigam** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: 30 June 2010

Chairman:

prof. dr. ir. Edoardo Charbon

Advisors:

dr. ir. Nick van der Meijs

dr. ir. Michel Berkelaar

Committee Members:

dr. ir. Said Hamdioui

Abstract

As we are moving toward nanometre technology, the variability in the circuit parameters and operating environment (*Process, Voltage and Temperature (PVT)*) are increasing, causing uncertainty in the circuit performance. *Statistical Static Timing Analysis (SSTA)* is a category of methodologies to analyse the variations in delay due to PVT variations. This thesis work is a part of the MODERN project, which is involved in developing a new SSTA methodology.

In this thesis, the variation of the delay in 45nm standard cells is analysed. In industry practice, the Monte Carlo method is often used to estimate the statistical moments. This method needs a large number of simulation iterations and these simulations are parameter distribution dependent. A fast statistical moment estimation method is proposed in this work. The proposed methodology is at least 100× faster than the Monte Carlo method and simulations are independent of the parameter distribution.

In the SSTA methodology of the MODERN project, the signal waveforms with their variations are preserved at each pin of the standard cell. The concept of a “set of waveforms” as a representation of a variable electrical signal is also developed in this thesis work. Possible methods to represent the set of waveforms and their integration with the timing analysis methodology are analysed. The pseudo circuit based representation turns out to be the most compact model. A methodology for the analysis of the accuracy and efficiency of the pseudo circuit model is proposed.

Acknowledgments

First and foremost I offer my sincerest gratitude to my supervisors dr. ir. Nick van der Meijs and dr. ir. Michel Berkelaar who have supported me throughout my thesis. They guided me towards the right direction at every difficult time during my thesis project. I attribute the level of my Masters degree to their encouragement and effort and without them this thesis, too, would not have been completed or written. I would also like to sincerely thank the other members of the MODERN project, Qin Tang, Amir Zjajo, and Kees-Jan van der Kolk. They spent many discussion hours to listen to my ideas and gave their feedback to improve the quality of work. They all encouraged and helped me to write the paper and supported to improve my thesis report. One simply could not wish for a better or friendlier supervisor and research group.

I would also like to thank prof. Alessandro Di Bucchianico, Eindhoven University of Technology, for useful discussions and contributions. His help in reviewing the mathematical models of the thesis work was a great help.

I would like to thank prof. dr. ir. Edoardo Charbon and dr. ir. Said Hamdioui for their valuable time spent in my thesis defence committee.

Antoon Frehe has provided me all kind of support for the compute machines and software which are critical in compute intensive thesis work. Laura Bruns and Judith Bukman-Vollering have helped me in all the office related work, which saved significant time and effort.

In my daily work I have been blessed with a friendly and cheerful group of fellow students. Chokalingam Veerappan, Saket Sakunia, Akansh Goyal, Shahzad Gishkori, and Maxim Volkov have helped me to develop and refine the methodology I used in my thesis work. Their healthy discussions and valuable feedback solved various problems of mine. There are many other friends in Delft who made my life filled with joy during my two years masters program. I sincerely thank all of them, without them life would not have been possible in Delft.

Finally, I thank my parents for supporting me throughout my studies at Delft University of Technology.

Ashish Nigam
Delft, The Netherlands
30 June 2010

Contents

Abstract	v
Acknowledgments	vii
List of Abbreviations	xv
1 Introduction	1
1.1 Motivation	1
1.2 Project Sketch	1
1.3 Thesis Overview	2
2 Digital Circuit and PVT Variations	3
2.1 Digital Circuit	4
2.2 Static Timing Analysis	9
2.2.1 Non-linear Delay Model	12
2.2.2 Composite Current Source Model	12
2.2.3 Effective Current Source Model	13
2.3 Statistical Static Timing Analysis	14
2.4 Circuit Simulation and Analysis Environment	15
2.5 Summary	16
3 Delay Variations	17
3.1 Variations in the PVT parameters	17
3.2 Circuit Simulation Configuration	20
3.3 Delay variation in an inverter	22
3.3.1 Individual Parameter Variation	24
3.3.2 Combination of two parameters variations	26
3.3.3 A realistic PVT variation	28
3.4 Higher order statistical moments	29
3.5 Delay variation in various standard cells	30
3.6 Summary	33
4 Statistical Moment Estimation and Probability Density Function	35
4.1 Motivation	35
4.2 Fast Statistical Moment Estimation Method	36
4.2.1 Circuit Simulation	37
4.2.2 Data Processing	37
4.3 Simulation Results and Comparison for FSME Method	42
4.4 Probability Density Function Estimation Method	48
4.4.1 PDF Estimation - Bin Method	50
4.4.2 PDF Estimation - Direct Method	51
4.5 Simulation Results for PDF Estimation Method	51

4.6	Summary	52
5	The Set of Waveforms	55
5.1	Concept of a Set of Waveforms	55
5.2	Representing Uncertainty with the Set of Waveforms	57
5.3	Representation of the Set of Waveforms	59
5.3.1	Lookup Table based Representation	60
5.3.2	Statistical Moments based Representation	60
5.3.3	Pseudo Circuit based Representation	63
5.4	Comparison of various Waveform Set Representations	65
5.5	Summary	65
6	Pseudo Circuit Model	67
6.1	Pseudo Circuit, Waveform Set and SSTA Engine	67
6.2	The Pseudo Circuit Model	67
6.2.1	The Pseudo Circuit	68
6.2.2	Database processing for waveform comparison	72
6.3	Waveform Comparison Methodology	76
6.4	Results	83
6.5	Summary	85
7	Conclusion	87
7.1	Summary	87
7.2	Future Work	88
A	Delay Variations	89
	Bibliography	97
	List of Publications	101

List of Figures

2.1	Full-custom design and semi-custom design	3
2.2	Data Path	5
2.3	Slew	5
2.4	Illustration of a circuit component delay	6
2.5	Master-Slave DFF with MUX	7
2.6	Master-Slave DFF with MUX as Pass Transistor Logic	7
2.7	Setup, Hold, and Clock to Q Delay in Waveform	8
2.8	DFF to DFF Path	9
2.9	Path Delay	9
2.10	CMOS Inverter	10
2.11	ON and OFF state of the PMOS and the NMOS in an Inverter	11
2.12	Measurement of $I_{out}(t)$ for a set of S_{in} and C_{out}	13
2.13	Measurement of $V_{out}(t)$ for a set of S_{in} and C_{out}	14
3.1	Spread estimation for PTM Model	19
3.2	Delay of INV_X1 vs W for $\mu_W = 90\text{nm}$ and $3\sigma_W = 58\text{nm}$	21
3.3	INV_X1 simulation configuration	23
3.4	Delay variation due to L	25
3.5	Delay variation due to W & L	27
3.6	Delay distribution pdf due to realistic PVT variations	29
4.1	Numerical Integration Method	39
4.2	Piecewise constant approximation of probability density function	40
4.3	The first four moment estimation vs simulation runs for MC and FSME with one parameter (L) and two parameters (L and W) variations in 45nm Inverter	44
4.4	The first four moment estimation vs simulation runs for MC and FSME with one parameter (L) and two parameters (L and W) variations in 32nm Inverter	45
4.5	Moment estimation vs simulation run in 45nm inverter chain using Gaussian (N), Lognormal (L), Gamma (G), and Beta (B) distributions.	48
4.6	Moment estimation vs simulation run in 45nm inverter chain using Gaussian (N), Lognormal (L), Gamma (G), and Beta (B) distributions with Relative Scale.	49
4.7	<i>pdf</i> of the delay of an 45nm inverter chain using Gaussian (N), Lognormal (L), Gamma (G), and Beta (B) distributions.	52
4.8	<i>pdf</i> of the delay of an 45nm inverter chain using Gaussian (N), Lognormal (L), Gamma (G), and Beta (B) distributions in one plot	53
5.1	A Inverter for Waveform Set	57
5.2	The Set of Waveforms for Inverter	58
5.3	A Inverter Chain for Waveform Set	58
5.4	The Set of Waveforms for Inverter Chain	59

5.5	SSTA Engine	59
5.6	SSTA Engine with Table Model	60
5.7	Vertical cross-section of Waveforms at time t	61
5.8	Vertical cross-section of Waveforms at time t with μ and $\pm\sigma$ of voltage	62
5.9	First four moments of the waveform set as a function of time	63
5.10	SSTA Engine with Moment Model	64
5.11	Example path for SSTA	64
5.12	SSTA Engine with Pseudo Circuit Model	65
6.1	An inverter chain with three inverters	68
6.2	Inverter chain with input source and output load	69
6.3	Inverter chain with capacitors at internal nodes	70
6.4	Inverter chain with time offset	70
6.5	The Pseudo Circuit Schematic	71
6.6	Quality factors for the waveform comparison	73
6.7	Target Waveform Set	78
6.8	S_{in} vs C_{load} intersection line for Q_{Slew} equals to T_{Slew}	79
6.9	S_{in} vs C_{load} intersection line for $Q_{ShiftMean}$ equals to $Q_{ShiftMean}$	80
6.10	Intersection of Slew and ShiftMean lines	81
6.11	Q_{Max} vs σ_L	82
6.12	Comparison of pseudo circuit model and target set of waveforms	84
6.13	Result comparison of pseudo circuit model and target waveform set	85
A.1	Delay variation due to L	89
A.2	Delay variation due to W	89
A.3	Delay variation due to V_{th}	90
A.4	Delay variation due to V_{DD}	90
A.5	Delay variation due to T	91
A.6	Delay variation due to V_{DD} & T	91
A.7	Delay variation due to V_{DD} & V_{th}	92
A.8	Delay variation due to V_{DD} & W	92
A.9	Delay variation due to V_{DD} & L	93
A.10	Delay variation due to T & V_{th}	93
A.11	Delay variation due to T & W	94
A.12	Delay variation due to T & L	94
A.13	Delay variation due to V_{th} & W	95
A.14	Delay variation due to V_{th} & L	95
A.15	Delay variation due to W & L	96
A.16	Delay distribution pdf for a realistic PVT variations	96

List of Tables

2.1	NLDM based cell delay (t_{delay}) lookup table	12
2.2	CCS based output current waveform ($I_{out}(t)$) lookup table	13
2.3	ECSM based output voltage waveform ($V_{out}(t)$) lookup table	13
3.1	Technology process parameter trends based on [3]	18
3.2	Spread of the PVT parameters	20
3.3	Nominal and 3σ range for Nangate INV_X1 inverter	21
3.4	Nominal and 3σ range for modified INV_X1 inverter	22
3.5	INV_X1 simulation configuration	23
3.6	Delay spread of INV_X1 due to individual parameter variation	24
3.7	Delay variation due to two parameter variations	28
3.8	Delay variation due to realistic PVT variation	29
3.9	First four statistical moments of delay variation in INV_X1 due to realistic PVT variation	30
3.10	List of standard cells	31
3.11	Delay variation in Nangate standard cells (INV, NAND2, and NOR2) due to realistic PVT variation	32
3.12	Delay variation in Nangate standard cells (BUF, AND2, OR2, XOR2 and XNOR2) due to realistic PVT variation	33
4.1	Error % comparison in the first four moments estimation of delay for one parameter (L) variation using Monte Carlo (5000 runs) and the proposed method (50 runs) in 45nm and 32nm PTM technologies	46
4.2	Error % comparison in the first four moments estimation of delay for two parameters (L and W) variation using Monte Carlo (10000 runs) and the proposed method (100 runs) in 45nm and 32nm PTM technologies	47
6.1	The Pseudo Circuit	71
6.2	Pseudo Circuit Parameters	71
6.3	Simulation output in database	72
6.4	Mean and SD curves for quality factor illustration	73
6.5	Database Structure	76
6.6	Quality factors of target waveform	77
6.7	Pseudo Circuit Parameters after waveform comparison methodology	83
6.8	Quality factors of target waveform set and pseudo circuit model	84

List of Abbreviations

ASIC	Application Specific Integrated Circuits
BSIM	Berkeley Short-channel IGFET Model
CCS	Composite Current Source Model
CMOS	Complementary Metal Oxide Semiconductor
DFF	Edge Triggered Flip-Flop
ECSM	Effective Current Source Model
EKV	Transistor Model given by C. C. Enz, F. Krummenacher and E. A. Vittoz
EPFL	Ecole Polytechnique Federale de Lausanne
FSME	Fast Statistical Moment Estimation Method
HP	High Performance
IC	Integrated Circuits
ITRS	International Technology Roadmap for Semiconductor
LHS	Latin Hypercube Sampling
LP	Low Power
MC	Monte Carlo
MODERN	MOdeling and DEsign of Reliable, process variation-aware Nanoelec- tronic devices, circuits and systems
MOSFET	Metal Oxide Semiconductor Field Effect Transistors
MUX	Multiplexer
NLDM	Non-Linear Delay Model
NMOS	n-type Metal Oxide Semiconductor Field Effect Transistors
PCA	Principal Component Analysis
PDF	Probability Density Function
PMOS	p-type Metal Oxide Semiconductor Field Effect Transistors
PTM	Predictive Technology Model
PVT	Process, Voltage and Temperature
PWC	Piecewise Constant
QMC	Quasi Monte Carlo
SH-QMC	Stratification + Hybrid Quasi Monte Carlo
SPICE	Simulation Program with Integrated Circuit Emphasis
SSTA	Statistical Static Timing Analysis
STA	Static Timing Analysis
UCB	University of California, Berkeley

Introduction

1.1 Motivation

Economical benefit is the main driving force for shrinking of CMOS circuit design technology. Other benefits of device shrinking are lower power consumption per transistor [1], enhancement in circuit performance [2] etc. These are possible because of faster devices and lower supply voltages in smaller technology nodes. However, as technology is moving into deep submicron dimensions, various complex device phenomena are playing an important role in digital circuit functionality and reliability. Additionally, variations in these devices are causing uncertainty in device behaviour. Delay test is one of the important analysis tools for circuit reliability and functionality check. The variations in the circuit parameters such as *Process, Voltage and Temperature (PVT)* are significantly affecting the delay of digital circuits [3]. In advanced technology nodes with increasing operating clock frequency in digital circuits, the relative error in the delay calculation is becoming very critical and significantly impacts yield [4, 5]. This is because the gate delay on silicon might be higher than the estimated gate delay, and cause the circuit not to work on the targeted clock frequency.

Traditionally, lookup table based gate models are used in the delay analysis. However, they are not sufficient to model the complex behaviour of the gate very accurately in sub 90nm technology nodes. Improved standard cell models have been proposed to overcome these problems. The *Composite Current Source (CCS)* Model from Synopsys and *Effective Current Source Model (ECSM)* from Cadence are industry standard modelling schemes among them. The problems due to complex standard cells are well addressed by the CCS and ECSM models, however, the problems due to PVT variations are not addressed.

The MODERN project is involved in developing an advanced delay analysis methodology which can address the problems mentioned above by mainly using fast circuit simulators and preserving the standard cell output waveforms along with their uncertainty. This thesis work is within the objectives of the MODERN project.

1.2 Project Sketch

The main contribution of this thesis work is to analyse the variations in the delay of standard cells caused by the PVT variations. Furthermore, variations in the output signal waveforms of the standard cells are analysed and a representation method is developed. State of the art 45nm predictive technology models are used in this work such that the complex behaviour of deep submicron technologies can be analysed. Monte Carlo is a well known industry standard method for statistical analysis of the circuit behaviour. However, there are two major limitations of Monte Carlo. Firstly, it needs

a very high number (thousands) of simulation runs and secondly, these circuit simulations are process variation dependent. A fast statistical circuit analysis method is proposed and its performance is compared with standard Monte Carlo methods.

The delay analysis methodology in the MODERN project is preserving various possible output waveforms of standard cells due to PVT variations. The concept of a “set of waveforms” as a representation of a variable output waveforms is also developed in this thesis work. Possible methods to represent the set of waveforms have been analysed. The pseudo circuit based representation turns out to be the most compact model. Therefore, a possible pseudo circuit model is proposed. Additionally, a methodology for the analysis of the accuracy and efficiency of the pseudo circuit model is developed.

The main contributions of this thesis are:

- Analysis of the variations in the delay of standard cells due to variations in the PVT parameters in a 45nm technology.
- Development of a fast statistical circuit analysis method and comparison of its performance with a standard Monte Carlo method. [6, 7]
- Development of the concept of a set of waveforms and possible methods for their representation and integration with the delay analysis methodology in the MODERN project.
- Development of a pseudo circuit based representation for the set of waveforms and a methodology for the analysis of the accuracy and efficiency of the pseudo circuit model. [8]

1.3 Thesis Overview

The organization of the report is as follows: The digital circuit, source of PVT variations and their impact on the circuit, and simulation environment are presented in Chapter 2. The variation in the delay of standard cells due to PVT variations is analysed in Chapter 3. The proposed method of fast statistical circuit analysis and the methodology of estimating probability density functions are discussed in Chapter 4. The concept of the set of waveforms, a method of uncertainty representation in these waveforms and its integration with the delay analysis tool is discussed in Chapter 5. A pseudo-circuit model representation for the set of waveforms and its performance and accuracy evaluation is given in Chapter 6. At the end, conclusions and possible future work is reported in Chapter 7.

Delay variations of an inverter (INV_X1) due to various possible combinations of PVT variations are given in Appendix A.

Digital circuit design is a stepwise approach with various levels of abstraction. The design process of a chip begins with the functional requirement, which can be first divided into functional units like control unit, processing unit etc. The functional units are designed with smaller blocks like adder, multiplier, selection logic etc. and these blocks are then implemented with logic gates. The logic gates are designed with *Metal Oxide Semiconductor Field Effect Transistors (MOSFET)*. Depending on the levels of abstraction, we can broadly classify digital circuit design into two major categories: full-custom design and semi-custom design. In full-custom design, end-to-end circuit design is owned by the designer with the MOSFET as the lowest level of abstraction; whereas in semi-custom design, designers use pre-designed library cells. These pre-designed library cells include commonly used gates such as NOT, AND, NAND, OR, NOR, XOR, MUX, D Flip Flop, Latch etc. The span of full-custom design and semi-custom design along with the various levels of abstraction in digital circuit design are shown in Figure 2.1.

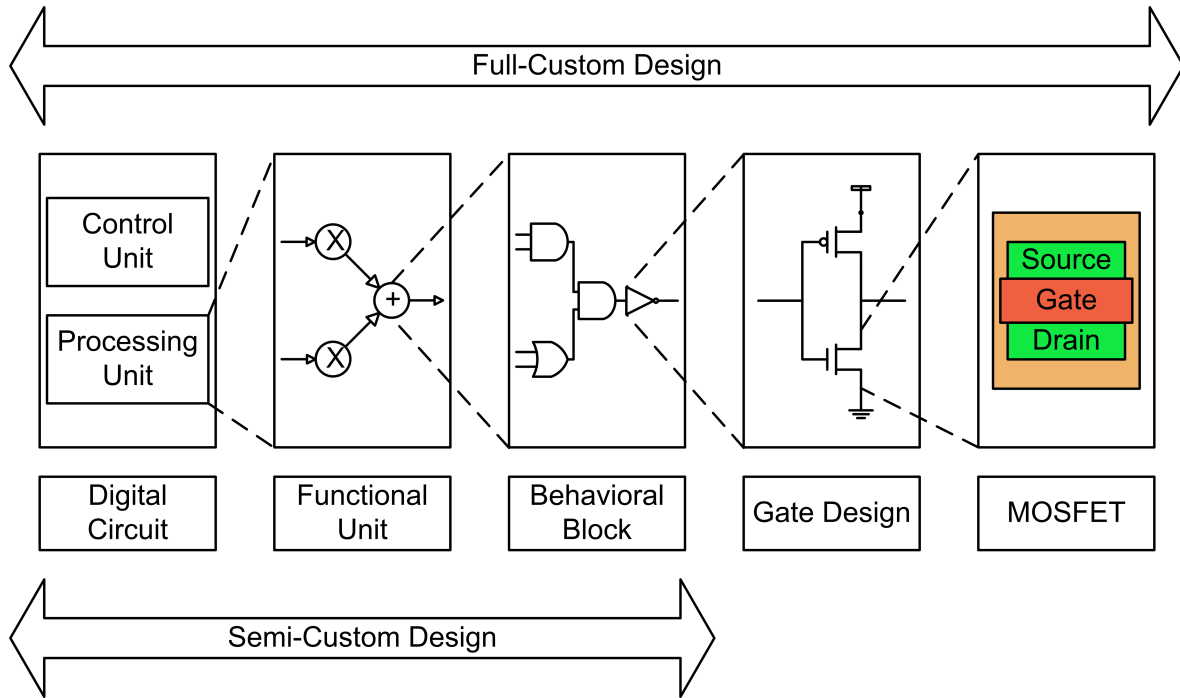


Figure 2.1: Full-custom design and semi-custom design

Full-custom design potentially maximizes performance and minimizes the area of the chip, however this is very labour intensive to implement. That is one of the main reason why full-custom design is limited to *Integrated Circuits (IC)* which are to be

fabricated in very high volume (like microprocessors) or require very high performance. The rest of the *Application Specific Integrated Circuits (ASIC)* use semi-custom design.

Performance of a digital circuit is primarily measured by the time it takes to process the input and execute the required operation. These circuits can either perform the operations on the clock tick or on the input signal availability. So, digital circuits can also be classified as synchronous circuits or asynchronous circuits. A synchronous circuit runs on the clock tick and uses edge triggered flip-flops or level sensitive latches as a memory elements to store the state. In an asynchronous circuit, dual rail encoded signals, and done signal acknowledgement mechanisms are used. In this chapter, we will only discuss synchronous circuits in detail. Edge triggered flip-flops are usually known as “DFFs” and level sensitive latches are known as “Latches”. However, in the following text we will use DFF instead of both edge triggered flip-flop and level sensitive latch for simplicity. Also, since we will be talking about ASIC synchronous design only, we will often use “circuit” instead of “ASIC synchronous circuit”.

The organization of the chapter is as follows: the basics of digital circuits are discussed in Section 2.1 and state of the art methodologies for *Static Timing Analysis (STA)* are discussed in Section 2.2. Furthermore, *Process, Voltage and Temperature (PVT)* variations, their impact on the digital circuit, and the need for *Statistical Static Timing Analysis (SSTA)* are discussed in Section 2.3. The MOSFET technology models, circuit simulation and data analysis tools are discussed in Section 2.4. At the end, the summary of the chapter is presented in Section 2.5.

2.1 Digital Circuit

In digital circuits, a signal moves from a gate to another gate and combines with other signals inside the gate, which results into a new signal at the output of the gate. These signals can be categorized based on the type of the signal travelling on it, e.g. data, clock, reset, test, etc. We will mainly focused on data signals and sometimes clock signals. A path of logic gates traversed by a data signal is also known as a “data path”. The signal transition quality is measured in terms of the “signal slew”. The time taken by a gate or an interconnect to reflect the change of the inputs to the output is called “delay”. The terms “data path”, “slew”, and “delay” are defined as:

Data Path: A data path starts from either a primary input pin of the circuit, an output pin of a DFF, or an output pin of a macro block. It ends with a primary output pin of the circuit, an input data pin of a DFF or an input of a macro block. An example of a data path is shown in Figure 2.2. Here, start points are shown on the left side of the logic gates and end points are shown on the right side of the logic gates.

Slew: The slew of a signal is a measure of the quality of the transition. Mathematically the slew is defined as a signal transition time between two slew threshold voltages, as given in (2.1) and depicted in Figure 2.3. Here V_1 and V_2 are the lower and upper threshold voltages, and t_1 and t_2 are the times when the signal reaches V_1 and V_2 respectively. The slew threshold voltages vary from design to design,

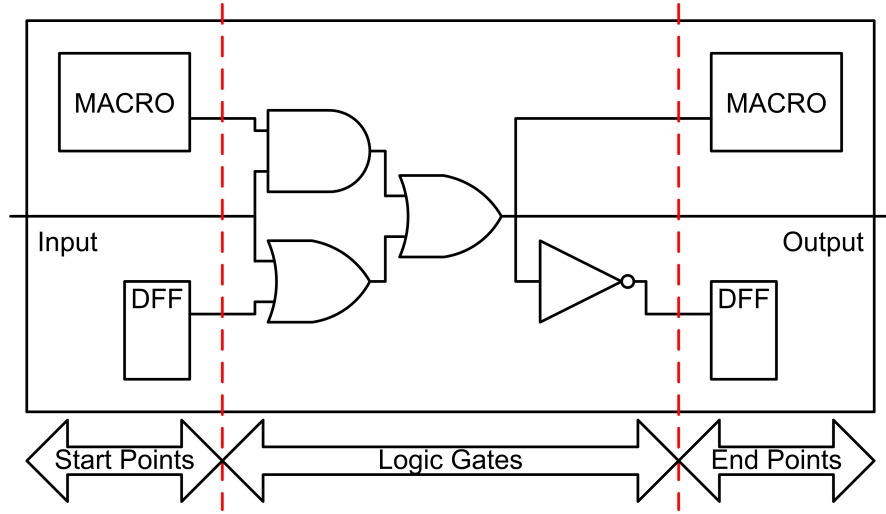


Figure 2.2: Data Path

however most of the designers use 40 / 60 or 30 / 70 % of the supply voltage (V_{DD}) as the threshold values. There are trade-offs in the threshold selection and the accuracy of the quality measurement. These trade-offs are not discussed in this chapter but interested readers can find details in [9].

$$Slew = |t_2 - t_1| \quad (2.1)$$

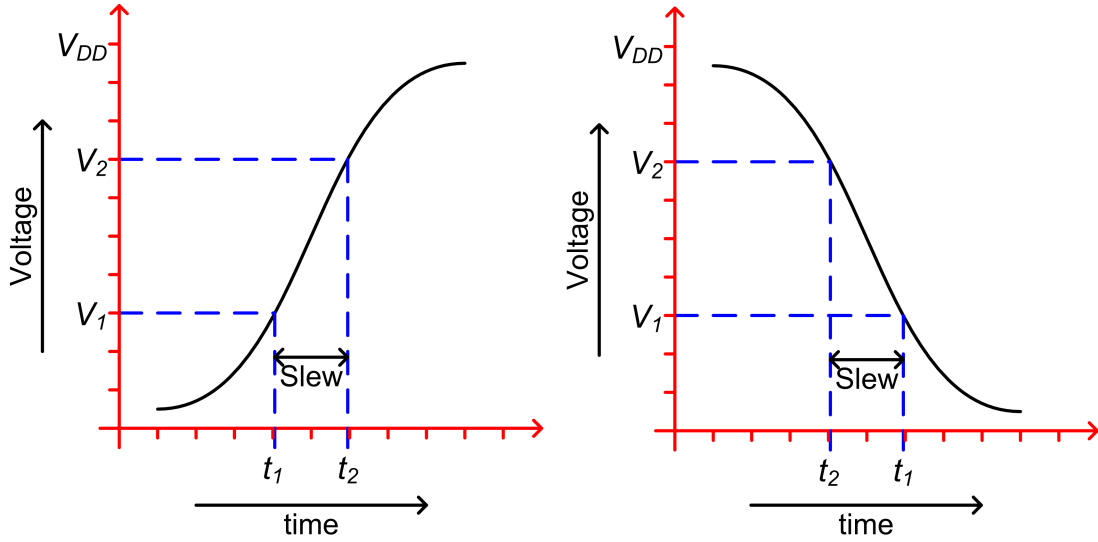


Figure 2.3: Slew

Sometimes, slew is also defined as the signal transition time between V_{DD} and lower supply voltage (V_{SS}) by fitting a straight line through two slew threshold voltage points. In this chapter, we will discuss the delay calculation methodology corresponding to

the first definition only. Appropriate changes can be made to incorporate the second definition into the delay calculation methodology.

Delay: Delay is defined as the time taken by a signal to travel from the input pin to the output pin of a circuit component. This circuit component could be an interconnect, a gate, or a collection of gates and interconnects. The mathematical equation for delay calculation is shown in (2.2). Here t_{delay} is the delay of the component, and t_{in} and t_{out} are the times when the input and output signals cross the delay threshold point which is usually $V_{DD}/2$. However, the delay threshold point can be changed if required. An example of the delay of an inverter along with the input and output waveforms is shown in Figure 2.4.

$$t_{delay} = t_{out} - t_{in} \quad (2.2)$$

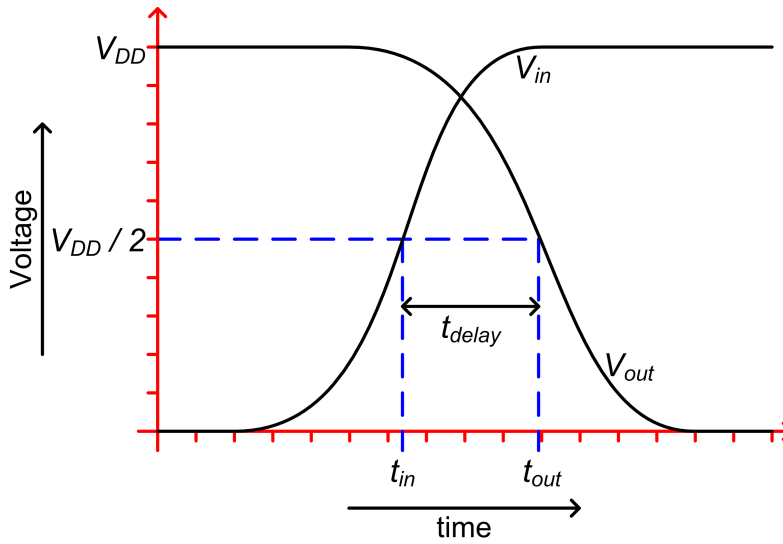


Figure 2.4: Illustration of a circuit component delay

A data signal traverses through the logic gates of a path and is captured by a DFF at the clock edge. Due to the internal delay of the DFF, some timing constraints are imposed on the input signals such as setup time, hold time, minimum clock width etc. Here we will concentrate on two major constraints: setup time and hold time, which are defined below as:

Setup Time: To store the data correctly into the DFF, the input signal needs to be stable before the clock edge. The minimum time before the clock edge for which the input signal needs to be stable is known as the setup time.

Hold Time: Similar to the setup time constraint, the input signal needs to be stable for some time after the clock edge to ensure correct functionality of the DFF. The minimum time after the clock edge for which the input signal needs to be stable is known as the hold time.

A master-slave positive edge triggered flip-flop design using *multiplexers (MUXes)* is shown in Figure 2.5. A design of the same DFF with a pass transistor logic implementation of the MUXes is shown in Figure 2.6. Here $I_1, I_2, I_3 \dots$ are the inverters and $T_1, T_2, T_3 \dots$ are the pass transistor logic. In these implementations, the internal node Q_m captures the input signal during the low level of the clock via I_1, T_1 , and I_3 . Q_m retains the signal value during the high level of the clock with the help of the feedback loop with I_2, T_2 , and I_3 . Similarly, during the high level of the clock, the stored signal at Q_m is passed to the output pin Q via I_4, T_3 , and I_6 and retains the output signal at the low level of the clock with the help of the loop with I_5, T_4 , and I_6 .

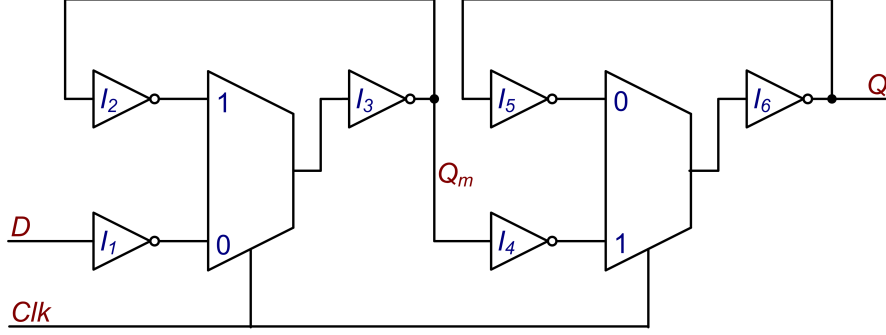


Figure 2.5: Master-Slave DFF with MUX

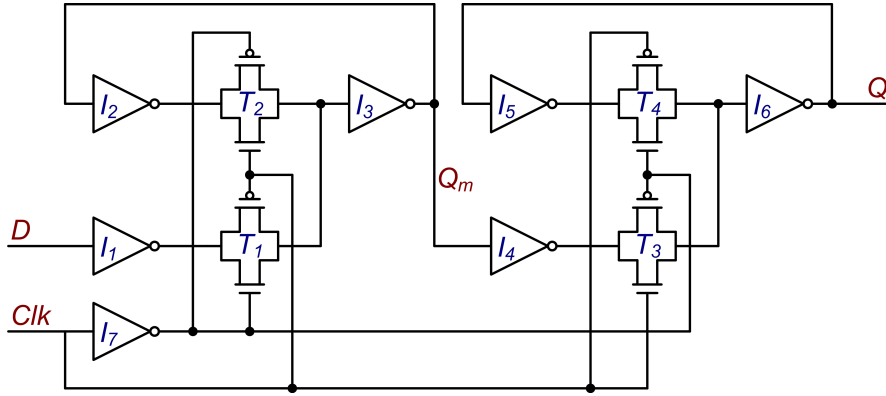


Figure 2.6: Master-Slave DFF with MUX as Pass Transistor Logic

To ensure that the correct signal value is stored into the DFF, the signal at the node Q_m and the loop I_2, T_2 , and I_3 should be stable before the rising edge of the clock. This constraint on the input signal gives the setup time of the DFF and is given in

$$t_{setup} = t_{d_{I_1}} + t_{d_{T_1}} + t_{d_{I_3}} + t_{d_{I_2}} \quad (2.3)$$

where t_{d_k} denotes the delay due to component k . Since the signal propagation from the node Q_m to the output of the inverter I_2 is considered in the setup time, the output of the inverter I_4 will also be stable at the rising edge of the clock. We should also consider that the input signal should not change before it reaches from the internal node Q_m to the output pin Q , which gives the hold time of the DFF expressed as

$$t_{hold} = t_{d_{T_3}} + t_{d_{I_6}} \quad (2.4)$$

Since a DFF stores new signal values only at the clock transition, the delay of a DFF is defined from the clock pin to the output pin and known as clock to Q delay (t_{ck-q}). Setup time (t_{setup}), hold time (t_{hold}), and clock to Q delay (t_{ck-q}) are depicted with the help of the voltage waveforms in Figure 2.7.

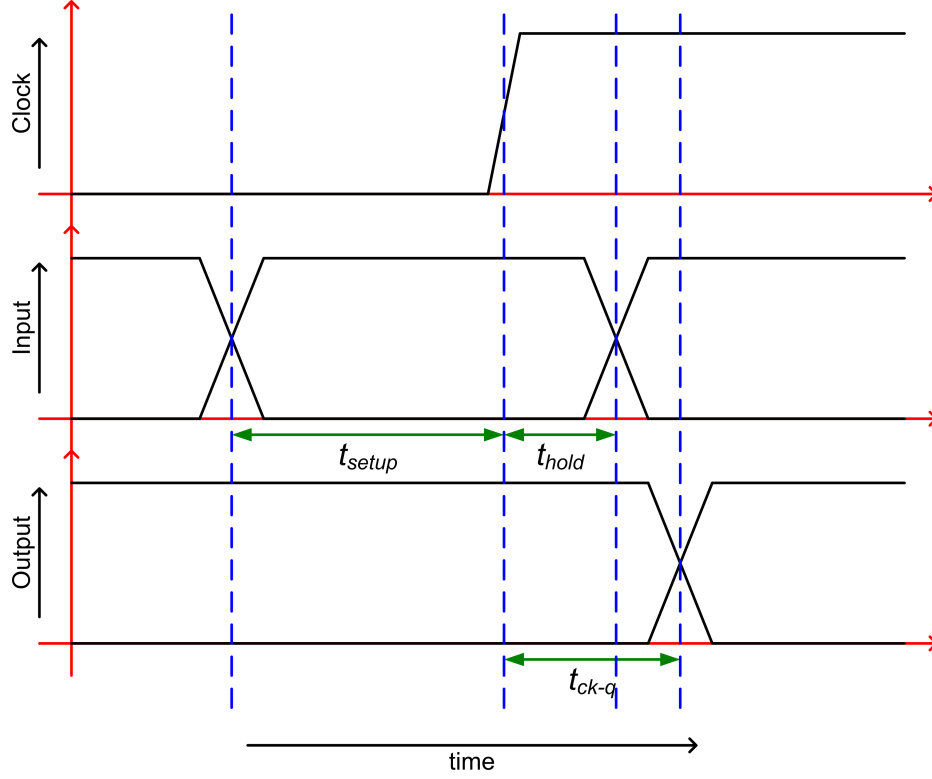


Figure 2.7: Setup, Hold, and Clock to Q Delay in Waveform

For a DFF to DFF data path, as depicted in Figure 2.8, DFF₁ is called the start point of the data path and DFF₂ is called the end point of the data path. In this path, the signal at the start point changes at the clock edge and the signal at the end point is captured at the next clock edge, which gives the relation between the maximum permissible delay of the data path and the minimum possible clock period as given in

$$\begin{aligned} t_{period} &\geq t_{path_delay} + t_{setup} + t_{hold} \\ \Rightarrow t_{period,min} &= t_{path_delay} + t_{setup} + t_{hold} \end{aligned} \quad (2.5)$$

Here t_{path_delay} is the delay of the logic gates and the interconnects between the start point and the end point of the data path, and $t_{period,min}$ is the minimum permissible time period of the clock. The data path which has the longest path delay in the entire circuit is also known as the critical path and it directly relates to the highest possible frequency of the circuit.

The path delay is comprised of the delay contribution due to each gate and the interconnects of the data path. An example of a data path is depicted in Figure 2.9. Path delay equations for the data path from DFF₁ to DFF₃ (t_{d1-3}) and DFF₂ to DFF₃

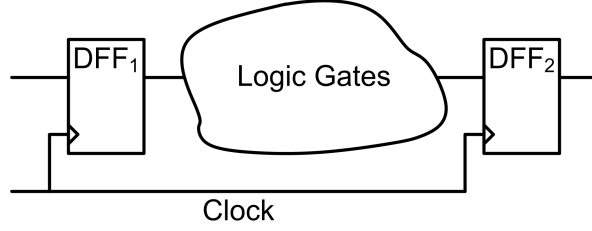


Figure 2.8: DFF to DFF Path

$(t_{d_{2-3}})$ are given in (2.6) and (2.7) respectively.

$$t_{d_{1-3}} = t_{DFF_1} + t_{n_1} + t_{g_1} + t_{n_2} + t_{g_2} + t_{n_3} \quad (2.6)$$

$$t_{d_{2-3}} = \text{MAX}((t_{DFF_2} + t_{n_6} + t_{n_4} + t_{g_1} + t_{n_2} + t_{g_2} + t_{n_3}), \\ (t_{DFF_2} + t_{n_6} + t_{n_5} + t_{g_3} + t_{n_7} + t_{g_2} + t_{n_3})) \quad (2.7)$$

Here t_{DFF_k} denotes the delay contribution due to DFF_k, t_{g_k} denotes the delay contribution due to gate g_k , and t_{n_k} denotes the delay contribution due to interconnect n_k . Path delay calculation is a stage-by-stage approach in which the input of one gate to the input of a successor gate is one stage. For example, DFF₁ to g_1 , g_1 to g_2 , and g_2 to DFF₃ are 3 stages of the data path from DFF₁ to DFF₃.

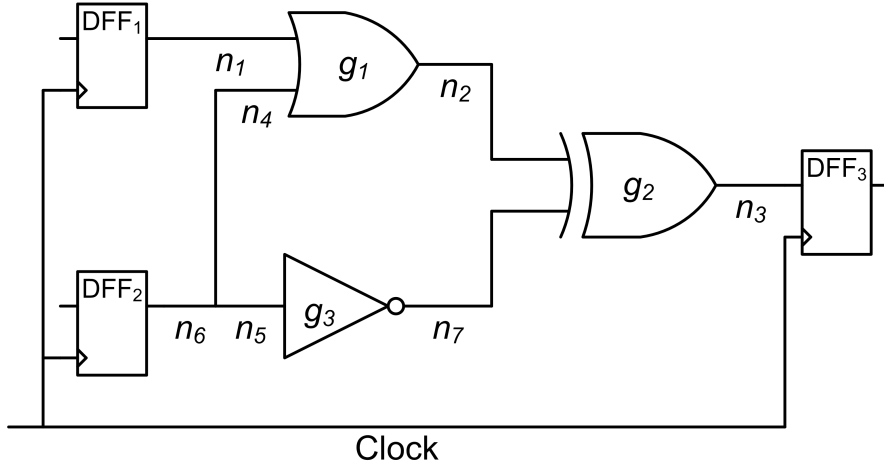


Figure 2.9: Path Delay

2.2 Static Timing Analysis

Up until this point we have understood that the critical path defines the maximum operating frequency of the circuit. However in the industry, the required operating frequency of the circuit is pre defined and the circuit needs to be designed to achieve the target operating frequency. In this case, a designer needs to estimate the delay of each possible data path and make sure that the critical path is within the target frequency. This delay analysis is known as “Timing Analysis”. Considering the size of

an ASIC design (in the order of millions of gates), transistor level simulation (such as *Simulation Program with Integrated Circuit Emphasis - SPICE*) analysis is not practical for the entire chip due to its high run time. This problem is addressed by modelling the gate behaviour mathematically and performing mathematical analysis instead of circuit simulation. This analysis is known as “*Static Timing Analysis (STA)*”. The gate behaviour of every standard cell is modelled and stored in a file, known as “Standard Cell Library File” and used as an input in the ASIC design flow stages such as synthesis, placement, routing, and timing analysis.

The models of the gates stored in the standard cell library file are used to estimate the delay of each gate of the data path. The delay of a gate depends on the input signal slew and the output load of the gate. This is discussed below in detail with the help of an inverter.

A MOSFET implementation of an inverter is shown in Figure 2.10. When the input signal (In) is at logic ‘0’ i.e. V_{SS} , the *n-type MOSFET (NMOS)* is in the cut-off mode and the *p-type MOSFET (PMOS)* is in the saturation mode which results in a path from V_{DD} to the output load via the PMOS and the output signal (Out) is at logic ‘1’ i.e. V_{DD} . When the input signal changes from logic ‘0’ to logic ‘1’ (i.e. rise transition), the PMOS is turned off which disconnects the path from V_{DD} to the output pin. The NMOS is turned on which opens a path from the output pin to V_{SS} .

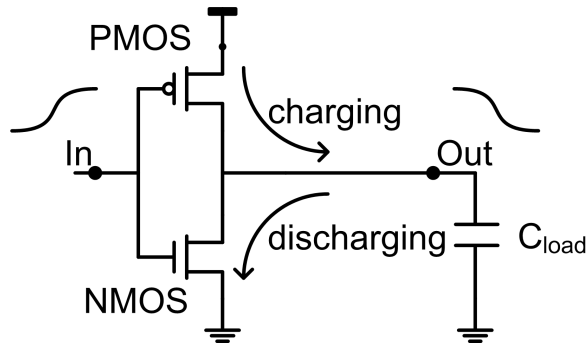


Figure 2.10: CMOS Inverter

The switching time of the output signal from V_{DD} to V_{SS} depends on two factors. Firstly, how quickly the NMOS is turned on and the PMOS is turned off and secondly, the time constant of the discharge circuit. The state of the NMOS and the PMOS of an inverter depends on its gate voltage. The cut-off condition of both the NMOS and the PMOS is given in (2.8) and (2.9), respectively.

$$V_{in} < V_{TN} \quad (2.8)$$

$$V_{in} > V_{DD} - |V_{TP}| \quad (2.9)$$

The same is shown on the rising edge input signal waveform in Figure 2.11. Here V_{TN} and V_{TP} are the threshold voltages of the NMOS and the PMOS respectively, t_n and t_p are the time when the NMOS and the PMOS switches between the stages (ON / OFF) respectively. The switching time of the output signal of the inverter depends on t_n and t_p which further depends on the input signal slew ($|t_2 - t_1|$).

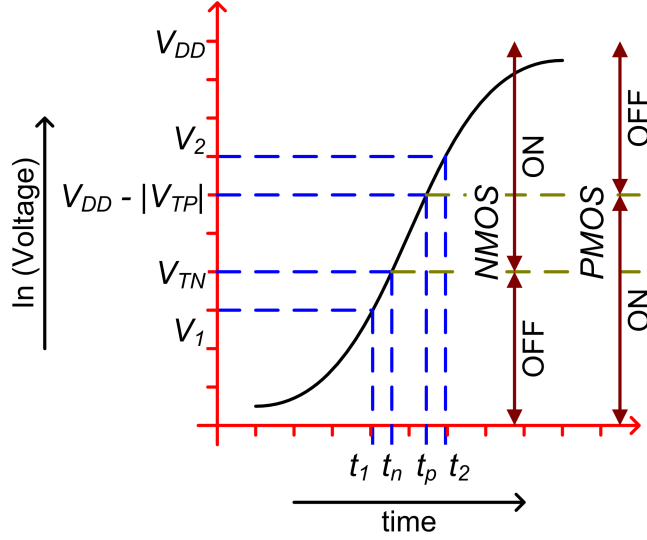


Figure 2.11: ON and OFF state of the PMOS and the NMOS in an Inverter

Output signal transition time also depends on the RC time constant (τ) of the discharge circuit. The charge accumulated on the output load (C_{load}) of the inverter is discharged through the on resistance of the NMOS (R_{on}). The RC time constant (τ) is defined as

$$\tau = R_{on} \times C_{load} \quad (2.10)$$

In this example, we understood that the delay of an inverter, for the input rise transition, depends on the input signal transition time (Slew) and the output load (C_{load}) of the inverter. A similar argument can be made for a fall transition of the input signal.

The total effective load at the output pin of a gate is due to the on resistance of the driver cell, interconnect impedance, and input impedance of the receiver cell [10, 11, 12]. We can see in Figure 2.10 that the effective output impedance of an inverter will have a contribution of the resistance from the output node to the V_{DD} and the V_{SS} via the PMOS and the NMOS respectively. This output resistance of the inverter is very high as compare to the interconnect resistance in higher technology nodes. Thus, the resistive component of the interconnect can be ignored, and this will introduce a negligible error in delay calculation. Also note that, due to the very small dimension of the interconnect, the inductance of the interconnect is small. In a low frequency circuit, the equivalent impedance due to this inductance is very small in comparison with the equivalent impedance due to the resistance and capacitance of the interconnect and can be ignored. These assumptions result in a simplified model of the total effective load at the output pin of a gate as an effective capacitance only for higher technology nodes. However, in the lower technology nodes, the interconnect resistance can no longer be ignored [13].

Now, we can say that the output voltage waveform of a driver can be defined for a given set of the input signal slew and output load, and the delay for STA can be estimated.

Various methodologies have been proposed to model the gate behavior for the delay estimation in STA. Traditionally, the *Non-Linear Delay Model (NLDM)* has been used for the delay estimation [14]. However due to increase in frequency and device shrinking, the relative error introduced by NLDM is no longer acceptable. Various improved modelling schemes have been proposed. The *Composite Current Source (CCS)* [15, 16, 17, 18] and the *Effective Current Source Model (ECSM)* [9, 19] are well known industry standard modelling schemes. We will briefly discuss NLDM, CCS and ECSM modelling schemes in the following subsections.

2.2.1 Non-linear Delay Model

NLDM is the traditionally used modelling scheme, in which the gate delay is modelled using lookup tables. If we neglect the resistive component of the impedance from the effective load of a cell (as discussed in Section 2.2), then the cell delay is dependent on two parameters: input signal slew and output effective capacitive load (for a specified *Process, Voltage, and Temperature (PVT)*). We also know that the MOSFET behaviour depends on the input signal transition i.e. rise transition and fall transition. In NLDM, cell delay is modelled in a two dimensional lookup table for each transition as shown in Table 2.1. Here S_{in} is the input signal slew, C_{out} is the effective output capacitive load of a cell and t_{delay} is the gate delay.

Table 2.1: NLDM based cell delay (t_{delay}) lookup table

Input Signal Slew	Effective Output Capacitance		
	C_{out_1}	C_{out_2}	C_{out_3}
S_{in_1}	t_{delay}	t_{delay}	t_{delay}
S_{in_2}	t_{delay}	t_{delay}	t_{delay}
S_{in_3}	t_{delay}	t_{delay}	t_{delay}

These gate delays are obtained from multiple SPICE simulations for every set of S_{in} and C_{out} . Since the lookup table contains only a small set of S_{in} and C_{out} , interpolation or extrapolation are required to evaluate t_{delay} and S_{out} for the desired set of S_{in} and C_{out} .

Being a very simple lookup table, NLDM makes delay calculation very fast at the cost of accuracy. Primary cause of the accuracy loss is due to interpolation, process variation, on-chip variation, and the complex behavior the MOSFET at the smaller technology nodes [15, 20]. Several more advanced modelling schemes have been developed in which the complex behaviour of the cell and the signal waveforms can be captured more accurately. These advanced modelling schemes reduce the relative error in the delay calculation. The *Composite Current Source (CCS)* Model from Synopsys and the *Effective Current Source Model (ECSM)* from Cadence are the most well known industry standard modelling schemes of this type.

2.2.2 Composite Current Source Model

Similar to NLDM, characterization in a CCS model is also carried out with a small set of combinations of input signal slew and output effective capacitive load. However, in

contrast with NLDM, CCS contains the output current waveform (see Figure 2.12) in the lookup table as shown in Table 2.2. Here $I_{out}(t)$ is the output current waveform. It is represented by a set of paired current and time values. In the library file, two vectors are used to store simulation time and corresponding output current value for every pair of S_{in} and C_{out} .

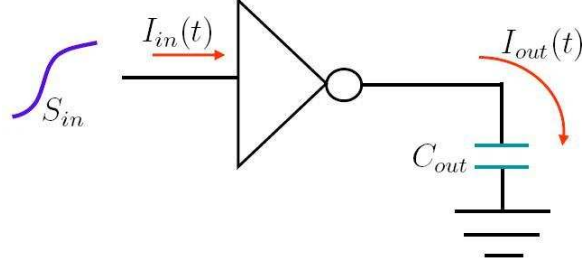


Figure 2.12: Measurement of $I_{out}(t)$ for a set of S_{in} and C_{out}

Table 2.2: CCS based output current waveform ($I_{out}(t)$) lookup table

Input Signal Slew	Effective Output Capacitance		
	C_{out_1}	C_{out_2}	C_{out_3}
S_{in_1}	$I_{out}(t)$	$I_{out}(t)$	$I_{out}(t)$
S_{in_2}	$I_{out}(t)$	$I_{out}(t)$	$I_{out}(t)$
S_{in_3}	$I_{out}(t)$	$I_{out}(t)$	$I_{out}(t)$

In contrast with NLDM, CCS does not define gate delay value for a given set of input signal slew and an effective output capacitive load. Instead of this, CCS has the output current waveform which needs to be processed to extract the delay value. The inaccuracy introduced by the current waveform interpolation in CCS is smaller than the inaccuracy introduced by the delay and slew interpolation in NLDM [15, 21].

2.2.3 Effective Current Source Model

ECSM is another well known industry standard modelling scheme for standard cell delay. In contrast with CCS, it contains the output voltage waveform of the cell ($V_{out}(t)$) (see Figure 2.13) in a lookup table instead of the current waveform as shown in Table 2.3. Similar to CCS, the voltage waveform in ECSM is represented by a set of paired voltage and time values. They are stored in the library file using two vectors, one for the simulation time and other one for corresponding output voltage magnitude.

Table 2.3: ECSM based output voltage waveform ($V_{out}(t)$) lookup table

Input Signal Slew	Effective Output Capacitance		
	C_{out_1}	C_{out_2}	C_{out_3}
S_{in_1}	$V_{out}(t)$	$V_{out}(t)$	$V_{out}(t)$
S_{in_2}	$V_{out}(t)$	$V_{out}(t)$	$V_{out}(t)$
S_{in_3}	$V_{out}(t)$	$V_{out}(t)$	$V_{out}(t)$

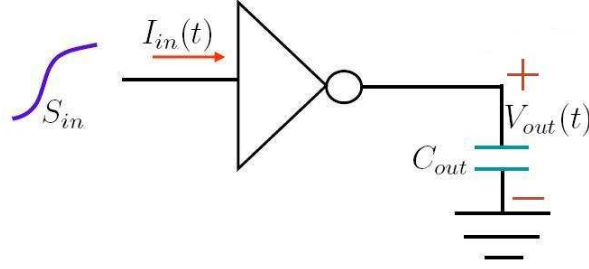


Figure 2.13: Measurement of $V_{out}(t)$ for a set of S_{in} and C_{out}

Similar to the delay calculation in CCS, the output voltage waveform of ECSM needs to be processed to extract the delay value. The inaccuracy introduced by the voltage waveform interpolation in ECSM is smaller than the inaccuracy introduced by the delay and the slew interpolation in NLDM [9, 21].

2.3 Statistical Static Timing Analysis

As discussed in Section 2.2, the delay of a gate depends on the input signal slew and output load of the cell. Along with this, the delay of a gate also depends on the current through the PMOS / NMOS transistors while charging or discharging the load. The current of the MOSFET further depends on its channel length (L), channel width (W), oxide thickness (t_{ox}), doping concentration, supply voltage, temperature, etc. Since the MOSFET is also called a device, the channel length, channel width, oxide thickness and doping concentration are also known as device parameters. There are many more device parameters which are not mentioned here. The value of these device parameters depends on the fabrication process of the integrated circuit. Therefore, the device parameters are also known as process parameters. It can be said that the delay of a gate depends on the *Process, Voltage, and Temperature (PVT)*.

The processing steps during integrated circuit fabrication are mainly mechanical, thermal, chemical and optical. Due to the uncertainty in these steps, the measured values of the process parameters might deviate from the expected values. This phenomenon is known as process variation. The resistive component of interconnect introduces voltage drop in the wire which results in variations in the supply voltage. Heat dissipation is also not uniform on the entire chip due to different switching activities in different regions of the chip. This causes temperature variations on the chip. All these effects result in variation of the PVT values which results in variation in the delay of a gate. Delay variations result in the variation of the maximum possible operating frequency of the circuit. It might also result in setup and hold violations and affect the correctness of the circuit functionality.

Delay estimation in the STA methodology does not consider the variations in the PVT parameters. Instead of this, STA is performed multiple times to estimate the delay in various PVT corners such as best case and worst case corners. In the best case corner, smallest delay is considered which implies fast process (low capacitance, fast transistors), high voltage and low temperature. Similarly, in the worst case corner,

highest delay is considered which implies slow process (high capacitance, slow transistors), low voltage and high temperature. This methodology is well known as corner case analysis. Additionally, the k-factor modelling approach is also used. K-factor modelling is a method to scale the delay of every cell in the library by a fixed factor k. The scaling factor k mostly varies between 0.9 and 1.1. This scaling is to protect the design from unknown PVT variations. This k-factor modelling introduces deviation from the SPICE simulation.

As we are moving towards nanometer technology, process variation is increasing, causing significant uncertainty in the delay estimation [3] which greatly impacts the yield [4, 5]. As a consequence, the accuracy of the conventional *Static Timing Analysis (STA)* with corner to estimate digital circuit performance in advance technology processes is a serious concern [22]. Due to these PVT variations, the delay really is a statistical parameter instead of a deterministic one. The methodology of estimating the timing of a data path with PVT variation is known as *Statistical Static Timing Analysis (SSTA)* [23, 24, 25, 26].

2.4 Circuit Simulation and Analysis Environment

The primary requirement for the implementation of the STA and SSTA methodologies is circuit simulation. Circuit simulation is a process of mathematically estimating the expected behavior of the physical circuit. Being a mathematical analysis tool, it needs a mathematical model of each circuit component to perform the circuit simulation. At earlier nodes of the IC technology development, simple MOSFET models have been used. However, due to device shrinking, various physical effects (like short channel effect, gate leakage etc.) are nowadays playing a significant role in the device behavior. Many detailed models have been developed to represent the complex behavior of the MOSFET. *EKV* (developed by C. C. Enz, F. Krummenacher and E. A. Vittoz, hence the initials EKV) [27] from EPFL (Ecole Polytechnique Federale de Lausanne) and *BSIM* (Berkeley Short-channel IGFET Model) [28] from UCB (University of California, Berkeley) are well known industry standard models. A variant of the BSIM model, known as BSIM4 [29] is predominantly used in state of the art integrated circuit development. The parameters of these models are extracted by characterization of the MOSFET developed by *fabrication plant (fab)*. These parameters are fab and technology dependent.

As technology is shrinking, various behaviors (like process variations) are significantly affecting the performance of the device. Research activities are necessary to understand the behavior of the future transistors in advance technology nodes. The *ITRS* (International Technology Roadmap for Semiconductor) [30] is actively involved in defining the future technology nodes. MOSFET models are also required for future technology nodes for use in research activities. The *PTM* (Predictive Technology Model) [31] is a well known set of technology models for future transistors as specified by ITRS. PTM provides SPICE based predictive BSIM4 models which can be used with simulation tools to analyse circuit behavior. Therefore, with PTM, research work can start well before the real development of the advance semiconductor technologies has completed. The Nangate Open Cell Library [32] is another well know name among the

researchers for the circuit design kit based on the PTM technology. It offers predictive standard cell libraries designed with PTM technologies.

For the 45nm technology node, two types of PTM models are available, namely PTM HP for high performance design and PTM LP for low power design. The 45nm BSIM4 model of the MOSFET based on the PTM HP technology is used in the work related to this thesis. Additionally, standard cell circuit design based on the Nangete open cell library is used for analyzing the delay variations due to PVT variations. Cadence Spectre [33] is used for circuit simulation and MATLAB [34] for data processing. Since the input technology file for Spectre are not quite the same as for SPICE, appropriate modifications have been made in the PTM model such that it can be used with Cadence Spectre.

The activities of this thesis work are under the umbrella of the MODERN Project [35]. One of the objectives of the MODERN project is to develop a SSTA engine and methodology flow to address the problems, which arise due to PVT variations. The MODERN project is targeting to address these problems at the root level. We know that the root of the inaccuracy in the delay estimation is due to the lookup based model with corner case analysis and k-factor approach. The highest accuracy can be achieved by using full SPICE-level simulation of each standard cell for the delay estimation. However, due to the high run time of the SPICE simulation, this is not a practical approach for digital circuits. Therefore, a fast circuit simulator is being developed in the MODERN project. The SSTA engine of the MODERN project will have this fast circuit simulator and delay of each gate will be estimated on the fly. This circuit simulation will increase the accuracy of the delay estimation due to the PVT variations. Furthermore, the accuracy of the delay variation estimation is increased by preserving various possible waveforms due to PVT variations at every input and output pins of the standard cells. These waveform collections are called “sets of waveforms”. Sets of waveforms as a representation of uncertainty in a waveform, and the possible approaches to preserve the waveforms are discussed in Chapter 5.

2.5 Summary

The basics of digital circuit design and its terminologies like Data Path, Delay, Slew, Setup Time, and Hold Time are introduced in this chapter. Following this, the state of the art methodologies for STA are presented. NLDM, CCS and ESCM are the industry practice standard cell models for STA. However, due to the variations in PVT, the standard cell behavior is not deterministic anymore. Therefore, there is a need of SSTA which is also discussed in this chapter. The MOSFET model, technology files, simulation and analysis environment are also introduced.

Delay Variations

As discussed in Chapter 2 that the maximum operating frequency depends on the delay of standard cells, there is a need to analyze the variation in the delay of standard cells due to PVT variations. There are two major objectives of this work. Firstly, understand the variation in the standard cell delay due to the variations in PVT parameters. Secondly, develop an environment to measure the variation of the delay, slew, and other circuit parameters of the standard cells due to the PVT parameters. Here five PVT parameters are considered to be varying; they are channel length (L), channel width (W), threshold voltage (V_{th}), supply voltage (V_{DD}), and temperature (T).

The organization of the chapter is as follows: the variations in the PVT parameters for the 45nm technology node are estimated in Section 3.1, followed by the circuit simulation configuration in Section 3.2. The variations in the delay of an inverter due to the variations in PVT parameters are discussed in Section 3.3. In this section, the delay variation analysis is subdivided into three categories, namely, variations in the delay due to each PVT parameter separately, combinations of two parameters, and a realistic scenario with all the parameters varying together. The delay variations do not follow a Gaussian distribution, therefore there is a need for higher order statistical moments as discussed in Section 3.4. The variations in the delay of various standard cells are discussed in Section 3.5. At the end, a summary of the chapter is presented in Section 3.6.

3.1 Variations in the PVT parameters

Since we are working on predictive technology models, fabrication data is not available to estimate the spread in the process variation. At this development stage, we have made an educated guess by extrapolating the spread of the process variations data from the existing technology nodes. Variation spread of the channel length (L), channel width (W), and threshold voltage (V_{th}) for technology nodes up to 70nm are reported in Table 3.1 [3]. The extrapolated spread data of the process parameters for the 45nm technology node is included in the last row of the same table. Here σ_X is the standard deviation of parameter X and $3\sigma_X$ is a measure of the spread of the parameter X .

The nominal value (μ) and spread (3σ) of the two physical parameters, L & W , are plotted in Figure 3.1a and Figure 3.1b respectively. Here nominal values of different technology nodes are on the X-axis and their spreads are on the Y-axis. These figures show a very strong linear relationship. Therefore, a linear best fit curve is used to extrapolate the spread of L and W to the 45nm technology node. A linear best fit curve and extrapolation point for the 45nm node are also shown in the same figures.

The threshold voltage (V_{th}) depends on various physical parameters, e.g. gate oxide thickness (t_{ox}), doping concentration (N) etc. Since the spread of these physical param-

Table 3.1: Technology process parameter trends based on [3]

L (nm)	$3\sigma_L$ (nm)	W (nm)	$3\sigma_W$ (nm)	V_{th} (mV)	$3\sigma_{V_{th}}$ (mV)
250	80	800	200	500	50
180	60	650	170	450	45
120	45	500	140	400	40
100	40	400	120	350	40
70	33	300	100	300	40
45	26	90	58	469	37

eters are not available for the existing technologies, the trend of the spread of the V_{th} is analysed against the technology node itself. Nominal values of the technology node (μ_L) and the spread of the threshold voltage ($\sigma_{V_{th}}$) are plotted in Figure 3.1c. There are two important observations with the threshold voltage variation. First, V_{th} in the 45nm PTM is relatively high, and second, the absolute value of the V_{th} spread ($3\sigma_{V_{th}}$) is constant in 120nm technology nodes and below. These observations are discussed here.

As technology is shrinking, the channel length and channel width are reducing. However, the gate oxide has reached to a very low thickness, and further reduction in the oxide thickness is not possible. High- k dielectric materials are being used instead of lowering the gate oxide thickness. Here k is the relative dielectric constant of the material filled between gate and channel. The relations between V_{th} , t_{ox} , and k are given below:

$$V_{th} \propto \frac{t_{ox}}{k} \quad (3.1)$$

Based on this theory, the threshold voltage should reduce while reducing each technology nodes. The reduction can be seen from 250nm to 70nm technology nodes (see Table 3.1). Following a similar trend, V_{th} in 45nm should reduce. However, in the 45nm PTM model, the threshold voltage is increased in comparison with previous technology nodes. There is not enough information available about the PTM model to understand this behaviour.

The variation in the threshold voltage also depends on the variation in doping concentration (N). As device dimensions are shrinking, the absolute number of dopant atoms in a device is reducing. In technology nodes above 90nm, the numbers of dopant atoms are still high and a variation in this number should have just a small impact on the threshold voltage. Therefore, the effect of variation in t_{ox} is the dominating factor for V_{th} variation and its spread should decrease with technology shrinking. However, in sub 90nm technology nodes, the numbers of dopant atom are already within several tens and the variation in the number of dopant atoms should significantly influence the threshold voltage variation. It implies that in sub 90nm technology nodes, the spread in V_{th} should increase. Based on the existing technology data, the spread of V_{th} is reducing for higher technology nodes and then it become constant for 120nm technology node and below.

As the dominating physical variation for the variation in V_{th} is different in higher and

lower technology nodes, any extrapolation on this data might lead to significant error in the spread estimation of the gate delay. Experiments (see Section 3.3) show that the contribution of the threshold voltage variation in the delay variation of INV_X1 is only 10% whereas the contribution of L and W are 53% and 63% respectively. Because of the very small impact of the threshold voltage variation in the delay variation, a linear best fit curve is used to extrapolate the spread of the V_{th} in 45nm technology node. The linear best fit curve and extrapolated spread of V_{th} in the 45nm technology node are plotted in Figure 3.1c.

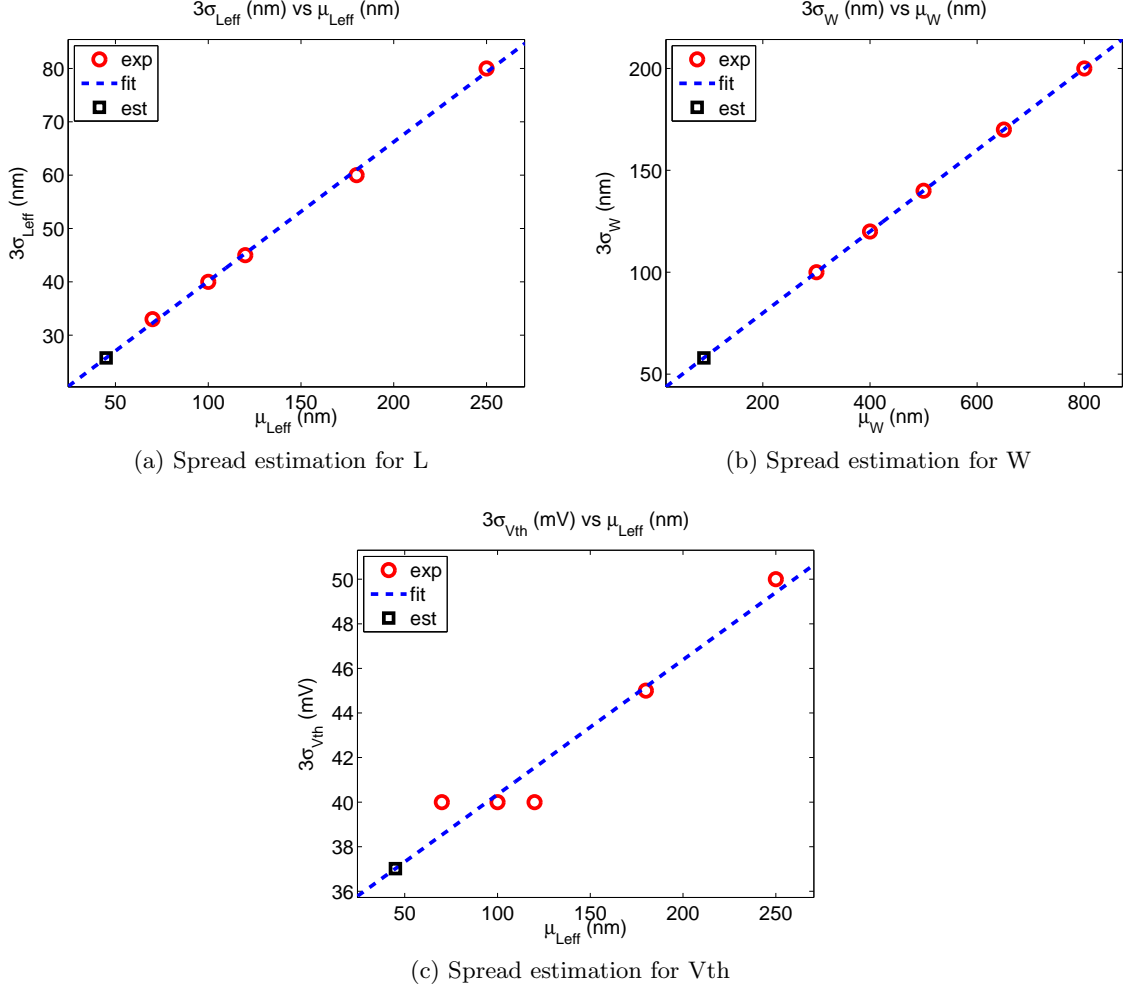


Figure 3.1: Spread estimation for PTM Model

Apart from L , W , and V_{th} , we need the spread for the supply voltage (V_{DD}) and the temperature (T) as well. We have considered a temperature variation from $15^\circ C$ to $75^\circ C$, which can be represented as a nominal value of $45^\circ C$ and $3\sigma_T$ of $30^\circ C$. Since the temperature variation is application dependent, the designer can change the variation during the analysis. For the supply voltage, we can assume that the spread is 15% of its nominal value. In this experimental setup, we have considered a supply voltage of 1.00V and $3\sigma_{V_{DD}}$ equal to 0.15V. The nominal values (μ), their spread (3σ), and the

spread variation in percentage with respect to its nominal value ($3\sigma\%$) for all the five PVT parameters are summarised in Table 3.2.

Table 3.2: Spread of the PVT parameters

Parameter	Nominal (μ)	Spread (3σ)	$3\sigma\%$
L	45nm	26nm	58%
W	90nm	58nm	64%
V_{th}	469mV	37mV	8%
T	45°C	30°C	67%
V_{DD}	1.0V	0.15V	15%

It is important to note that the spread of the channel length L and channel width W are IC fabrication process dependent. Therefore, it can be safely assumed that the spread is independent of the absolute value of the L and W . This implies that the 3σ values of L and W remain constant for various sizes of the standard cells and their values can be taken from the Table 3.2.

The process parameters are either physical parameters (e.g. L , W) or they depend on physical parameters (e.g. V_{th} depends on gate oxide thickness, doping concentration etc.). These physical parameters further depend on fabrication processing steps and their parameters. In general, fabrication process parameters have a non-linear relationship with the physical parameters [22]. It can be assume that the fabrication processing parameters may follow a Gaussian distribution due to central limit theorem [36]. However, due to the non-linear relationship between fabrication parameters and physical parameters, the physical parameters might not follow a Gaussian distribution. Therefore, in general, the process parameters do not follow a Gaussian distribution. However, it is very complicated to measure the exact distribution of each process parameters. In industrial practice, all parameters are usually assumed to follow a Gaussian distribution. Following this industrial practice, for the rest of this chapter, each parameter is considered to follow a Gaussian distribution.

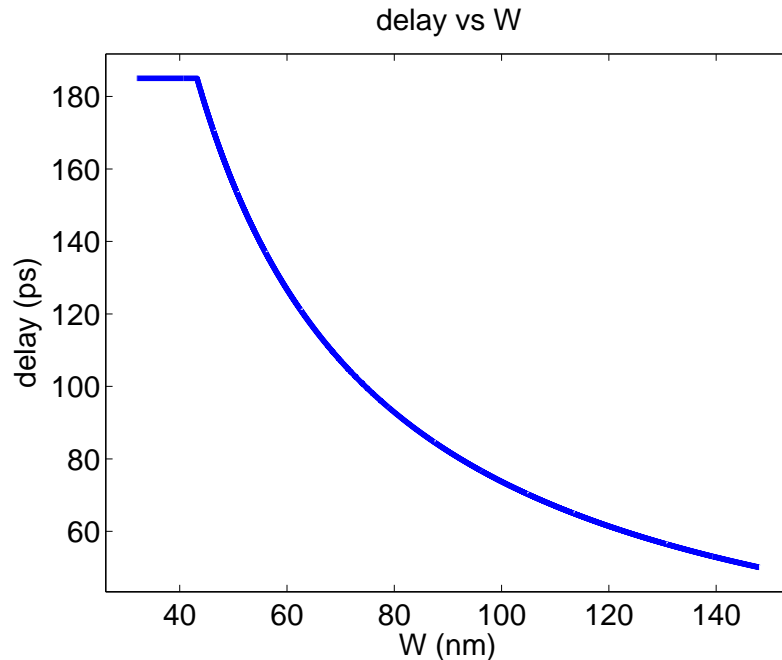
3.2 Circuit Simulation Configuration

Based on the 45nm Nangate open cell library, the nominal size of the channel length (L_n) and channel width (W_n) of an NMOS in the *smallest inverter* (INV_X1) should be 50nm and 90nm respectively. Similarly, the nominal size of the channel length (L_p) and channel width (W_p) of a PMOS in the same inverter should be 50nm and 135nm respectively. The 3σ spread of the L and W for both NMOS and PMOS can be taken from Table 3.2, and the values are 26nm and 58nm respectively. For a Gaussian distribution of any random variable, a spread of $\pm 3\sigma$ around its mean (μ) covers 99.8% probability of the random variable. Due to this, in industrial practice, $(\mu \pm 3\sigma)$ coverage is used for all the process parameters in the SSTA methodology. Therefore, in this analysis, the circuit simulation is carried out in the range of $(\mu_L \pm 3\sigma_L)$ for channel length L and $(\mu_W \pm 3\sigma_W)$ for channel width W . The nominal values, 3σ spread and the required range of L and W of the NMOS and PMOS in INV_X1 are given in Table 3.3.

Table 3.3: Nominal and 3σ range for Nangate INV_X1 inverter

Name	Nominal (μ)	Spread (3σ)	Range ($\mu \pm 3\sigma$)
L_n	50nm	26nm	24nm to 76nm
W_n	90nm	58nm	32nm to 148nm
L_p	50nm	26nm	24nm to 76nm
W_p	135nm	58nm	77nm to 193nm

During simulation of the INV_X1 using the Spectre circuit simulator and the BSIM4 model from PTM, we encountered the following two problems. First, Spectre exits with a fatal error while simulating the inverter with a channel length smaller than 32nm while keeping the rest of the parameters at their nominal values. This fatal error is due to the fact that the *effective channel length* of the given MOSFET in the PTM technology is becoming negative. Since the required range for L has a lowest value of 24nm, the inverter can not be simulated for the entire required range. Second, Spectre can simulate the inverter with the specified range of the channel width, but it has been observed in the simulation experiments that the delay of the inverter does not change for a channel width smaller than 45nm while keeping the rest of the parameters at their nominal values. This might be because the MOSFET model in 45nm PTM technology is not well defined and tested below 45nm of channel width. The variation in the delay of an INV_X1 due to channel width variation is shown in Figure 3.2.

Figure 3.2: Delay of INV_X1 vs W for $\mu_W = 90\text{nm}$ and $3\sigma_W = 58\text{nm}$

These problems demonstrate that the available 45nm PTM model is not very well defined for very small MOSFETs. In this chapter, these problems are addressed by increasing the nominal values of L and W such that the $(\mu \pm 3\sigma)$ ranges of these

parameters can be simulated meaningfully in Spectre. Therefore, the minimum size of L is taken as 60nm and the minimum size of W is taken as 105nm. The nominal values, 3σ spread and the required range of L and W for the proposed MOSFET size are given in Table 3.4. The minimum sizes of L and W reported in this table are used in various other standard cells as well.

Table 3.4: Nominal and 3σ range for modified INV_X1 inverter

Name	Nominal (μ)	Spread (3σ)	Range ($\mu \pm 3\sigma$)
L_n	60nm	26nm	34nm to 86nm
W_n	105nm	58nm	47nm to 163nm
L_p	60nm	26nm	34nm to 86nm
W_p	157.5nm	58nm	99.5nm to 215.5nm

The delay variation analysis is carried out for various standard cells. Based on the number of parameter variations, the analysis is classified into three categories, viz. delay variation due to individual parameter, due to a combination of two parameters, and due to variations of all PVT parameters together. Delay variation in an INV_X1 inverter is discussed in detail first, followed by the delay variation of other standard cells.

To make a consistent circuit environment among all the gates, the input signal slew and output capacitive load of each standard cell is kept the same. Here, slew threshold voltages are 10% and 90% of V_{DD} . Based on the Nangate open cell library of 45nm, an approximated average of the signal slew is 8ps and effective load is 10fF. These signal slew and capacitive load values are used in every standard cell simulation for the delay variation analysis.

3.3 Delay variation in an inverter

The circuit configuration for the INV_X1 inverter is given in Figure 3.3. Due to the very close physical location of L_n and L_p , they are assumed to be highly correlated. Similarly, W_n and W_p are also assumed to be highly correlated. Threshold voltage variation in PMOS and NMOS are also assumed to be highly correlated because of two main reasons. First, at any transition, signal value can change either from low to high value or high to low value. Therefore, either PMOS or NMOS will be active to charge or discharge the effective capacitive load. Since most of the times during transition either NMOS or PMOS is active, keeping high correlation in their threshold voltage does not introduce a significant error. Second, keeping high correlation in the threshold voltages reduces the total number of independent parameters in the circuit simulation. This in turn reduces the complexity and number of simulation iterations required during circuit simulation.

The relation between channel length, channel width, and threshold voltages of both NMOS and PMOS are given below. The relation between W_n and W_p are taken from the Nangate library. The nominal threshold voltage of NMOS (V_{th_n}) is positive whereas the nominal threshold voltage of PMOS (V_{th_p}) is negative. Additionally, V_{th_n} increases

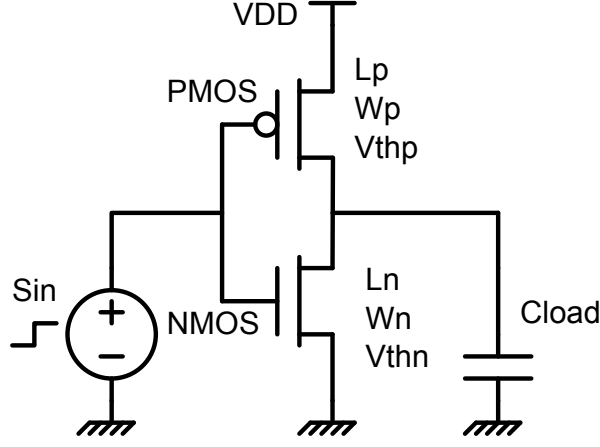


Figure 3.3: INV_X1 simulation configuration

with increase in doping concentration whereas V_{thp} decreases with the doping concentration. Therefore, a negative correlation among them is considered here. All the circuit parameters used in the simulation of the INV_X1 and their ranges are summarised in Table 3.5.

$$L_n = L + \Delta L \quad (3.2)$$

$$W_n = W + \Delta W \quad (3.3)$$

$$L_p = L + \Delta L \quad (3.4)$$

$$W_p = 1.5 \cdot W + \Delta W \quad (3.5)$$

$$\Delta V_{thn} = \Delta V_{th} \quad (3.6)$$

$$\Delta V_{thp} = -\Delta V_{th} \quad (3.7)$$

Table 3.5: INV_X1 simulation configuration

Name	Value	Range
L	60nm	± 26 nm
W	105nm	± 58 nm
V_{DD}	1V	± 0.15 V
ΔV_{th}	0	± 37 mV
T	45°C	± 30 °C
S_{in}	8ps	N.A.
C_{load}	10fF	N.A.

The delay variation of this inverter due to individual parameters, combinations of two parameters, and due to all the parameters together is discussed in the following subsections.

3.3.1 Individual Parameter Variation

In this setup, it has been assumed that only one parameter is varying and rest of the parameters are at their respective nominal values. The purpose of this experiment is to understand the impact in the delay variation due to the individual PVT parameters separately. A plot of the delay of an inverter INV_X1 as a function of the channel length (L) is shown in Figure 3.4a. Here, channel length is on the X-axis and delay is on the Y-axis. This figure shows that the delay is not a linear function of L in the interested range of variation. Due to this non-linear relation, the delay variation will not follow a Gaussian distribution, even if the channel length follows a Gaussian distribution.

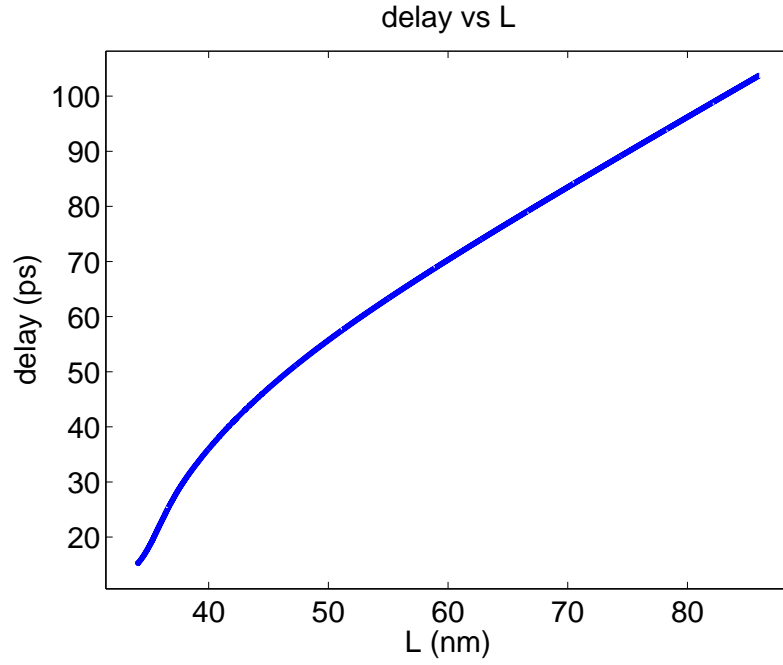
The probability density function (pdf) of delay variation and its Gaussian approximation are shown in Figure 3.4b. Here delay is on the X-axis and probability is on the Y-axis. The circle markers are the pdf of the delay from the simulation experiment and the solid line is the approximated Gaussian distribution. The mean (μ) and standard deviation (σ) of the delay are also reported in the same figure. The approximated Gaussian distribution is generated while keeping the same value of μ and σ of the delay distribution.

The plot of the pdf of the delay shows that the delay variation does not follow a Gaussian distribution exactly. Similarly, the pdf of the delay due to the variations in W and V_{DD} also does not follow a Gaussian distribution. However, the pdf of the delay due to the variations in V_{th} and T are very close to a Gaussian distribution. Since there is nothing new to learn from these other plots, the delay function and pdf plots due to each parameter variation are added in Appendix A. The non-linear relation of the delay with respect to each parameter can be seen in these plots. Additionally, the pdf of the delay variation does not follow their approximated Gaussian distribution. The methodology to estimate μ , σ and the pdf of the simulation output (e.g. delay of an standard cell) is discussed in Chapter 4. The mean (μ), spread (3σ) and spread percentage with respect to its mean value ($3\sigma\%$) of the delay due to each parameter variation are reported in Table 3.6.

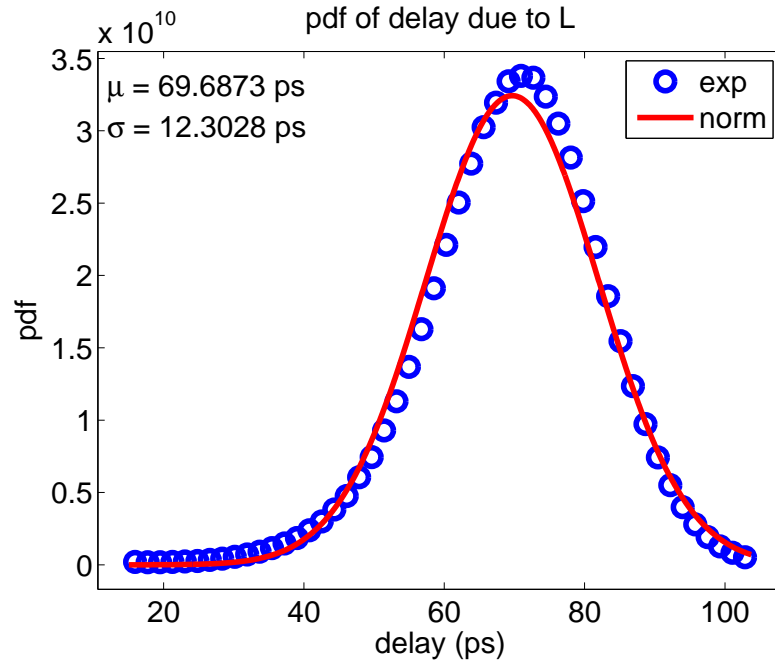
Table 3.6: Delay spread of INV_X1 due to individual parameter variation

Parameter	Mean (μ) (ps)	Spread (3σ) (ps)	$3\sigma\%$
L	69.69	36.90	52.95
W	73.16	45.81	62.62
V_{th}	70.36	7.14	10.15
T	70.41	12.69	18.02
V_{DD}	70.71	12.48	17.65

It is important to note that the mean of the delay (μ) due to the variation in channel width (W) is relatively high when compared to the delay mean due to other parameter variations. The reason for this behaviour is as following. In the specified range of the 3σ variations in the PVT parameters, the delay variation is highly non-linear with respect to W as compared to other parameters. The variations of the delay as a function of each individual parameter are given in Appendix A. This highly non-linear relation is due to the fact that the drain-source current (I_{ds}) in a MOSFET is directly proportional



(a) Delay as a function of L



(b) Delay distribution pdf due to L

Figure 3.4: Delay variation due to L

to W and the delay is inversely proportional to I_{ds} . Therefore, the delay is inversely proportional to W . Due to this inverse relation between delay and W , the mean of the delay is shifted from the nominal delay value. It can also be observed that W is the primary cause for delay not following a Gaussian distribution.

Based on this experiment, we can say that the channel length (L) and channel width (W) are having the highest impact on the delay. This implies that the SSTA simulator should have higher accuracy for the modelling of L and W than that of V_{DD} and V_{th} . This information is very useful for the development of the SSTA engine.

3.3.2 Combination of two parameters variations

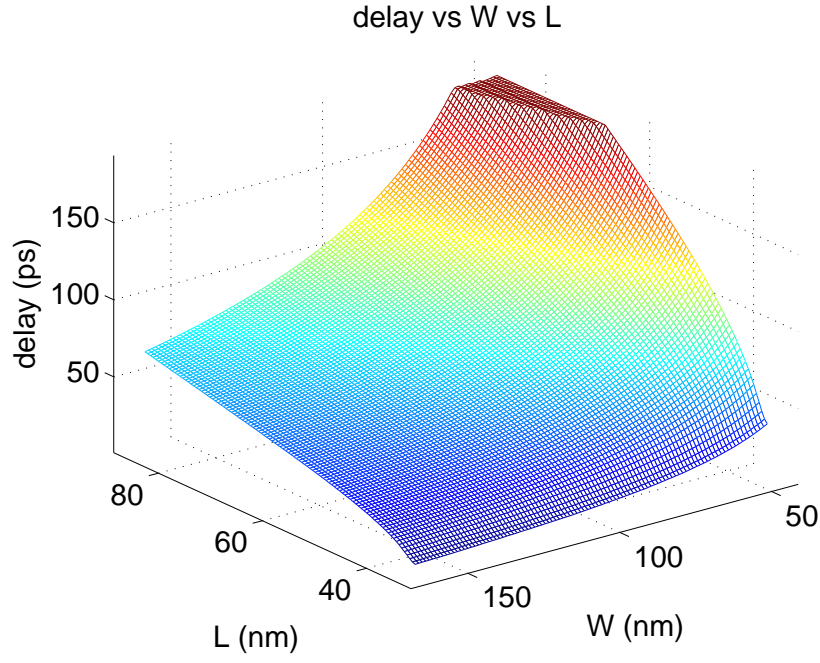
After understanding the effect of the individual parameter variations on the delay of an inverter, pairs of two parameters variations have been used to understand their joint influence on the delay variation. Due to unavailability of the covariance information, it has been assume that the parameters are independent.

In this work, five PVT parameters (L , W , V_{th} , T , V_{DD}) are considered to have variation which results in ten paired combinations. All these possible pairs have been used to analyze their effect on the delay variation. However, only the variation in the delay due to W and L are discusses in detail. Since the delay variation due to other parameter pairs also shows similar behaviour, they have not been discussed in detail. However, a summary of all other combinations of variables is discussed here. The plots of the delay variation due to other pair of parameters can be found in Appendix A.

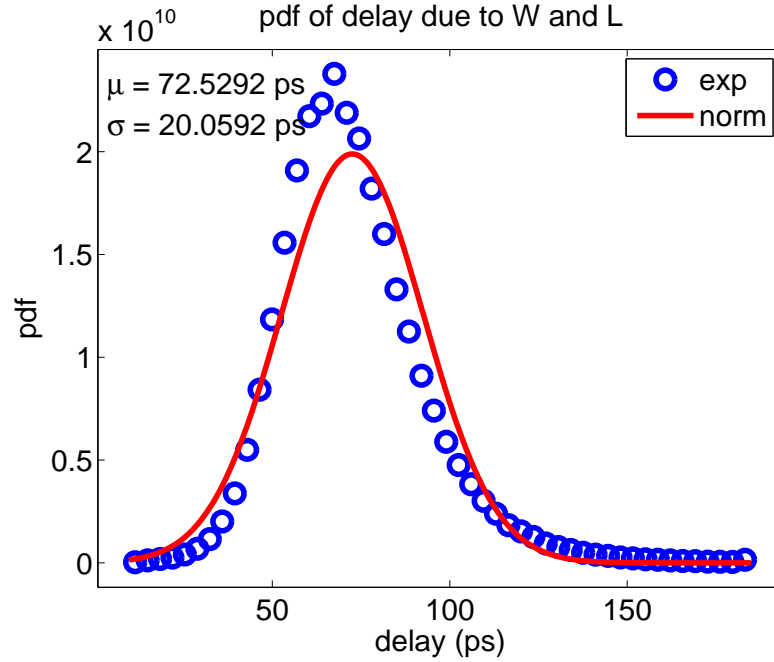
The delay as a function of L and W is shown in Figure 3.5a. Here W is on the X-axis, L is on the Y-axis and the delay is on the Z-axis. There are two main observations in the plot. First, the delay values are clipped for high L and low W . This indicates that the 45nm PTM model is not well defined in this region. However, due to very low probability of this segment, the inaccurate behaviour of the PTM model has been ignored. The second observation is about the linearity of the delay variation due to variation in L and W . Since the plot is not a linear plane, the delay distribution should be deviating from a Gaussian distribution, even if the input parameters (L and W) are following a Gaussian distribution.

The pdf of delay variation and its Gaussian approximation are shown in Figure 3.5b. Similar to the pdf plot of the delay due to one parameter variation, in this figure the delay is on the X-axis and probability is on the Y-axis. The circle markers are the pdf of the delay from the simulation experiment and the solid line is the approximated Gaussian distribution. The mean (μ) and standard deviation (σ) of the delay are also reported in the same figure. Again, the approximated Gaussian distribution is generated while keeping the same value of μ and σ of the delay distribution. The mean (μ), spread (3σ) and spread percentage with respect to its mean value of the delay due to each pair of parameter combinations are reported in Table 3.7a, Table 3.7b, and Table 3.7c respectively.

There are two important observations in the result of this simulation experiment. First, in each pair where W is one of the parameters, the mean of the delay is higher than the mean of the delay due to other pairs. This is due to the fact that delay is inversely proportional to W in the interested range of variation. A similar behaviour has been observed in individual parameter variations. The second observation is in the pdf curve of the delay. It can be seen that there are a few outlier points in the pdf curve; therefore the curve is not very smooth. This is because limited simulation iterations have been used in these experiments. The smoothness of the curve can be improved by increasing the number of simulation iterations. However, it is important



(a) Delay as a function of W & L



(b) Delay distribution pdf due to W & L

Figure 3.5: Delay variation due to W & L

to note that the pdf curve is mainly used to visualize the distribution of the delay, and their statistical moments (μ and σ) are used in SSTA methodology. The proposed fast statistical moment estimation method has been used to estimate these statistical moments and this methodology is presented in Chapter 4. Furthermore, it has been

Mean (μ) (ps)	L	W	V _{th}	T
W	72.53			
V _{th}	69.75	73.24		
T	69.79	73.28	70.47	
V _{DD}	70.10	73.60	70.79	70.83

(a) Mean (μ) of delay variation due to two parameter variations

Spread (3σ) (ps)	L	W	V _{th}	T
W	60.18			
V _{th}	37.77	46.62		
T	39.48	48.00	14.61	
V _{DD}	39.45	48.12	14.55	17.91

(b) Spread (3σ) of delay variation due to two parameter variations

3 σ %	L	W	V _{th}	T
W	82.97			
V _{th}	54.15	63.65		
T	56.57	65.50	20.73	
V _{DD}	56.28	65.38	20.55	25.29

(c) Spread % (3σ %) of delay variation due to two parameter variations

Table 3.7: Delay variation due to two parameter variations

shown that only a very small number of iterations is needed to estimate the statistical moments in the proposed methodology. Having more simulation iterations does not change the value of the statistical moments of the delay variation. Therefore, the non smoothness of the pdf curve does not affect the accuracy of the estimation of the statistical moments.

The results of one parameter variation and two parameter variations are consistent, i.e. the variation in the L and W are the primary contributors of the delay variation. The pdf plots show that the distribution of the delay with two parameters varying is also not following a Gaussian distribution.

3.3.3 A realistic PVT variation

After analyzing the effect of the variation in individual parameters and pairs of two parameters on the delay of an inverter, an experiment is carried out with all five parameters variations together. This scenario is used to produce delay variations closely matching reality in real silicon. Since the delay data in this experiment has six dimensions, the delay as a function of the PVT parameters can not be plotted.

The pdf of the delay variation and its Gaussian approximation is plotted in Figure 3.6. Similar to the previous pdf plots, the delay is on the X-axis and probability is on the Y-axis. The circle markers are the pdf of the delay from the simulation experiment and the solid line is the approximated Gaussian distribution. The mean (μ) and standard deviation (σ) of delay are also reported in the same figure. The approximated Gaussian distribution is using the same value of μ and σ of the delay distribution. The

mean (μ), spread (3σ) and spread percentage with respect to its mean value of the delay are reported in Table 3.8.

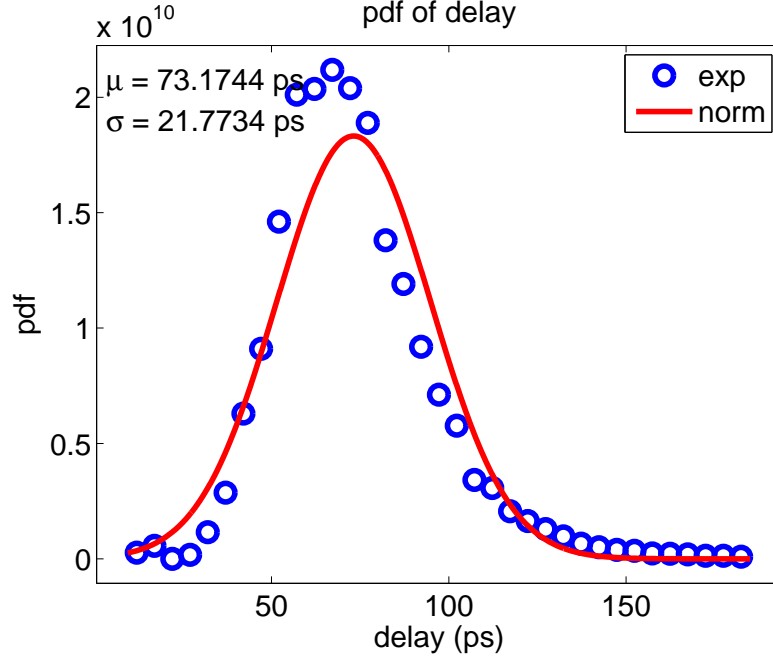


Figure 3.6: Delay distribution pdf due to realistic PVT variations

Table 3.8: Delay variation due to realistic PVT variation

	Value
Mean (μ) (ps)	73.17
Spread (3σ) (ps)	65.31
$3\sigma\%$	89.26

The conclusion of this experiment is that the delay distribution in a realistic PVT variation scenario does not follow a Gaussian distribution. Therefore, there is a need to quantitatively measure the deviation of the delay distribution from its corresponding Gaussian distribution. The quantitative measure of the distribution deviation is discussed in the next section.

3.4 Higher order statistical moments

The pdf of delay in a realistic scenario of PVT variations is significantly deviated from its corresponding Gaussian distribution. This indicates that imposing a Gaussian distribution on the delay variation during SSTA is introducing an error in the delay variation estimation. The deviation of the pdf of delay from its corresponding Gaussian distribution can be captured with higher order statistical moments. Therefore, it is important to include higher order statistical moments in SSTA for accurate analysis.

The first four statistical moments have been used in this thesis work for delay analysis. These moments are mean (μ), standard deviation (σ), skewness (γ), and kurtosis (κ) [37]. They are also known as first moment, second moment, third moment, and fourth moment. We have seen earlier in this chapter that the mean measures the location of the distribution and standard deviation measures its spread. It is important to mention that the skewness measures the symmetry of the distribution and kurtosis measures the flatness or peakedness of the distribution.

A Gaussian distribution is perfectly symmetric, therefore its skewness is always zero. Furthermore, the kurtosis of a Gaussian distribution is always exactly 3. Therefore, a Gaussian distribution does not need third and fourth moments, and it is completely defined with the first two moments. Since the third and fourth moments of a Gaussian are constant, the third and fourth moments of any distribution can be used to quantify their deviation from the corresponding Gaussian distribution.

Let us take an example of the delay variation in an inverter (INV_X1) with the realistic scenario of the PVT variations to understand the quantitative measure of the non-Gaussian distribution. As seen in Figure 3.6, the delay distribution shows a longer tail in the direction of the higher delay values. Therefore, the skewness (γ) of the delay distribution should be positive. Furthermore, it can be seen that the peak of the delay distribution is higher than its corresponding Gaussian distribution. Therefore, the kurtosis (γ) of the delay distribution should be more than three. The values of the first four moments of the delay are reported in Table 3.9. The skewness of the distribution is 0.97 and the kurtosis is 5.02. These values match with above theoretical analysis and confirm that the delay distribution is significantly deviated from its corresponding Gaussian distribution.

Table 3.9: First four statistical moments of delay variation in INV_X1 due to realistic PVT variation

Statistical Moments	Value
Mean (μ) (ps)	73.17
Standard Deviation (σ) (ps)	21.77
Skewness (γ)	0.97
Kurtosis (κ)	5.02

3.5 Delay variation in various standard cells

Up until this point, the delay variation in an inverter (INV_X1) due to various PVT parameter variations has been discussed. In continuation of the delay variation analysis experiment, various other standard cells have been simulated. Since standard cells might have more than one input signals, multiple input signal combinations are possible. A limited set of the input signal combinations have been used in this experiment. Simulation run time is the main driving force to limit the experiments for a smaller set of input signal combinations. This set of signal combinations are based on the following rules:

Rule I. Only rising input transitions.

Rule II. Transition on each pin while keeping other pins at constant signal value.

Rule III. Multi-input switching with simultaneous signal transition on all pins.

The transistor gate sizes for each of the standard cells have been taken from the 45nm PTM based Open Cell Library developed by Nangate Inc. However, as discussed in Section 3.2, the PTM model is not very well developed for very low device sizing. Therefore, a similar compensation has been used to adjust the value of channel length and channel width in each standard cell. In this compensation process, channel length is scaled by a factor of 60nm/50nm and channel width is scaled by a factor of 105nm/90nm.

In the open cell library, each gate is available with various drive strengths. All possible drive strengths have been used in this experiment. However, these experiments are limited to only one and two input gates. This is because more input pins results into more signal transition combinations and high run time. Limiting the experiment to only one and two input gates is mainly due to run time concerns.

A list of standard cells used in this experiment, their symbolic name, drive strengths and input signal transitions are summarised in Table 3.10. Here, X1, X2, X4, ... are indicating the drive strength of the gate. X1 is the smallest gate, X2 has double driving capability as compared to X1 and so on. A, A1, and A2 are the input pin names of the gate. The gate name, pin name for input signal rising transition, the first four statistical moments of the delay distribution, and spread percentage are reported in Table 3.11 and Table 3.12. Because of the very large table size, the complete data is split into two tables.

Table 3.10: List of standard cells

Standard Cell	Symbolic Name	Driving Strength	Signal Transition
Inverter	INV	X1, X2, X4, X8, X16, X32	A
Buffer	BUF	X1, X2, X4, X8, X16, X32	A
2 Input NAND	NAND2	X1, X2, X4	A1, A2, A1+A2
2 Input NOR	NOR2	X1, X2, X4	A1, A2, A1+A2
2 Input AND	AND2	X1, X2, X4	A1, A2, A1+A2
2 Input OR	OR2	X1, X2, X4	A1, A2, A1+A2
2 Input XOR	XOR2	X1, X2	A1, A2
2 Input XNOR	XNOR2	X1, X2	A1, A2

The following observations have been made from this experiment:

Observation I. The mean of the delay of each standard cell is decreasing with increasing drive strength. However, the delay of buffer is increasing. A similar trend is observed in the NLDM delay model in the standard cell library file of the 45nm PTM model in the open cell library. It is because of increased loading of the first inverter when upsizing the second one to drive high loads

Table 3.11: Delay variation in Nangate standard cells (INV, NAND2, and NOR2) due to realistic PVT variation

Cell	Transition	μ (ps)	σ (ps)	γ	κ	$3\sigma\%$
INV_X1	A	73.19	21.77	0.97	5.02	89.25
INV_X2	A	36.82	8.11	0.23	3.87	66.09
INV_X4	A	20.66	4.18	0.03	3.79	60.67
INV_X8	A	12.94	2.56	0.00	3.71	59.46
INV_X16	A	9.00	1.75	-0.03	3.60	58.25
INV_X32	A	6.95	1.33	-0.08	3.64	57.34
NAND2_X1	A1	91.01	22.97	0.52	3.66	75.71
NAND2_X1	A2	93.09	23.34	0.49	3.58	75.23
NAND2_X1	A1+A2	94.19	23.40	0.51	3.51	74.53
NAND2_X2	A1	47.67	10.56	0.30	3.67	66.45
NAND2_X2	A2	49.74	11.04	0.29	3.66	66.56
NAND2_X2	A1+A2	50.81	11.16	0.34	3.60	65.91
NAND2_X4	A1	27.55	6.02	0.27	3.63	65.55
NAND2_X4	A2	29.63	6.51	0.24	3.64	65.89
NAND2_X4	A1+A2	30.55	6.64	0.32	3.60	65.25
NOR2_X1	A1	75.00	22.31	0.90	4.91	89.22
NOR2_X1	A2	78.30	23.16	0.85	4.66	88.72
NOR2_X1	A1+A2	39.34	11.58	1.07	6.46	88.33
NOR2_X2	A1	38.55	8.59	0.16	4.11	66.82
NOR2_X2	A2	41.56	9.37	0.17	4.07	67.63
NOR2_X2	A1+A2	21.05	4.57	0.06	4.27	65.06
NOR2_X4	A1	22.23	4.58	-0.07	4.21	61.79
NOR2_X4	A2	25.01	5.26	-0.02	4.11	63.08
NOR2_X4	A1+A2	12.89	2.58	-0.28	4.86	60.14

Observation II. The spread percentage of the delay in NAND, NOR, INV, and XNOR is decreasing with the increase in the gate driving strength. This is because the minimum MOSFET size is increasing while increasing the driving strength of the gate.

Observation III. In contrast to the previous observation, the spread percentage of the delay in AND, OR, and XOR is not varying with the increase in the gate driving strength. This is because these cells are internally composed of two stages. The first stage is an inverting NAND, NOR or XNOR, followed by an inverter. The output signal is therefore buffered by a high drive strength inverters while minimum size MOSFETs with high variability are used in the first stage.

Observation IV. The third and fourth statistical moments indicate that the delay distribution is often significantly non-Gaussian.

Table 3.12: Delay variation in Nangate standard cells (BUF, AND2, OR2, XOR2 and XNOR2) due to realistic PVT variation

Cell	Transition	μ (ps)	σ (ps)	γ	κ	$3\sigma\%$
BUF_X1	A	103.28	32.32	0.13	2.73	93.90
BUF_X2	A	62.44	19.39	0.37	3.61	93.16
BUF_X4	A	49.85	15.44	0.42	3.76	92.93
BUF_X8	A	55.08	17.65	0.55	4.11	96.12
BUF_X16	A	76.62	25.14	0.57	3.82	98.43
BUF_X32	A	119.62	33.34	-0.18	2.59	83.62
AND2_X1	A1	108.20	32.93	0.06	2.62	91.31
AND2_X1	A2	110.10	32.95	0.00	2.59	89.79
AND2_X1	A1+A2	110.33	32.86	0.02	2.55	89.35
AND2_X2	A1	68.81	21.11	0.36	3.53	92.04
AND2_X2	A2	70.97	21.62	0.35	3.51	91.39
AND2_X2	A1+A2	71.45	21.71	0.38	3.51	91.16
AND2_X4	A1	58.07	17.66	0.36	3.58	91.25
AND2_X4	A2	60.20	18.16	0.35	3.58	90.52
AND2_X4	A1+A2	60.90	18.31	0.38	3.57	90.17
OR2_X1	A1	105.33	32.74	0.08	2.75	93.25
OR2_X1	A2	109.10	33.08	-0.01	2.69	90.97
OR2_X1	A1+A2	101.24	32.23	0.16	2.86	95.50
OR2_X2	A1	64.70	20.09	0.35	3.67	93.15
OR2_X2	A2	68.75	21.21	0.35	3.67	92.53
OR2_X2	A1+A2	57.76	17.97	0.33	3.68	93.32
OR2_X4	A1	52.16	16.13	0.39	3.85	92.76
OR2_X4	A2	56.04	17.23	0.40	3.88	92.25
OR2_X4	A1+A2	40.62	12.40	0.31	3.75	91.59
XOR2_X1	A1	139.62	30.72	-0.96	4.02	66.00
XOR2_X1	A2	140.76	30.22	-1.01	4.21	64.40
XOR2_X2	A1	97.21	29.34	0.20	3.09	90.54
XOR2_X2	A2	99.01	29.53	0.18	3.06	89.47
XNOR2_X1	A1	93.34	23.44	0.51	3.55	75.34
XNOR2_X1	A2	97.34	24.32	0.48	3.39	74.96
XNOR2_X2	A1	49.87	11.04	0.31	3.62	66.39
XNOR2_X2	A2	53.71	12.01	0.33	3.57	67.10

3.6 Summary

The 3σ spread of the process parameters is estimated using an educated guess by extrapolating the spread of the process variation data from the existing technology nodes. It has been observed that the desired ($\mu \pm 3\sigma$) range of L and W in the 45nm PTM technology model can not be simulated. This is mainly because the available 45nm PTM model is not well defined for very small MOSFETs. Therefore, the nominal values of L and W are increased, such that the standard cells can be simulated in the

desired $(\mu \pm 3\sigma)$ range.

The experiments show that L and W are having the highest impact on the delay variation. Additionally, it has also been observed that the delay variation does not follow a Gaussian distribution. Similar trends have been observed for combinations of two parameters varying together. In a realistic PVT variation scenario, when all parameter vary together, the delay variation is significantly deviated from a Gaussian distribution. Therefore, there is a need for higher order statistical moments. The third and fourth statistical moments have been used to measure the deviation of the pdf of the delay from a Gaussian distribution.

The delay variation in various standard cells has been estimated. In these experiments, the transistor gate sizes and circuit schematics of each gate of the standard cell have been taken from the 45nm PTM based Open Cell Library developed by Nangate Inc. The delay variation of these gates are estimated for single input switching as well as for multi input switching in various driving strength of the gates. The third and fourth statistical moments indicate that the delay variation is deviated from a Gaussian distribution. The gate delay is highly spread due to the PVT variations as indicated by the 3σ of the delay variation which is between 57% and 100%.

Statistical Moment Estimation and Probability Density Function

4

Monte Carlo methods and simulation are often used to estimate the mean, variance, and higher order statistical moments of signal properties like delay and slew. The main issues with Monte Carlo methods are the required long run time and the need for prior detailed knowledge of the distribution of the variations. Additionally, most of the available circuit simulation tools can run Monte Carlo analysis for Gaussian, lognormal and uniform distributions only. In this chapter, in order to estimate the statistical moments, we propose a new method based on a uniform sampling technique using a weighted sample estimator. The proposed method needs significantly fewer simulation runs, and does not need detailed prior knowledge of the variation distributions. Furthermore, it can be used for any type of probability distribution irrespective of the circuit simulation tool used for the analysis. The results obtained show that the proposed method needs at least $100\times$ fewer simulations iterations than Monte Carlo runs for accurate moments estimation of the delay for standard cells in 45nm and 32nm technologies.

The organization of the chapter is as follows: the motivation to develop a fast statistical moment estimation method is presented in Section 4.1. Following this, the proposed method for fast statistical moment estimation is described in Section 4.2 [6, 7]. The simulation results and a comparison of the proposed method with standard Monte Carlo method are discussed in Section 4.3. Furthermore, the proposed method to estimate the probability density function from the simulation data is also described in Section 4.4 and the probability density functions of the simulation results are discussed in Section 4.5. At the end, a summary of the chapter is presented in Section 4.6.

4.1 Motivation

In SSTA, the standard cell delay and signal slew are stochastic parameters, and their properties are often specified with their statistical moments. Practically, *Monte Carlo (MC)* is the dominant method of choice for statistical moment estimation of these parameters [38, 39]. However, standard Monte Carlo has the following two limitations.

First, due to the underlying principle of MC analysis, a large number (thousands to tens of thousands) of simulation iterations are required for moment estimation with a high confidence bound. Due to the large number of cells in standard cell libraries and long simulation times for advanced transistor models, the necessity of thousands of simulation iterations results into very long circuit simulation run times. Practically, the long run time required for SSTA library characterization, limits its usability for large scale circuits.

Second, due to the nature of semiconductor manufacturing processes and circuit

behaviour, the PVT parameters typically do not follow a Gaussian distribution [22]. Furthermore, their non-linear relationship with delay and slew may result into non-Gaussian distribution of the delay and slew. However, the state of the art circuit simulation tools (e.g. Cadence Spectre [33]) can only run MC with Gaussian, lognormal and uniform distributions, and, unfortunately, forcing any non-Gaussian PVT into these distributions can lead to large errors. To deal with these issues, several non-Gaussian SSTA methodologies have been proposed [40]. These methodologies require higher order moments for accurate modelling of the variations. Additionally, the higher order moments further increase the simulation iterations required in MC iterations.

Several research efforts have been made to speed up the standard Monte Carlo method by improving the random sampling method of the parameters, e.g. *Latin Hypercube Sampling (LHS)* [41], *Quasi Monte Carlo (QMC)* [42], and *Stratification + Hybrid QMC (SH-QMC)* [43]. However, the parameter sampling in the circuit simulations in these methods are still dependent on their distributions and are not applicable for various types of probability density function.

In this chapter, we propose a *Fast Statistical Moment Estimation (FSME)* method, which provides two major advantages over standard MC: first, the FSME method can use any probability density function (*pdf*) irrespective of the simulation tools, and second, for the same accuracy as MC, the FSME method requires two orders of magnitude fewer simulation iterations which results into 100× speedups in the library characterization. The application of the FSME method is not only limited to digital circuit design and SSTA; it is equally applicable in analog circuit design.

4.2 Fast Statistical Moment Estimation Method

The standard MC method is based on random sampling of the parameters of interest based on their *pdf*. This procedure takes more samples around the parameter values with higher probability than around the less probable values. Since the sampling method depends on the *pdf* of the parameters, a large number (thousands) of samples are normally required to generate enough samples for less probable values. Additionally, the dependence on the *pdf* of the corresponding process parameter makes it necessary to provide the statistical details of the parameter variation before the start of the simulation. The circuit simulation is repeated for each set of sampled parameters. This results into long run times and high memory requirement to store all the data. The desired simulation output is measured in each simulation, leading to the sample set of measured values. The moments of the circuit simulation outputs are calculated using standard moment estimation equations on the sample set.

In contrast, by using the FSME method, the probability distribution of the process parameters and the circuit simulation are decoupled. In the proposed method, instead of randomly sampling, the space is sampled with a uniform distribution. Moreover, to accurately estimate the statistical moments, a weighted sample estimator is used. The processes involved in circuit simulation and data processing are discussed in detail below.

4.2.1 Circuit Simulation

Unlike a MC method, the FSME method runs the simulation with a uniformly spaced parameter sweep, which ensures the required coverage of each simulation parameter, e.g. if a parameter is following a Gaussian distribution then $\pm 3\sigma$ spread around its mean value is sufficient. This implies that the range of the parameter sweep in the simulation needs to be close to the spread of the real parameter distribution. Note that this is the only link required between the real parameter distribution and the data needed to perform simulations.

Let us assume that X are the process parameters (e.g. effective channel length L , channel width W , threshold voltage V_{th} , etc.), where X is a set of vectors X_i , with each vector X_i corresponding to the sampled points $X_i[j_i]$ of the i^{th} parameters, and that Y is a vector of the simulation output $Y[k]$ (e.g. delay, slew, etc.). X and Y will be used in the data processing step to estimate the statistical moments of the output.

4.2.2 Data Processing

The statistical moments of the circuit simulation output (Y) depend on the probability of each simulation run, which in turn depends on the probability of each process parameter (X_i) used in the simulation. As a result, the *pdf* of each process parameter is required in the data processing step. The probability of each simulation is estimated first, followed by the moment estimation of the output.

In probability space of the proposed method, each simulation is considered as a discrete random event and the probability of each simulation event is equal to the joint probability of the process parameters. Additionally, the simulation is carried out only for sampled values of the process parameter X_i , and each sampled value of the process parameter ($X_i[j_i]$) is associated with a certain probability. In the data processing stage, the probability of each discrete process parameter $X_i[j_i]$ is estimated from the given *pdf* of X_i . Following this, the probability of each discrete experiment event k is estimated from the probability of the discrete process parameter values. Thereafter, the statistical moments of the circuit simulation outputs (Y) are estimated from the probability of each discrete experiment event.

4.2.2.1 Probability of Process Parameter

The following notation for the probability and the *pdf* function will be further used in the chapter

- $P_d() \rightarrow$ Probability of discrete variable
- $P_c() \rightarrow$ Probability of continuous variable
- $P_s() \rightarrow$ Probability of discrete simulation event
- $f_i() \rightarrow$ Probability density function of X_i

For a stochastic process parameter X_i , its probability and its *pdf* are related with

$$P_c(X_i[m] < X_i \leq X_i[n]) = \int_{X_i[m]}^{X_i[n]} f_i(x) dx \quad (4.1)$$

Let us assume that $X_i[\]$ is a vector of a uniformly sampled process parameter X_i with a sampling step of ΔX_i , under the constraint that ΔX_i is much smaller than the standard deviation σ_{X_i} . $X_i[j_i]$ are the sampled values of the X_i , which are used in the circuit simulation with uniform sampling. The vector $X_i[j_i]$ can be written as:

$$X_i[j_i] = [\cdots, -2\Delta X_i, -\Delta X_i, 0, \Delta X_i, 2\Delta X_i, \cdots] \quad (4.2)$$

where

$$\Delta X_i \ll \sigma_{X_i} \quad (4.3)$$

Let us define the probability of a discrete variable $X_i[j_i]$ to be equal to the probability of a continuous variable X_i varying from $(X_i[j_i] - \Delta X_i/2)$ to $(X_i[j_i] + \Delta X_i/2)$

$$P_d(X_i[j_i]) = P_c\left(X_i[j_i] - \frac{\Delta X_i}{2} < X_i \leq X_i[j_i] + \frac{\Delta X_i}{2}\right) \quad (4.4)$$

$$= \int_{X_i[j_i] - \Delta X_i/2}^{X_i[j_i] + \Delta X_i/2} f_i(x) dx \quad (4.5)$$

A numerical integration method is required to estimate $P_d(X_i[j_i])$. Various interpolation functions can be used in the numerical integration. Two well known fast integration methods are based on the rectangle rule and the trapezoidal rule [44]. In the rectangle rule, a constant interpolation function (a polynomial of degree zero, piecewise constant approximation) is assumed which passes through the mid point of the integral bound (see Figure 4.1a)

$$\int_a^b f(x) dx \approx (b - a) f\left(\frac{a + b}{2}\right) \quad (4.6)$$

Whereas, in the trapezoidal rule, a linear interpolation function (a polynomial of degree one, piecewise linear approximation) is assumed which passes through the end points of the integral bound (see Figure 4.1b)

$$\int_a^b f(x) dx \approx (b - a) \frac{f(a) + f(b)}{2} \quad (4.7)$$

Since ΔX_i is much smaller than σ_{X_i} , the numerical integration method based on rectangle rule with *piecewise constant (PWC)* approximation is simple enough to evaluate the integration of the *pdf* while keeping minimum computation cost

$$P_d(X_i[j_i]) = \int_{X_i[j_i] - \Delta X_i/2}^{X_i[j_i] + \Delta X_i/2} f_i(x) dx \quad (4.8)$$

$$\approx f_i(X_i[j_i]) \cdot \Delta X_i \quad \text{PWC approx.} \quad (4.9)$$

$$\Rightarrow P_d(X_i[j_i]) = f_i(X_i[j_i]) \cdot \Delta X_i \quad \text{if } \Delta X_i \ll \sigma_{X_i} \quad (4.10)$$

Thus, the probability of the discrete process parameter $X_i[j_i]$ is equal to the integral of the *pdf* around $X_i[j_i]$ within the bound of $\pm \Delta X_i/2$, and piecewise constant approximation can be used to simplify the integration.

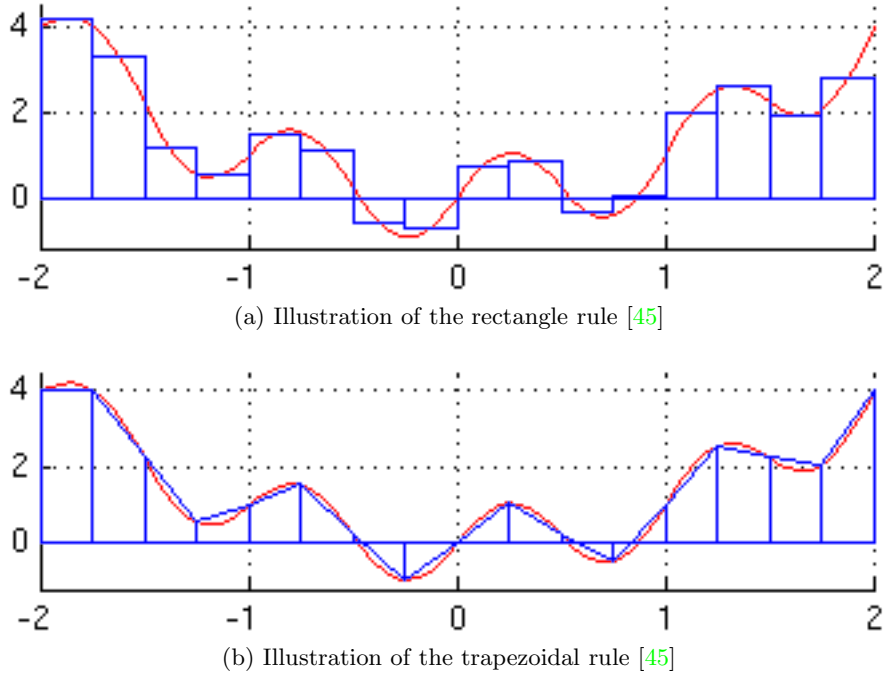


Figure 4.1: Numerical Integration Method

To illustrate it with an example, consider a PWC approximation of a Gaussian distributed random variable Z with zero mean and unit variance as shown in Figure 4.2a. Integration of this *pdf* around some $Z[l]$ and the PWC approximation for integration around the same $Z[l]$ are shown with a filled bar in Figure 4.2b and Figure 4.2c, respectively. In this approximation, the *pdf* values higher than the *pdf* at $Z[l]$ are decreased to $pdf(Z[l])$, and the *pdf* values lower than the *pdf* at $Z[l]$ are increased to $pdf(Z[l])$. The errors introduced by these changes have opposite sign and this neutralization effect reduces the error due to the approximation. The total error can also be reduced by increasing the number of samples during circuit simulation.

If X_i is sampled from $-\infty$ to $+\infty$, then the sum of the probability of all discrete values will be equal to one.

$$\sum_{\text{all } j_i} P_d(X_i[j_i]) \approx \int_{-\infty}^{\infty} f_i(x) dx = 1 \quad (4.11)$$

In our example, if Z follows a Gaussian distribution, then $\pm 3\sigma_Z$ spread of Z around its mean covers 99.8% of the probability space.

$$\int_{-3\sigma_Z}^{3\sigma_Z} pdf(z) dz = 0.998 \quad (4.12)$$

Consequently, if $Z[l]$ is sampled within a range of $\pm 3\sigma_Z$ around its mean, then the sum of the discrete values $Z[l]$ will cover 99.8% of the probability space.

$$\sum_{\text{all } l} P_d(Z_l) \approx 0.998 \quad (4.13)$$

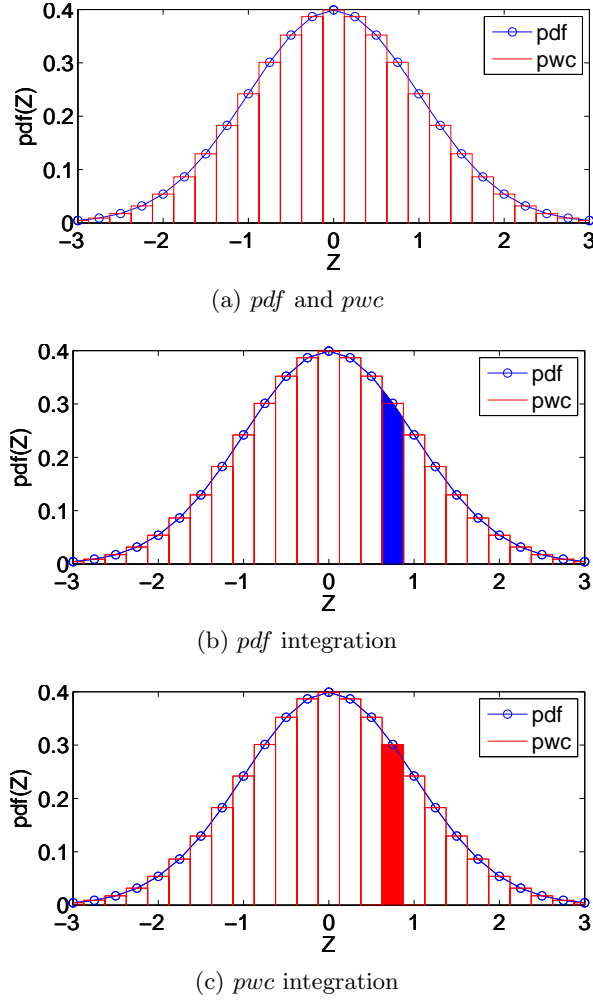


Figure 4.2: Piecewise constant approximation of probability density function

The range of the process parameter sweep in the simulation can be changed based on the requirement of the probability coverage.

4.2.2.2 Probability of Simulation

The probability of each discrete simulation event (k) is equal to the joint probability of all process parameters

$$P_s(k) = P_d(X_1[j_1], X_2[j_2], \dots) \quad (4.14)$$

In general, the process parameters are not independent. In order to simplify the data processing step, *Principal Component Analysis (PCA)* [46] can be used to convert the correlated process parameters into uncorrelated simulation parameters. Hence, without loss of generality, the parameters X_i used in this chapter are assumed to be independent after PCA. The joint probability of independent random variables is equal to the product of the probability of each random variable, so the probability of each

discrete simulation event can be rewritten as

$$P_s(k) = P_d(X_1[j_1]) \cdot P_d(X_2[j_2]) \dots \quad (4.15)$$

Let us define $P_d(Y[k])$ as the probability of the output $Y = Y[k]$ due to experiment k only. Since the probability of simulation event k is $P_s(k)$, we can define $P_d(Y[k])$ as follows:

$$P_d(Y[k]) = P_s(k) \quad (4.16)$$

$$\Rightarrow P_d(Y[k]) = P_d(X_1[j_1]) \cdot P_d(X_2[j_2]) \dots \quad (4.17)$$

Each experiment k gives an outcome $Y[k]$. The unknown probability of obtaining this outcome, $P_d(Y[k])$, is estimated from the known joint probability of the process parameters in the sample point k . Because of the assumption of independence, this joint probability is the product of the probabilities of each individual parameter.

Note that the $P_d(Y[k])$ is not the probability of $Y = Y[k]$, as more than one experiments could produce the same value of the output $Y[k]$ in case of a non-monotonous function of Y . Each $P_d(Y[k])$ will have a different probability value depending on the probability of the experiment k .

4.2.2.3 Moment Estimation

A weighted sample estimator is used here for estimating the moments of the output parameter Y [38]. To illustrate this process, consider the circuit simulation run N times. Let us assume that the probability of each simulation output $Y[k]$, i.e. $P_d(Y[k])$, is already estimated in the previous step. The probability of each output $Y[k]$ implies that the output event $Y[k]$ should repeat itself $\lceil N \cdot P_d(Y[k]) \rceil$ times. To obtain a high accuracy, each simulation output $Y[k]$ should occur at least once, implying that the lower bound of N should be defined as

$$N \geq \frac{1}{\min(P_d(Y[k]))} \quad (4.18)$$

Now, let us define a vector $R(Y[k])$ as a set of experiment outputs $Y[k]$ which, repeats $\lceil N \cdot P_d(Y[k]) \rceil$ times, i.e.

$$R(Y[k]) = [Y[k], Y[k], \dots] \quad \lceil N \cdot P_d(Y[k]) \rceil \text{ times} \quad (4.19)$$

where the outcome (O) can be written as

$$O = [R(Y[1]), R(Y[2]), R(Y[3]), \dots] \quad (4.20)$$

Once the outcomes (O) have been generated, the statistical moments of simulation output Y can be evaluated using standard moment estimation equations on the sample set, where the mean (μ), variation (σ^2), and normalized n^{th} central moments (μ_n) are given as

$$\mu_y = E(O) \quad (4.21)$$

$$\sigma_y^2 = E((O - \mu_y)^2) \quad (4.22)$$

$$(\mu_n)_y = E((O - \mu_y)^n) / \sigma_y^n \quad (4.23)$$

These equations can be rewritten using (4.19) and (4.20) as

$$\mu_y = \frac{\sum_k \{Y[k] \cdot P_d(Y[k])\}}{\sum_k \{P_d(Y[k])\}} \quad (4.24)$$

$$\sigma_y^2 = \frac{\sum_k \{Y[k]^2 \cdot P_d(Y[k])\}}{\sum_k \{P_d(Y[k])\}} - \left[\frac{\sum_k \{Y[k] \cdot P_d(Y[k])\}}{\sum_k \{P_d(Y[k])\}} \right]^2 \quad (4.25)$$

$$(\mu_n)_y = \sum_{m=0}^n \frac{\binom{n}{m} (-\mu_y)^m \sum_k \{Y[k]^{n-m} \cdot P_d(Y[k])\}}{(\sigma_y)^n \cdot \sum_k \{P_d(Y[k])\}} \quad (4.26)$$

Note that N is only used to develop the outcome O during the illustration of the process of estimating the moments. When rewriting (4.21), (4.22), and (4.23) using (4.19) and (4.20), N appears in both the numerator as well as denominator, and cancels out. Hence, N is not required in the final moment estimation equations.

Using the method described above, the statistical moments of various non-Gaussian probability density functions can be estimated irrespective of the simulation tool. The proposed sampling approach of parameter values requires fewer simulations leading to a much faster conversion of the moments. Moreover, since the exact process variation distribution is not required during the simulation run, a change in the process variation spread can be analyzed without rerunning the circuit simulation.

4.3 Simulation Results and Comparison for FSME Method

To evaluate the accuracy of the FSME method, extensive Spectre circuit simulations have been carried out with the FSME method as well as with the standard MC method. The results of both simulation methods are reported and compared below.

In the experimental setup, 45nm and 32nm *Predictive Technology Models (PTM)* have been used for all simulations [31]. Five different circuits (Inverter, Buffer, NAND, NOR, 5 Inverters Chain) have been used and all these standard cells were sized according to their corresponding predictive technology model [32]. The process variations are considered to be a Gaussian distribution such that the results can be compared with the standard MC results. The proposed method is scalable for any number of parameter variations and various process parameters can be used, e.g. L , W , V_{th} , etc. However, due to space limitations, only two sets of variations are discussed here.

Set I. In the first set, the variation considered is the effective channel length (L) of the MOSFET with $3\sigma_L$ equal to 20% of the nominal value of L .

Set II. In addition to the first set, the variation in the effective channel width (W) is also considered in the second set with $3\sigma_W$ equal to 20% of the nominal value of W .

In the output, the first four statistical moments (mean $[\mu]$, standard deviation $[\sigma]$, skewness $[\gamma]$ and kurtosis $[\kappa]$) of the delay of the standard cell have been estimated. Cadence Spectre was used for circuit simulation and Matlab for data processing.

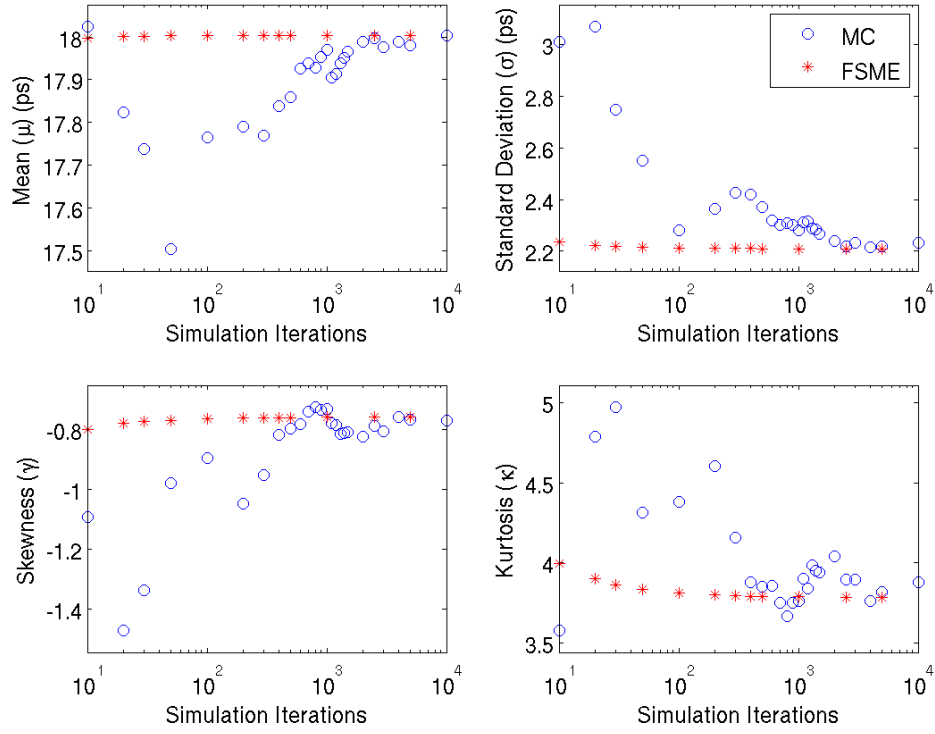
The first four moments (μ , σ , γ , and κ) of the delay vs simulation runs for the 45nm inverter with first set of variation using MC and FSME are shown in Figure 4.3a. Similarly, these moments of the delay vs simulation runs for the 45nm inverter with second set of variations are shown in Figure 4.3b. It is clear from these plots that FSME converges much faster than MC. The scattered plot of MC is due to its random sampling nature. As a result of the better convergence of FSME, the best available moments estimates from the FSME are taken as a golden reference value from both sets of variations for FSME and MC run comparison.

The error for the first set of variations after *five thousand iterations in MC* and *fifty iterations in FSME* with respect to the respective reference value is reported below. The error in the mean estimation using five thousand iterations of MC with reference value is 0.133% whereas fifty iterations of FSME lead to an error of only 0.006%. Similarly, the error for standard deviation estimation using MC is 1.059% whereas the FSME has an error of only 0.245%. The MC error in skewness estimation is 1.598% and FSME gives an error of only 1.257%. Lastly, kurtosis estimation has an error of 2.489% in MC where as FSME is at a 1.304% error.

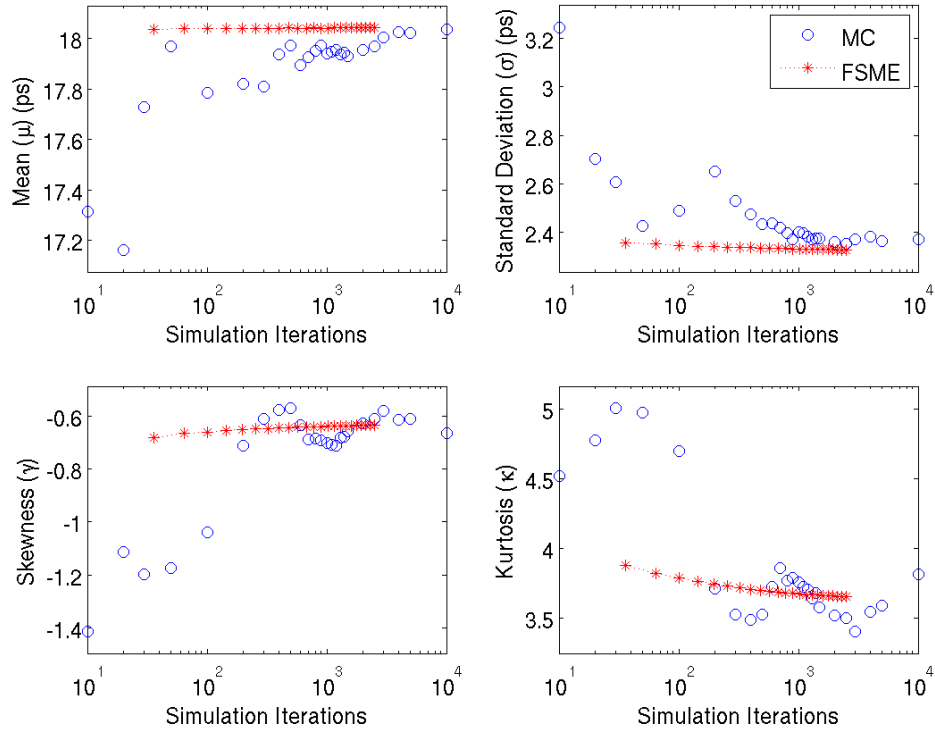
Since two parameter variations need more simulation iterations, *ten thousand iterations in MC* and *hundred iterations in FSME* are used to estimate the error with the respective reference value. The error in the mean estimation using ten thousand iterations of MC with reference value is 0.03% whereas one hundred iterations of FSME show only 0.019% error. Similarly, the error for standard deviation estimation using MC is 1.772% whereas the FSME has an error of only 0.817%. The MC error in skewness estimation is 4.637% and FSME gives an error of only 4.038%. Lastly, kurtosis estimation has an error of 4.432% in MC where as FSME is at a 3.663% error.

It is clear from above experimental results that MC with five thousand iterations produces more inaccurate results in comparison to the respective golden reference than FSME with only fifty simulations for one parameter variations. Equally, for two parameter variations, MC with ten thousand iterations has more error in comparison to the respective golden reference than FSME with only hundred simulations for one parameter variations. Similar behaviour is observed in all the five test circuits in both the 45nm and the 32nm technologies. The reference values of these four moments along with the error in the moment estimation for MC with five thousand runs and FSME with fifty runs using 45nm and 32nm technology with first variation set are reported in Table 4.1. A similar table for the second set of variations is reported in Table 4.2. The plots of the moment estimation vs simulation runs for an inverter in 32nm PTM using the first and second set of variations are shown in Figure 4.4a and Figure 4.4b, respectively. Furthermore, Buffer, NAND, NOR, and Inverter Chain have similar behaviour, thus their plots are not included here.

In the results above, we assumed that the variations are following a Gaussian distribution. Now, four different probability density functions (Gaussian, Lognormal, Gamma, and Beta) with the same mean and standard deviation have been considered for the first set of variations. The first four moments of the delay vs simulation runs for a 45nm inverter chain using these probability density functions are shown in Figure 4.5. In the process of these moment estimations, only the mathematical implementation of the *pdf* function is changed, and rerunning of the simulation is not required.

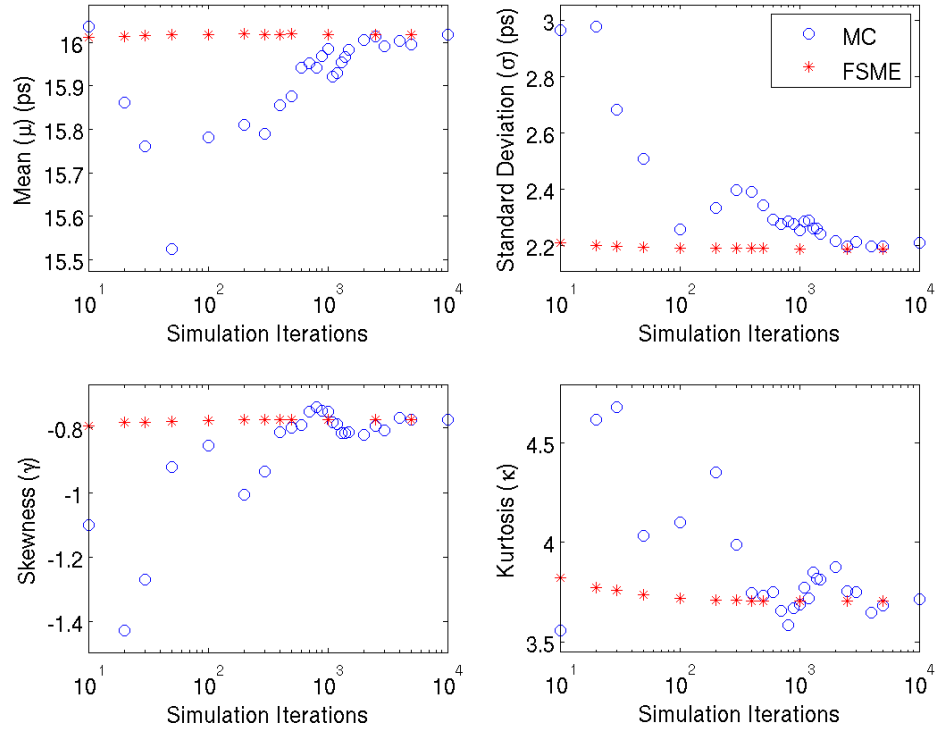


(a) 45nm Inverter - One parameter variation

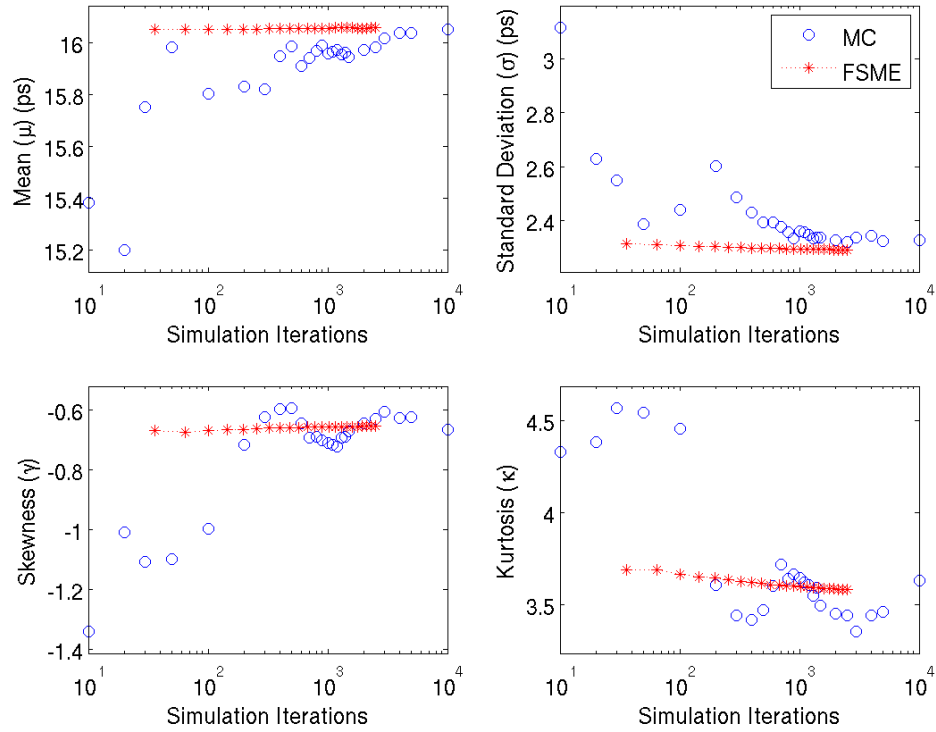


(b) 45nm Inverter - Two parameter variations

Figure 4.3: The first four moment estimation vs simulation runs for MC and FSME with one parameter (L) and two parameters (L and W) variations in 45nm Inverter



(a) 32nm Inverter - One parameter variation



(b) 32nm Inverter - Two parameter variations

Figure 4.4: The first four moment estimation vs simulation runs for MC and FSME with one parameter (L) and two parameters (L and W) variations in 32nm Inverter

Circuits	Mean (μ)			Standard Deviation (σ)		
	Ref (ps)	MC %	New %	Ref (ps)	MC %	New %
45nm PTM Technology						
Inverter	18.004	0.133	0.006	2.210	1.059	0.245
Buffer	21.570	0.167	0.001	3.338	0.991	0.173
NAND	25.913	0.103	0.004	2.473	1.093	0.267
NOR	20.773	0.127	0.020	2.369	1.176	0.304
Inverter Chain	37.558	0.158	0.001	5.530	1.055	0.207
32nm PTM Technology						
Inverter	16.019	0.148	0.005	2.187	0.915	0.233
Buffer	18.082	0.203	0.003	3.409	0.922	0.215
NAND	23.624	0.115	0.004	2.513	0.999	0.251
NOR	17.723	0.141	0.006	2.298	1.158	0.310
Inverter Chain	29.928	0.184	0.005	5.155	0.986	0.189

(a) Error % comparison in Mean and Standard Deviation

Circuits	Skewness (γ)			Kurtosis (κ)		
	Ref	MC %	New %	Ref	MC %	New %
45nm PTM Technology						
Inverter	-0.759	1.598	1.257	3.786	2.489	1.304
Buffer	-0.305	1.853	1.359	2.958	2.128	0.674
NAND	-0.682	2.015	1.381	3.685	3.218	1.336
NOR	-0.804	3.659	1.384	3.997	5.659	1.506
Inverter Chain	-0.253	2.176	1.144	2.981	2.622	0.775
32nm PTM Technology						
Inverter	-0.773	0.117	0.811	3.704	0.527	0.824
Buffer	-0.186	3.481	0.743	2.729	2.125	0.452
NAND	-0.607	0.836	1.154	3.580	1.165	1.041
NOR	-0.875	2.817	1.203	4.100	4.490	1.413
Inverter Chain	-0.231	3.324	0.418	2.864	2.097	0.591

(b) Error % comparison in Skewness and Kurtosis

Table 4.1: Error % comparison in the first four moments estimation of delay for one parameter (L) variation using Monte Carlo (5000 runs) and the proposed method (50 runs) in 45nm and 32nm PTM technologies

It can be observed from the figure that the higher order moments of the delay vary with the distribution of the process parameters. The same data with relative axis for the statistical moments are plotted in Figure 4.6a. In this figure, the best available statistical moments, i.e. data corresponding to the maximum simulation iterations, of the delay with the Gaussian distribution as an input parameter distribution is taken as a reference value, assigned to 100. The Y-axis is transformed with respect to the reference value and axis is scaled for the best fit of the plot. The scaled data with fixed Y-axis is plotted in Figure 4.6b. The plot with fixed Y-axis can be used to compare the relative differences among the statistical moments.

Circuits	Mean (μ)			Standard Deviation (σ)		
	Ref (ps)	MC %	New %	Ref (ps)	MC %	New %
45nm PTM Technology						
Inverter	18.042	0.030	0.019	2.328	1.772	0.817
Buffer	21.592	0.001	0.010	3.374	1.445	0.692
NAND	25.965	0.022	0.011	2.653	1.813	0.810
NOR	20.810	0.031	0.033	2.477	1.982	0.983
Inverter Chain	37.598	0.006	0.007	5.574	1.500	0.700
32nm PTM Technology						
Inverter	16.056	0.029	0.027	2.293	1.530	0.635
Buffer	18.104	0.018	0.030	3.447	1.311	0.638
NAND	23.691	0.013	0.008	2.713	1.659	0.769
NOR	17.761	0.039	0.017	2.398	1.992	0.897
Inverter Chain	29.972	0.012	0.008	5.205	1.397	0.649

(a) Error % comparison in Mean and Standard Deviation

Circuits	Skewness (γ)			Kurtosis (κ)		
	Ref	MC %	New %	Ref	MC %	New %
45nm PTM Technology						
Inverter	-0.633	4.637	4.038	3.656	4.432	3.663
Buffer	-0.290	4.593	3.689	2.976	3.492	2.134
NAND	-0.523	5.476	4.875	3.535	5.099	3.831
NOR	-0.699	8.667	5.072	3.880	9.717	4.932
Inverter Chain	-0.247	4.332	3.114	3.001	4.181	2.422
32nm PTM Technology						
Inverter	-0.655	1.512	2.055	3.581	1.371	2.306
Buffer	-0.167	2.425	3.535	2.753	2.985	1.758
NAND	-0.410	1.232	4.072	3.463	2.283	2.744
NOR	-0.764	6.888	4.936	3.968	8.065	4.895
Inverter Chain	-0.219	2.070	1.448	2.882	3.118	1.907

(b) Error % comparison in Skewness and Kurtosis

Table 4.2: Error % comparison in the first four moments estimation of delay for two parameters (L and W) variation using Monte Carlo (10000 runs) and the proposed method (100 runs) in 45nm and 32nm PTM technologies

It can be observed from the figure that the relative difference in the mean of the delay is within 0.1% and the relative difference in the standard deviation is within 2.5%. The small differences indicate that the *pdf* of the delay distribution shall be very close to each other. Although, the absolute difference in the skewness is within 0.15, the relative difference is around 60%. The very high relative difference in skewness is because of very small absolute values of it. Being very small absolute skewness values, the delay *pdfs* are expected to have similar symmetry. The relative difference in kurtosis is 15% and the absolute difference is 0.4. However it is important to note that the differences in the kurtosis of the delay due to Gaussian, Log-normal, and Gamma distribution are

within 3% only, and the kurtosis due to Beta distribution is around 12% than others. This observation indicates that the *pdf* of the delay due to Beta distribution is expected to be relatively flat as compared to other *pdf*s. The *pdf* of the delay due to various input *pdf* variations are plotted and analyzed in Section 4.4.

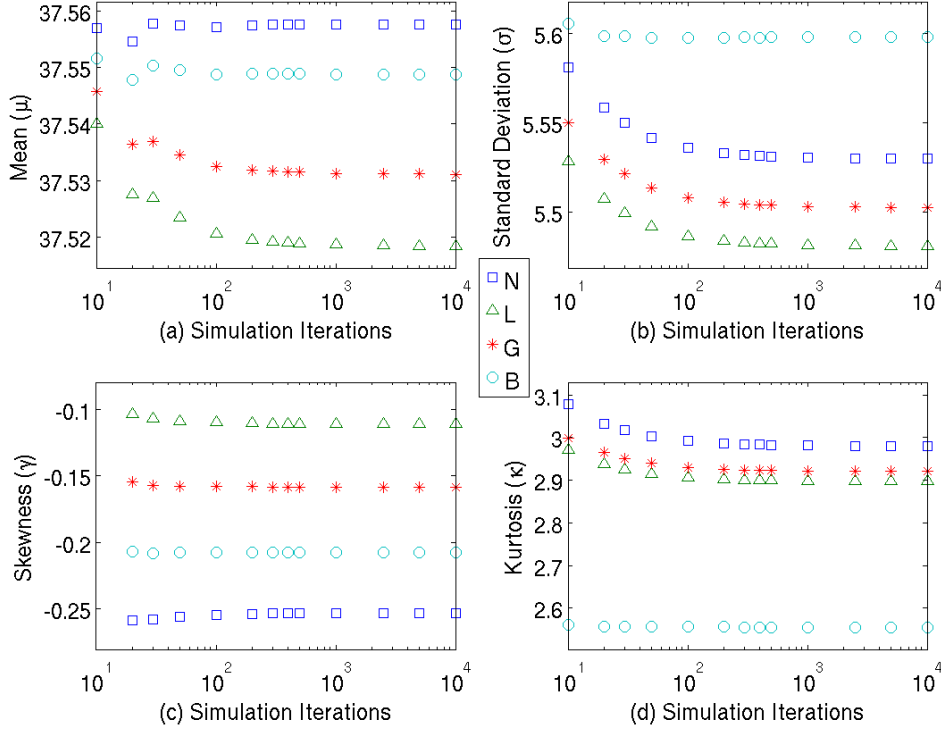


Figure 4.5: Moment estimation vs simulation run in 45nm inverter chain using Gaussian (N), Lognormal (L), Gamma (G), and Beta (B) distributions.

The above results show that the simulation iterations required in FSME to estimate the moments differ from the simulation iterations required in MC by two orders of magnitude. This results into 100 \times speedup during library characterization. Furthermore, a different parameter spread can be analyzed in FSME by just changing the parameter of the *pdf* function in the data processing stage. Moreover, any type of probability density function can be used with FSME by changing the implementation of the *pdf* function only. This extra data processing does not require rerunning of the circuit simulator, which results into faster run times and smaller memory requirement to store all the data.

4.4 Probability Density Function Estimation Method

The probability density function (*pdf*) of the circuit simulation output (Y) depends on the probability of each simulation run. These probabilities ($P_d(Y[k])$) are calculated in Section 4.2.2.2, and will be used to estimate the *pdf* of the Y . It has been discussed in Section 4.2.2.1 that if any variable is uniformly sampled, then its *pdf* and probabilities

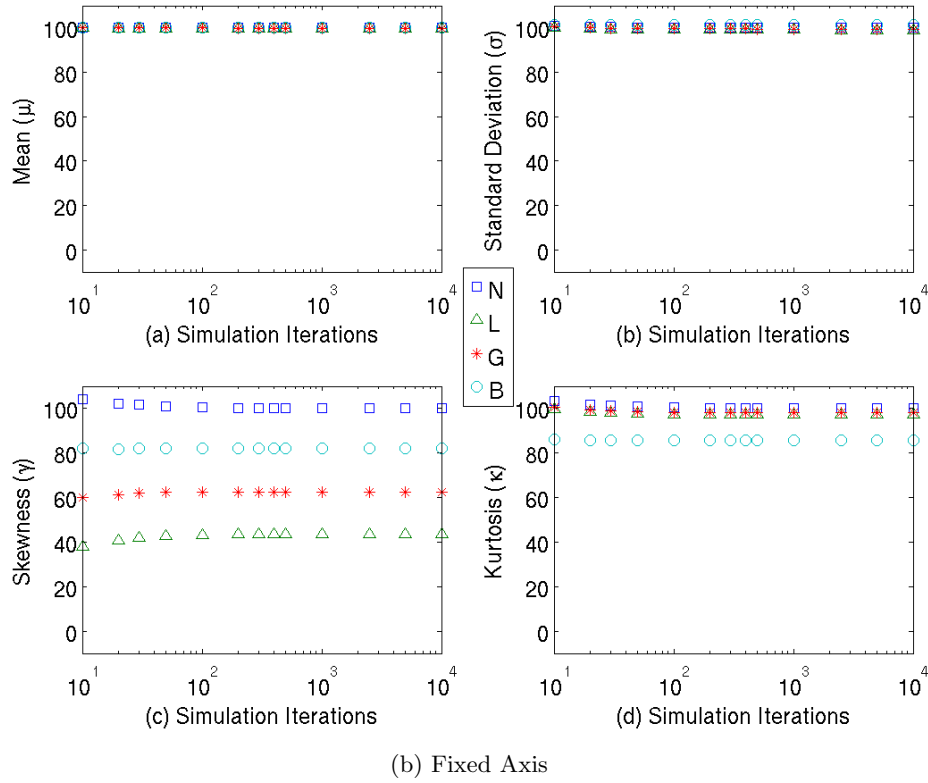
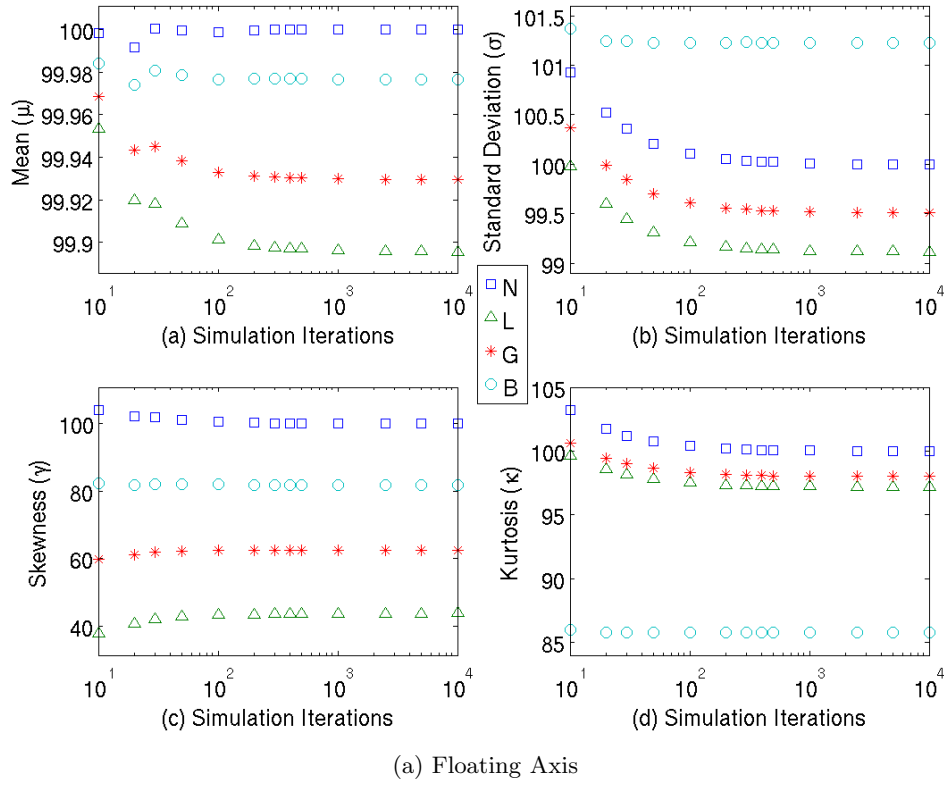


Figure 4.6: Moment estimation vs simulation run in 45nm inverter chain using Gaussian (N), Lognormal (L), Gamma (G), and Beta (B) distributions with Relative Scale.

are related by (4.10). The relation is given below with an appropriate change in the variable for convenience.

$$P_d(Y[k]) = pdf(Y[k]) \cdot \Delta Y \quad (4.27)$$

Since the simulation output Y is not necessarily a linear function of the PVT parameters (X_i), the sampled values of the simulation output Y need not to be uniformly spaced. Therefore, (4.27) can not be used directly to estimate the *pdf* of Y . There are two approaches to convert the probability of Y to its *pdf*. The first approach is to divide the output Y into uniform brackets and calculate the delay of each bracket by adding up the delays of all the sampled outputs $Y[k]$ within the brackets. In the following text, each bracket is called as a bin and call this method the Bin Method. The second approach is to make appropriate changes in (4.27) to incorporate the non-uniform samples of $Y[k]$. In the following text, this method is call the Direct Method.

4.4.1 PDF Estimation - Bin Method

In the bin method, the sample space of Y is divided into uniformly spaced bins, followed by estimating the probability of each bin. Since the newly sampled probabilities are uniformly spaced, (4.27) can be used to estimate the *pdf* of Y . Here, the bin method is discussed in detail.

The sampled simulation outputs $Y[k]$ are spread between Y_{min} and Y_{max} such that

$$Y_{min} = \min(Y[k]) \quad (4.28)$$

$$Y_{max} = \max(Y[k]) \quad (4.29)$$

and

$$Y_{min} \leq Y[k] \leq Y_{max} \quad (4.30)$$

Let us create uniformly spaced sample points in Y and name them Y_b . Here subscript b stands for bin. For n uniformly sampled points in Y_b , ΔY_b can be calculated as

$$\Delta Y_b = \frac{Y_{max} - Y_{min}}{n} \quad (4.31)$$

and the probability of each bin can be calculated as

$$P_d(Y_b[l]) = \sum_{Y_b[l] - \frac{\Delta Y_b}{2} < Y[k] \leq Y_b[l] + \frac{\Delta Y_b}{2}} P_d(Y[k]) \quad (4.32)$$

Now, the uniformly sampled points of simulation output Y and their probability is available. Therefore, the *pdf* of Y can be calculated based on the sample points of $Y_b[l]$.

$$pdf(Y_b[l]) = \frac{P_d(Y_b[l])}{\Delta Y_b} \quad (4.33)$$

Since the probability of each bin highly depends on the bin boundary, a slight shift in the bin boundary could results the sample points close to the bin boundary to fall into different bins. It might lead to considerable inaccuracy if the probability of each bin is comparable to the probability of the sample points close to the boundaries. The

error can reduce by increasing the probability of each bin, which results in to larger bin size and fewer numbers of bins. Therefore, number of bins should be lesser than the number of sample points by more than an order of magnitude. The error can further reduce by lowering the probability of the boundary points which causes the error in the *pdf* estimation due to shift from one bin to another. This leads to the reduction in the probability of each sample points $Y[k]$, which can be achieve by increasing number of samples in the simulation. Therefore, more simulation runs and larger bin size could results into smaller error in the *pdf* estimation using bin method.

4.4.2 PDF Estimation - Direct Method

In (4.27), ΔY is a constant and independent of the sampled values $Y[k]$. If different ΔY can be calculated for each value of the simulation output $Y[k]$ based on their local separation of $Y[k]$ from their neighboring points ($Y[k - 1]$ and $Y[k + 1]$), then the equation will be valid for non uniform sampled points as well. Let us call different ΔY for each sample of $Y[k]$ as $\Delta Y'[k]$. Therefore (4.27) can be rewritten as

$$P_d(Y[k]) = pdf(Y[k]) \cdot \Delta Y'[k] \quad (4.34)$$

Now, using $P_d(Y[k])$ and $\Delta Y'[k]$, the *pdf* of Y sampled at $Y[k]$ points can be estimate using

$$pdf(Y[k]) = \frac{P_d(Y[k])}{\Delta Y'[k]} \quad (4.35)$$

It is important to note that the direct method is applicable only for one parameter variation with monotonous function of simulation output with respect to the input parameter. This is because of the following reasons. In a non-monotonous function, two sample pointes of the simulation output Y can attain the same absolute value. Furthermore, $P_d(Y[k])$ is the probability of the sample event k . Thus, the probability of $Y[k]$ need not to be equal to the probability of $Y = Y[k]$ which is the primary requirement in (4.34). Additionally, due to non-monotonous samples of Y , it is not possible to estimate $\Delta Y'[k]$ by knowing the neighboring points. Similarly, for more than one parameter variations, two samples of Y can attain the same absolute value. Therefore the direct method is not applicable for any non-monotonous functions or multi parameter variations.

However, the sum operation in bin method, while calculating the probability of each bin ($P_d(Y_b[l])$) in (4.32), resolve the problem arises by the non-monotonous function and multi parameter variations. Thus the bin method is applicable to any simulation output.

4.5 Simulation Results for PDF Estimation Method

As discussed earlier, the direct method is not applicable in various scenarios, whereas, the bin method can be use with any simulation data. It has also been discussed that inaccuracy in the bin method can reduce by increasing the number of simulation iterations. Therefore, the bin method with very high number of simulation iterations have been used to estimate the *pdf* of the simulation output Y .

The *pdf* of the delay in 45nm inverter chain is estimated here. This circuit scenario is same as the one used to analyze the statistical moments of the delay due to various *pdfs* for the input parameter variation (see Figure 4.5). In this simulation, input parameters are assume to follow four different *pdf* functions, namely, Gaussian, Lognormal, Gamma, and Beta distributions. The *pdf* of the delay due to each input parameter variations are shown in Figure 4.7. These four *pdf* curves are plotted together in Figure 4.8 for convenience in the comparison.

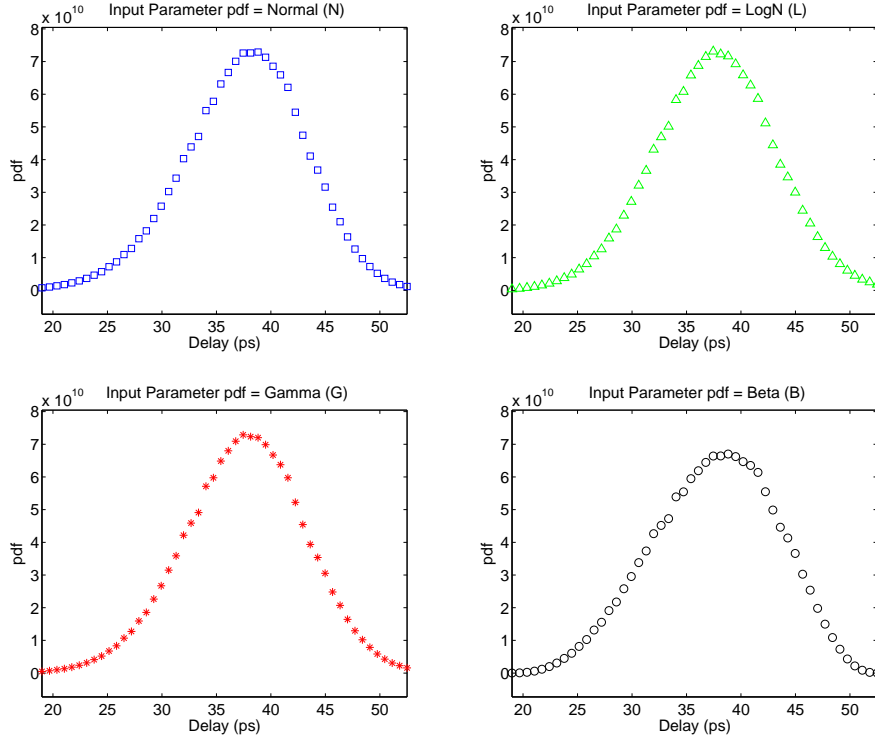


Figure 4.7: *pdf* of the delay of an 45nm inverter chain using Gaussian (N), Lognormal (L), Gamma (G), and Beta (B) distributions.

It can be observed from the plots, that these four *pdf* are very close to each other. Furthermore, the pdf due to Beta distribution is more flat than other *pdf* distributions. These observations confirm the discussion about the *pdf* of delay in Section 4.3.

4.6 Summary

This chapter proposes a simulation and analysis method based on the uniform sampling technique and a weighted sample estimator, which requires fewer simulation runs for statistical moment estimation. The number of simulation iterations required by this *Fast Statistical Moment Estimation (FSME)* method is at least two orders of magnitude lower than the number of simulation runs required in the Monte Carlo method. This results into a 100 \times speedup in SSTA library characterization. Along with this, changes in parameter spread and/or probability density function do not require rerunning of

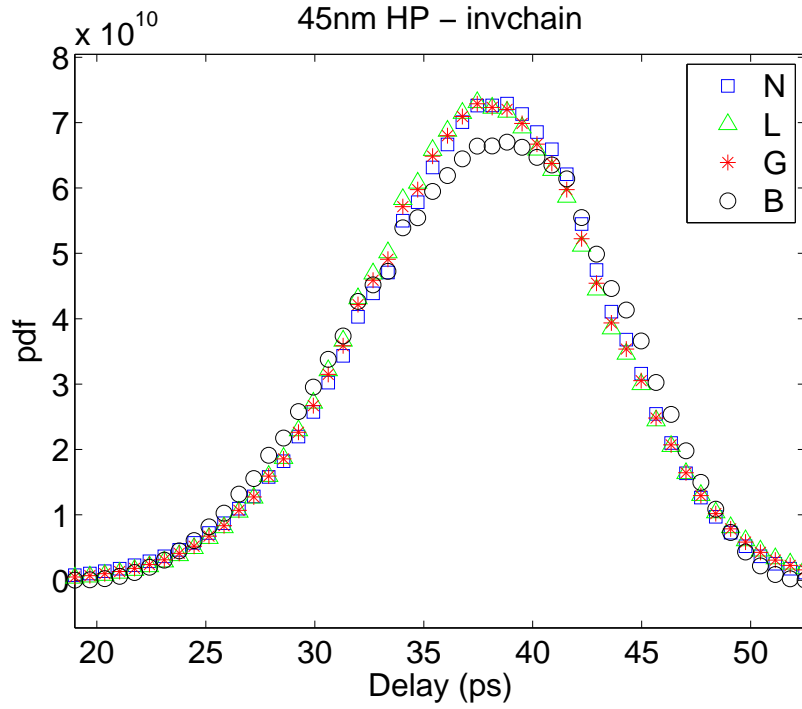


Figure 4.8: *pdf* of the delay of an 45nm inverter chain using Gaussian (N), Lognormal (L), Gamma (G), and Beta (B) distributions in one plot

the circuit simulations, which results into faster run time and smaller memory requirement. The state of the art circuit simulation tools can run Monte Carlo with Gaussian, lognormal and uniform distribution only whereas any distribution can be used in the proposed method.

Furthermore, a method to estimate the probability density function from the simulation data is proposed. The Bin Method is applicable to simulation data with any function and multi parameter variations whereas the Direct Method is applicable only to monotonous functions and one parameter variation.

The Set of Waveforms

Due to the variations in the PVT parameters, static timing analysis has limited applicability when analyzing the performance and correct functionality of a circuit. In Chapter 2, it has been discussed that static timing analysis may be used to estimate the delay in a circuit but variation is not taken care of in this analysis. Furthermore, in Chapter 3 we have analyzed the delay variations of a standard cell due to PVT variations. Similarly it has also been noted that researchers are developing a statistical static timing analysis methodology to estimate the delay variation due to PVT variations. The MODERN project team is developing a new approach within the SSTA methodology in which a SPICE like simulation will be carried out for the gates instead of making use of a lookup table based gate model [47, 48, 49, 50]. Although the transistor-level SPICE simulation is quite accurate, SPICE is much slower in comparison to the gate-level models. Therefore, the team is developing a compact transistor model for the digital circuit. In this SSTA approach, the output voltage waveforms of each standard cell and their variations due to the PVT variations are preserved.

The organization of the chapter is as follows: the concept of a set of waveforms and its usefulness for representing uncertainty during timing analysis is discussed in Section 5.1. Furthermore, the set of waveforms of standard cells are presented in Section 5.2. Additionally, the possible methods to represent the waveforms during timing analysis are presented in Section 5.3 and a comparison of various representations is discussed in Section 5.4. At the end, a summary of the chapter is presented in Section 5.5.

5.1 Concept of a Set of Waveforms

The proposed concept of a set of waveforms in SSTA is introduced to cache the various output waveforms which may arise due to PVT variations in the circuit. The main goal is to improve the accuracy of the SSTA methodology. The idea of using a set of waveforms is similar to the development of a waveform based model (e.g. CCS and ECSM) from a delay based model (e.g. NLDM) in STA. In this section, the limitations of NLDM will be discussed first, followed by the advantage of CCS and ECSM. The concept of the set of waveforms will be developed based on the analogy with STA.

Traditionally, NLDM models have been used in industry for static timing analysis. As discussed in Chapter 2, an NLDM model is a very simple lookup table based model which stores the delay of each standard cell for a set of the input signal slew and output effective capacitive load values. Being a very simple lookup table based method, NLDM makes the STA run very fast; however the accuracy is compromised. In general there are two major sources of inaccuracy in NLDM based STA. First, in NLDM, a waveform is described by only two points in the form of a signal slew. A single slew value may

correspond to an infinite number of waveforms. Since the input signal waveform affects the output of a gate, using only the two-point slew-based representation for each waveform may introduce a substantial error in the STA. Additionally, only delay and slew are not sufficient to estimate the behavior of the gates in sub 90nm technology. Further, the gate in the NLDM is characterized for a small set of the input signal slew and output effective capacitive load, and in practice, the slew and load values used during STA need not match with these characterized values. Therefore, interpolation or extrapolation is required to estimate the delay. A practical interpolation method considers four neighboring points and assumes a linear variation in the gate delay between the selected points. This interpolation method is known as bilinear interpolation. The error introduced by the interpolation can be reduced by considering more points and a polynomial fit within the region, which is also known as spline interpolation. However, the calculation will be more complicated, which increases runtime and makes it unsuitable for multi-million gate designs. In contrast with interpolation, four neighbors are not available for extrapolation. Only, some of the boundary points from the lookup table are considered for extrapolation and a linear variation of the gate delay and output signal slew is assumed outside the table. This assumption is very inaccurate and introduces large deviation from a SPICE simulation.

The error introduced by the sources mentioned above is relatively small for larger technology nodes and low frequency devices. However, the error at advanced technology nodes and higher operating frequency devices is relatively high and cannot be ignored. Due to this increase in the relative error during delay calculation, NLDM does not remain an attractive solution for standard cell modeling in STA. Many advanced modeling schemes have been developed in which the complex behavior of the MOSFET and the signal waveforms can be captured more accurately. These advanced modeling schemes reduce the relative error in the delay calculation. The *Composite Current Source (CCS)* Model and the *Effective Current Source Model (ECSM)* are well known industry standard modeling schemes. In contrast with NLDM, CCS model stores the output current waveform instead of the signal slew. During the STA analysis, the output voltage waveform is computed from the output current waveform and the delay is estimated. Similarly ECSM stores the output voltage waveform. Both CCS and ECSM preserve detailed waveforms which results into more accurate delay estimation as compared to NLDM at the cost of extra computation time. Details of NLDM, CCS and ECSM models were discussed in Chapter 2.

Being waveform-based technologies, CCS and ECSM are more accurate in estimating the delay of standard cells. However the problems that arise due to variations in the PVT parameters are not addressed efficiently. As discussed in Chapter 2, the *Statistical Static Timing Analysis (SSTA)* methodology has been developed to address the error in the delay estimation due to the process variations. In a broad sense, delay and its variation is estimated for each standard cell in SSTA. The spread in the delay estimated by SSTA helps in estimating the effect of PVT variations. However, similar to the issues of STA, only the spread of the delay is not sufficient to accurately estimate the effect of PVT variations. As waveform-based technologies like CCS and ECSM have improved the delay estimation accuracy over the delay-based technologies like NLDM, there is a need to model uncertainty in the waveform instead of just modeling delay and

slew. The uncertainty in the waveform will improve the accuracy of the delay variation estimation. Since the delay of a gate is highly dependent on the output waveform of previous gates, an accurate modeling in the variation in output waveform will improve the accuracy in timing analysis. This will eventually results in an improvement of the yield of ASIC chip manufacturing.

The proposed method of modeling the uncertainty in each waveform requires us to store various possible output waveforms of a standard cell due to the PVT variations. This collection of the output waveforms is called a set of waveforms. Since various possible waveforms and their probability will be available in SSTA, an accurate delay variation and output waveform variation can be estimated.

5.2 Representing Uncertainty with the Set of Waveforms

As we discussed in Section 5.1, a set of waveforms is a collection of possible output waveforms due to PVT variations. It can be generated by varying the PVT parameters of the gate. A simple inverter gate will be used here to illustrate the concept of a set of waveforms.

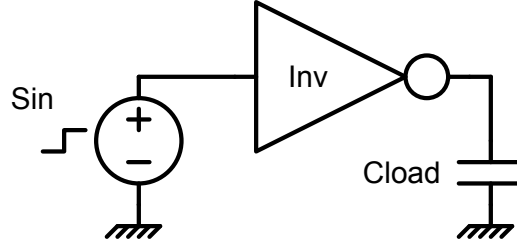


Figure 5.1: A Inverter for Waveform Set

A standard inverter is shown in Figure 5.1. The PVT variations are taken from Table 3.1. If the MOSFET channel length (L) is varied from $(\mu_L - 3\sigma_L)$ to $(\mu_L + 3\sigma_L)$ while keeping the rest of the PVT parameters, the input signal and the output load constant, various possible output waveforms can be generated. The set of input and output waveforms due to variations in L are shown in Figure 5.2a and Figure 5.2b respectively. In these figures, time is on the X-axis and voltage is on the Y-axis. As the channel length (L) varies, the gate capacitance of the MOSFET and the drain current varies. This variation results in an output waveform variation (See Figure 5.2b). Since variation in the input waveform is not considered in this illustration, a single input waveform is used here (See Figure 5.2a).

A rising input signal is considered in this example. When the input is logic '0', the PMOS is ON and NMOS is OFF, and when the input is logic '1', the PMOS is OFF and NMOS is ON (See Figure 2.11 in Section 2.2). At the beginning of the rising transition, when the input signal is less than V_{TN} , the signal remains steady at the output. During the input signal transition, i.e. when the input signal is between V_{TN} and V_{TP} , there is a path from the output node to V_{SS} through the NMOS. This results into the fall of the output signal. Since very little discharge occurs, waveforms corresponding to various channel lengths are very close to each other. This results in a very narrow spread at

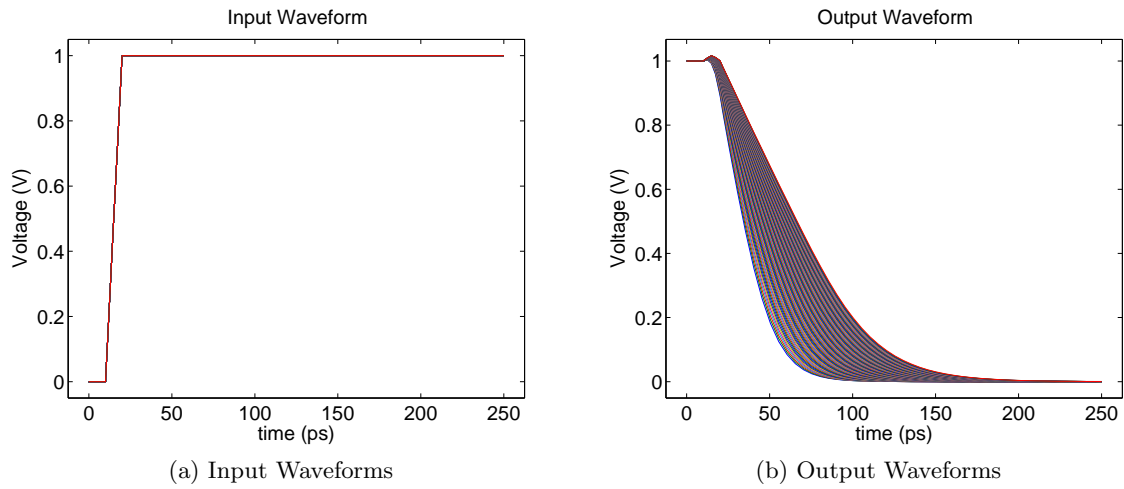


Figure 5.2: The Set of Waveforms for Inverter

the beginning of the output signal transition. As the input voltage increases further, the output signal discharge is continued. Since the discharge drain current varies with the channel length, each waveform deviates from the nominal output waveform. This results into the wide spread at the ending of the output signal waveform.

In a real circuit, the output of a gate is the input of the next gate. This implies that the input waveform will also have a variation. Therefore, let us take an inverter chain circuit as shown in Figure 5.3. Here we have provided a single ramp input to the first inverter. Additionally, a correlated variation in the channel length of each MOSFET is assumed.

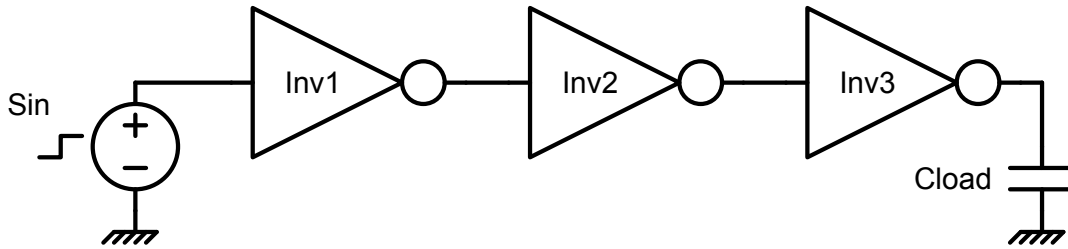


Figure 5.3: A Inverter Chain for Waveform Set

The output of the first inverter is similar to Figure 5.2b. However, in this circuit, the varying signal is given as an input to the second inverter. The input of the last inverter is more realistic and close to practical circuit signal variations. The input and the output of the last inverter are shown in Figure 5.4a and Figure 5.4b respectively. Similar to Figure 5.2, time is on the X-axis and voltage is on the Y-axis in these figures.

Due to variations in the input signal, the output is no longer narrow at the beginning. This analysis shows that the real circuit waveforms are expected to be like Figure 5.4b. Preservation of the waveforms will increase the accuracy in the estimation of the delay and its variation. Additionally, it will allow us to more accurately estimate the output waveform spread, which will serve as the input of the next gate. This will results in a

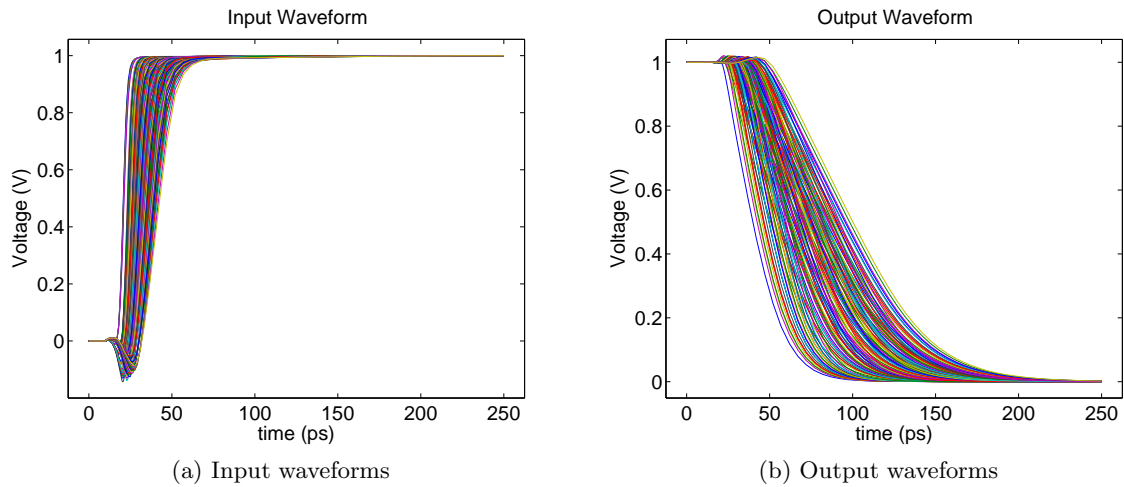


Figure 5.4: The Set of Waveforms for Inverter Chain

better estimation of standard cell behavior.

Each process variation is associated with a probability density function (pdf). In industrial practice, each process variation is assumed to follow a Gaussian distribution. In the discussion above, we can assume that the MOSFET channel length (L) is following a Gaussian distribution. It implies that there is a higher probability that the channel length will be close to its nominal mean value than values away from the nominal value. Since each waveform in the set is due to some value of L , each waveform is associated with a probability value. In general, for more than one parameter variation, the probability of each waveform should be equal to the probability of the corresponding variable combination. These probabilities can be estimated as discussed in Section 4.2.2.2.

5.3 Representation of the Set of Waveforms

In Section 5.1, we described the necessity of modeling a set of waveforms and in Section 5.2 we presented some real sets of waveforms. The block diagram of the SSTA engine using a set of waveforms is shown in Figure 5.5.

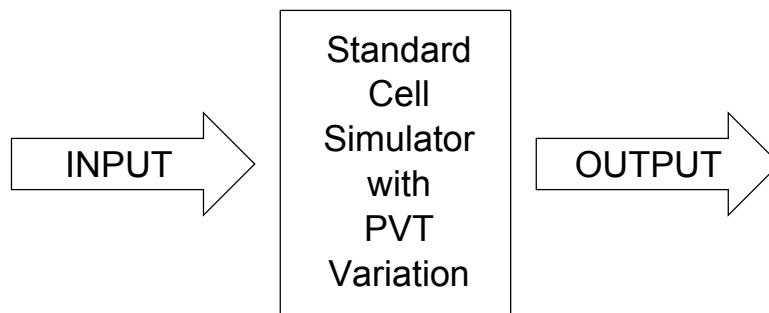


Figure 5.5: SSTA Engine

Here INPUT and OUTPUT are the set of waveforms and the center box is the SSTA engine. In our SSTA methodology, it is required to represent and store these sets of waveforms. The main challenge of this approach is the huge amount of data required to represent the waveforms. The large data set at every standard cell output results in a large memory requirement and high run time. There are three possible approaches to represent the set of waveforms. These three approaches and a comparison with respect to memory requirement are discussed in the following subsections.

5.3.1 Lookup Table based Representation

A lookup table based method is a very straightforward method for a representation of a set of waveforms. In this representation method, each waveform is stored as an array of time-value pairs. The probability of each waveform is also stored along with the waveform value. The data can be compressed by storing a down-sampled waveform and imposing a piecewise linear approximation on it. Similarly to CCS and ECSM models, this representation stores the entire waveform set in a raw data format. A block diagram of an SSTA engine with a lookup table based representation is shown in Figure 5.6.

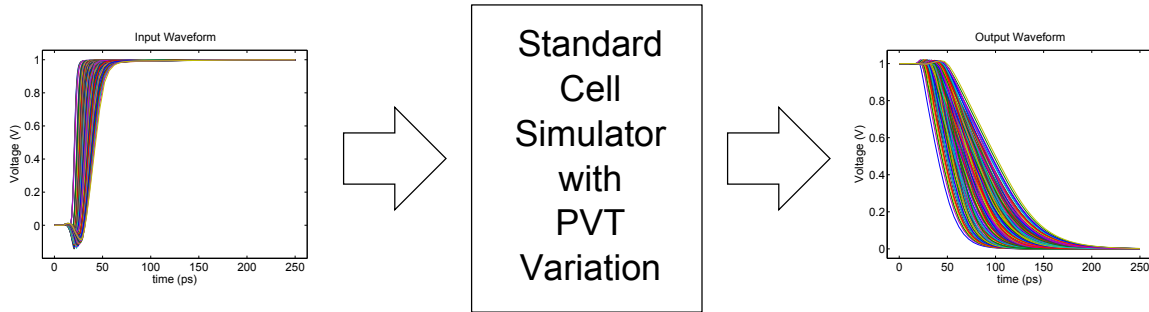


Figure 5.6: SSTA Engine with Table Model

5.3.2 Statistical Moments based Representation

Each waveform in the waveform set model is associated with a probability value. If the set of waveforms is cross-sectioned vertically at any time value, we can observe various possible output voltages at that time. If there are n waveforms in the set, n possible voltage values can be measured. The voltage values and their probabilities can be used to generate the probability distribution curve of the voltages at that time. An example of such a vertical cross-section is shown in Figure 5.7a. Here X-axis is time and Y-axis is voltage. A vertical line shows the cross-section of the waveform set at a particular time. The probability density curve for the voltages at that particular time is shown in Figure 5.7b. Here, voltage is shown on the X-axis and its probability density function is shown on the Y-axis.

The voltage value and its probability curve can be used to estimate the mean, standard deviation and higher order statistical moments of the voltage at this particular time. The mean (μ) and spread (3σ) of the voltage at the same time is annotated on

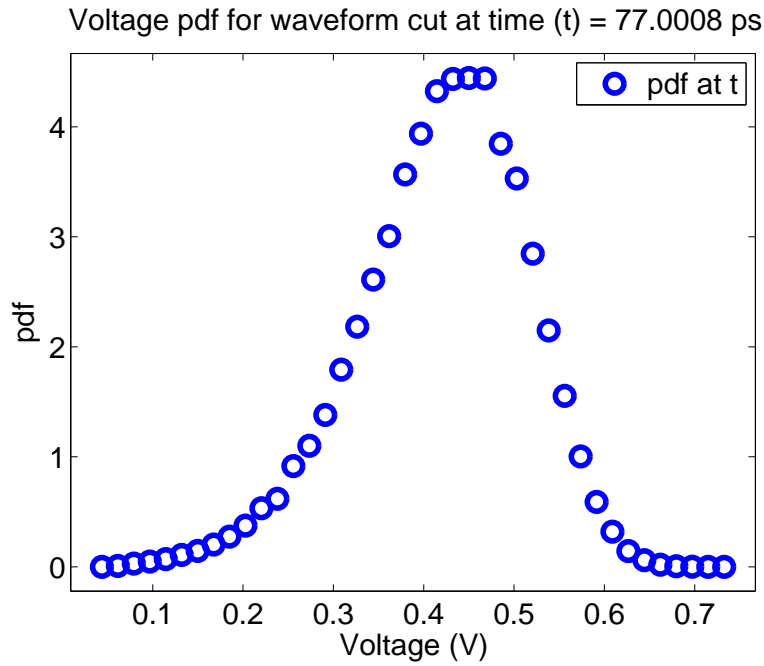
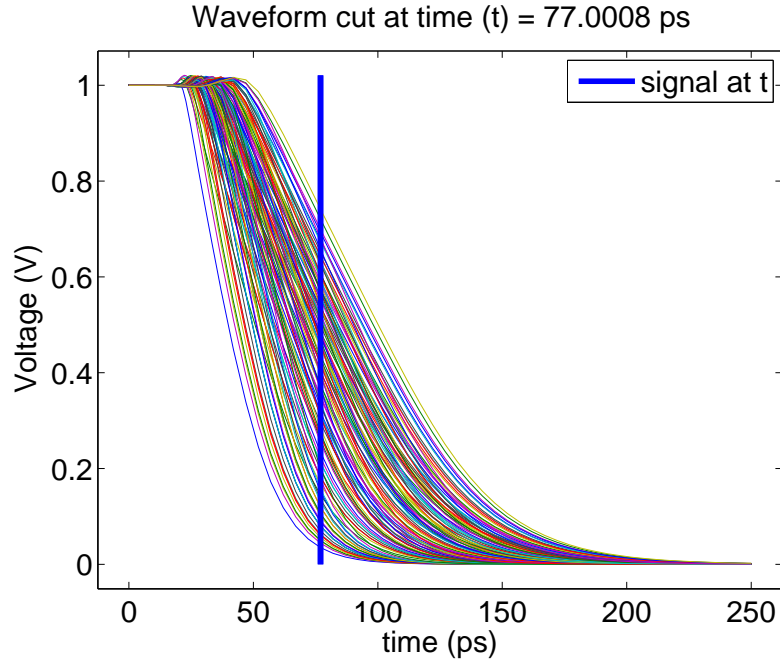
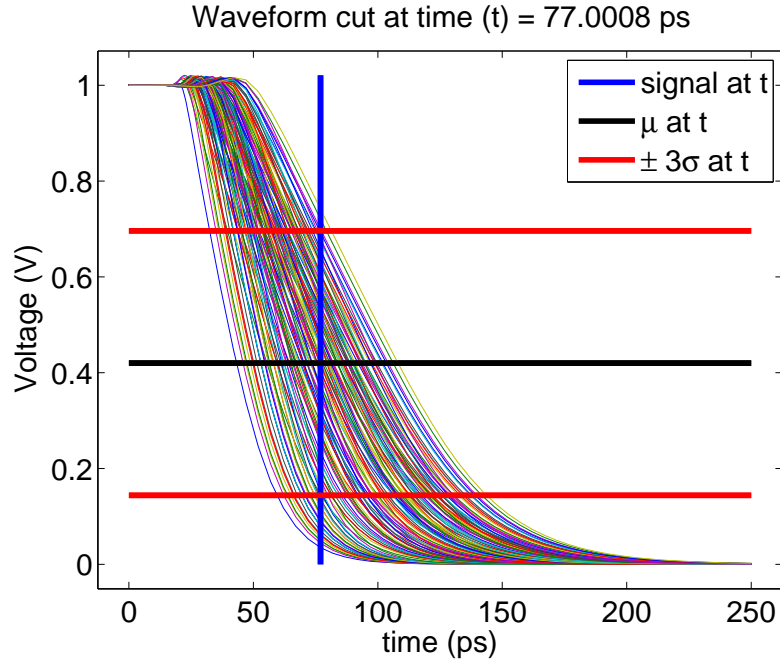
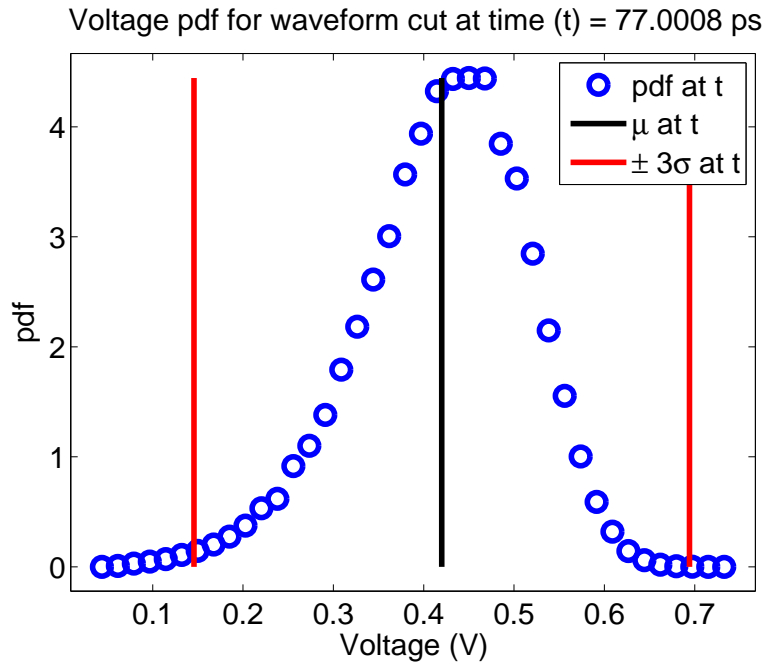


Figure 5.7: Vertical cross-section of Waveforms at time t

the waveform set as shown in Figure 5.8a. Here the centre horizontal line corresponds to the mean and the top and the bottom lines represent the $\pm 3\sigma$ bound of the voltage at this time. Annotation of the same μ and $\pm 3\sigma$ on the probability density curve are shown in Figure 5.8b. Here the centre vertical line corresponds to the mean and the



(a) The Set of Waveforms



(b) Probability of Voltage

Figure 5.8: Vertical cross-section of Waveforms at time t with μ and $\pm\sigma$ of voltage

right and the left lines represent the $\pm 3\sigma$ bound of the voltage at this time.

The same process can be repeated for the entire time axis. This results into the statistical moments of the voltage as a function of time. The first four moments, Mean (μ), Standard Deviation (σ), Skewness (γ), and Kurtosis (κ), of the voltage as a

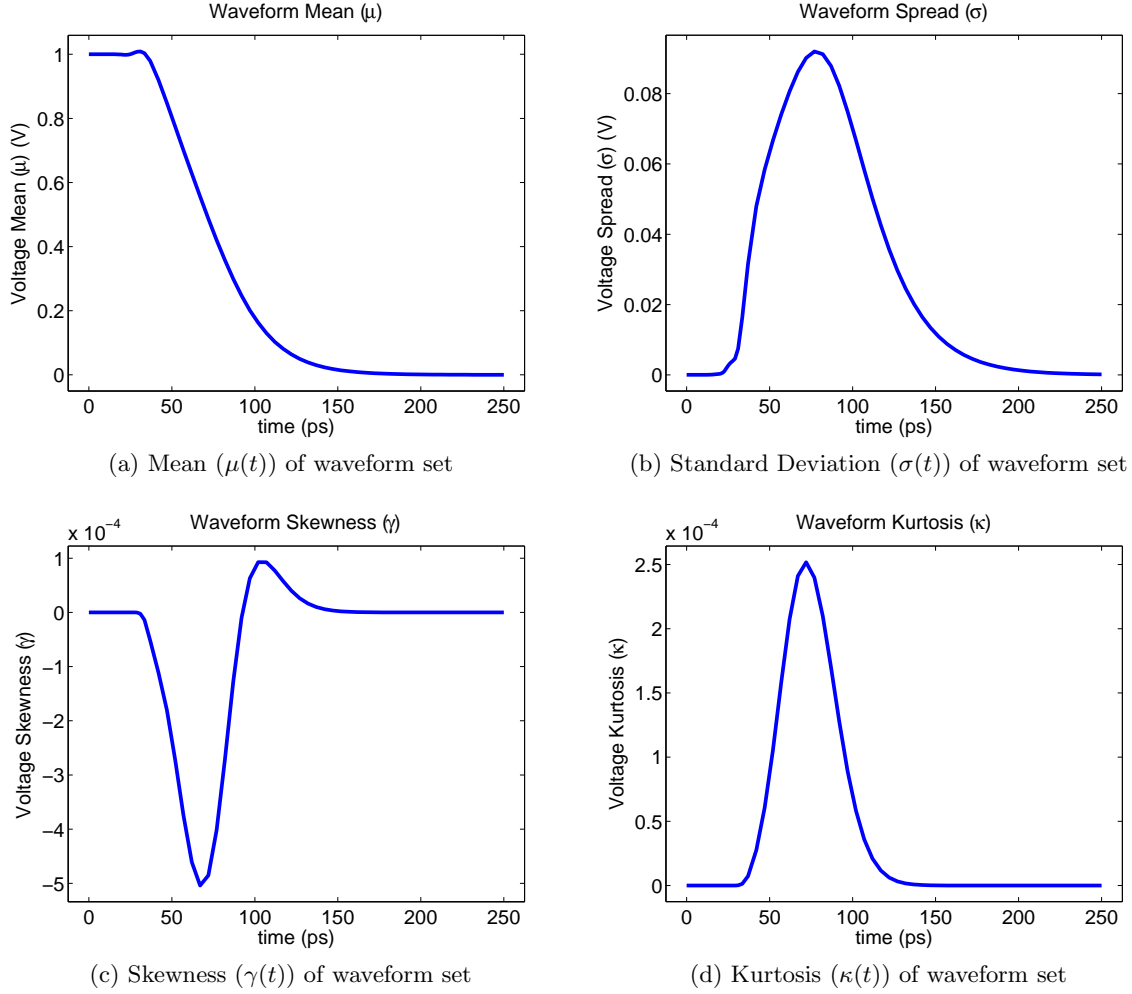


Figure 5.9: First four moments of the waveform set as a function of time

function of the time are shown in Figure 5.9a, Figure 5.9b, Figure 5.9c, and Figure 5.9d respectively. In these figures, time is on the X-axis and the respective moments are on the Y-axis.

Instead of storing the entire waveform set and the corresponding probability values, only the moment curves just described can be stored. The amount of higher order moment curves depends on the SSTA engine and its requirements. As compared to the lookup table based method, this approach requires much less memory, as only few curves are sufficient to store the entire waveform set. The block diagram of an SSTA engine using a statistical moments based representation is shown in Figure 5.10. Here we assume that only the first two moments (μ and σ) are required for the SSTA engine.

5.3.3 Pseudo Circuit based Representation

In SSTA, we need to store the set of output waveforms of any standard cell. Thus, a set of waveforms can be represented by a circuit with the PVT variation in it. Having

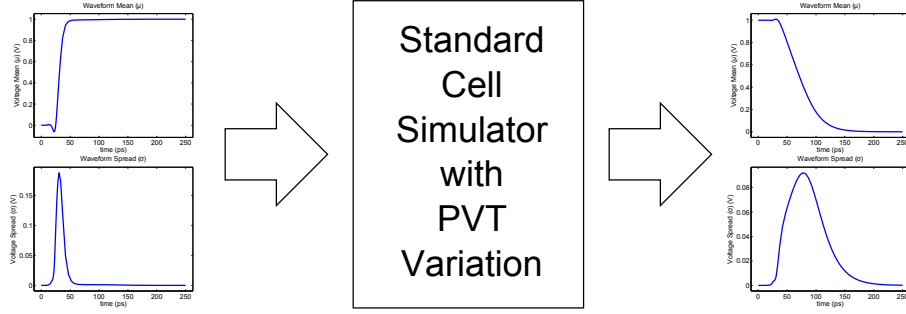


Figure 5.10: SSTA Engine with Moment Model

the circuit and its PVT variations is sufficient to regenerate the voltage waveforms.

We learned in Chapter 2 that the SSTA methodology may be used to estimate the path delay in a digital circuit. Since SSTA is a stage-by-stage approach, the signal at each node is mainly influenced by the part of the circuit between the start point of the data path and that node. An example of a data path is depicted in Figure 5.11. Here, the signal at the second input of the gate g_2 depends on the circuit components from the input of the DFF_2 up to the interconnect n_7 .

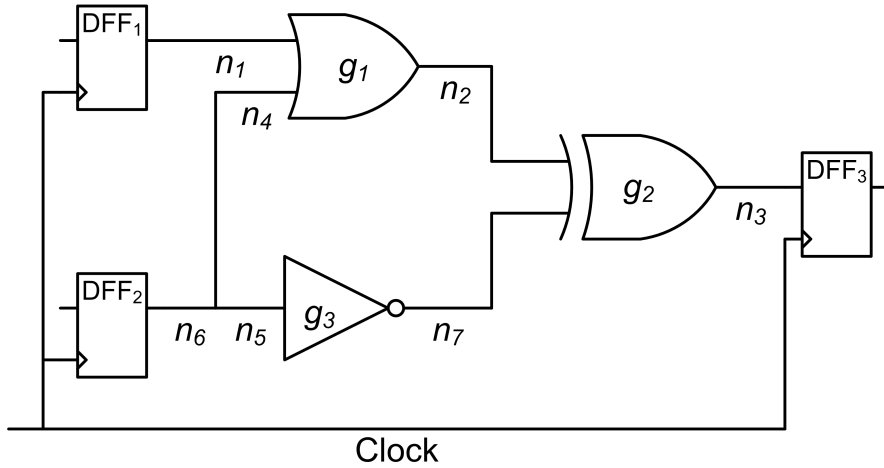


Figure 5.11: Example path for SSTA

In this example, instead of storing the entire circuit up to interconnect n_7 , a reference circuit can be used. This reference circuit can be used to regenerate the similar output waveform. This reference circuit will act as a pseudo circuit for everything up to the input of the gate g_2 . The block diagram of an SSTA engine with this pseudo circuit based representation is shown in Figure 5.12.

Here, the input signals of the gate are replaced by pseudo circuits. Since the MODERN project is using fast SPICE like simulation, this pseudo circuit will be merged with the standard cell circuit during simulation. Thus, the input waveforms are regenerated while doing SSTA.

The proposed idea is to use one pseudo circuit with few parameters associated with it. Since the pseudo circuit is fixed by only a few parameters instead of the entire

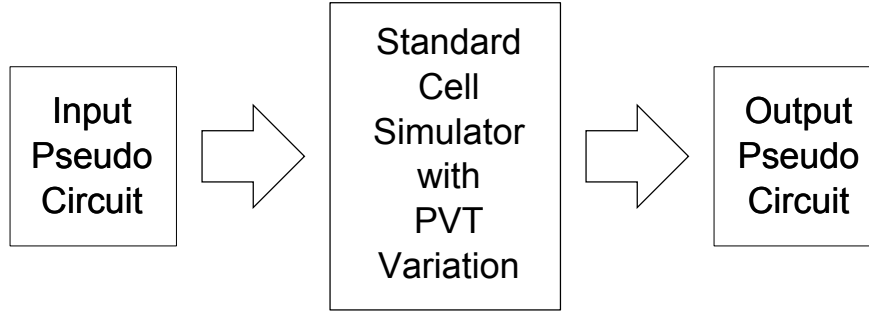


Figure 5.12: SSTA Engine with Pseudo Circuit Model

waveform set and their probabilities, it is a very compact model and able to regenerate a similar set of waveforms as needed at the input of the following standard cell.

5.4 Comparison of various Waveform Set Representations

As discussed in Section 5.3, there are three possible approaches to represent a set of waveforms, namely: lookup table, statistical moments, and pseudo circuit based representation. In the lookup table based approach, the complete waveform set along with their probability density is stored in the lookup table. Considering the large number of waveforms, a very large memory space is required. In the statistical moments based approach, the statistical moments as a function of time are stored. This approach requires a smaller memory space. The SSTA engine in the MODERN project requires higher order statistical moments of the waveforms. The need of higher order moments increases the memory requirement. In the pseudo circuit based approach, a pseudo circuit is used to regenerate the waveform set similarly to the expected waveform set at the input of standard cell during SSTA. A standard reference circuit with few parameters is used to model the pseudo circuit. Due to the very small number of parameters, the pseudo circuit representation is the most compact modeling approach among the three proposed approaches. The details of the pseudo circuit will be discussed in Chapter 6.

5.5 Summary

Similar to the development of the waveform based model from the point based model in STA, there is a need to have a waveform based uncertainty representation in SSTA. Therefore, the MODERN project is preserving various possible output waveforms due to PVT variations at each node. This is called “a set of waveforms”. The concept of the set of waveforms and the uncertainty representation with it are discussed in this chapter. There are three possible methods to represent the set of waveforms in SSTA methodology, namely a lookup table based representation, a statistical moments based representation and a pseudo circuit based representation. The pseudo circuit based representation turns out to be the most compact model because the set of waveforms can be regenerated using a reference circuit with only a few parameters in it. The proposed pseudo circuit is discussed in Chapter 6.

Pseudo Circuit Model

As discussed in Chapter 5, the pseudo circuit based representation for the set of waveforms is the most compact model among the three proposed models. Additionally, the pseudo circuit model can be efficiently integrated with the SSTA engine used in the MODERN project. In this chapter, the concept of the pseudo circuit, its integration with the SSTA engine of the MODERN project, the approach and challenges for the pseudo circuit modelling, a waveform comparison methodology and results are discussed.

The organization of the chapter is as follows: the waveform set, pseudo circuit, and its integration with SSTA engine is discussed in Section 6.1. The proposed pseudo circuit and the simulation database required for waveform comparison are described in Section 6.2. Furthermore, the proposed waveform comparison methodology with an example is described in Section 6.3. The results of the pseudo circuit model are presented in Section 6.4. At the end, a summary of the chapter is presented in Section 6.5.

6.1 Pseudo Circuit, Waveform Set and SSTA Engine

In the SSTA engine of the MODERN project, each input of a gate is a set of waveforms. This set of waveforms can be represented in various ways (see Chapter 5), and the pseudo circuit based representation turns out to be the most compact approach to preserve the set of waveforms. The main purpose of the pseudo circuit is to reconstruct the desired set of waveforms at the input of the standard cell. The SSTA engine in the MODERN project will estimate the delay variation of each gate by simulating the circuit of the gate. The simulation of the pseudo circuit can be carried out along with the gate circuit for reconstruction of the waveforms. Therefore, the pseudo circuit needs to merge with the gate circuit while estimating the delay of the gate.

6.2 The Pseudo Circuit Model

The pseudo circuit model can be divided into two parts. First, we look at the selection of the pseudo circuit and its simulation using Spectre simulator to evaluate the accuracy and efficiency of the pseudo circuit model, which could be embedded to MODERN project. Second, we study processing of the simulation output and constructing a database such that the waveforms can be compared. These two parts of the pseudo circuit model are discussed in the following subsections.

6.2.1 The Pseudo Circuit

The design space of the pseudo circuit is large because it has a very high degree of freedom in the circuit selection. Since the pseudo circuit is not a part of the real digital circuit which needs to be fabricated, some non trivial circuits can also be selected for the design. This further increases the design space of the problem. There are five constraints imposed on the pseudo circuit:

1. The output of the pseudo circuit should be similar to the output of a gate in the digital circuit.
2. The pseudo circuit should be small.
3. The output space of the pseudo circuit should cover the entire possible output range of gates.
4. The absolute transition time of the waveform should match.
5. The pseudo circuit should have PVT variations.

The pseudo circuit will be developed while addressing the constraints mentioned above. First, the output of the pseudo circuit should be similar to the output of the digital circuits. This constraint is required to ensure that the reconstructed waveforms at the input of the gates are similar to the expected input waveforms. The output signal transition of a gate can be explained by charging or discharging of the effective capacitive load at the output pin of the gate. Therefore, an RC circuit is one of the simplest possible pseudo circuits. However a small digital gate can generate a more accurate waveform than an RC circuit. Since the accuracy in the delay estimation is the primary objective of the MODERN project, a digital gate is selected for the pseudo circuit design.

The pseudo circuit is included during simulation of each gate. Therefore, constraining the size of the small circuit will help to ensure a low overhead on the simulation run time. The smallest digital circuit is an inverter. Therefore, an inverter can be selected as a pseudo circuit. As discussed in Chapter 5, the set of waveforms at the output of an inverter chain is more realistic than at the output of a single inverter. Therefore, an inverter chain with three inverters is selected as a pseudo circuit. An inverter chain with three inverters is shown in Figure 6.1. Here Inv_1 , Inv_2 and Inv_3 are the three inverters of the inverter chain.

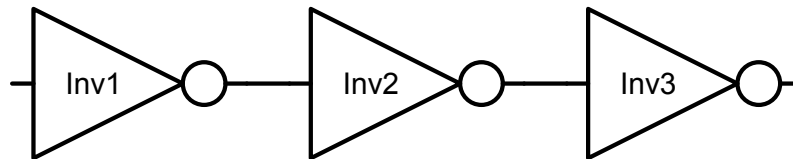


Figure 6.1: An inverter chain with three inverters

The pseudo circuit will be used to represent any possible target waveform set. This target waveform set could be output of any digital cell in the design. Therefore, as

mentioned in the third constraint, the pseudo circuit must be able to generate all such possible sets of waveforms. Quantitative measurement of the waveform set to ensure its full coverage is very complex due to the large number of waveforms. At this stage of the development, only the nominal waveform is considered for quantitative measurement to reduce the complexity of the problem. The nominal waveform is the output of the circuit when all the PVT parameters are at their nominal values. The pseudo circuit should be able to generate all possible nominal waveforms. A single waveform can be qualitatively measured by its slew. Therefore, it can be said that the nominal waveform of the pseudo circuit should be able to ensure the coverage of all possible slew values in the digital circuit. As we know, the output of a gate primarily depends on two external parameters, input signal slew and effective output capacitive load (see Chapter 2). The pseudo circuit should have input signal slew (S_{in}) and output effective capacitive load (C_{load}) as two parameters of the pseudo circuit. An ideal ramp voltage source and a capacitive load are added in the pseudo circuit as shown in Figure 6.2.

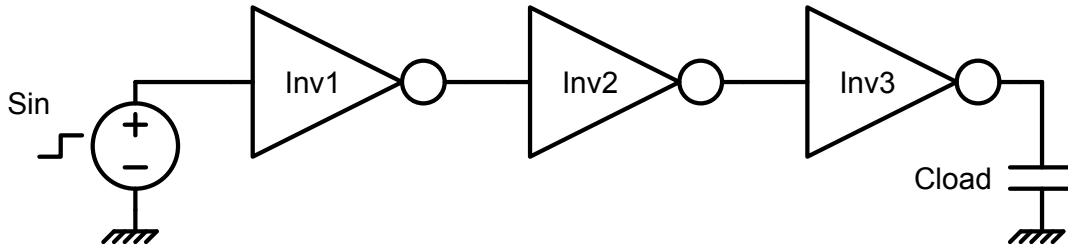


Figure 6.2: Inverter chain with input source and output load

The circuit shown in Figure 6.2 has one limitation. The output load of Inv_1 and Inv_2 is not varying. For a constant capacitive load, the sensitivity of the output signal slew with respect to the input signal slew is very low for a digital circuit. Due to this phenomenon, the variation of the signal slew at the input of Inv_1 is not completely reflected to the input of Inv_2 . Due to a similar phenomenon, the input signal slew of Inv_3 is almost constant. It means that the input signal slew of Inv_3 is not controllable in this circuit. It has been discussed earlier that the pseudo circuit should be able to generate all possible nominal waveforms. Furthermore, the output of a gate depends on the input signal slew and effective output capacitive load. However, in this pseudo circuit, the input signal slew of Inv_3 is not controllable. This limits the generation of all possible waveforms. The problem is addressed by adding additional capacitors at each intermediate node between the inverters. The signal slew at the input of each inverter can be controlled by varying the value of the capacitor. Therefore, both the input signal slew and the intermediate capacitor value will change together. Two independent capacitance values increase the design space by two dimensions. However, the exactly same capacitance value (C_{in}) for these intermediate capacitors can reduce the design space by one dimension while keeping full control on the output capacitance of Inv_2 . Therefore, the same capacitance value (C_{in}) for these intermediate capacitors is used. For each pair of S_{in} and C_{load} , the value of C_{in} is selected such that the input signal slew of Inv_3 is exactly same as the input slew of Inv_1 . The modified circuit is shown in Figure 6.3.

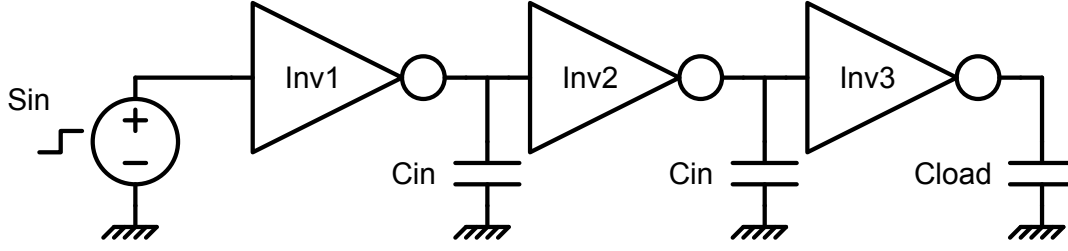


Figure 6.3: Inverter chain with capacitors at internal nodes

The absolute start time of the transition of the output waveform in the circuit shown in Figure 6.3 strongly depends on the start time of the transition of the input signal. The target waveform set does not necessarily have the same start time of transition. Therefore, the fourth constraint is required. The synchronization of the target waveform with the output of pseudo circuit is achieved by delaying the transition of the input ramp signal in the pseudo circuit. The delay in the voltage source is incorporated by introducing a parameter for the signal offset in the voltage source as shown in Figure 6.4. This time is called as T_{start} .

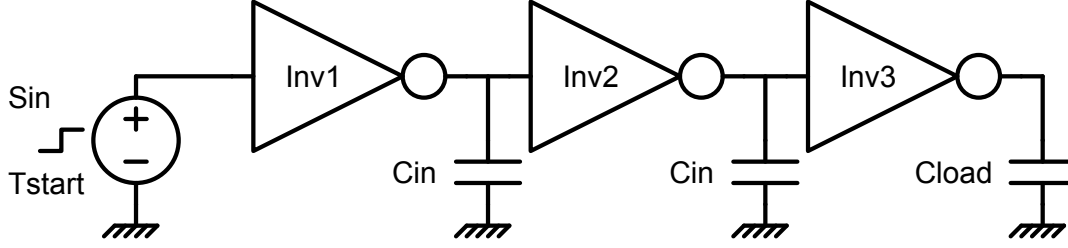


Figure 6.4: Inverter chain with time offset

Now, the circuit shown in Figure 6.4 is capable of generating all possible nominal output waveforms. This circuit has four parameters: input signal slew, output capacitive load, internal capacitors, and start time of the ramp signal. However, the value of internal capacitors are dependent on the input signal slew and output capacitive load. Therefore, the circuit has only three independent parameters.

The fifth constraint for the pseudo circuit design is that it must have PVT variations. This constraint is needed to generate variation in the waveforms. Since the PVT variations have many parameters, the selection of the parameters for the PVT variation has a very high degree of freedom. The intuitive approach would be to add variations in all the parameters because this scenario will increase the accuracy of the variations in the output waveforms. However, the complexity of the simulation run will increase. Additionally, the required iterations for the circuit simulation will increase exponentially. Therefore, only one parameter variation is selected to represent the variation in the pseudo circuit. This choice increases the simulation run time at the cost of limited variations in the output waveforms. In this phase of the development of the pseudo circuit model, such approximations are very useful for the analysis of the model. However, in future work, multiple parameters can be considered to increase the accuracy. In this pseudo circuit, the channel length of both the PMOS (L_p) and

the NMOS (L_n) in the last inverter of the inverter chain is considered to have process variations. The channel lengths L_n and L_p are assumed to have a correlation equal to one. The variation in the channel length is represented by its standard deviation (σ_L). The complete pseudo circuit is shown in Figure 6.5 and the circuit parameters are reported in Table 6.1.

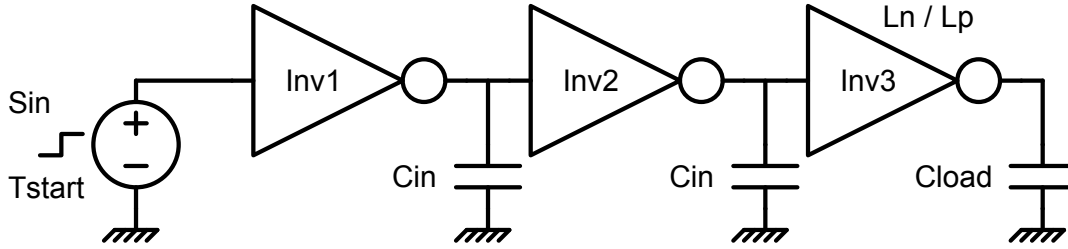


Figure 6.5: The Pseudo Circuit Schematic

Table 6.1: The Pseudo Circuit

Name	Description
Circuit	Inverter chain with three inverters and capacitive load at each node
Input Signal	Ideal Ramp
Output Load	Capacitive Load
Sweep	Input Slew and Output Capacitive Load
Variations	Channel length in last inverter
Starting Time	Based on the transition time of the target waveform

The final pseudo circuit has four independent parameters and one dependent parameter. The names of all the parameters, their range and the sizes of the MOSFETs are given in Table 6.2. These values are selected after various experiments while keeping the design constraints in mind.

Table 6.2: Pseudo Circuit Parameters

Symbol	Parameter	Value / Range	Remark
S_{in}	Input Signal Slew	[50 ps , 500 ps]	Independent
C_{load}	Output Capacitive Load	[1 fF , 15 fF]	Independent
C_{in}	Internal Capacitive Load	[0.1 fF , 14 fF]	Dependent
σ_L	Spread of Channel Length	[3.33% , 16.67%]	Independent
T_{start}	Offset in input signal	0 , > 0	Independent
L_n / W_n	NMOS length and width	100 nm / 90 nm	Fix
L_p / W_p	PMOS length and width	100 nm / 135 nm	Fix

The spread in the channel length (σ_L) directly controls the spread in the output waveform. Therefore, the channel length spread (σ_L) can attain various possible values in the specified range to generate various possible target waveform sets. Multiple simulations of the pseudo circuit are needed to cover the entire range of σ_L , which results in very high run time. This problem is addressed by using the proposed method

of Fast Statistical Moment Estimation (FSME) (see Chapter 4). The FSME method gives the flexibility to select the pdf of the parameters after the circuit simulation. Due to this feature, the pseudo circuit is simulated only once with the maximum possible parameter spread (σ_L) and the output is stored in a database. The actual spread of the channel length variation is used during waveform comparison. The FSME method reduces the rerun of the circuit simulation for various possible values of σ_L .

The pseudo circuit is simulated for the sampled values of S_{in} and C_{load} from their respective range using the highest value of σ_L in the specified range. The output waveforms of each simulation along with their circuit configurations are stored in a database. This database is used during the SSTA flow to estimate the pseudo circuit parameters such that the target waveforms can be generated. The simulation output and its structure in the database are given in Table 6.3. Here index i is used for input slew, j for output load, k for the time, and l for simulation iteration due to uniform sampling of the channel length as given in FSME method.

Table 6.3: Simulation output in database

Sampled Data	Function	Remark
$S_{in}[i]$	S_{in}	Input Slew
$C_{load}[j]$	C_{load}	Output Load
$C_{in}[i, j]$	$C_{in}(S_{in}, C_{load})$	Internal Capacitor
$Wave[i, j, k, l]$	$Wave(S_{in}, C_{load}, t, \sigma_L)$	Waveform Output

6.2.2 Database processing for waveform comparison

The comparison of the waveform sets is in itself a very challenging task due to the fact that a waveform set is a dataset with five dimensions (S_{in} , C_{load} , t , σ_L , $Voltage$). To reduce the complexity of the problem, only mean and standard deviation curves of the waveforms are compared. Additionally, instead of comparing the entire mean and standard deviation curve, only their quality factors are compared. Here, quality factors are the specific parameters which can quantitatively measure the shape of waveforms. Various possible quality factors have been analysed to compare mean and standard deviation curves. Since we have four independent parameters in the pseudo circuit, at least four quality factors are required to estimate these independent parameters. The selected quality factors of the mean and standard deviation curves are discussed below.

1. The slew of the mean curve (Q_{Slew})
2. The separation of mean and standard deviation curves ($Q_{ShiftMean}$)
3. The peak height of the standard deviation curve (Q_{Max})
4. The $V_{DD} / 2$ crossing time of the mean curve (Q_{Tmid})

Here the prefix Q stands for quality factor. The quality factors are demonstrated with an example of the mean and standard deviation curve in Figure 6.6. The details of the quality factors are discussed below using a set of waveforms. The notations of

the mean and standard deviation curves of the waveform set used in this illustration are given in Table 6.4. Here k and l are the integral index of the sampled array of the continuous functions.

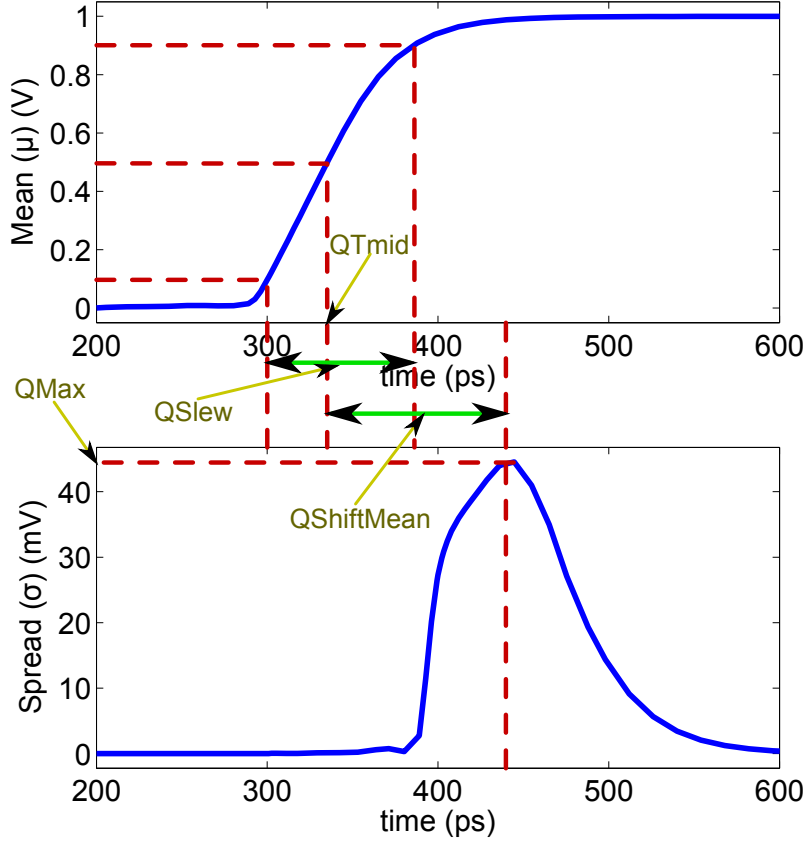


Figure 6.6: Quality factors for the waveform comparison

Table 6.4: Mean and SD curves for quality factor illustration

Sampled Data	Function	Remark
$M[k, l]$	$Mean(t, \sigma_L)$	Waveform Mean
$S[k, l]$	$Sigma(t, \sigma_L)$	Waveform Standard Deviation
$MU[k]$	$MaxMean(t)$	Waveform Mean with maximum spread
$SD[k]$	$MaxSigma(t)$	Waveform SD with maximum spread
$T[k]$	t	time

The total number of samples in the time axis, i.e. the length of the vector $T[k]$ is n .

$$n = \text{length}(T) \quad (6.1)$$

Q_{Slew} is the slew of the mean curve of the waveform set. Since the effect of the positive and negative variation in the channel length are opposite, it keeps the nominal

and mean curve very close to each other. Therefore the Q_{Slew} of the mean curve due to various values of σ_L remains consistent. In the waveform comparison methodology, the slew of the mean curve due to maximum spread ($MU[k]$) is used to define the Q_{Slew} of the set of waveforms. Therefore, the Q_{Slew} is independent of the spread used in the pseudo circuit.

Mathematically, Q_{Slew} can be define as follows:

$$Q_{Slew} = |t_{90} - t_{10}| \quad (6.2)$$

Here, t_{10} and t_{90} are the 10% and 90% voltage crossing time of $MU[k]$ respectively. The crossing time t_{10} and t_{90} can be defined as follows:

$$t_{10} = t \quad \text{such that } MaxMean(t) = 0.1 \times V_{DD} \quad (6.3)$$

$$t_{90} = t \quad \text{such that } MaxMean(t) = 0.9 \times V_{DD} \quad (6.4)$$

QShiftMean is a measure of separation between mean and standard deviation curves. The position of the mean curve is defined by its 50% voltage crossing time. Similar to Q_{Slew} , the mean curve due to maximum spread in the channel length is used to define the position of the mean curve. The position of the standard deviation curve is defined by the weighted mean of time while considering the standard deviation curve as a weight profile. The weight of a sampled time can be estimated by integrating the weight profile. Since the standard deviation of the output waveform is directly proportional to the spread of the channel length, the weighted mean of time does not change due to change in the spread of the channel length. Therefore, the standard deviation curve due to maximum spread is used to measure the position of the standard deviation curve.

Mathematically, $Q_{ShiftMean}$ can be defined as follows:

$$Q_{ShiftMean} = t_{sd} - t_{mu} \quad (6.5)$$

Here, t_{sd} and t_{mu} are the positions of the $SD[k]$ and $MU[k]$ curves respectively. The position of $MU[k]$ can be defined as follows:

$$t_{mu} = t \quad \text{such that } MaxMean(t) = 0.5 \times V_{DD} \quad (6.6)$$

Here, t_{mu} is the 50% voltage crossing time of $MU[k]$. The position of $SD[k]$ can be defined as follows:

$$t_{sd} = \frac{\sum_k (T[k] \times W[k])}{\sum_k W[k]} \quad (6.7)$$

Here, t_{sd} is the weighted sum of time ($T[k]$) with its weight $W[k]$. Since $SD[k]$ is the weight profile of time ($T[k]$), weight ($W[k]$) can be calculated by integrating the weight profile ($SD[k]$) in the adjacent non-overlapping windows (local windows) at every sampled time. If $MaxSigma(t)$ is the weight profile function, then $MaxSigma(t)$ and $SD[k]$ can be related as:

$$SD[k] = MaxSigma(T[k]) \quad (6.8)$$

Here, $SD[k]$ contains the sampled values of the continuous function $MaxSigma(t)$. Now, if the lower bound of the local window for $T[k]$ is $t_l[k]$ and the upper bound is $t_h[k]$, then the weight of time $T[k]$ can be define as:

$$W[k] = \int_{t_l[k]}^{t_h[k]} MaxSigma(t) dt \quad (6.9)$$

Integration is a compute intensive process. Therefore, a piecewise constant approximation is used to speedup the integration for calculation of weight ($W[k]$). The modified equation with piecewise constant approximation is given below:

$$W[k] = SD[k] \times (t_h[k] - t_l[k]) \quad (6.10)$$

$$\Rightarrow W[k] = SD[k] \times \Delta T[k] \quad (6.11)$$

Here, $\Delta T[k]$ is the length of the local window around the time sample $T[k]$. The lower and upper bound of this local window can be calculated as follows:

$$t_l[k] = \begin{cases} T[1] - \frac{T[2]-T[1]}{2} & \text{if } k = 1 \\ \frac{T[k-1]+T[k]}{2} & \text{if } k \neq 1 \end{cases} \quad (6.12)$$

and

$$t_h[k] = \begin{cases} \frac{T[k]+T[k+1]}{2} & \text{if } k \neq n \\ T[n] + \frac{T[n]-T[n-1]}{2} & \text{if } k = n \end{cases} \quad (6.13)$$

Therefore, $\Delta T[k]$ can be define as:

$$\Delta T[k] = t_h[k] - t_l[k] \quad (6.14)$$

$$\Rightarrow \Delta T[k] = \begin{cases} t[2] - t[1] & \text{if } k = 1 \\ \frac{t[k+1]-t[k-1]}{2} & \text{if } k \neq 1 \& k \neq n \\ t[n] - t[n-1] & \text{if } k = n \end{cases} \quad (6.15)$$

Q_{Max} is the maximum spread of the waveform. Since the spread in the output waveform is directly proportional to the spread of the channel length (σ_L), Q_{Max} is a function of σ_L . For each value of the spread, Q_{Max} is defined as:

$$Q_{Max}(\sigma_L) = Max(Sigma(t, \sigma_L)) \quad (6.16)$$

Q_{Max} on the sampled database can be designed as follows:

$$Q_{Max}[l] = Max(S[k, l]) \quad (6.17)$$

During the experiments, it has been found that Q_{Max} is a linear function of σ_L . Therefore, Q_{Max} can be decomposed as follows:

$$Q_{Max} = Q_{MaxM} \times \sigma_L + Q_{MaxC} \quad (6.18)$$

Here, Q_{MaxM} is the slope of the line and Q_{MaxC} is the intersection of line with the Q_{Max} axis.

Q_{Tmid} is the 50% voltage crossing time of the mean curve. Q_{Tmid} is the same as T_{mu} in (6.6) which is redefined as

$$Q_{Tmid} = t \quad \text{such that } MaxMean(t) = 0.5 \times V_{DD} \quad (6.19)$$

This parameter is used to measure the absolute start time of the transition of the waveform set. The database of the simulation outputs of the pseudo circuit is generated with T_{start} equal to zero.

$$T_{start} = 0 \quad \text{for pseudo circuit database} \quad (6.20)$$

Therefore, Q_{Tmid} corresponds to $T_{start} = 0$. The difference in the absolute start time of the transition in the target waveform set is compensated by changing the T_{start} parameter of the pseudo circuit model which will be used in the SSTA flow.

The quality factors of one set of waveforms are discussed here. However, the simulation database has waveforms for each pair of input signal slew (S_{in}) and output capacitive load (C_{load}). Therefore the quality factors are calculated for each pair of S_{in} and C_{load} . The extended structure of the database is given in Table 6.5. Index i is used for input slew, j for output load, k for the time, and l for simulation iteration due to process variation.

Table 6.5: Database Structure

Sampled Data	Function	Remark
$S_{in}[i]$	S_{in}	Input slew
$C_{load}[j]$	C_{load}	Output load
$C_{in}[i, j]$	$C_{in}(S_{in}, C_{load})$	Internal capacitor
$T[k]$	t	Time
$Wave[i, j, k, l]$	$Wave(S_{in}, C_{load}, t, \sigma_L)$	Waveform output
$Q_{Slew}[i, j]$	$Q_{Slew}(S_{in}, C_{load})$	Slew of mean curve
$Q_{ShiftMean}[i, j]$	$Q_{ShiftMean}(S_{in}, C_{load})$	Separation of mean and sigma curve
$Q_{MaxM}[i, j, l]$	$Q_{MaxM}(S_{in}, C_{load}, \sigma_L)$	Peak of sigma curve
$Q_{MaxM}[i, j]$	$Q_{MaxM}(S_{in}, C_{load})$	Slope of the Q_{Max} vs σ_L
$Q_{MaxC}[i, j]$	$Q_{MaxC}(S_{in}, C_{load})$	Q_{Max} axis intersection of Q_{Max} vs σ_L
$Q_{Tmid}[i, j]$	$Q_{Tmid}(S_{in}, C_{load})$	$V_{DD}/2$ crossing time of mean curve

6.3 Waveform Comparison Methodology

A database of the pseudo circuit simulation outputs and a targeted waveform set are the input for the pseudo circuit model generation flow. The generated pseudo circuit from this flow can be used by the SSTA engine to reconstruct the target waveform set at the input of the gate. It is discussed earlier in this chapter that the output waveform of the pseudo circuit should closely match with the target waveform. It has also been discussed that the quality factors will be used to estimate the configuration of the pseudo circuit. The quality factors and circuit parameters are given below:

- Quality Factor
 1. Q_{Slew}
 2. $Q_{ShiftMean}$
 3. Q_{Max}
 4. Q_{Tmid}
- Circuit Parameters
 1. S_{in}
 2. C_{load}
 3. C_{in}
 4. σ_L
 5. T_{start}

The quality factors of the targeted waveform set are calculated first. Let us name them T_{Slew} , $T_{ShiftMean}$, T_{Max} , and T_{Tmid} . Here the prefix T stands for target waveform. The methodology to match the quality factors and circuit parameter estimation is illustrated with an example.

A target waveform set, its mean and standard deviation curves are shown in Figure 6.7 and the quality factors are reported in Table 6.6.

Table 6.6: Quality factors of target waveform

Quality Factor	Value
T_{Slew}	76.1 ps
$T_{ShiftMean}$	9.7 ps
T_{Max}	89.6 mV
T_{Tmid}	307.0 ps

The quality factor Q_{Slew} of the database is a three dimensional dataset of the slew of the mean curve of the waveforms for each pair of input signal slew (S_{in}) and output capacitive load (C_{load}) as shown in Figure 6.8a. In this figure C_{load} is on the X-axis, S_{in} is on the Y-axis, and Q_{Slew} is on the Z-axis. Each node of the plane is the Q_{Slew} due to corresponding pair of S_{in} and C_{load} . The target slew (T_{Slew}) can also be represented in a three dimensional space using a plane parallel to the S_{in} - C_{load} plane. Both the Q_{Slew} and T_{Slew} planes are shown in Figure 6.8b.

Since the T_{Slew} plane is parallel to the S_{in} - C_{load} plane, the intersection line of Q_{Slew} and T_{Slew} planes will be a line in a two dimensional space of S_{in} - C_{load} as shown in Figure 6.8c. In this figure, C_{load} is on the X-axis and S_{in} is on the Y-axis. The red points in this figure are the intersection points of two planes. These points are estimated with the help of a numerical analysis method. The black line is a linear best fit curve of the intersection points. The linear relation of S_{in} and C_{load} for this intersection line is given in the title of the figure. Each point in this line can generate a configuration of the pseudo circuit which will produce the Q_{Slew} equal to T_{Slew} .

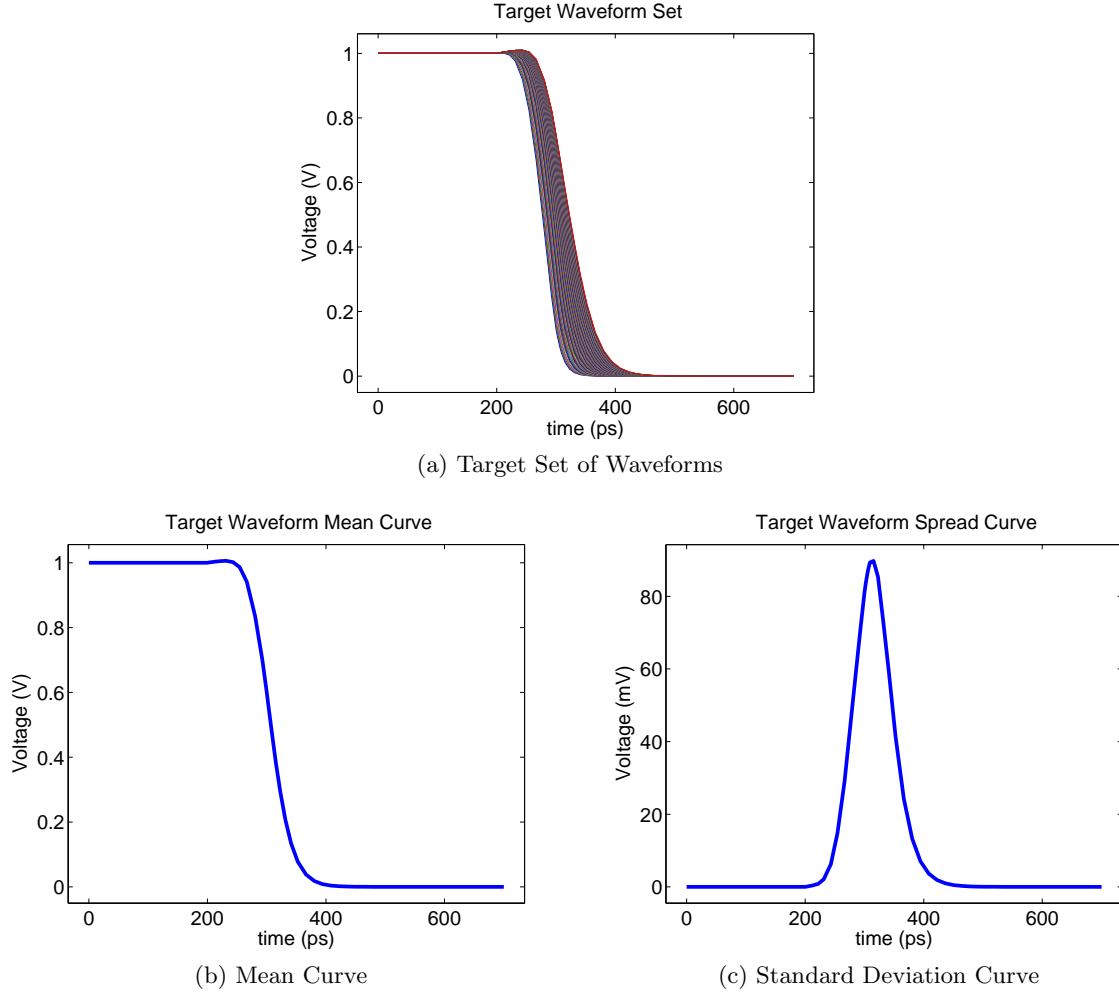
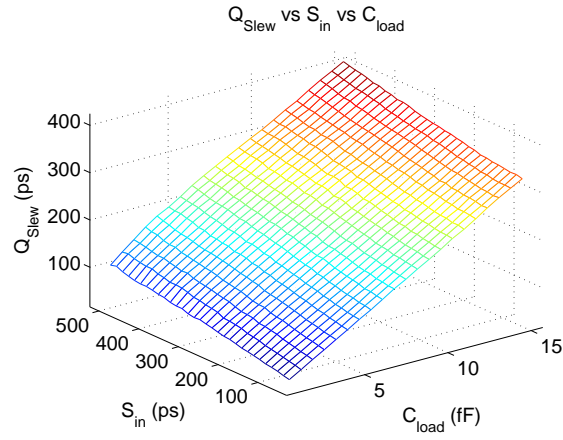


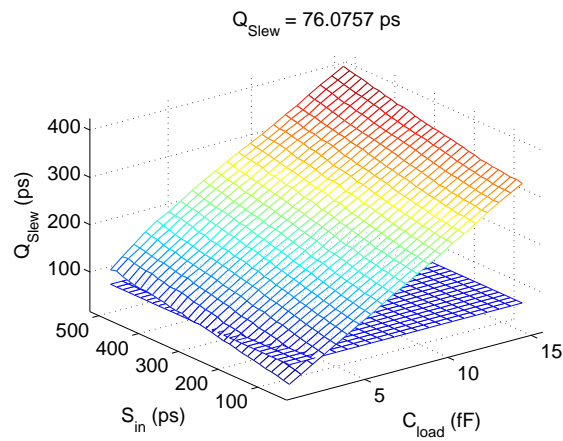
Figure 6.7: Target Waveform Set

Similar to Q_{Slew} , $Q_{ShiftMean}$ is also a three dimensional dataset of the separation between mean curve and standard deviation curve of the waveform set for each pair of C_{in} and C_{load} as shown in Figure 6.9a. In this figure, C_{load} is on the X-axis, S_{in} is on the Y-axis, and $Q_{ShiftMean}$ is on the Z-axis. Each node of the plane is the $Q_{ShiftMean}$ due to corresponding pair of S_{in} and C_{load} . The target separation between mean and standard deviation curve ($T_{ShiftMean}$) can also be represented in a three dimensional space using a plane parallel to the S_{in} - C_{load} plane. Both the $Q_{ShiftMean}$ and $T_{ShiftMean}$ planes are shown in Figure 6.9b.

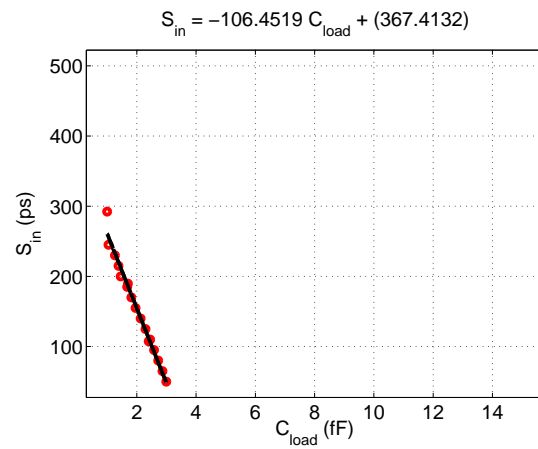
Similar to the T_{Slew} plane, the $T_{ShiftMean}$ plane is also parallel to the S_{in} - C_{load} plane, its intersection with $Q_{ShiftMean}$ plane will be a line in a two dimensional space of S_{in} - C_{load} as shown in Figure 6.9c. In this figure, C_{load} is on the X-axis and S_{in} is on the Y-axis. The red points in this figure are the intersection points of two planes. These points are estimated with the help of a numerical analysis method. The black line is a linear best fit curve of the intersection points. The linear relation of S_{in} and C_{load} for this intersection line is given in the title of the figure. Each point in this line can



(a) Q_{Slew} vs S_{in} vs C_{load}



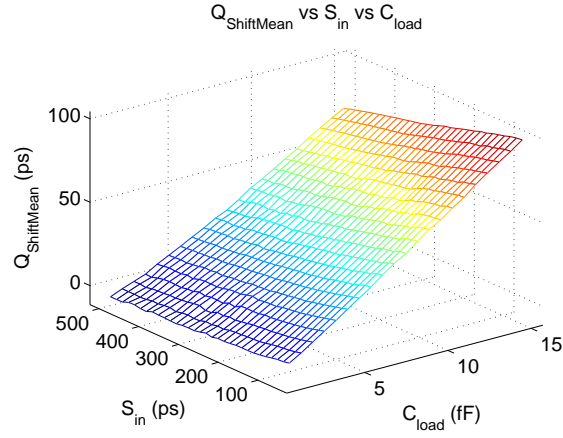
(b) Q_{Slew} and T_{Slew} Planes



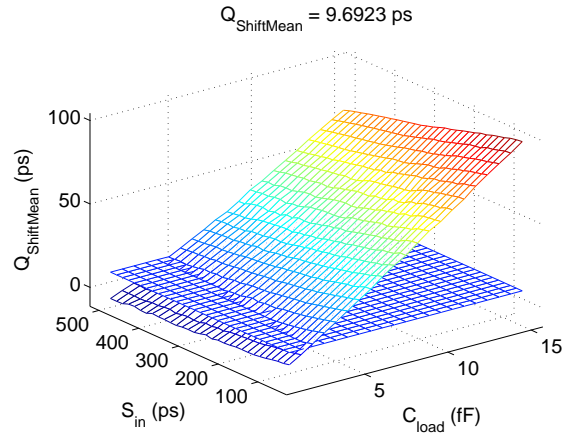
(c) Intersection of Q_{Slew} and T_{Slew} Planes

Figure 6.8: S_{in} vs C_{load} interaction line for Q_{Slew} equals to T_{Slew}

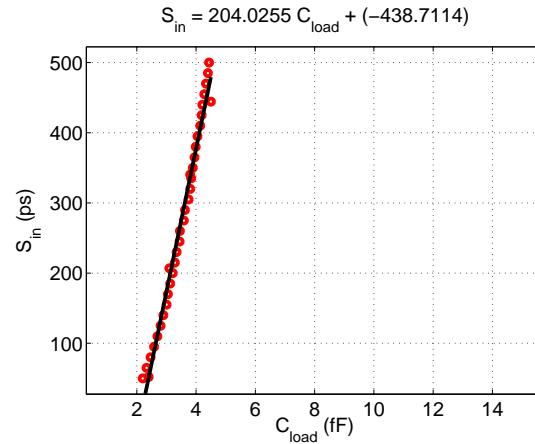
generate a configuration of the pseudo circuit which will produce the $Q_{ShiftMean}$ equal to $T_{ShiftMean}$.



(a) $Q_{ShiftMean}$ vs S_{in} vs C_{load}



(b) $Q_{ShiftMean}$ and $T_{ShiftMean}$ Planes



(c) Intersection of $Q_{ShiftMean}$ and $T_{ShiftMean}$ Planes

Figure 6.9: S_{in} vs C_{load} intersection line for $Q_{ShiftMean}$ equals to $T_{ShiftMean}$

The intersection of Q_{Slew} with T_{Slew} and $Q_{ShiftMean}$ with $T_{ShiftMean}$ gives two lines in the S_{in} - C_{load} plane which satisfy the individual quality factors. The intersection of these two lines will give a pair of S_{in} - C_{load} which will satisfy both the quality factors simultaneously as shown in Figure 6.10. In this figure, the lines of Figure 6.8c and Figure 6.9c are plotted together. The intersection point is given in the title of the figure. Let us call the intersection value of S_{in} and C_{load} MS_{in} and MC_{load} respectively. Here the prefix M stands for pseudo circuit model. These MS_{in} and MC_{load} are the two parameters of the pseudo circuit which satisfy the quality factors T_{Slew} and $T_{ShiftMean}$.

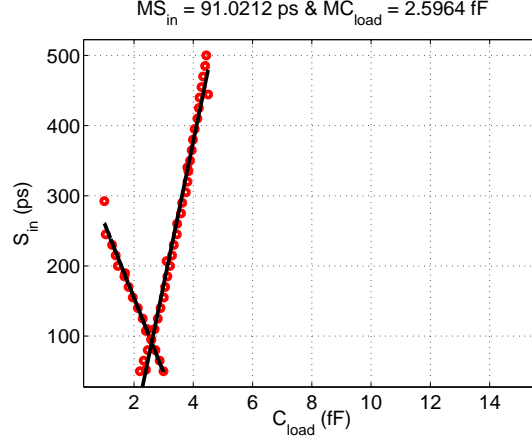


Figure 6.10: Intersection of Slew and ShiftMean lines

The datasets C_{in} , Q_{MaxM} , Q_{MaxC} , and Q_{Tmid} are a function of S_{in} and C_{load} as shown in Table 6.5. These functions are stored in the database using three dimensional matrix formats. The values of C_{in} , Q_{MaxM} , Q_{MaxC} , and Q_{Tmid} for the corresponding model parameters MS_{in} and MC_{load} can be estimated using an interpolation function. Let us call these interpolated values MC_{in} , MQ_{MaxM} , MQ_{MaxC} , and MT_{mid} such that

$$MC_{in} = C_{in}(MS_{in}, MC_{load}) \quad (6.21)$$

$$MQ_{MaxM} = Q_{MaxM}(MS_{in}, MC_{load}) \quad (6.22)$$

$$MQ_{MaxC} = Q_{MaxC}(MS_{in}, MC_{load}) \quad (6.23)$$

$$MT_{mid} = Q_{Tmid}(MS_{in}, MC_{load}) \quad (6.24)$$

As we discussed earlier, Q_{MaxM} and Q_{MaxC} are the coefficients of the linear function of Q_{Max} vs σ_L . The equation is rewritten below.

$$Q_{Max} = Q_{MaxM} \times \sigma_L + Q_{MaxC} \quad (6.25)$$

The value of σ_L for the target T_{Max} can be estimated using the following equation. Here the estimated value of σ_L is called $M\sigma_L$.

$$M\sigma_L = \frac{T_{Max} - MQ_{MaxC}}{MQ_{MaxM}} \quad (6.26)$$

$$M\sigma_L = (T_{Max} - MQ_{MaxC})/MQ_{MaxM} \quad (6.27)$$

The linear relation of Q_{Max} versus σ_L is plotted in Figure 6.11. Here, σ_L is on the X-axis and Q_{Max} is on the Y-axis. The values of T_{Max} and $M\sigma_L$ are given in the title of the figure.

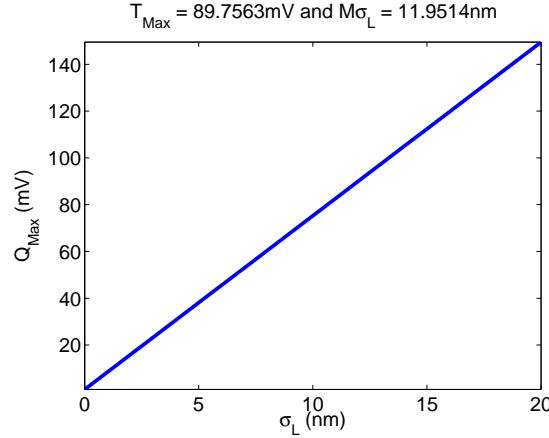


Figure 6.11: Q_{Max} vs σ_L

The 50% crossing time of the mean curve of the waveform set corresponding to the selected MS_{in} and MC_{load} is M_{Tmid} . The same for the target waveform is T_{Tmid} . Since zero offset in the absolute starting time of the ideal ramp is used in the pseudo circuit, M_{Tmid} should be smaller than T_{Tmid} . The delay compensation, required for the synchronization of the absolute time, is equal to the difference of the T_{Tmid} and M_{Tmid} . This time value is used to offset the start time of the input signal transition and called MT_{start} .

$$MT_{start} = T_{Tmid} - M_{Tmid} \quad (6.28)$$

The pseudo circuit configuration parameters for the target waveform are estimated using the methodology discussed above. The output of the pseudo circuit with this circuit configuration is expected to match the specified quality factors. However, there is one problem which is not yet addressed. This is about the loading of the pseudo circuit by the standard cell in the design. This limitation is addressed by the compensation method discussed below.

Load Compensation During the pseudo circuit database generation, the pseudo circuit was loaded by the C_{load} only. Whereas, during the simulation of the pseudo circuit in the SSTA flow, the input impedance of the gate is added in parallel with C_{load} . The extra load is compensated by reducing the C_{load} by the effective input capacitance of the standard cell.

Mathematically, the load compensation is described below. Let us say that TC_{load} is the effective input capacitive load of the standard cell. The MC_{load} is reduced by TC_{load} value to compensate the extra load due to the gate as follows:

$$MC_{load} = MC_{load} - TC_{load} \quad (6.29)$$

This ends the methodology for the waveform set comparison and selection of the pseudo circuit parameters. The list of all the pseudo circuit parameters with their

values is given in Table 6.7.

Table 6.7: Pseudo Circuit Parameters after waveform comparison methodology

Parameter	Value	Remark
MS_{in}	91.0 ps	Slew of the input ramp signal
MC_{load}	2.6 fF	Capacitive load of the pseudo circuit
MC_{in}	1.7 fF	Internal capacitive load of the inverters
$M\sigma_L$	12.0 nm	Variation spread of the channel length
MT_{start}	47.6 ps	Offset in the start time of the simulation

6.4 Results

Firstly, the output of the pseudo circuit model for the example target waveform set used in Section 6.3 will be discussed. Following this, results from various other sets of waveforms will be presented briefly.

The target set of waveforms shown in Figure 6.7 is used in the previous section to develop the waveform comparison methodology. The quality factors corresponding to the target waveform are given in Table 6.6 and the pseudo circuit parameters based on the discussed waveform comparison methodology are given in Table 6.7.

The output of the pseudo circuit with estimated circuit parameters is generated. The target set of waveforms is plotted in Figure 6.12a. The output of the pseudo circuit model output with corresponding pseudo circuit parameters is plotted in Figure 6.12b. Here time is on the X-axis and mean voltage (μ) in on the Y-axis. The output of the pseudo circuit model is called the model set of waveforms. It can be observed that the target waveform set and model waveform set do not match exactly. This is primarily because only quality factors are compared to select the parameters of pseudo circuit. Therefore, it is expected that the quality factors should match very closely. Furthermore, the mean curve and standard curve of these two waveform sets are expected to match closely, which is discussed further below.

The mean curve of the target waveform set and pseudo circuit model output are plotted in Figure 6.13a. Here time is on the X-axis and mean voltage (μ) in on the Y-axis. The blue curve is the mean curve of the target waveform set and red curve is the output of the pseudo circuit model. It can be seen that the output of the pseudo circuit model matches quite well. The error in the mean curve of the pseudo circuit with respect to the mean curve of the target waveform set is plotted in Figure 6.13c. Here time is on the X-axis and error in the mean curve (ε_μ) is on the Y-axis. The supply voltage is used as a reference to calculate the error percentage of the difference in the mean curves, which is given as

$$\varepsilon_\mu = 100 \times \frac{\mu_T - \mu_M}{V_{DD}} \quad (6.30)$$

Here, ε_μ is the error percentage and V_{DD} is the supply voltage. The mean curve of the target waveforms is represented by μ_T and μ_M represents the mean curve of the pseudo circuit model waveforms. It can be seen that the error is within 1.6%.

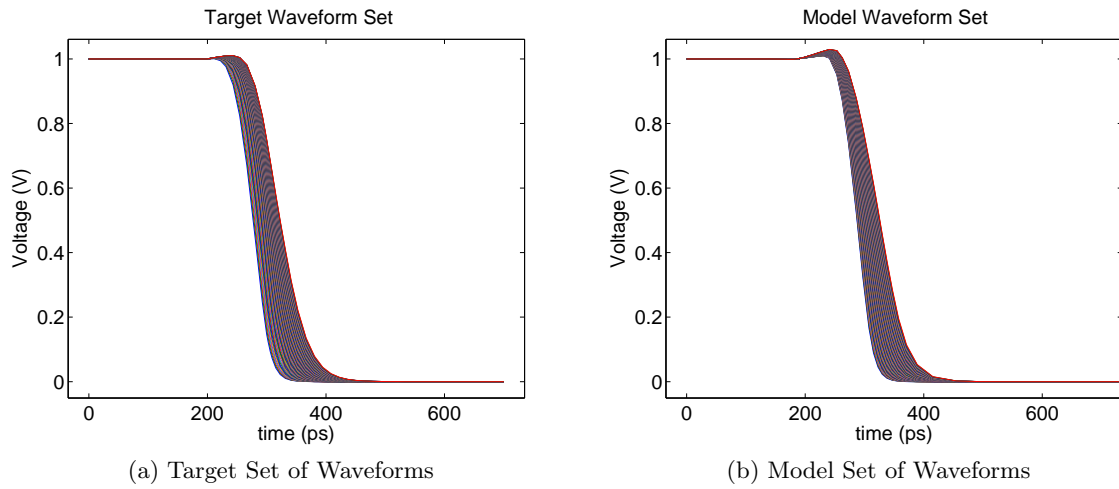


Figure 6.12: Comparison of pseudo circuit model and target set of waveforms

Similarly to the mean curve, the standard deviation curve of the model waveform and target waveform are plotted in Figure 6.13b. Here time is on the X-axis and standard deviation (σ) is on the Y-axis. The blue curve is the σ curve of the target waveform and red curve is corresponding to the pseudo circuit model. The σ curves also fit very well. The error in the σ curve is plotted in Figure 6.13d. Here time is on the X-axis and error in the σ curve (ε_σ) is on the Y-axis. The Q_{Max} (maximum value of the σ curve) in the target waveform is used as a reference for error percentage calculation, which is given as

$$\varepsilon_\sigma = 100 \times \frac{\sigma_T - \sigma_M}{Q_{Max}} \quad (6.31)$$

Here, ε_σ is the error percentage, σ_T is the σ curve for the target waveforms and σ_M is the σ curve for the pseudo circuit model waveforms. It can be seen that the error is within 8.2%.

The four quality factors of the target waveform and pseudo circuit model waveform along with the error percentages are given in Table 6.8.

Table 6.8: Quality factors of target waveform set and pseudo circuit model

Quality Factor	Target Waveform	Pseudo Circuit Model Waveform	Error %
Q_{Slew}	76.08 ps	74.82 ps	1.66 %
$Q_{ShiftMean}$	9.70 ps	10.41 ps	7.42 %
Q_{Max}	89.76 mV	88.83 mV	1.03 %
Q_{Tmid}	306.93 ps	307.34 ps	0.13 %

Various other waveform sets have been evaluated and the error percentages in the quality factors are within 10%.

However, it has been observed that that some sets of waveforms generate negative values of the pseudo circuit parameters. This is due to the fact that the line of intersection of Q_{Slew} and the T_{Slew} plane intersect with the line of intersection of $Q_{ShiftMean}$

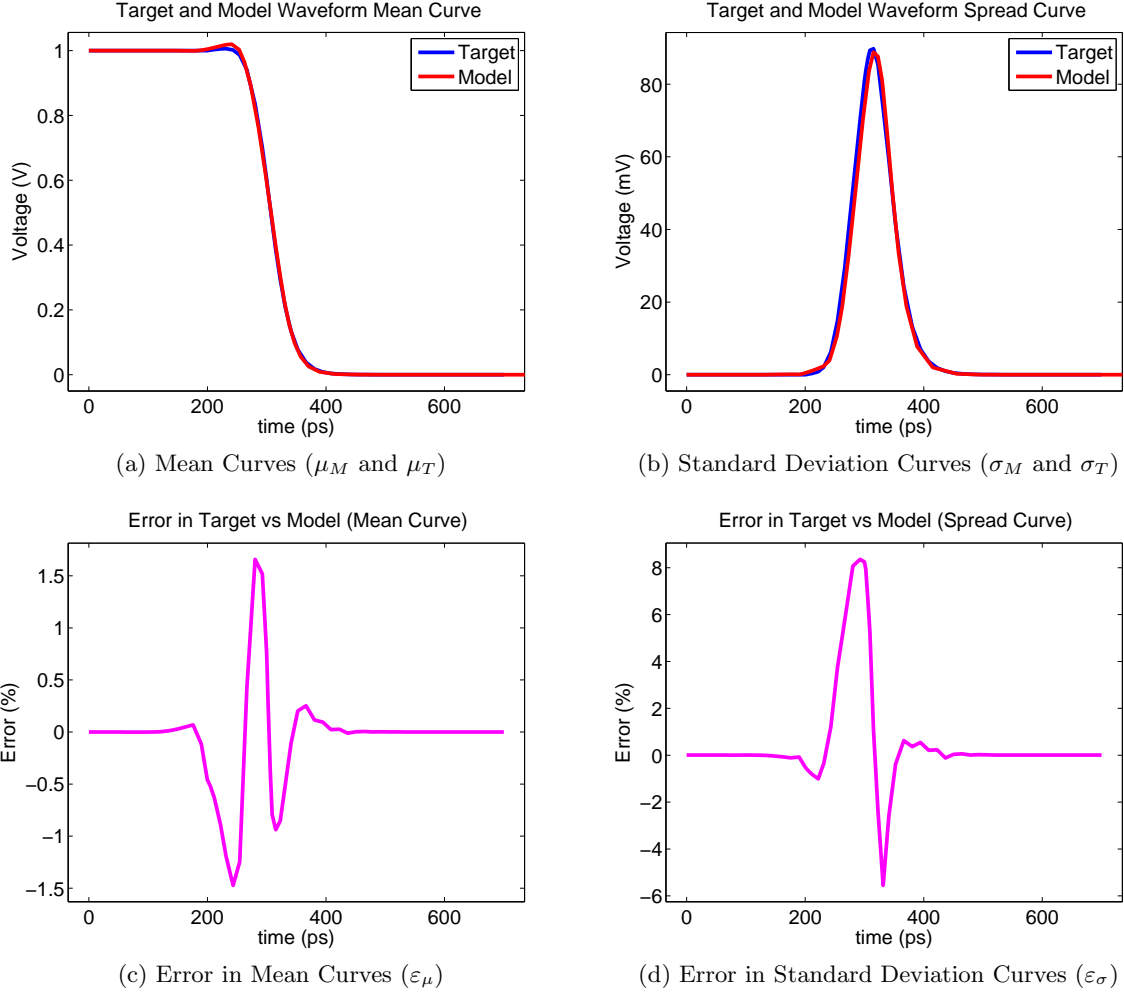


Figure 6.13: Result comparison of pseudo circuit model and target waveform set

and $T_{ShiftMean}$ in negative values of S_{in} or C_{load} (see Figure 6.10). The failing waveform sets are mainly the output of the standard cells where input signal slew or output load is very low or very high. More investigation and development of the pseudo circuit model could be a possible extension of the presented work.

6.5 Summary

This chapter proposes a pseudo circuit model which can be used to reconstruct the target set of waveforms. The main components of the pseudo circuit are an inverter chain of three inverters with process variations in the last inverter, an ideal ramp voltage source with variable start time of the signal transition, a capacitive load, and internal capacitors. The pseudo circuit has four independent parameters and one dependent parameter. The rest of the circuit parameters are kept fixed.

The pseudo circuit is simulated with various circuit configurations of its parameters

and a database is built. The target waveform is compared with this database of the collection of waveforms and the pseudo circuit model parameters are estimated. Since the set of waveforms contains a huge amount of data points, direct comparison is not possible. Therefore, only a few quality factors of the mean and standard deviation curves of the set of waveforms are compared. The pseudo circuit has four independent parameters, therefore four independent quality factors have been proposed for the comparison.

The four quality factors, pseudo circuit database generation for quality factor comparison, and the waveform comparison methodology are described in this chapter. The complete flow of waveform comparison and pseudo circuit parameter estimation is illustrated with an example of the target waveform set.

The mean and standard deviation curves of the target waveform set and the output of the pseudo circuit model are compared. The error between the quality factors of both waveform sets are within 7% in the given example. Various other target waveform sets have been used to generate the pseudo circuit model and the errors in the quality factors are within 10%. However, a few waveform sets generate negative pseudo circuit model. More investigation and development of the pseudo circuit model is required.

Conclusion

After one year of hard and dedicated work, some progress has been made in the thesis project. Summary and the future work related to this project are discussed here.

7.1 Summary

A detailed introduction about the state of the art in timing analysis methodologies has been discussed in Chapter 2. Available circuit simulation and analysis tools and MOSFET models are also discussed there.

In Chapter 3, a discussion about the variation in the delay of the standard cells due to PVT variations with the help of the first two statistical moments (Mean and Standard Deviation) is presented. The standard cells in this analysis are taken from the 45nm PTM model based Nangate open cell library. It was found that the variations in the channel length and channel width affect the variation of delay the most. Furthermore, the primary cause of the non-Gaussian distribution of the delay variation is due to the inverse relation of the delay with the channel width. Higher orders of statistical moments (Skewness and Kurtosis) have been used to quantify the non-Gaussian distributions.

In Chapter 4, a simulation and analysis method based on the uniform sampling technique and weighted sample estimator have been proposed, which require fewer simulation runs for statistical moment estimation. The number of simulation iterations required by this *Fast Statistical Moment Estimation (FSME)* method is at least two orders of magnitude lower than the number of simulation runs required in the standard Monte Carlo method. Along with this, changes in parameter spread and/or probability density function do not require rerunning of the circuit simulations, which results into faster run time and smaller memory requirement. State of the art industrial circuit simulation tools can run Monte Carlo with Gaussian, lognormal and uniform distribution only whereas any distribution can be used in the proposed method. Along with this, two methods have been discussed to estimate the probability density of the simulation result. The probability density function of the standard cell output with different probability density functions of the input PVT parameters is also discussed.

The concept of the set of waveforms and its integration with the timing analysis tool are presented in Chapter 5. The uncertainty representation with the set of waveform with the help of simulated waveform sets are discussed here. Furthermore, three possible methods of representation of the set of waveforms are developed and compared. These representation methods are based on lookup table, statistical moments and pseudo circuits. The pseudo circuit based representation is the most compact modelling approach among them.

The pseudo circuit based waveform set representation is developed in Chapter 6.

The requirements and constraints to develop the pseudo circuit model are discussed first. Followed by this, the pseudo circuit model is developed by addressing the design constraints. Thereafter, waveform comparison methodology has been developed to generate the pseudo circuit model for any given set of waveforms.

7.2 Future Work

The waveform set model has not been evaluated in detail for various possible waveform sets. The detailed evaluation of the performance and accuracy of the pseudo circuit model is a possible extension of the presented work. In this thesis work, Spectre is used to develop the pseudo circuit model. However, the application of this pseudo circuit model is within SSTA engine which does not have Spectre as a circuit simulator tool. Therefore, the performance and accuracy analysis of the proposed pseudo circuit model with the circuit simulation engine of the SSTA flow (e.g. fast circuit simulator developed in the MODERN project) is also a possible extension to this work.

Delay Variations

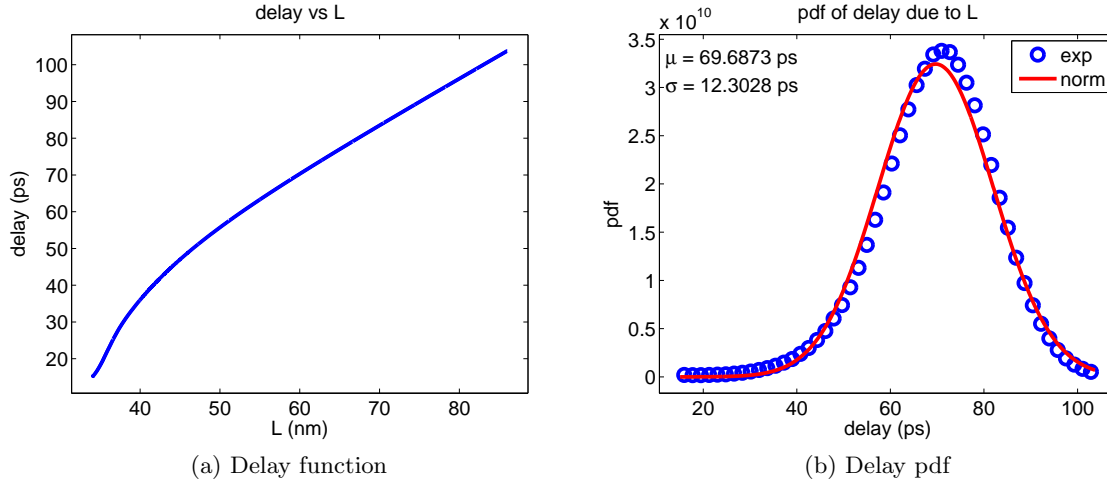
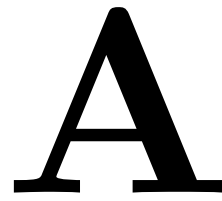


Figure A.1: Delay variation due to L

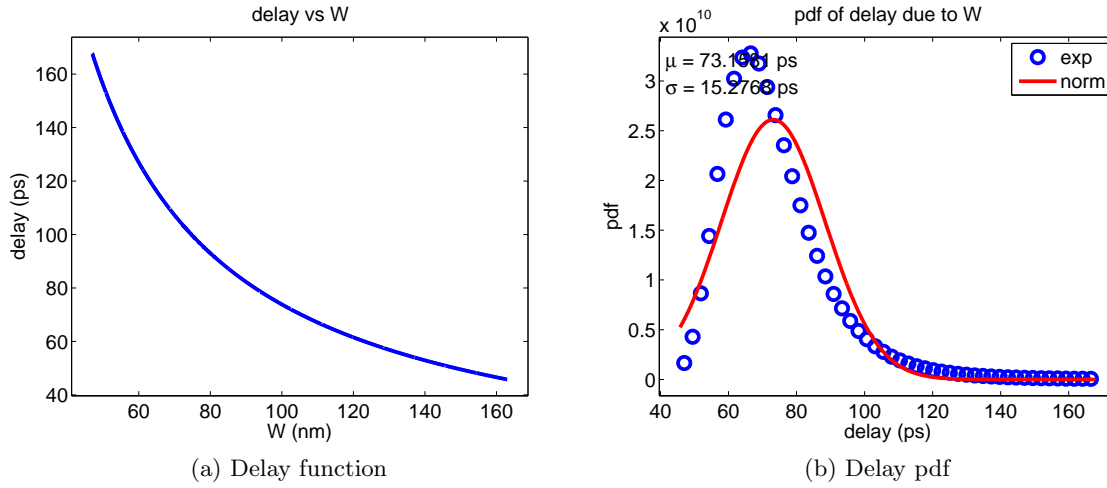


Figure A.2: Delay variation due to W

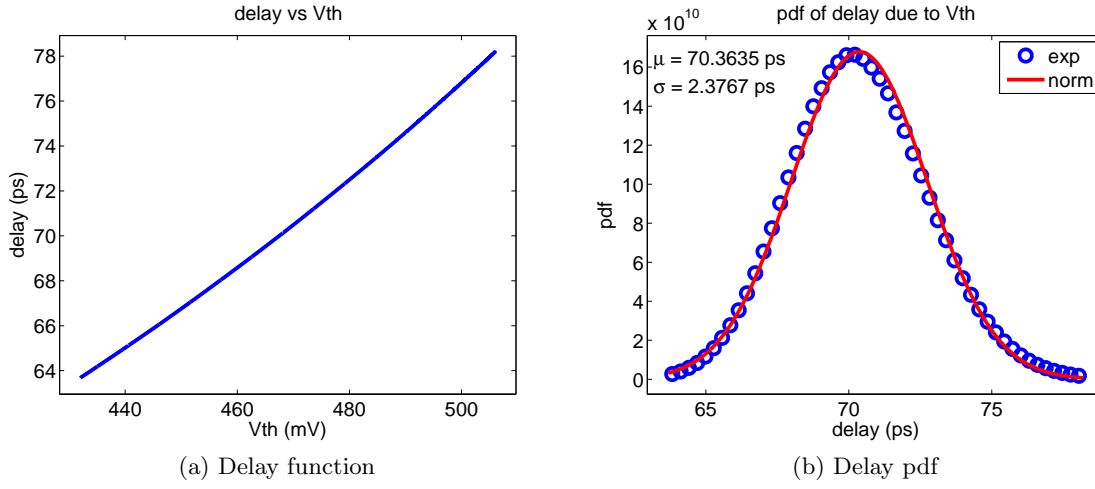


Figure A.3: Delay variation due to V_{th}

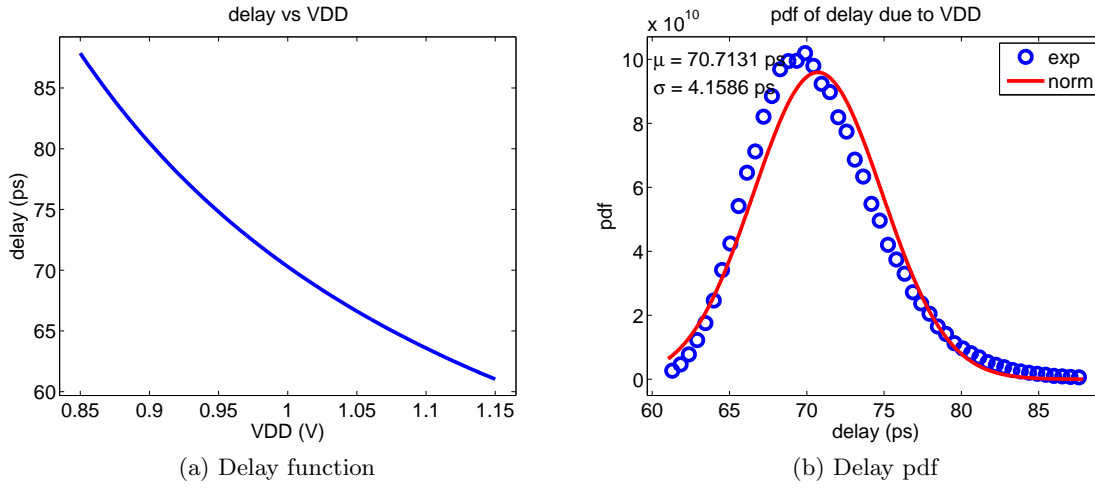
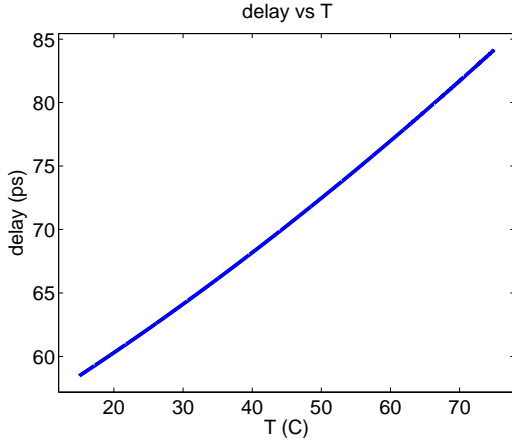
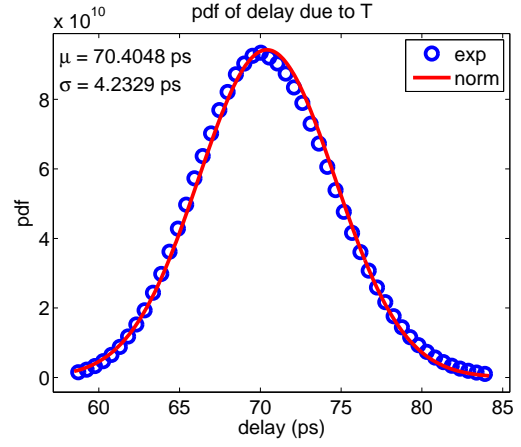


Figure A.4: Delay variation due to V_{DD}

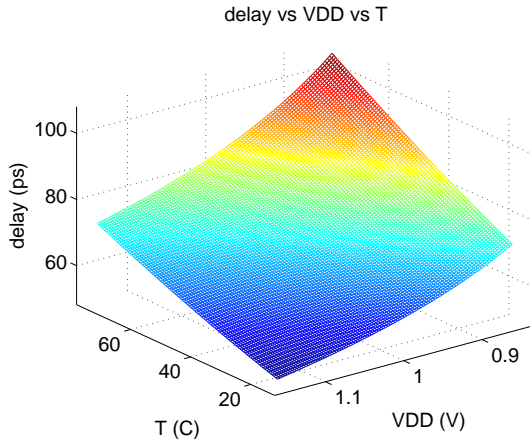


(a) Delay function

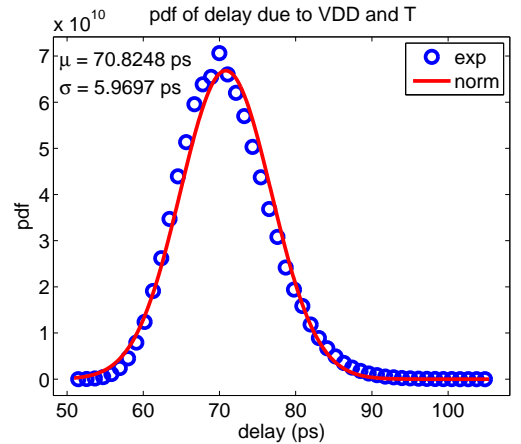


(b) Delay pdf

Figure A.5: Delay variation due to T



(a) Delay function



(b) Delay pdf

Figure A.6: Delay variation due to V_{DD} & T

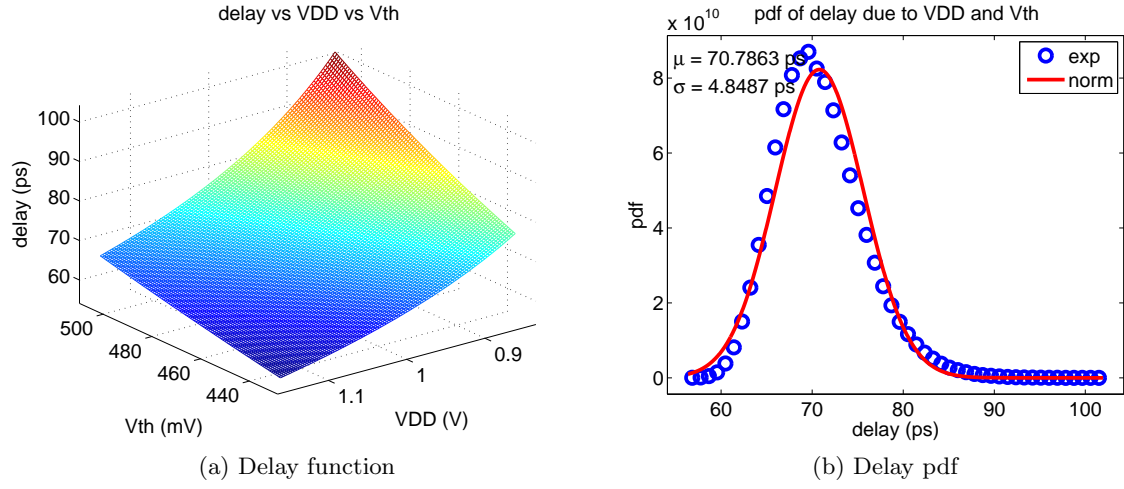


Figure A.7: Delay variation due to V_{DD} & V_{th}

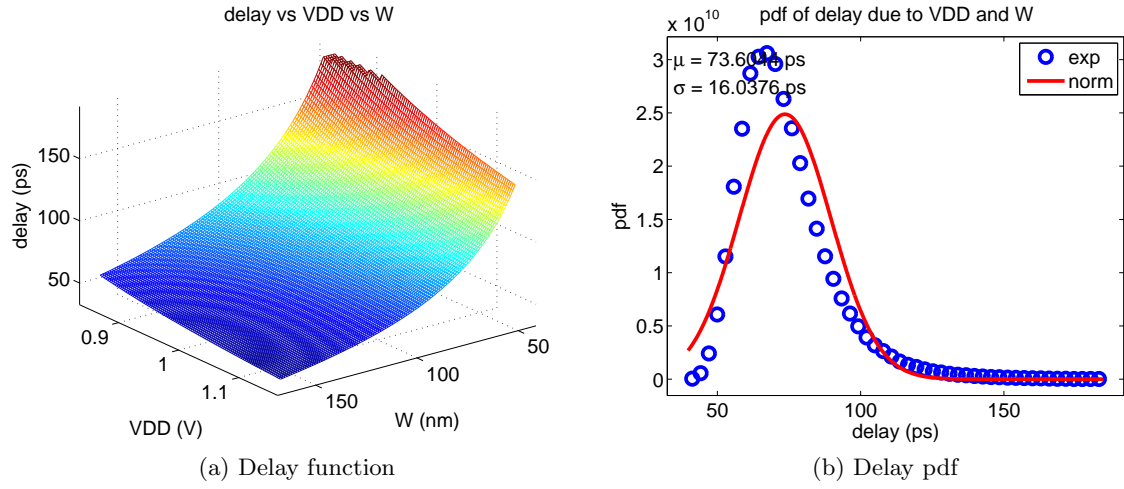
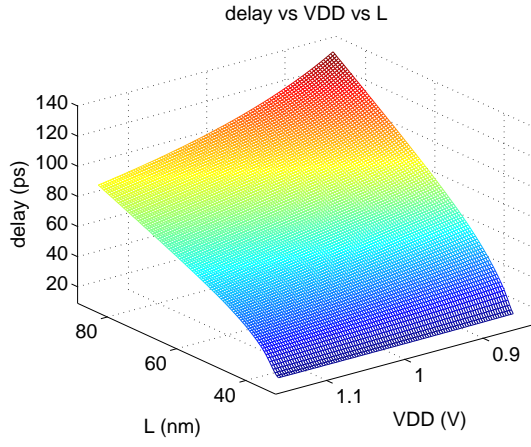
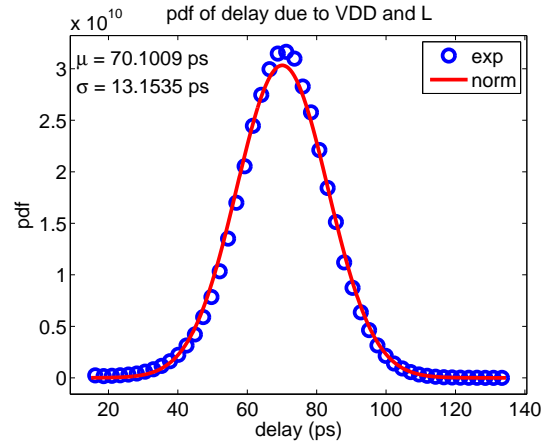


Figure A.8: Delay variation due to V_{DD} & W

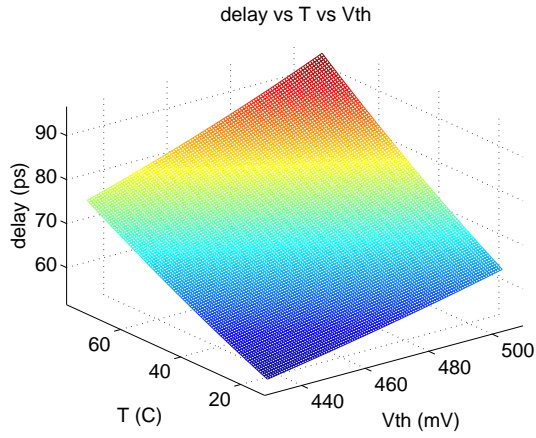


(a) Delay function

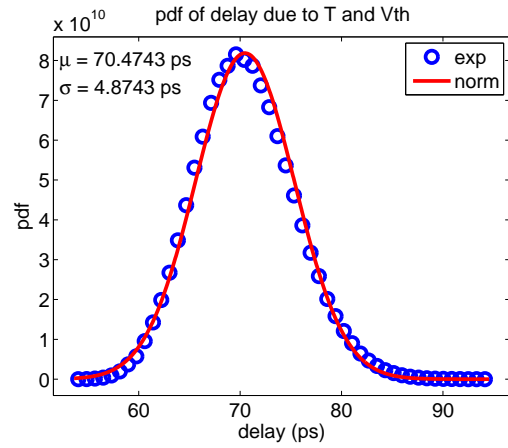


(b) Delay pdf

Figure A.9: Delay variation due to V_{DD} & L



(a) Delay function



(b) Delay pdf

Figure A.10: Delay variation due to T & V_{th}

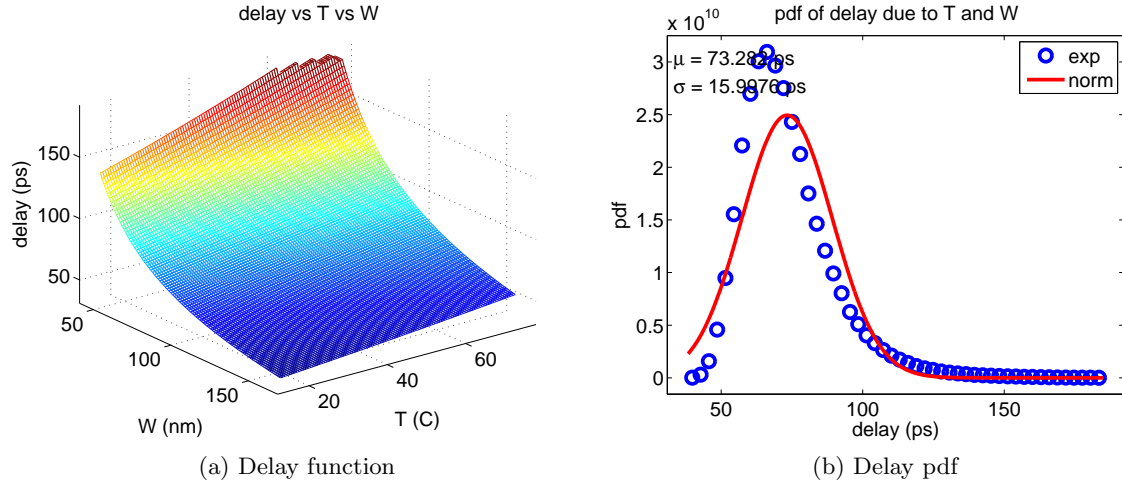


Figure A.11: Delay variation due to T & W

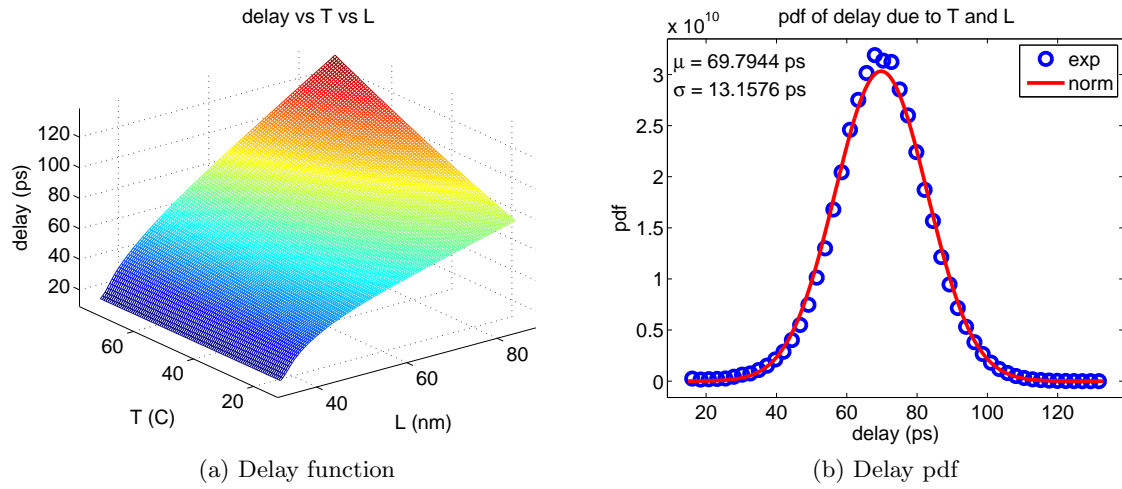


Figure A.12: Delay variation due to T & L

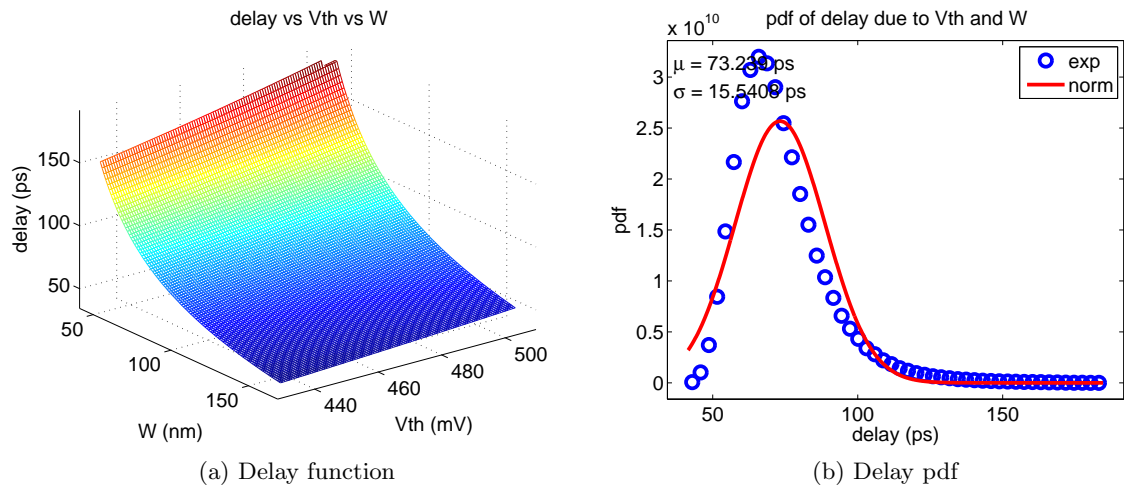


Figure A.13: Delay variation due to V_{th} & W

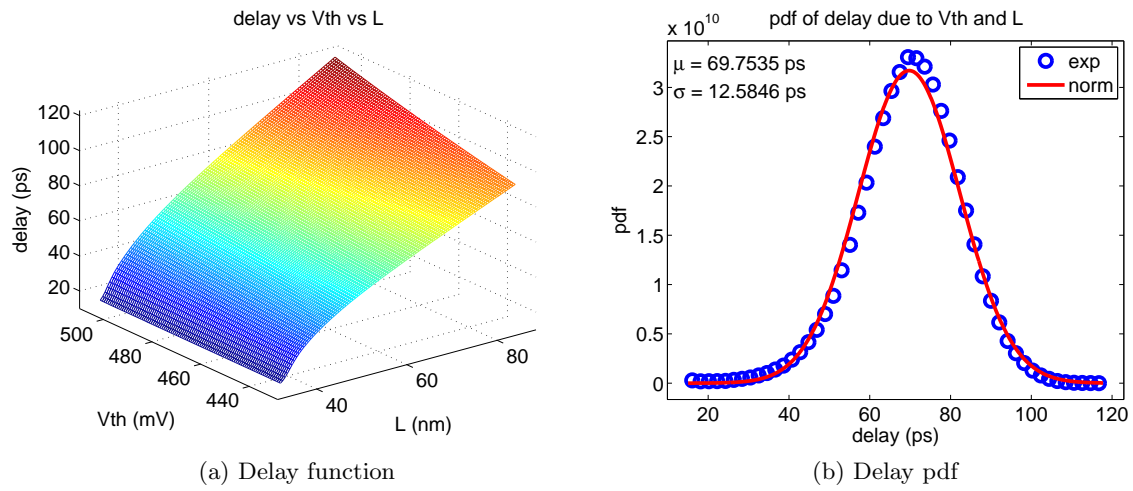


Figure A.14: Delay variation due to V_{th} & L

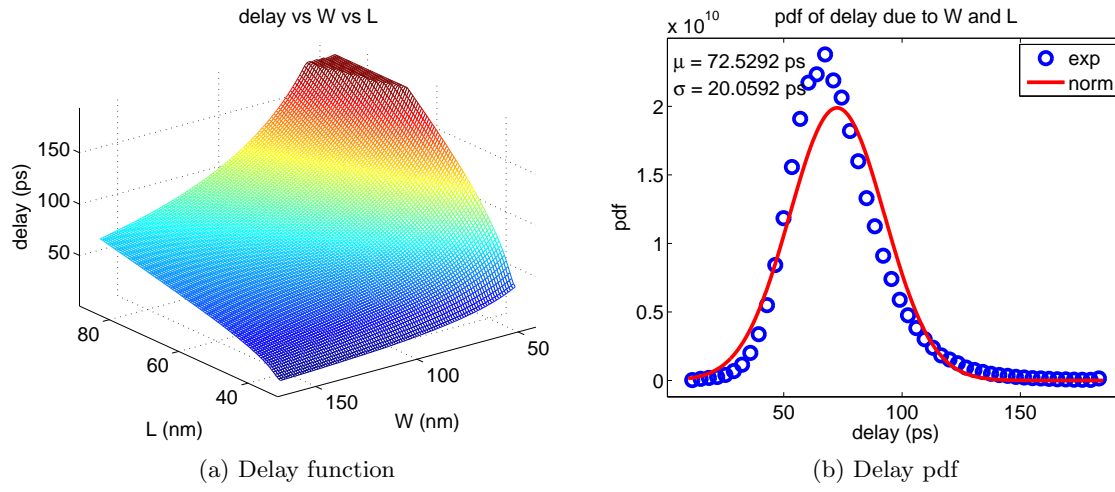


Figure A.15: Delay variation due to W & L

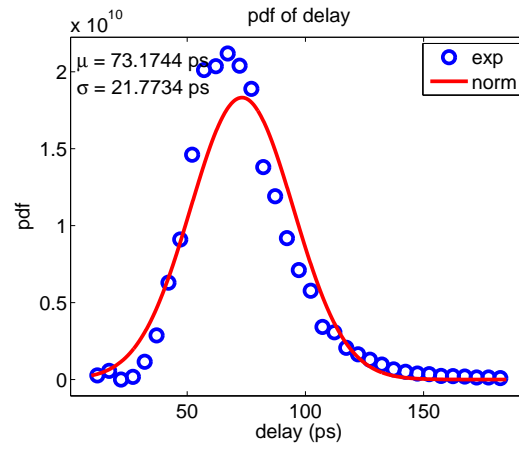


Figure A.16: Delay distribution pdf for realistic PVT variations

- [1] Jan M. Rabaey, Anantha P. Chandrakasan, and Borivoje Nikolia, “The Devices,” *Digital Integrated Circuits: A Design Perspective*, pp. 73–130, 2003, ISBN: 8178089912.
- [2] B. Davari, R.H. Dennard, and G.G. Shahidi, “CMOS Scaling for High Performance and Low Power - The Next Ten Years,” *Proceedings of the IEEE*, vol. 83, no. 4, pp. 595–606, 1995.
- [3] C. Forzan and D. Pandini, “Statistical static timing analysis: A survey,” *Integration, the VLSI Journal*, vol. 42, no. 3, pp. 409–435, June 2009.
- [4] S. Nassif, “Within-chip variability analysis,” *IEDM Technical Digest*, p. 283, 1998.
- [5] S. R. Nassif, “Modeling and analysis of manufacturing variations,” in *IEEE Conference on Custom Integrated Circuits*, 2001, pp. 223–228.
- [6] A. Nigam, Q. Tang, A. Zjajo, M. Berkelaar, and N.P. van der Meijs, “Statistical Moment Estimation in Circuit Simulation,” in *VARI, The European workshops on CMOS Variability*, May 2010, 6 pages.
- [7] A. Nigam, Q. Tang, A. Zjajo, M. Berkelaar, and N.P. van der Meijs, “Statistical Moment Estimation in Circuit Simulation,” *Journal of Low Power Electronics*, vol. 6, no. 4, December 2010, Invited Paper, yet to submit.
- [8] A. Nigam, Q. Tang, A. Zjajo, M. Berkelaar, and N.P. van der Meijs, “Pseudo Circuit Model for Representing Uncertainty in Waveforms,” in *DATE, Design, Automation and Test in Europe*, March 2011, Planned to submit.
- [9] Cadence, “Characterization Guidelines for ECSM Timing Libraries,” Tech. Rep. Version 1.0, Cadence Design Systems, Inc., <https://www.si2.org/openeda.si2.org/projects/omcdistrib>, December 2006.
- [10] CL Ratzlaff, S. Pullela, and LT Pillage, “Modeling The RC-Interconnect Effects In A Hierarchical Timing Analyzer,” in *Custom Integrated Circuits Conference, 1992., Proceedings of the IEEE 1992*, May 1992, pp. 15.6.1–15.6.4.
- [11] J. Qian, S. Pullela, and L. Pillage, “Modeling the “Effective Capacitance” for the RC Interconnect of CMOS Gates,” *IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems*, vol. 13, no. 12, pp. 1526–1535, December 1994.
- [12] P.R. O Brien and T.L. Savarino, “Modeling the Driving-Point Characteristic of Resistive Interconnect for Accurate Delay Estimation,” *IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems*, pp. 512–515, November 1989.

- [13] A. Korshak and J.C. Lee, “An Effective Current Source Cell Model for VDSM Delay Calculation,” in *Quality Electronic Design, 2001 International Symposium on*, March 2001, pp. 296–300.
- [14] H. Bhatnagar, “Synopsys Technology Library,” *Advanced ASIC Chip Synthesis: Using Synopsys Design Compiler, Physical Compiler, and PrimeTime*, pp. 63–80, 2002, ISBN: 0792376447.
- [15] Synopsys, “CCS Timing Technical White Paper,” Tech. Rep. Version 2.0, Synopsys, Inc., http://www.opensourceliberty.org/ccspaper/ccs_timing_wp.pdf, December 2006.
- [16] Synopsys, “CCS Timing Library Characterization Guidelines,” Tech. Rep. Version 3.2, Synopsys, Inc., http://www.opensourceliberty.org/resources_ccs.html, March 2008.
- [17] Synopsys, “CCS Timing Library Validation Guidelines,” Tech. Rep. Version 2.0, Synopsys, Inc., http://www.opensourceliberty.org/resources_ccs.html, February 2006.
- [18] Synopsys, “CCS Timing Liberty Syntax,” Tech. Rep. Version 1.2, Synopsys, Inc., http://www.opensourceliberty.org/resources_ccs.html, June 2006.
- [19] Cadence, “Si2 Effective Current Source Model (ECSM) Timing and Power Specification,” Tech. Rep. Version 1.3, Cadence Design Systems, Inc., <https://www.si2.org/openeda.si2.org/projects/omcdistrib>, September 2007.
- [20] I. Keller, K.H. Tam, and V. Kariat, “Challenges in Gate Level Modeling for Delay and SI at 65nm and Below,” in *Design Automation Conference*. ACM, June 2008, pp. 468–473.
- [21] R. Trihy, “Addressing Library Creation Challenges from Recent Liberty Extensions,” in *Design Automation Conference*. ACM, June 2008, pp. 474–479.
- [22] D. Blaauw, K. Chopra, A. Srivastava, and L. Scheffer, “Statistical timing analysis: From basic principles to state of the art,” *IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems*, vol. 27, no. 4, pp. 589–607, April 2008.
- [23] C.S. Amin, N. Menezes, K. Killpack, F. Dartu, U. Choudhury, N. Hakim, and Y.I. Ismail, “Statistical static timing analysis: how simple can we get?,” in *Design Automation Conference*. ACM, 2005, pp. 652–657.
- [24] C. Visweswariah, K. Ravindran, K. Kalafala, SG Walker, and S. Narayan, “First-order incremental block-based statistical timing analysis,” in *Design Automation Conference*. ACM, 2004, pp. 331–336.
- [25] C. Visweswariah, K. Ravindran, and K. Kalafala, “First-order parameterized block-based statistical timing analysis,” *TAU’04*, 2004, 6 pages.

- [26] H. Chang and S.S. Sapatnekar, “Statistical timing analysis considering spatial correlations using a single PERT-like traversal,” in *International Conference on Computer-Aided Design*. IEEE Computer Society Washington, DC, USA, 2003, pp. 621–625.
- [27] “EKV MOSFET Model,” Ecole Polytechnique Federale de Lausanne, <http://legwww.epfl.ch/ekv>, Retrieved 1 June, 2010.
- [28] “BSIM MOSFET Model,” BSIM Research Group, UC Berkeley, <http://www-device.eecs.berkeley.edu/~bsim3>, Retrieved 1 June, 2010.
- [29] “BSIM4 MOSFET Model,” BSIM Research Group, UC Berkeley, <http://www-device.eecs.berkeley.edu/~bsim3/bsim4.html>, Retrieved 1 June, 2010.
- [30] “International Technology Roadmap for Semiconductor (ITRS),” International Technology Roadmap for Semiconductors, <http://www.itrs.net>, Retrieved 1 June, 2010.
- [31] “Predictive Technology Model (PTM),” Arizona State University, <http://ptm.asu.edu>, Retrieved 1 June, 2010.
- [32] “Open Cell Library, Nangate Inc.,” Nangate - Design Optimization Company, <http://www.nangate.com>, Retrieved 1 June, 2010.
- [33] Cadence, *Virtuoso Spectre Circuit Simulator Reference*, Cadence Design Systems, Inc., http://www.cadence.com/products/cic/spectre_circuit, product version 5.1.41 edition, November 2004.
- [34] “MATLAB - The Language Of Technical Computing,” The MathWorks, Inc., <http://www.mathworks.com/products/matlab>, Retrieved 1 June, 2010.
- [35] “MOdeling and DEsign of Reliable, process variation-aware Nanoelectronic devices, circuits and systems (MODERN),” <http://www.eniac-modern.org>, Retrieved 1 June, 2010.
- [36] Oliver Johnson and Oliver Thomas Johnson, *Information Theory and The Central Limit Theorem*, Imperial College Press, 2004, ISBN: 1860944736.
- [37] TR Jain, SC Aggarwal, and Dr. RK Rana, *Basic Statistics for Economists*, VK Publications, 2009, ISBN: 8188597783.
- [38] R.Y. Rubinstein and D.P. Kroese, *Simulation and the Monte Carlo Method*, Wiley-Interscience, 2008, ISBN: 0470177942.
- [39] P. Glasserman, *Monte Carlo methods in Financial Engineering*, Springer Verlag, 2003, ISBN: 0387004513.
- [40] L. Zhang, J. Shao, and C.C.P. Chen, “Non-Gaussian statistical parameter modeling for SSTA with confidence interval analysis,” in *International Symposium on Physical Design*. ACM, 2006, pp. 33–38.

- [41] V. Veetil, D. Sylvester, and D. Blaauw, “Criticality aware latin hypercube sampling for efficient statistical timing analysis,” *ACM/IEEE TAU*, February 2007.
- [42] A. Singhee and R.A. Rutenbar, “From finance to flip flops: A study of fast quasi-Monte Carlo methods from computational finance applied to statistical circuit analysis,” in *International Symposium on Quality Electronic Design*, 2007, 8 pages.
- [43] V. Veetil, D. Sylvester, and D. Blaauw, “Efficient Monte Carlo based incremental statistical timing analysis,” in *Design Automation Conference*. ACM, 2008, pp. 676–681.
- [44] Arnold R. Krommer and Christoph W. Ueberhuber, *Numerical Integration on Advanced Computer Systems*, Springer, 1994, ISBN: 3540584102.
- [45] “Numerical Integration,” Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/wiki/Numerical_integration, Retrieved 1 June, 2010.
- [46] I. T. Jolliffe, *Principal Component Analysis*, Springer, 2002, ISBN: 0387954422.
- [47] Q. Tang, A. Zjajo, M. Berkelaar, and N.P. van der Meijs, “A Simplified Transistor Model for CMOS Timing Analysis,” in *ProRISC*, November 2009, 6 pages.
- [48] Q. Tang, A. Zjajo, M. Berkelaar, and N.P. van der Meijs, “Transistor Level Waveform Evaluation for Timing Analysis,” in *VARI, The European workshops on CMOS Variability*, May 2010, 6 pages.
- [49] Q. Tang, A. Zjajo, M. Berkelaar, and N.P. van der Meijs, “RDE-Based Transistor-Level Gate Simulation for Statistical Static Timing Analysis,” in *Design Automation Conference*, June 2010, pp. 787–792.
- [50] Q. Tang, A. Zjajo, M. Berkelaar, and N.P. van der Meijs, “Transistor-Level Gate Modeling for Nano CMOS Circuit Verification Considering Statistical Process Variations,” in *International Workshop on Power And Timing Modeling, Optimization and Simulation (PATMOS)*, September 2010, 6 pages, accepted for presentation and publication.

List of Publications

- [1] A. Nigam, Q. Tang, A. Zjajo, M. Berkelaar, and N.P. van der Meijs, “Statistical Moment Estimation in Circuit Simulation,” in *VARI, The European workshops on CMOS Variability*, May 2010, 6 pages.
- [2] A. Nigam, Q. Tang, A. Zjajo, M. Berkelaar, and N.P. van der Meijs, “Statistical Moment Estimation in Circuit Simulation,” *Journal of Low Power Electronics*, vol. 6, no. 4, December 2010, Invited Paper, yet to submit.
- [3] A. Nigam, Q. Tang, A. Zjajo, M. Berkelaar, and N.P. van der Meijs, “Pseudo Circuit Model for Representing Uncertainty in Waveforms,” in *DATE, Design, Automation and Test in Europe*, March 2011, Planned to submit.