TIMING MODELS FOR EFFICIENT CHARACTERIZATION OF NANOSCALE VLSI SINGLE STAGE STANDARD CELLS

Ph.D THESIS

by

BALJIT KAUR

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE-247667, INDIA
SEPTEMBER, 2014

TIMING MODELS FOR EFFICIENT CHARACTERIZATION OF NANOSCALE VLSI SINGLE STAGE STANDARD CELLS

A THESIS

Submitted in partial fulfilment of the requirements for the award of the degree

of

DOCTOR OF PHILOSOPHY

in

ELECTRONICS AND COMMUNICATION ENGINEERING

by

BALJIT KAUR

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE-247667, INDIA
SEPTEMBER, 2014

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled "**TIMING MODELS FOR EFFICIENT CHARACTERIZATION OF NANOSCALE VLSI SINGLE STAGE STANDARD CELLS** " in partial fulfilment of the requirements for the award of the Degree of Doctor of Philosophy and submitted in the Department of Electronics and Communication Engineering of the Indian Institute of Technology Roorkee, Roorkee is an authentic record of my own work carried out during a period from December, 2009 to September, 2014 under the supervision of **Dr. Anand Bulusu**, Associate Professor and **Dr. Sanjeev Manhas**, Associate Professor, Department of Electronics and Communication Engineering, Indian Institute of Technology Roorkee, Roorkee.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other Institute.

(BALJIT KAUR)

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

(Anand Bulusu)
Supervisor

(Sanjeev Manhas)
Supervisor

Date:

ABSTRACT

In nanometer range VLSI technologies, semi-custom chip design approach using pre-designed and pre-characterized standard cells is popular because of increasing complexity. For efficient circuit design, these standard cells are pre-characterized for delay, transition times, terminal capacitances, power dissipation, noise and area using SPICE simulations. The traditional Non Linear Delay Models (NLDM) based Lookup Table (LUT) approach for standard cell characterization and Static Timing Analysis (STA) is facing serious challenges in nanometer technologies, because it does not account for the nature of input and output terminal voltage transitions. Because of voltage dependent values of effective capacitances of device and interconnects, it becomes important to consider the nature of output transition. To overcome these limitations, researchers introduced Current Source Modeling (CSM), in which for a given value of load capacitance, the value of output voltage and equivalent circuit parameters of the standard cell are given as a function of input voltage. This increases the complexity of the model and the amount of data to be stored for standard cell characterization.

To solve these issues, vendors found the models as a middle path (between NLDM and CSM), known as vendor CSM formats. These vendor CSM formats are Effective CSM (ECSM) and Composite CSM (CCSM) . For a given input transition time (T_R) and load capacitance (C_l=C_{eff}) values, ECSM stores the times at which the output voltage waveform crosses certain predefined threshold points, whereas CCSM stores the output current values at specified voltage level points. Both the vendor models are equivalent and one can be derived from the other. Vendor CSMs use Lookup Table (LUT) based format for representing characterization data. The major issue with ECSM characterization is that it requires re-characterization of standard cells with variation in cell size, layout dependent parameters, temperature, supply voltage and device model updates. This re-characterization is highly time consuming. Therefore, there is a need for a model which is more efficient in standard cell characterization, thus saving time and effort. In this thesis, we undertake a detailed study of existing timing/delay models of CMOS inverter and NAND gate standard cells. We find that these delay models are unsuitable for use in standard cell characterization, because the region of validity in T_R , C_l space is not clear. Apart from this cell size, layout dependent parameters, power supply voltage and temperature variations are also not inputs to such models. In addition, it is seen that the intermediate node voltage transition of the series stack nMOS devices of a NAND gate, which plays an important role in sub-nanometer technology nodes, has not been considered appropriately in earlier models. We show that considering these issues in standard cell characterization, the re-characterization effort would increase significantly.

For an efficient ECSM characterization, we developed the timing models for CMOS inverter and 2-input NAND gate (therefore, for 2-input NOR gate also), to reduce the re-characterization effort. All the multistage combinational cells can be derived from these basic cells. Firstly, we propose the analytical timing models relating all the Threshold

Crossing Points (TCPs) of output transition with T_R, C_l values, for CMOS inverter and 2-input NAND gate. We then identify the region of validity of the model in T_R, C_l space used in characterization LUTs. It makes the timing models useable in reducing the HSPICE simulations in ECSM library characterization. Further, we identify the relationships of model coefficients with cell size, process induced mechanical stress, temperature and supply voltage variation. Our NAND gate timing models are robust because of an appropriate and detailed consideration of voltage transition at the intermediate node of the series stack of nMOS devices. For this, we consider the input to intermediate node capacitive coupling effect, parasitic capacitances at the intermediate node and the regions of operation of the two nMOS devices placed in series stack. We present the timing models for 2-input NAND gate, considering the following two cases : 1) When upper nMOS transistor in the series stack switches; and 2) When lower nMOS transistor in the series stack switches. In this thesis, we show that the use of these models in standard cell characterization reduces the number of SPICE simulations by 50% and 67% for CMOS inverter and 2-input NAND gate, respectively. We also show that our timing models remain valid with PVT variations. This would help in reducing the re-characterization effort significantly (nearly 85% reduction in SPICE simulations) for standard cell libraries. Further, we present an analytical overshoot timing model for CMOS inverter and NAND gate for accurate timing analysis. For NAND gate overshoot modeling, we see that an inclusion of intermediate node voltage transition with appropriate assumptions lead to accurate estimation (max. error is 2.5% with respect to HSPICE simulations) of overshoot timing values for Case 1 and 2, respectively.

ACKNOWLEDGEMENT

Completion of this doctoral dissertation was possible with the support of several people. I would like to express my sincere gratitude to all of them. First and foremost, I would like to express my deepest gratitude to my supervisors Dr. Anand Bulusu and Dr. Sanjeev Manhas for their inspiration, encouragement and moral support in the successful completion of this work.

I am also thankful to Dr. A.K. Saxena, Dr. Sudeb Dasgupta, Dr. B.K. Kaushik of Microelectronics and VLSI (MEV) Group for their valuable guidance, scholarly inputs and consistent encouragement I received throughout my research work. I am grateful to Prof. M.V. Kartikeyan, Head of the department, for the academic support and the facilities provided to carry out the research work. I express my gratitude to the board of research studies and the members of my Student Research Committee (SRC) Prof. Dharmendra Dingh, Chairman (DRC), and Prof. R. Nath, Department of Physics, for the support at various phases of the programme and sparing their precious time in spite of all their busy schedule for carefully examining my research work and providing valuable suggestions. I gratefully acknowledge the funding sources that made my Ph.D. work possible. I was funded by the Ministry of Human Resource Development, Government of India for 4 years and 3 months. I would like to thank research scholars Manoj Kumar Majumder, Asutosh Nandi, Menka, Pankaj Kumar Pal, V. Ramesh Kumar, Shivam Verma, Om Prakash, Savitesh Sharma and Ruchi, of my lab for helping me with technical discussions. I admire the support of my colleagues Satish Maheshwaram, Archana Pandey and Arvind Kumar Sharma, and thank them for their concern and good wishes. My sincere thanks to my Ph.D. seniors of Microelectronics and VLSI group Dr. S. S. Rathod, Dr. R. Vaddi, Dr. Jitendra Kanungo, Dr. Naushad Alam and Dr. Gaurav Kaushal, for their support and encouragement throughout my work. I sincerely appreciate the lab staffs, Mr. Naveen Kanwar and Mr. Dinesh Sharma , for their help and support during the research work. I would like to thank my friends Nitasha Bansal, Deepti Chaudhary, Deepika Agarwal, Surabhi Singh, Mamta Saxena, Sujata Parida and Divya Madhuri for their support during my Ph.D.

In particular, I am profoundly indebted to my guide Dr. Anand Bulusu, a person with an amicable and positive disposition. I consider it as a great opportunity to do my doctoral programme under his guidance and to learn from his research expertise. He has always been responsible and supportive throughout all the highs and lows during the journey of my Ph.D. He has been a tremendous mentor for me. Words are not enough to express my gratitude to him. Last but not the least, I am deeply indebted to my parents for their unconditional love, sacrificial giving and support to me. Without their teachings and blessings, I could not have come to this stage of life. I owe a lot to my whole family who encouraged and helped me at every stage of my personal and academic life, and longed to see this achievement come true.

Above all, I owe it all to Almighty God for granting me the wisdom, health and strength to undertake this research task and enabling me to its completion.

Date:

Place: (Baljit Kaur)

Contents

List of Figures

List of Tables

Abbreviations and Symbols

CCSM	Composite Current Source Model
CSM	Current Source Model
DELVT0	Threshold Voltage Shift
ECSM	Effective Current Source Model
LUT	Lookup Table
MFGSs	Multi-Fingered Gate Structures
MULU0	Mobility Multiplier
NF	Number of Fingers
NLDM	Non-Linear Delay Model
PTM	Predictive Technology device Model
PVT	Process, Voltage and Temperature
STA	Static Timing Analysis
TCP	Threshold Crossing Point
VLSI	Very Large Scale Integration

Chapter 1

Introduction

1.1 Motivation

Due to aggressive technology scaling, the number of logic gates in VLSI chips continue to grow fast. A semi-custom design with pre-designed and pre-characterized standard cell library is essential in this scenario. In this regard, it is prerequisite to characterize basic standard cells, such as inverter, NAND, NOR, latch, flip-flop etc. The objective of cell characterization is to design a high quality model that accurately and efficiently predicts the cell behavior of a standard cell library. Digital design tools use these characterization models for different purposes. Generally, the characterization of standard cells is for parameters related to timing, area and power [5].

In nanometer range technologies, digital design tools need to account for several complex phenomena such as short channel effects, input to output coupling capacitance, interconnect coupling, power supply noise and process variations, etc. Because of the smaller delays and transition times of nanometer range technology standard cells, a minor variation due to these effects influences timing parameters significantly [6]. This increases the need for re-characterization of standard cell library at different process, voltage and temperature corners. Therefore, designing high performance VLSI standard cell library and its characterization has become more challenging than ever in deep sub-micron era. The following two methods are used to measure delays and thus eliminate timing violations in a data path :

1. Circuit simulations using SPICE can be used to estimate the delay of a circuit accurately. However, SPICE simulations need large CPU times to process an entire circuit having large number of transistors. SPICE takes few seconds to process individual transistors in a circuit, so the processing of an entire circuit takes large time [7].

2. An alternative method to measure delay is Static Timing Analysis (STA) method. STA makes use of the simple gate delay models to find the delay of the entire data path, hence takes lesser time [8].

In order to find the delay of an entire combinational circuit using STA, we must determine the delay of its logic gates. This delay is expressed as a function of the load

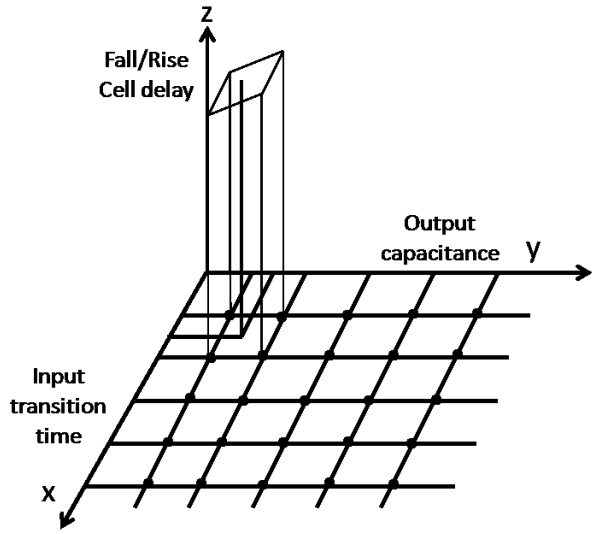


Figure 1.1: Variation of cell delay with input transition time and output capacitance [1].

capacitance (C_l) and input transition time (T_R) of logic gates . These gate delay models are classified as:

1. Analytical Delay Models: The delay of a logic gate is found from the output voltage transition of logic gate across the load capacitance. They make use of the current equation of the MOSFET. The accuracy of these models depends on accuracy of the current equation. Alpha power law delay model is a typical example [9].

2. Empirical Delay Models: These models are based on the curve fitting on the simulation data obtained using SPICE. Scalable polynomial delay model is a typical example [10].

3. Look Up Table (LUT) Method: Here we have representation of the delay variation with T_R and C_l values. Having known T_R and C_l values, we can pick the delay of that particular gate from the table. A typical example is Synopsys ".lib" format Non-Linear Delay Model (NLDM) representation [5, 1], shown in Fig. 1.1.

Conventionally, NLDM based standard cell library characterization was used in STA. In this, delay is a non-linear function of C_l and T_R and it is expressed in an LUT with respect to several values of C_l and T_R. However, NLDM doesn't capture the nature of terminal voltage transition and variation of capacitance with terminal voltage [11, 12]. To address these issues, recently Current Source Model (CSM) has become important in the standard cell characterization and STA [13]. CSMs ideally support arbitrary input waveforms and output loads since their model parameters are waveform and load independent[14]-[17]. However, vendor CSMs presently impose some more constraints; they provide CSM data for a set of T_R and effective load capacitance C_{eff}. Two popular vendor CSMs are known as Effective CSM (ECSM) and Composite CSM (CCSM). For a given set of values of T_R

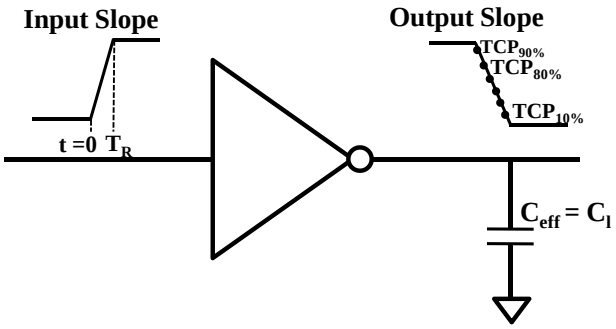


Figure 1.2: ECSM characterization overview.

and C_{eff}, ECSM stores the time at which the output voltage waveform crosses certain predefined α% Threshold Crossing Points (TCPs) (shown in Fig.1.2) while CCSM stores the output current values at different threshold crossing points [13]. Both the vendor models are equivalent and one can be derived from the other. In ECSM, LUT of threshold crossing points for several values of C_l and T_R is used.

	Input Transition Time (T_R)				
C_l	T_{R1}	T_{R2}	.	.	T_{Rn}
C_{l1}	V_{11}	V_{21}			V_{n1}
C_{l2}	V_{12}	V_{22}			V_{n2}
.					
.					
C_{ln}	V_{1n}	V_{2n}			V_{nn}

Figure 1.3: An LUT of ECSM vectors. For each set of C_l and T_R, the characterization data is represented as vector V consisting of time of $TCPs$.

An LUT used for STA is a two dimensional table (shown in Fig.1.3), in which parameter to be characterized is stored for various C_l and T_R values. To minimize the storage, values of parameter to be extracted is stored for limited set of C_l and T_R and linear interpolation is used to obtain delay or time of a TCP for other values of C_l, T_R. ECSM characterization is a computationally tedious task. In addition, Process, Voltage and Temperature (PVT) variations and frequent device model updates require a lot of re-characterization. In this work, we address the problem of reducing the number of SPICE simulation in ECSM library characterization. We consider the input to output coupling capacitance effect for CMOS inverter and NAND gate to accurately measure the timing values for ECSM characterization.

1.2 Previous Work

Several delay models have been proposed to characterize the behavior of CMOS logic

gates. Some commonly used delay models are discussed here.

The lumped RC delay model is proposed by Ousterhout [18]. In this model, the delay is computed by lumping all of the resistances and capacitances together through a stage. All of the resistances are summed separately, as are all the capacitances, and the product gives the delay through the stage as

$$delay = \left(\sum R\right)\left(\sum C\right) \tag{1.1}$$

The drawback of this model includes overestimation of delay due to lumping of resistance and capacitance values. It does not take into account the influence of input waveform on delay since it is limited to the step input waveform only. Later on, Hedenstierna and Jeppson [19] presented an analytical delay model for CMOS inverter, derived using Shockley's model [20]. The model included the input waveform slope effect. Although it is quite simple to use, but the model proved to be unsuccessful for short channel devices [9]. It does not consider the short channel effects dominating at sub-65nm technology nodes. Sakurai *et al.* [9] proposed an alpha power law model. This model is an extension of Shockley's model considering the velocity saturation effect in short channel devices. The model remains valid for the fast input ramp where input slope crosses one-third of the output slope [21]. In [21], authors reported that the α-power law model is not valid for slow input ramps and presented an analytical delay model for both fast and slow input rise time.

In 1998, Sutherland *et al.* [22] presented a logical effort method to measure the delay efficiently. This model enables us to find the least delay of the circuit. It finds that how many stages would then be required and what should be size of the transistors in the gates to get the least delay. The method is quite useful for optimizing the circuit speed. The author presented the delay of a logic gate as

$$d = f + p \tag{1.2}$$

Where, f represents the effort delay proportional to the gate's output load and p represents the parasitic delay. The parasitic delay is fixed for a logic gate and independent of cell size and load capacitance it drives. The model is very simple to use but it neglects the input transition time effects and secondary effects like velocity saturation, body effect etc. Traditional method (empirical model) to characterize delay is to use an equation of the form $k_1C_l + k_2$, where k_1 is the input transition slope and k_2 is the intrinsic delay. In [10], researchers working with Synopsys presented the Scalable Polynomial Delay Model (SPDM) to characterize the cell delay. The model uses a product of polynomials to fit the delay data. For example, to characterize the delay for two parameters C_l and T_R, a product of m^{th} order polynomial in C_l with an n^{th} order polynomial in T_R may be used in the following form:

$$delay = (a_0 + a_1C_l + \ldots + a_mC_l^m)(b_0 + b_1T_R + s\ldots + b_nT_R^n) \tag{1.3}$$

To overcome this complexity, people in industry are now making use of LUTs to obtain

the delay of a logic gate. LUTs with delays tabulated for several values of T_R and C_l values are today used in STA. These LUTs are popular due to limitations of analytical and polynomial delay models as discussed above. LUT used for STA is a 2-D table, where the standard cell delay is characterized with several T_R and C_l values. There are several problems with the existing LUTs, for example T_R and C_l values are selected in an ad-hoc manner, need of re-characterization due to PVT variations, accuracy of the delay values obtained depends on the size of the LUT.

All the aforementioned delay models measure only the propagation delay. These models do not consider the PVT variations and input to output coupling capacitance which are very important at nanometer range technologies. To address the variation in delay values in presence of PVT variations, several delay models are reported [23]-[45]. However, they did not derive the model coefficients considering PVT variation with physical reasoning for the same. While considering process variations, they considered the change in process parameters such as oxide thickness, threshold voltage, doping concentration, gate length, etc. However, they didn't consider the effect of mechanical induced stress as a function of Number of Fingers (NF) on the timing values while the stress engineering is widely being incorporated at nanometer range technologies to enhance the device performance. In this work, we show the dependence of timing values on cell size, load capacitance, input slew, process (mechanical induced stress as a function of NF), voltage and temperature variations. Beside this, we also investigated the importance of overshoot modeling in accurate timing analysis. We found that there are very few researchers who worked on overshoot timing models [46, 47, 48, 49]. In [49], authors show that the conventional delay models [9, 21, 50, 51], which ignore overshoot effect face serious issues of accuracy at these technology nodes. Huang *et al.* [49] proposed an analytical overshoot model at nanometer technology node. This work gives an overshoot time (t_{ov}) expression according to which t_{ov} is a function of the logic gates capacitance C_l. The authors assumed that a linearly time varying current discharges (charges) C_l. This model has been verified for only large values of input transition time T_R and C_l. The minimum value of C_l used in [49] are typically fan-out 100. They also assume that t_{ov} varies with C_l which is not validated by the simulation results shown in their reported research work.

1. Therefore, the delay models which can be used in nanometer range technologies are either inaccurate or fully empirical and cumbersome to use.

2. They neither consider PVT variations through physical reasoning nor the important second order phenomenon such as overshoot due to gate to drain coupling capacitance (C_{gd}).

3. They are not amenable to be used easily in standard cell characterization for STA.

1.3 Problem Definition

The aim of this thesis is to propose the timing models to reduce the computational effort in characterization of single stage standard cells at nanometer range technologies. Such a work can easily be extended to multistage standard cells. In order to accomplish this, following approach have been taken:

- Derivation of timing models and their region of validity as a function of T_R and C_l for standard cell CMOS inverter
- Verification of the model coefficients' behavior with cell size and technology node for standard cell CMOS inverter
- Derivation of timing models and their region of validity with PVT variability for standard cell CMOS inverter
- Derivation of timing models and their region of validity as a function of T_R and C_l for 2-input CMOS NAND gate
- Verification of the model coefficients' behavior with cell size for NAND gate
- Derivation of timing models and their region of validity with PVT variability for NAND gate
- CMOS inverter and NAND gate overshoot modeling

1.4 Contributions

The objective of this thesis is to develop an accurate timing model for ECSM characterization of all the TCPs of output voltage at nanometer range technologies. For this, we first develop a timing model for CMOS inverter at 32nm technology node. The model is formulated based on current equations considering velocity saturation. All the factors which affects delay, namely: cell size, T_R and C_l are considered in the model. The proposed model matches with the HSPICE simulated results closely for all the TCPs. The model remains valid with the technology scaling.

At nanometer technology node, there is a need for library re-characterization due to on-chip PVT variations. To improve the performance, it is important to accurately measure the timing values in circuits while considering voltage, temperature and process induced variability. We derive relationships of the coefficients of our ECSM timing models with PVT variation. For voltage variability, we consider $\pm 10\%$ variation in nominal supply voltage. For temperature variability, we vary the temperature range from $298K$ to $423K$. For stress variability, we consider the variation in device channel mechanical stress as a function of NF in inverter layout. This is because, often the size of inverter cell is increased by increasing

NF, therefore the consequent variation in channel mechanical stress is important. Stress induced in the channel of pMOS (nMOS) using various sources like Compressive/Tensile Etch Stop Liner (c/t-ESL), embedded SiGe (eSiGe) and Stress Memorization Technique (SMT) [52] has been considered in this work.

We derive the timing models for 2-input CMOS NAND gate, considering single input switching and also derive their region of validity in T_R, C_l space. We observe that the improper consideration of intermediate node voltage transition leads to significant percentage error in delay/timing values (we later discuss this in detail). Based on this observation, we consider the intermediate node voltage transition for accurate timing analysis. Further, we determine the relationships of the model coefficients with the cell size, power supply voltage, carrier mobility, threshold voltage and temperature. We also consider layout dependent effects due to mechanical stress in deriving these relationships.

Next, we propose an analytical overshoot timing model for very wide range of T_R values for CMOS inverter and 2-input NAND gate. For NAND gate standard cell, it becomes important to consider the intermediate node voltage transition in accurate timing analysis. Therefore, we first model the behavior of intermediate node voltage for large values of T_R. Later on, we derive the relationships of overshoot time with cell size, T_R and C_l values. We find that the proposed model is independent of cell size and C_l values.

The proposed model reduces the number of HSPICE simulations in ECSM characterization of CMOS inverter and NAND gate standard cell by nearly 50% and 67%, respectively. The need to re-characterize the timing models with PVT variation, has thus been reduced.

1.5 Thesis Organization

This thesis is organized into 7 chapters as follows:

Chapter 1: In this chapter, an introduction to the standard cell characterization and challenges it faces due to the need for more accurate cell characterization are presented.

Chapter 2: This chapter provides a detailed literature review on propagation delay models, overshoot timing models, along with the need of an efficient timing model. Technical gaps in the existing literature on timing models and digital circuit performance are discussed. The chapter is concluded with a brief summary of technical gaps to be addressed in the thesis.

Chapter 3: This chapter focuses on modeling of timing values of threshold crossing points ($t_{TCP}s$) as a function of T_R and C_l for a minimum sized CMOS inverter. Further, the region of validity of the model in T_R, C_l space is derived. The relationship between cell size and model coefficient is also derived. We also analyzed, the impact of technology scaling on these model coefficients. The results depict that the proposed model is in good agreement with HSPICE simulations with a maximum error of 2.5%. The contribution of these models in reducing the number of HSPICE simulations in ECSM characterization of the inverter standard cell is nearly half.

Chapter 4: In this chapter, we consider the impact of process parameters (mechanical induced stress as a function of Number of Fingers (NF)), supply voltage and temperature variations on the proposed (t_{TCP}) models. We derive the relationships of model's coefficients and their region of validity with the process, supply voltage and temperature variations in T_R, C_l space. Therefore, the models considering PVT variation helps in reducing the re-characterization effort in standard cell library characterization. In this chapter, we demonstrate that the inclusion of PVT variation in our t_{TCP} models reduces the number of HSPICE simulations by about half.

Chapter 5: This chapter focuses on modeling of $t_{TCP}s$ as a function of T_R and C_l for 2-input CMOS NAND gate. The NAND gate ECSM characterization done for the following two cases.

- Case 1: When the upper nMOS transistor in series-stack switches
- Case 2: When the lower nMOS transistor in series-stack switches

The region of validity of the t_{TCP} models in both the cases of 2-input NAND gate is derived. The relationship between cell size and model coefficient is also derived. Further, the impact of PVT variation on the model coefficient's is observed. The results depict that the proposed models are in good agreement with HSPICE simulations with a maximum error of 3%.

Chapter 6: This chapter focuses on overshoot timing model for CMOS inverter and 2-input NAND gate. For NAND gate, the boundary conditions are identified based on operating regions of the nMOS series-stack transistors. The relationship between cell size and model coefficient is also derived. The results depict that the overshoot time is independent of C_l and proportional to T_R. The proposed model gives a highly accurate estimation of delay values at nanometer range technologies.

Chapter 7: Finally, a summary of the presented research work along with the major conclusions of the work and future scope are presented in this chapter.

Chapter 2

Literature Review

2.1 Overview

This chapter starts with the study of existing delay models of CMOS inverter and NAND gate standard cell. There are several delay/timing models used in digital system design, some are briefly mentioned in the previous chapter. Further, we discuss the technical gaps present in the existing models. We also discuss the reason for their unsuitability in standard cell characterization.

2.2 Introduction

Semi-custom design approach at sub nanometer technologies requires efficient computer-aided design tools. Designers need to consider several aspects of the chip design such as timing, area and power. In this regard, designers need an accurate and efficient timing model in order to adequately optimize the standard cell based designs. For this an accurate and efficient method for characterizing standard cells is required. In this context, we discuss the limitations of the existing delay models in the following section.

2.3 Literature survey on existing delay models

As we discussed in previous section, down scaling and high performance circuits demand accurate timing analysis at sub-nanometer regime. At such technology nodes, a minor variation due to several factors like short channel effects, input to output coupling capacitance, power supply noise and process variations, etc. influences timing parameters significantly [6]. Standard cell or logic gate timing characterization must therefore account for all these parameters. Therefore, designing accurate and high performance circuits has become more challenging than ever in deep sub-micron era, in semi-custom domain. This section presents a brief description of past researcher's timing models for CMOS combinational logic gates.

The first delay model for CMOS inverter was introduced by Burns [53]. The model presented the differential equation based closed-form expression for the output waveform of CMOS inverter using a step input. Burns also derived a closed-form expressions for the

rise and fall time of CMOS inverter. Later on, the lumped RC delay model was proposed by Ousterhout [54]. In this model, the delay is computed by lumping all of the resistances and capacitances together through a stage. But, this model gives an overestimation of the delay due to lumped resistances and capacitances. And the major drawback of this model is that it doesn't deal with the shape of input waveform. To improve the delay analysis, Hedenstierna and Jeppson [19] presented an analytical delay model derived using Shockley's model [20], that includes the input waveform slope effect. Since, the Shockley's model is quite simple and has been used by many researchers. However, the model is not suitable for short channel devices as it does not consider the second order effects .

In [55], Jeppson presented an improved semi-empirical delay model by considering gate-to-drain coupling capacitance. Soon, Nabavi *et. al* [56] presented an empirical model for computing the inverter delay. In [56] , authors discussed the transient behavior for the two extreme cases i.e. very fast and very slow input transition times. For slow inputs, authors assumed that the negligible current flows through the load, whereas for fast inputs, the change in output voltage is negligibly small during the input transition. Hence, they ignored the effect of the output load on the delay for the extremely slow and fast inputs. All the models described above have a limitation that they ignored velocity saturation effect which is prominent at nanometer range technology nodes.

In [9], Sakurai *et. al* presented the "α- power law delay model" considering the velocity saturation effect for short channel devices. To derive this model authors neglected the short circuit current and gate-to-drain coupling capacitance. And, the model remains valid for fast input ramps only. Using the α-power law model, Embabi *et. al* [57] presented the delay model considering the short-circuit current into account. But, the model assumed the output voltage and the currents through the transistors to be piecewise linear. For slow input ramps, Dutta *et. al* [21] presented a delay model for CMOS inverter, which is an extension of Sakurai's delay model. Soon, Choi *et. al* [58] presented a delay model for CMOS logic gates to overcome the disadvantage of RC delay model. The model considered the MOSFET as a resistor or current source depending on the input and output voltage of the inverter. For NAND/NOR gate, authors assumed that N number of transistors in the series stack can be modeled as a single transistor. This assumption leads to significant error in the delay values. In [22], Sutherland *et al.* presented the method of "logical effort" which enables us to choose the topology and also tells us which topology is better for the circuit, how many stages would then be required to get the least delay and what should be size of the transistors in the gates. The method is quite simple and accurate if the input slope effect is ignored. Since, input transition time is one of the important parameters in the standard cell characterization which can not be ignored.

Traditional method (empirical models) to characterize delay is to use an equation of the form $k_1 C_l + k_2$, where k_1 is the input transition slope and k_2 is the intrinsic delay. Soon, Synopsys [10], presented the Scalable Polynomial Delay Model (SPDM) to characterize the cell delay. The model uses a product of polynomials to fit the delay data. For example, to characterize the delay for two parameters C_l and T_R, a product of m^{th} order polynomial in

C_l with an n^{th} order polynomial in T_R may be used in the following form:

$$delay = (a_0 + a_1 C_l + \ldots + a_m C_l^m)(b_0 + b_1 T_R + s \ldots + b_n T_R^n) \tag{2.1}$$

Which is a purely empirical expression and has to be extracted for each variation in any parameter affecting circuits. To avoid such complex expressions, industry people are now using LUTs to represent the delay of a logic gate. LUTs with delays tabulated for several values of T_R and C_l values are today used in STA. These LUTs are popular due to limitations of analytical and polynomial delay models as discussed above. LUT used for STA is a 2-D table, where the standard cell delay is characterized with several T_R and C_l values. Delays for the T_R and C_l values which are not listed in the LUTs are obtained using linear interpolation between the nearest two T_R and C_l values. There are several problem with the existing LUTs, for example T_R and C_l values are selected in an ad-ho c manner, need of re-characterization due to PVT variations, accuracy of the delay values obtained depends on the size of the LUT.

In [59], authors proposed a linear timing model to characterize delay and power dissipation of cells. The 50% delay is used to characterize the linear delay parameters. In [60], Patel presented a method to characterize the cell delay and capacitance parameters. However, these models are inconsistent when a cell drives different type of cell. In [61], authors proposed an LUT based approach to simplify the characterization of complicated cells. The method is particularly useful for the cases when the internal structure of standard cell is known. Further, in [62], Cirit used an LUT based approach to characterize the cells, where a cell being characterized is considered as a black box. It gave us the flexibility to consider any standard cell for characterization whose internal structure is not known. The model uses interpolation method to compute the delay values for those T_R and C_l values which are not given in the LUT. It requires software to perform mathematical analysis, thus makes the cell characterization process slow.

There are several delay models [63]-[73], proposed for CMOS inverter, and very few for the NAND gate standard cell. Recently, Gummalla *et al.* [2] presented an analytical timing model for 2-input NAND gate. The authors used Elmore delay model to consider the intermediate node voltage transition of the series stack for the switching of lower nMOS transistor in the series stack of the NAND gate, which may lead to significant error. This is because the upper nMOS transistor in the series stack operates in the saturation region for the C_l and T_R values typically found in circuits (explained in detail in Chapter 5). However, since they use Elmore delay model, they assumed it to be operating in linear region which results in a gross underestimation of delay values (as shown in Fig. 2.1).

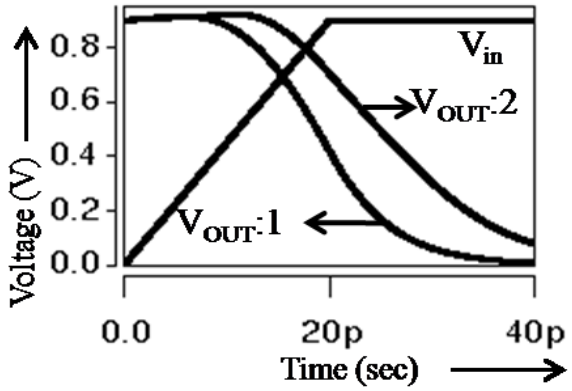


Figure 2.1: Response of the NAND gate for fixed input transition time; V_{out} : 1 represents the O/P voltage when the upper nMOS transistor is replaced by an equivalent resistor and V_{out} : 2 represents the O/P voltage for the realistic case where the upper nMOS of 2-input NAND gate is not replaced (For $V_{out:1}$, we obtain the equivalent resistance value from the linear current equation (1) of [2]).

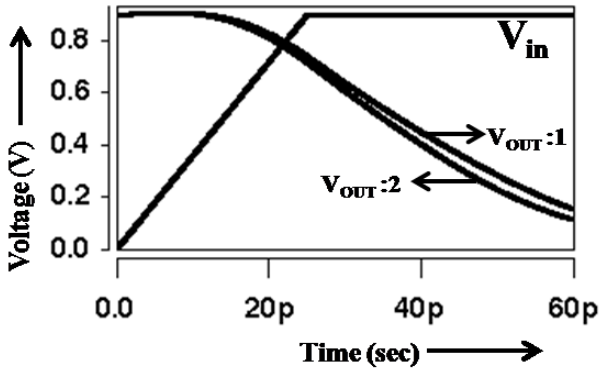


Figure 2.2: Response of the NAND gate for fixed input transition time; V_{out} : 1 represents the O/P voltage when the pull down part is replaced with the half width single nMOS transistor and V_{out} : 2 represents the O/P voltage for the conventional 2-input NAND gate.

When upper nMOS transistor in the series stack switches, Gummalla did not consider the effect of intermediate node of the series stack of nMOS transistors and replaced the series stack of nMOS transistors with the half width single nMOS transistor. We observe that this assumption leads to an overestimation of delay values as shown in Fig. 2.2.

Due to simplicity and accuracy, NLDM based on LUT approach is used widely for standard cell characterization by Synopsys. At sub-nanometer technologies, there are limitations of the NLDM in cell characterization which makes it less accurate. These limitations are the inaccurate shape of input waveform, undesirable value of nonlinear capacitance, etc [55]. Therefore, in modern CMOS technology, it becomes increasingly important to model the complex input waveforms, nonlinear capacitance and process variations [74]. In this scenario, the conventional standard cell characterization approach isn't found as an efficient technique to address problems. This conventional technique is useful to model the signal transitions as saturated ramps with known arrival and transition times. Therefore, researchers have introduced an alternative modeling technique known as Current Source

Modeling (CSM) which becomes increasingly important for use in standard cell characterization and static timing analysis (STA).

2.3.1 CSM based standard cell timing analysis

CSMs ideally support arbitrary input waveforms and output loads since their model parameters are waveform and load independent [75]. A current-based gate model includes a 2-D lookup table $I_0(V_i, V_0)$ which gives gate output current for a pair of gate input and output voltages, and voltage-controlled capacitor at the gate output. The CSM model proposed by Tutuianu *et al.* [76] is similar to [77, 78]. Croix *et al.* [14], proposed a CSM model which is independent of input waveform and output load as shown in Fig. 2.3. This "Blade" model consists of a voltage-controlled current source, an internal capacitance ($C_{internal}$), and a time shift of the output waveform. The model is essentially a $V_i - V_o$ based (input voltage, output voltage) current source with transient effects modeled by a linear capacitance at the output. A linear capacitance to model the active input load is assumed because the capacitances have a linear relationship with respect to device dimension for a given technology.

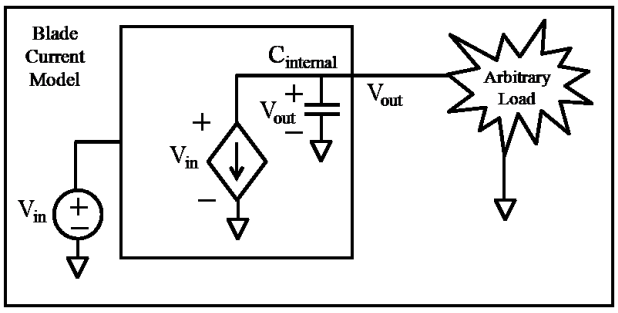


Figure 2.3: The Blade model consists of a voltage-controlled current source with a constant internal capacitance and input waveform time shift driving an arbitrary load.

It was the first CSM of a logic cell in which a pre-characterized current source is utilized to capture the non-linear behavior of the cell with respect to the input and output voltage values. The single output capacitance does not capture non-linearity. The miller effect between input and output nodes was ignored in this model. The ignorance of miller capacitance resulted in an under-estimation of delay. Keller et. al. [15], presented a CSM for the purpose of crosstalk noise analysis. The authors used a pre-characterized current source for the noise analysis. The parasitic components, namely the output and miller capacitances are assumed to be constant regardless of input and output voltage values. In practice, these capacitive effects can vary by orders of magnitude depending on cell input and output voltage values [79, 80].

In [16], Li and Acar has resolved this weakness by introducing a non-linear output capacitance model. Soon, Fatemi et.al. [3] used non-linear input, output and miller capacitances along with an output current source for the delay analysis, these all are function of the input and output voltages as shown in Fig. 2.4.

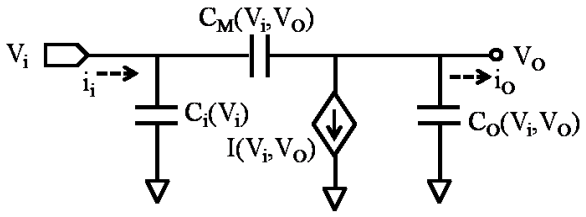


Figure 2.4: Current-based circuit model of a logic cell proposed by [3].

In [81], Kashyap *et. al.* presented a CSM in which input and output pins as well as several chosen internal pins of the cell are modeled with a voltage dependent current source and a non-linear capacitance. Veetil *et. al.* [82], investigated the importance of various modeling decisions on the accuracy and complexity of CSMs. The authors reported the bi-cubic spline based DC current source model for accurate and efficient timing analysis. For transient analysis, authors assumed that a cell can be replaced with simple parasitic capacitance model and a time shift parameter. These models require the pre-characterization of standard cells. Its very time consuming and cumbersome, as each parameter is dependent on input and output voltages.

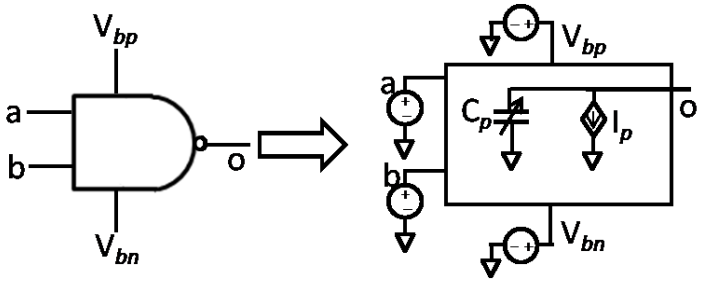


Figure 2.5: Example of a CSM: the output port is modeled as a nonlinear voltage controlled current source, dependent on all input port voltages, in parallel with a nonlinear capacitance [4]).

In 2010, Gupta *et al.* [4] developed a new approach to capture compactly the body bias effects within a mainstream CSM framework (shown in Fig. 2.5). The model is based on Blade model (given in [14]) except that the output port is replaced by a nonlinear voltage controlled current source, I_p, in parallel with a nonlinear capacitance, C_p. The mathematical framework for this new approach consists of two key steps. First approach is made by adapting an existing scheme to enable the compact storage of look-up tables for the sensitivities of CSM components to body bias, over the range of allowable values of the body bias. Second approach is on the basis of development of a novel waveform sensitivity model for evaluating the impact of applied body bias that provides accurate waveforms at the output of the cell under any body bias with minimal computation. Challenges of CSM [83] :

1. Each parameter of the cell is dependent on input and output node voltages which results in much larger libraries. Hence, library size has increased due to the development of CSM.

2. The complexity of the models has increased significantly over the NLDM models used for standard cell characterization.

EDA tool vendors found a middle path to solve these issues, known as vendor CSM formats. Two available CSM vendor formats are ECSM [84] and CCSM [85]. For a given input slew and load capacitance (C_l=C_{eff}) values, ECSM stores the times at which the output voltage waveform crosses certain predefined threshold points whereas CCSM stores the output current values at specified voltage level points. Both the vendor models are equivalent and one can be derived from the other. Vendor CSMs use LUT based format for representing characterization data. It overcomes the problems of the voltage-based models which is not compatible with the arbitrary shapes of voltage waveforms and falls short when dealing with crosstalk-induced noisy waveforms. ECSMs have recently received increased attention with major EDA vendors for supporting the noise model or power droop model [77, 78].

2.3.2 Overshoot timing Model for CMOS inverter and 2-input NAND gate

With the continuous scaling of devices in nanometer regime, the overshoot time becomes dominating component of gate delay for CMOS inverter standard cell. Due to the input-to-output coupling capacitance, the output voltage of a CMOS gate is beyond the power supply range at the beginning of the transition. This phenomenon is referred as overshoot effect and the time corresponds to the output voltage at power supply range is known as overshoot time. Several researchers [55, 56, 63, 64, 68, 72] accounted for the non linearity induced by the input-to-output coupling capacitance in their proposed gate delay model. Earlier, Turgis *et al.* [86] and Rossello *et al.* [72] estimated the power consumption of CMOS buffers under the consideration of the influence of input-to-output coupling capacitance in sub-micrometer technologies. Several empirical models have also been proposed to estimate overshoot time. In [63], the authors derived a closed form expression to compute the CMOS gate delay time. Using an empirical model, Rossello *et al.* [72] analyzed the CMOS gate power consumption. Recently, Huang *et al.* [49] modeled the overshoot effect of CMOS inverter delay in nanometer technologies. The proposed model accurately takes into account the input-to-output coupling capacitance of the CMOS inverter. The authors have verified the proposed model with 32nm PTM high-k/metal gate model [87, 46]. The overshoot effect in multi-input CMOS gate is also an important issue in current nanometer regime technologies. Recently, the authors in [46], presented the overshoot timing model for multi-input gate (NAND and NOR) taking the miller capacitance into consideration. The overshoot timing models presented in [49, 46], for CMOS inverter and multi-input gates are dependent on the load capacitance according to their model equation. But it is not validated by their own simulation results.

2.4 Technical Gaps

Based on literature survey, the importance of efficient and accurate delay model on device performance with technology scaling is observed. However, the accuracy of delay models is highly dependent on shape of waveform, parasitic capacitances, load capacitance and input transition time. Though, to enhance the circuit performance in standard cell characterization, extensive work has been done by the several authors, but the following major gaps and important issues are still there which have not been addressed.

1. Using NLDM or ECSM/CCSM based standard cell characterization, industry people extract all the delay point in C_l, T_R space using fully HSPICE generated characterization LUT. This approach takes a lot of time to obtain all the delay values in given C_l, T_R space. However, none of the above authors reported the region of validity of their proposed model.

2. Now a days, LUT based NLDM being used for standard cell characterization. To consider the PVT variations at such technology node, there is need of re-characterization of these LUTs resulting in huge characterization effort and time as well. None of the above author, addressed the method to reduce this re-characterization effort.

3. Below 65nm technology node, stress engineering is being incorporated to enhance the device performance. None of the above author reported the delay model which can be used in standard cell characterization, considering the variation in device channel mechanical stress as a function of NF.

4. There are several delay models for CMOS inverter (as we discussed above), but very few for the NAND gate. The latest model for NAND gate is reported by Gummalla [2]. The model considers the intermediate node voltage transition to obtain the propagation delay. The author makes use of Elmore delay model to consider the intermediate node voltage transition. This assumption leads to significant percentage error in delay values.

5. There are very few overshoot timing models, recently Huang *et al.* reported the overshoot timing model for CMOS inverter and multi-input gates (NAND, NOR). They reported in the paper that overshoot time is function of C_l, which it is not validated by their own results. No one has reported earlier that overshoot time is independent of C_l.

Chapter 3

Timing model for CMOS inverter standard cell and its region of validity

3.1 Overview

This chapter focuses on modeling of $t_{TCP}s$ as a function of T_R and C_l for a CMOS inverter. Further, the region of validity of the model in T_R, C_l space is derived. The relationship between cell size and model coefficients are also derived. While developing the model, we make appropriate assumptions and later justify the use of all our assumptions. Further, the impact of technology scaling on these model coefficients is considered. The results depict that the proposed model is in good agreement with HSPICE simulations with a maximum error of 2.5%. We then propose a method to use our t_{TCP} models in reducing the number of HSPICE simulations in ECSM characterization of standard cells. The contribution of these models in reducing the number of HSPICE simulations in ECSM characterization of the inverter standard cell is nearly half.

The chapter is organized as follows. In Section 6.2, we describe our simulation setup. In Section 3.3, we derive our t_{TCP} models. In Section 3.4, we use these models to reduce the number of simulations for ECSM characterization. In Section 3.5, we verify the validity of proposed models with respect to technology node.

3.2 Simulation Setup

In this chapter, we use HSPICE simulations at 32nm CMOS technology node. In these simulations, we use BSIM 4.0 Predictive Technology device Model (PTM)[1]. We keep $W_p/W_n = 2.5$ to obtain equal inverter's rise and fall transition times. Therefore, W_n can represent the size of inverter standard cell. We verify our models with technology scaling using HSPICE simulations at 22nm CMOS technology node. In this chapter, the value of parasitic capacitance (C_p) of CMOS inverter, is extracted using the integral of the difference of the currents through the sources of pMOS and nMOS for $80\% - 20\%$ of output transition.

[1]Obtained from http://ptm.asu.edu/

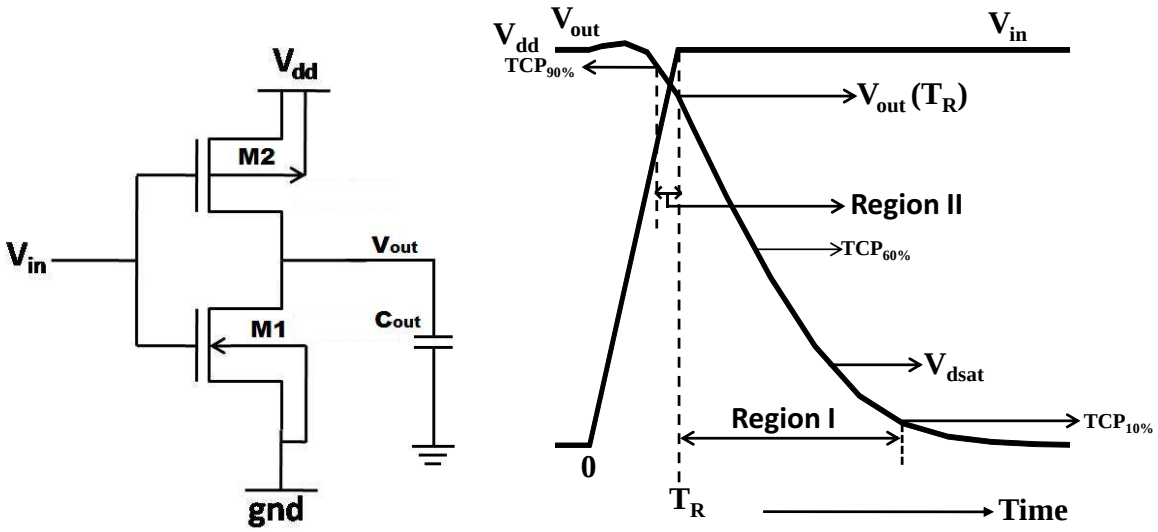


Figure 3.1: CMOS inverter with input and output waveform.

In this chapter, we represent the $0 - V_{dd}$ input transition as T_R and $20 - 80\%$ of V_{dd} of input transition as T_{Rin}. When rising input transition starts to increase, respective time is represented as $t = 0$.

3.3 Timing model for CMOS inverter standard cell

In this section, we first develop a timing model for CMOS inverter. We assume the rising input transition for the derivation of $TCPs$ model, similar analysis is valid for falling input transition. For our derivation, we classify the $TCPs$ into two regions (shown in Fig. 3.1):

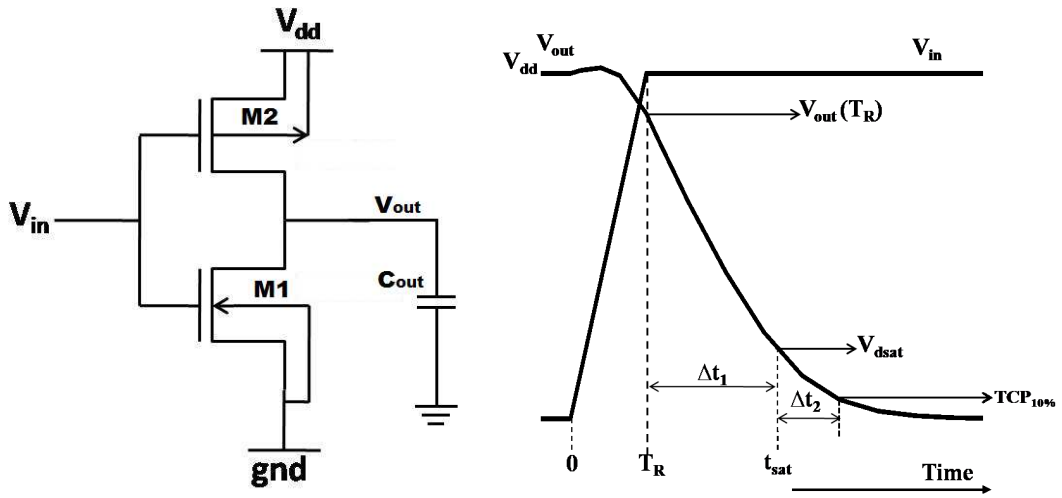
- Region I: When $V_{in} = V_{dd}$ (for $t_{TCPs} > T_R$)
- Region II: When $V_{in} < V_{dd}$ (for $t_{TCPs} < T_R$)

3.3.1 Derivation of the model in Region I

In this subsection, we derive the relationship of t_{TCPs} with T_R, C_l values and size of inverter (W_n) for Region I. As we depict in Fig. 3.1, Region I contain all the TCPs having value V_{TCP} smaller than $V_{out}(T_R)$. The proposed model remains valid for all the cases where $V_{out}(T_R) \geq V_{dsat}$, where V_{dsat} is the for the inverter's (discharging) nMOS device [2]. The output discharge comprises of two regions: First, when the input transitions from 0 to V_{dd} and second, when the input voltage reaches V_{dd}.

In this derivation, we assume that the discharging nMOS operates either in saturation or in linear region. To derive the model, we first integrate the saturation current through the nMOS during the input transition $V_{in}(t) = V_{dd}(\frac{t}{T_R})$ for $0 \leq t \leq T_R$. We equate this

[2]We assume that the value of V_{dsat} is very weakly dependent on the values of V_{ds}, as in [9].



(a) (b)

Figure 3.2: (a) CMOS inverter schematic. (b) I/O waveform of CMOS inverter for Region I.

integral to $(C_l + C_p)(V_{dd} - V_{out}(T_R))$ to obtain $V_{out}(T_R)$. The output voltage discharge $\triangle Q(T_R)$ from 0 to T_R is given as (refer Fig. 3.2):

$$\triangle Q(T_R) = \int_0^{T_R} I_{M1} dt = (C_l + C_p)(V_{dd} - V_{out}(T_R)) \tag{3.1}$$

Where, C_l is the load capacitance and C_p is the parasitic capacitance. The values of I_{ON}, I_{lin}, V_{dsat} and R used in this model are obtained from alpha power law model [9] as:

$$I_{ON} = v_{sat}\, W_n\, P_s\, (V_{gs} - V_{th})^{\alpha} \tag{3.2}$$

$$I_{lin} = \mu \frac{W_n}{L_{eff}} P_l\ (V_{gs} - V_{th})^m\, V_{ds} \tag{3.3}$$

$$V_{dsat} = \frac{v_{sat}\, P_s\ (V_{gs} - V_{th})^m\, L_{eff}}{\mu\, P_l} \tag{3.4}$$

$$R = \frac{1}{\mu \frac{W_n}{L_{eff}} P_l\ (V_{gs} - V_{th})^m} \tag{3.5}$$

Where P_s, P_l are technology dependent parameters, v_{sat} is the saturation velocity and μ is the mobility of the nMOS device. The exponents m and α are velocity saturation indices, which are also technology dependent. In our case, we use $m = \alpha = 1$, which we verify through HSPICE simulations. We now explain our derivation of t_{TCP} in Region I:

1. We integrate the saturation current through the nMOS from $t = 0$ to T_R and equate this integral to $(C_l + C_p)(V_{dd} - V_{out}(T_R))$ to find the expression of $V_{out}(T_R)$.

2. Next, we proceed as follows:

- If $V_{TCP} > V_{dsat}$, we integrate the discharging nMOS current, I_{ON} ($V_{gs} = V_{dd}$), from $t = T_R$ to $t = t_{TCP}$. We equate the sum of these integrals to $(C_l+C_p)(V_{out}(T_R)-V_{TCP})$ to obtain t_{TCP}.

- If $V_{TCP} < V_{dsat}$, we integrate the discharging nMOS current, I_{ON} ($V_{gs} = V_{dd}$), from $t = T_R$ to $t = t_{sat}$, represented as Δt_1 in Fig. 3.2, (where $t = t_{sat}$ when $V_{out} = V_{dsat}$). We equate the sum of these integrals to $(C_l + C_p)(V_{out}(T_R) - V_{dsat})$ to obtain $\triangle t_1$. From $t = t_{sat}$ to $t = t_{TCP}$, represented as Δt_2, shown in Fig. 3.2, where, the nMOS transistor operates in linear region. We find that Δt_2 proportional to time constant RC. We add Δt_1 and Δt_2, to obtain the t_{TCP} model.

From $t = t_{sat}$ to $t = t_{TCP}$, the nMOS device acts as a resistance (R) (given in (3.5)) and its value is obtained by differentiating the linear region current equation of M_1 with V_{ds}, equating $V_{gs} = V_{dd}$. Therefore, the time duration from t_{sat} to t_{TCP} is proportional to time constant $R(C_l + C_p)$.

We observe that the same assumptions are valid for all the $TCPs$ under Region I if the constraint $V_{out}(T_R) \geq V_{dsat}$ is followed. We derive the model for $t_{TCP_{10\%}}$(because, $t_{TCP_{10\%}} > T_R$), since it lies somewhere in the time span $\triangle t_2$ (refer to Fig. 3.2). The term $t_{TCP_{10\%}}$ defines the time at which output waveform crosses a voltage level of $0.1V_{dd}$. Thus, the $t_{TCP10\%}$ model would be a representative for all the TCPs when the nMOS device goes into linear region.

To derive the $t_{TCP_{10\%}}$ model, we follow the procedure, as discussed in the previous paragraph. We need to find out the expressions for $\triangle t_1$ and $\triangle t_2$. The timing model for $t_{TCP10\%}$ can be written as:

$$t_{TCP10\%} = T_R + \triangle t_1 + \triangle t_2 \tag{3.6}$$

Please note that we measure the $t_{TCP}s$ from $t = \frac{V_{th}}{V_{dd}} T_R$ to $t = t_{TCP10\%}$, as we assume that nMOS starts operating in saturation region at $t = \frac{V_{th}}{V_{dd}} T_R$. In (3.6), $\triangle t_1$ is represented as the time taken by output voltage to discharge from $V_{out}(T_R)$ to V_{dsat}. The output voltage discharge $\triangle Q(t_1)$ from $V_{out}(T_R)$ to V_{dsat} is given as:

$$\triangle Q(t_1) = \int_{T_R}^{t_{sat}} I_{M1} dt = (C_l + C_p)(V_{out}(T_R) - V_{dsat}) \tag{3.7}$$

Solving (3.1) and (3.7), we get the expression for $\triangle t_1$. As we discussed earlier, from $t = t_{sat}$ to $t = t_{TCP}$ *i.e.* $\triangle t_2$ is proportional to time constant $R(C_l + C_p)$. Using all the above equations, we get the $t_{TCP10\%}$ as:

$$t_{TCP_{10\%}} = K_1 C_l + K_2 T_R + K_3 \tag{3.8}$$

Where, K_1, K_2 and K_3 are the coefficients extracted by fitting (4.1) into HSPICE simulated data.

Where,

$$K_1 = \frac{(V_{dd} - V_{dsat})}{I_{ON}} + R \tag{3.9}$$

$$K_2 = \left[0.8 - \frac{S_T}{I_{ON}}\right] \tag{3.10}$$

Where,

$$S_T = v_{sat} W_n P_s (\frac{V_{dd}}{2} - V_{th})$$

$$K_3 = C_p \left[\frac{(V_{dd} - V_{dsat})}{I_{ON}} + R\right] \tag{3.11}$$

We obtain a linear relationship between t_{TCP} and C_l and T_R as given in (4.1). We observe that the same assumptions are valid for all the $TCPs$ under Region I since the constraint $V_{out}(T_R) \geq V_{dsat}$ is followed. In this chapter, we show the model validation results for $t_{TCP60\%}$, since $t_{TCP60\%} > T_R$. Where, $t_{TCP_{60\%}}$ defines the time at which output waveform crosses a voltage level of $0.6V_{dd}$. For $t_{TCP60\%}$, the form of model remains same. The coefficient values also remains same, the only difference is that the term R will not be present. The same model (*i.e.* $t_{TCP60\%}$) is valid for t_{sat}. The following observations have been made from the derivation of (4.1):

- Observation 1: K_1 is a linear function of $1/W_n$
- Observation 2: K_2 and K_3 both are independent of W_n

As explained earlier, (4.1) is valid if $V_{out}(T_R) \geq V_{dsat}$, this imposes the following constraint on the region of validity:

$$\triangle Q(T_R) = S_T T_R \leq (C_l + C_p)(V_{dd} - V_{dsat}) \tag{3.12}$$

Where $\triangle Q(T_R)$ is the output discharge from 0 to T_R, S_T is a constant proportional to W_n, C_p is the inverter's parasitic capacitance proportional to W_n, V_{dd} is the power supply voltage. Further we use the term t_{rb}, it denotes the maximum value of T_R which satisfy (3.12). Equation (3.12) shows the linear relationship between t_{rb} and C_l. The following observations are made regarding the derived linear relationship:

- Observation 3: Slope of t_{rb} vs C_l plot (*i.e.* S_{trb}) is proportional to $1/W_n$
- Observation 4: Intercept is a constant with W_n

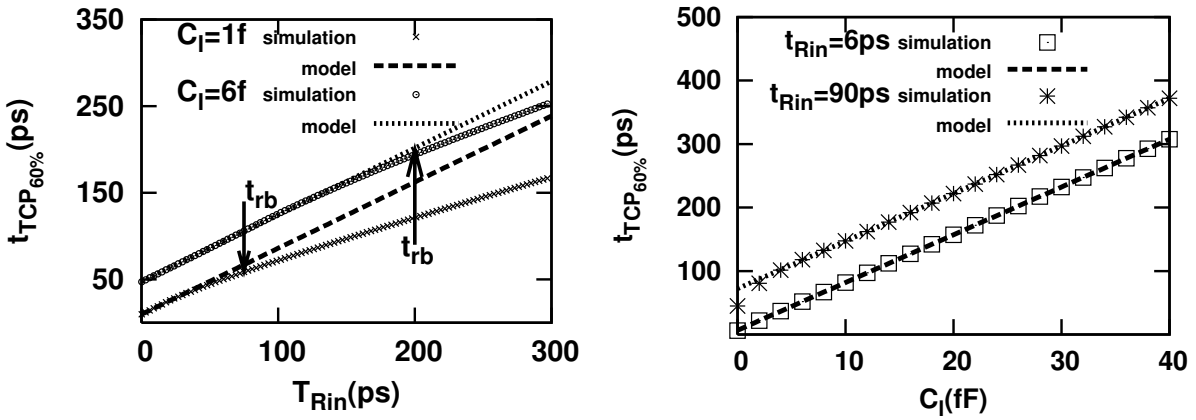


Figure 3.3: Variation of $t_{TCP_{60\%}}$ with respect to T_{Rin} and C_l values.

3.3.1.1 Verification of the model

We verified Observation 1-4 with HSPICE $32nm$ BSIM PTM shown in Fig. 3.4 to Fig. 3.5. Fitting (4.1) on the simulated values of $t_{TCP_{60\%}}$ (as shown in Fig. 3.3), we extracted the coefficients (K_1, K_2, K_3) of (4.1).

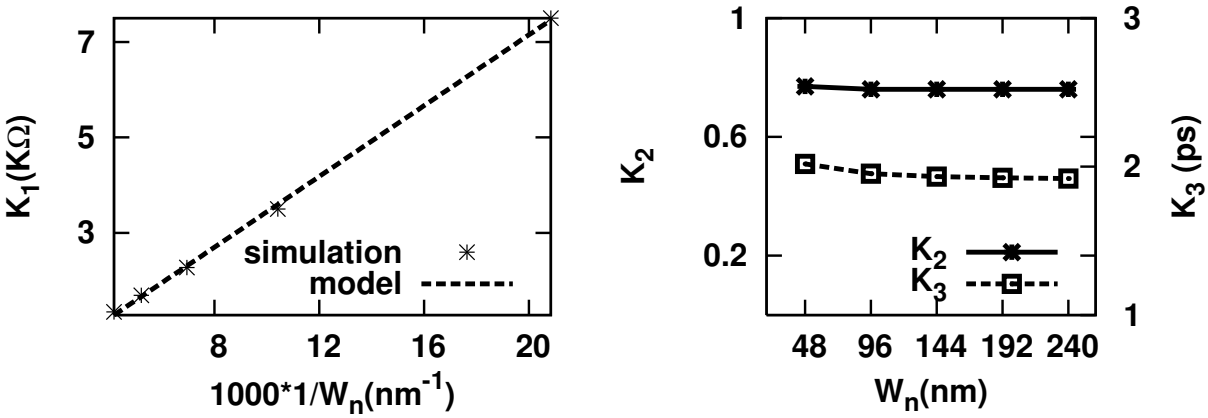


Figure 3.4: Variation of K_1, K_2 and K_3 with W_n.

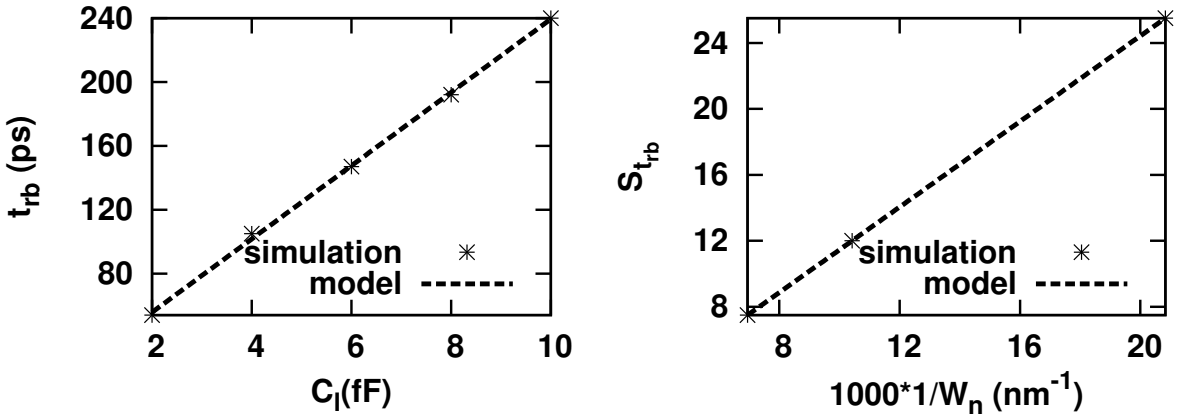


Figure 3.5: S_{trb} variation with W_n .

Here, the points are the simulated data and the discontinuous lines are the curve fitting

of the proposed model. The t_{TCP} model given by (4.1) for Region I has thus been verified using HSPICE simulated data for $TCP_{60\%}$.

3.3.2 Derivation of the model in Region II

In this section, we derive the relationship of t_{TCPs} with T_R, C_l and size of inverter (W_n) for Region II when $V_{out}(T_R) \geq V_{dsat}$ (Refer to Fig. 3.6). As a representative, we derived model for $t_{TCP_{90\%}}$ shown in Fig. 3.6. This is because $t_{TCP_{90\%}} < T_R$ in the whole T_R, C_l space used in characterization LUTs.

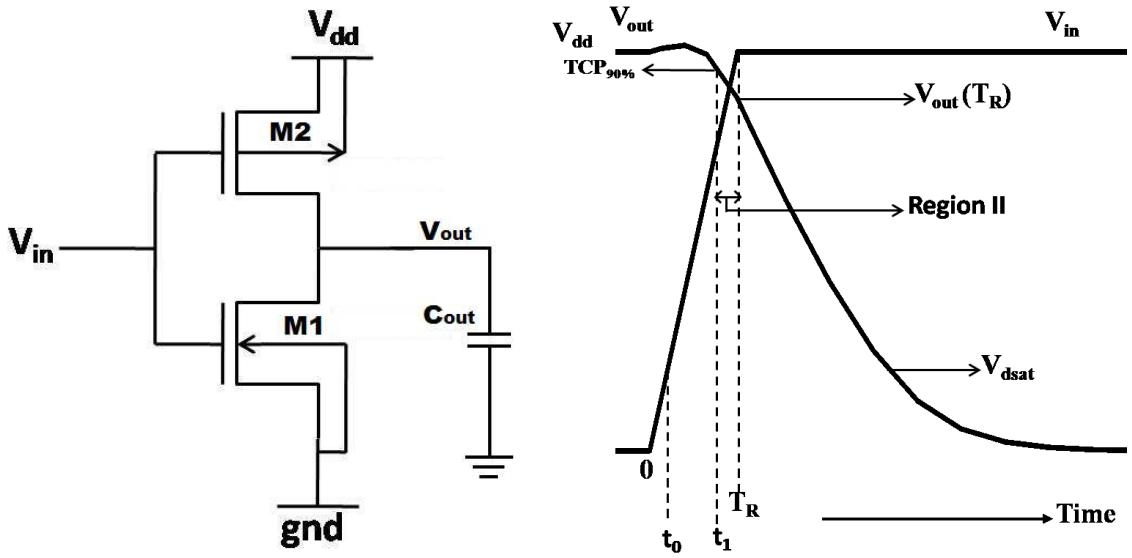


Figure 3.6: I/O waveform of CMOS inverter for Region II.

We assume the rising input transition for the derivation of $t_{TCP90\%}$ model and it is given as;

$$V_{GS} = \frac{V_{dd}}{T_R} \times t \qquad for\,(0 \leq t \leq T_R)$$

$$= V_{dd} \qquad for\,(t > T_R) \tag{3.13}$$

Where V_{GS} is the gate source voltage of inverter's nMOS device, t is the time and T_R is the transition time of input rising from 0 to V_{dd}. In this derivation, we assume that the discharging nMOS operates in velocity saturation, respective current is represented as I_{ON} given in (3.3) derived from alpha power law model [9]. The output voltage discharge $\triangle Q(t_1)$ through nMOS device from $t_0 = \left(\frac{V_{th}}{V_{dd}} T_R\right)$ to $t_1 (= t_{TCP90\%})$ (refer to Fig. 3.6) is:

$$\Delta Q(t_1) = \int_{t_0}^{t_1} I_{M1}\, dt \tag{3.14}$$

The value of $\left(\frac{V_{th}}{V_{dd}}\right)$is much smaller than 1 in CMOS technology. In our case it is $0.2V$.

$$\frac{\Delta Q(t_1)}{\beta_s} = \int_{t_0}^{t_1} \left(\frac{V_{dd}.t}{T_R} - V_{th}\right) dt \tag{3.15}$$

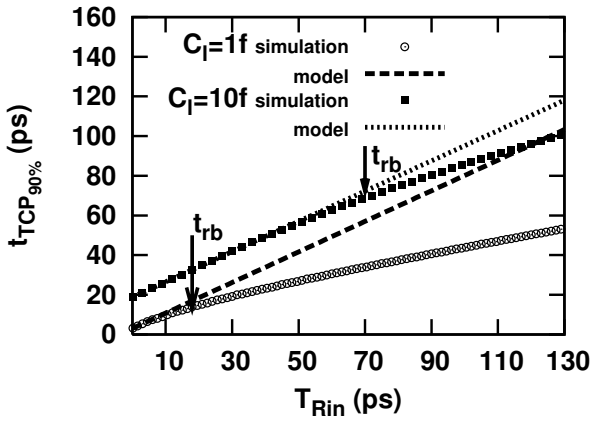


Figure 3.7: Variation of $t_{TCP_{90\%}}$ with respect to T_{Rin}.

Where, for $t_{TCP90\%}$, $\Delta Q(t_1) = 0.1V_{dd}(C_l + C_p)$ and $\beta_s = v_{sat}\, W_n\, P_s$. Solving (3.15), we get an expression for t_1 as follows:

$$t_{TCP90\%} = A_1 T_R + A_2 \sqrt{T_R} \tag{3.16}$$

Where,

$$A_1 = \left(\frac{V_{th}}{V_{dd}}\right) \tag{3.17}$$

and

$$A_2 = \left(\sqrt{\frac{0.2(C_l + C_p)}{\beta_s}}\right) \tag{3.18}$$

Where, $t_{TCP_{90\%}}$ defines the time at which output waveform crosses a voltage level of $0.9V_{dd}$ and A_1, A_2 are the model's coefficients (extracted by fitting (4.2) into HSPICE simulated data). The following observations have been made from the (4.2) :

- Observation 5: A_1 is independent of W_n and C_l
- Observation 6: A_2 is proportional to $\sqrt{C_l}$
- Observation 7: A_2 is a linear function of $\sqrt{\frac{1}{W_n}}$

Similar to subsection 3.3.1, the region of validity of proposed model for Region II is:

$$\triangle Q(T_R) = S_T T_R > (C_l + C_p)(V_{dd} - V_{dsat}) \tag{3.19}$$

Equation (5.20) shows a linear relationship between t_{rb} and C_l. Therefore, Observation 3 and 4 remains valid for the proposed model in Region II. The same assumptions (as discussed in this subsection) are valid for all the $TCPs$ under Region II since the constraint $V_{out}(T_R) \geq V_{dsat}$ is followed.

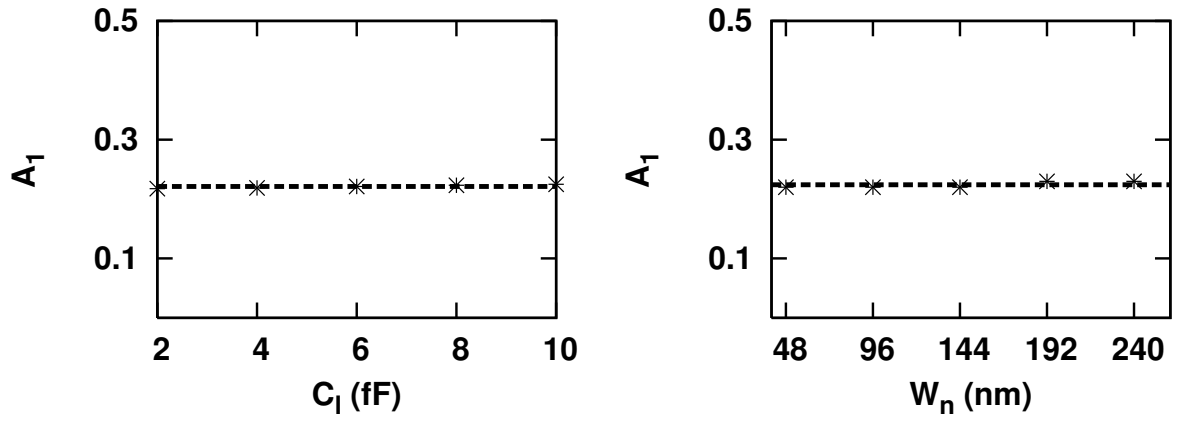


Figure 3.8: A_1 is independent of C_l and W_n.

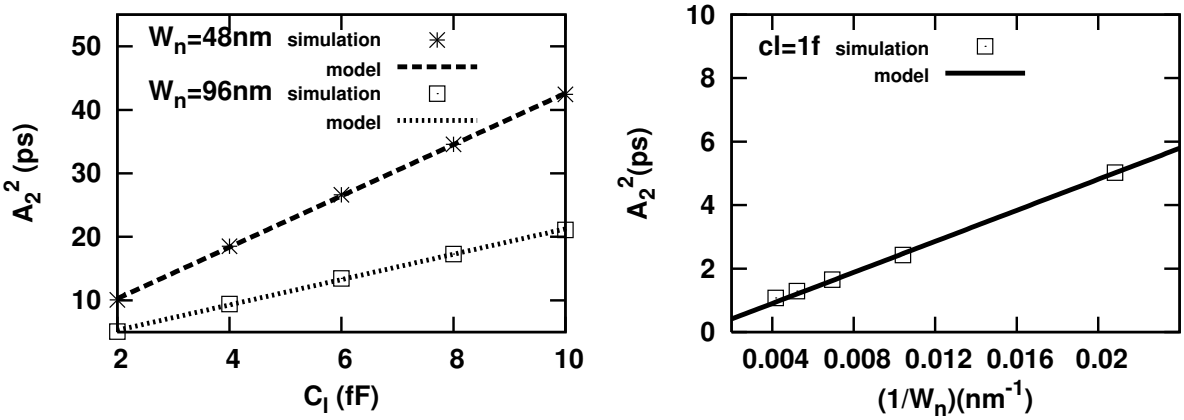


Figure 3.9: A_2 varies with C_l and $\frac{1}{W_n}$ as predicted by (4.7).

3.3.2.1 Verification of the model

We verify Observation 5-7 using HSPICE simulations of $32nm$ BSIM PTM. The t_{TCP} model given by (4.2) for Region II has also been verified using HSPICE simulated data for $TCP_{90\%}$. Please note that we have not included the inverse narrow width effects in the proposed work. However, the corresponding changes in the values of model's coefficients can be predicted using well known equations in [88] and our models.

3.4 Efficient ECSM Characterization

In this section, we use the models derived in Section 3.3 to reduce the number of HSPICE simulations required in ECSM (or CCSM) characterization of an inverter standard cell. Using (4.1) and (4.2), within their regions of validity, we can get the values of all $t_{TCP}s$ without simulations. Hence, it saves HSPICE simulations in standard cell characterization. On the other hand, the t_{TCP} values which are out of validity bound in Region I and II, will be calculated from HSPICE simulations. For an inverter standard cell, we first extracted the values of K_1, K_2 and K_3 for TCPs in Region 1 using 7 HSPICE simulations and A_1, A_2 using 4 HSPICE simulations. Later on, we calculate the values of $t_{TCP}s$ (entries shown by numeric values in Table 3.1 and 3.4) for C_l, T_R values lying within the region of validity

Table 3.1: LUT of $TCP_{60\%}$ for minimum size CMOS inverter using our Region I model.

$\mathbf{C_l}$ (fF)	$\mathbf{T_R}$ (ps)						
	2.20	**4.84**	**10.66**	**23.46**	**51.62**	**113.60**	**250.00**
1.51	14.35	15.57	18.26	24.17	HSPICE	HSPICE	HSPICE
2.28	20.14	21.36	24.05	29.96	42.97	HSPICE	HSPICE
3.45	28.89	30.11	32.79	38.70	51.72	HSPICE	HSPICE
5.22	42.10	43.32	46.01	51.92	64.93	93.57	HSPICE
7.88	62.07	63.29	65.98	71.89	84.90	113.54	HSPICE
11.91	92.25	93.47	96.16	102.07	115.08	143.72	206.74
18.00	137.86	139.08	141.77	147.68	160.69	189.33	252.34

Table 3.2: LUT of $TCP_{60\%}$ for minimum size CMOS inverter obtained using HSPICE simulations.

$\mathbf{C_l}$ (fF)	$\mathbf{T_R}$ (ps)						
	2.20	**4.84**	**10.66**	**23.46**	**51.62**	**113.60**	**250.00**
1.51	14.49	15.70	18.37	24.32	37.48	61.71	106.50
2.28	20.30	21.51	24.18	30.11	43.43	70.28	118.65
3.45	29.09	30.30	32.98	38.89	52.09	81.07	134.17
5.22	42.36	43.58	46.25	52.15	65.28	94.65	153.74
7.88	62.42	63.64	66.31	72.21	85.26	114.34	178.27
11.91	92.72	93.94	96.62	102.52	115.52	144.40	209.10
18.00	138.52	139.74	142.41	148.31	161.29	190.03	254.07

of (4.1) and (4.2).

Table 3.3: Percentage error in proposed model's LUT with respect to fully HSPICE generated ECSM LUT of $TCP_{60\%}$ for CMOS inverter. Entries shown by '×' correspond to the values obtained from HSPICE simulations (not through our models).

$\mathbf{C_l}$ (fF)	$\mathbf{T_R}$ (ps)						
	2.20	**4.84**	**10.66**	**23.46**	**51.62**	**113.60**	**250.00**
1.51	0.97	0.83	0.60	0.62	×	×	×
2.28	0.79	0.70	0.54	0.50	1.06	×	×
3.45	0.69	0.63	0.58	0.49	0.71	×	×
5.22	0.61	0.60	0.52	0.44	0.54	1.14	×
7.88	0.56	0.55	0.50	0.44	0.42	0.70	×
11.91	0.51	0.50	0.48	0.44	0.38	0.47	1.13
18.00	0.48	0.47	0.45	0.42	0.37	0.37	0.68

In Table 3.1 and Table 3.4, $t_{TCP}s$ values (shown by 'HSPICE') of C_l, T_R which are out of region of validity for (4.1) and (4.2), will be extracted by HSPICE simulations. Table 3.2 and Table 3.5 shows the $t_{TCP}s$ values obtained using conventional HSPICE simulations. Whereas Table 3.3 and 3.6 shows percentage error in LUT generated by proposed models with respect to fully HSPICE generated ECSM LUT of $TCP_{60\%}$ and $TCP_{90\%}$ for CMOS inverter. For inverter standard cells having a different size, we use the relationships of K_1, K_2, K_3 and A_1, A_2 with W_n (discussed in Section 3.3) to re-generate TCP LUTs as discussed above. Thus, for these cells with a different size, we need not extract the model's

Table 3.4: LUT of $TCP_{90\%}$ for minimum size CMOS inverter using our Region II model.

$\mathbf{C_l}$ (fF)	$\mathbf{T_R}$ (ps)						
	2.20	**4.84**	**10.66**	**23.46**	**51.62**	**113.60**	**250.00**
1.51	HSPICE	HSPICE	8.48	13.54	22.21	37.61	66.06
2.28	HSPICE	HSPICE	9.88	15.63	25.30	42.20	72.87
3.45	HSPICE	HSPICE	HSPICE	18.19	29.10	47.84	81.23
5.22	HSPICE	HSPICE	HSPICE	21.34	33.77	54.77	91.51
7.88	HSPICE	HSPICE	HSPICE	HSPICE	39.51	63.28	104.15
11.91	HSPICE	HSPICE	HSPICE	HSPICE	46.57	73.76	119.68
18.00	HSPICE	HSPICE	HSPICE	HSPICE	HSPICE	86.63	138.78

coefficients using simulations.

Table 3.5: LUT of $TCP_{90\%}$ for minimum size CMOS inverter obtained using HSPICE simulations.

$\mathbf{C_l}$ (fF)	$\mathbf{T_R}$ (ps)						
	2.20	**4.84**	**10.66**	**23.46**	**51.62**	**113.60**	**250.00**
1.51	5.05	6.26	8.70	13.86	22.25	37.20	64.86
2.28	6.40	7.62	9.10	15.79	25.19	41.65	71.51
3.45	8.45	9.66	12.34	18.21	28.84	47.14	79.79
5.22	11.54	12.75	15.42	21.33	33.34	53.89	89.99
7.88	16.2	17.42	20.09	25.98	38.91	62.20	102.50
11.91	23.24	24.46	27.13	33.03	46.04	72.45	117.84
18.00	33.88	35.10	37.77	43.67	56.65	85.12	136.67

Table 3.6: Percentage error in proposed model's LUT with respect to fully HSPICE generated ECSM LUT of $TCP_{90\%}$ for CMOS inverter. Entries shown by '×' correspond to the values obtained from HSPICE simulations (not through our models).

$\mathbf{C_l}$ (fF)	$\mathbf{T_R}$ (ps)						
	2.20	**4.84**	**10.66**	**23.46**	**51.62**	**113.60**	**250.00**
1.51	×	×	2.50	2.31	0.18	1.10	1.85
2.28	×	×	2.40	1.01	0.44	1.32	1.90
3.45	×	×	×	0.11	0.90	1.48	1.80
5.22	×	×	×	0.05	1.29	1.63	1.69
7.88	×	×	×	×	1.54	1.74	1.61
11.91	×	×	×	×	1.15	1.81	1.56
18.00	×	×	×	×	×	1.77	1.54

Table 3.7 shows the percentage saving in HSPICE simulations using our method of generating LUTs explained above. We find that the values of $t_{TCP}s$ in our LUTs differ by a maximum of 2.5% from those in fully HSPICE generated conventional LUTs. In Table 3.7, we see a 50% saving in required number of HSPICE simulations for $TCP_{90\%}$. Therefore, we observe that standard cell characterization can be done with a significantly lesser number

Table 3.7: Percentage saving in HSPICE simulation for ECSM characterization using our models.

$TCPs$	LUT's size	Number of matrix elements obtained using		% saving
		Proposed model	HSPICE	
$TCP_{60\%}$	7×7	40	9	81.63
	8×8	52	12	81.25
	9×9	65	16	80.25
$TCP_{90\%}$	7×7	26	23	53.06
	8×8	32	32	50.00
	9×9	42	39	51.85

of HSPICE simulations (approximately 50% reduction).

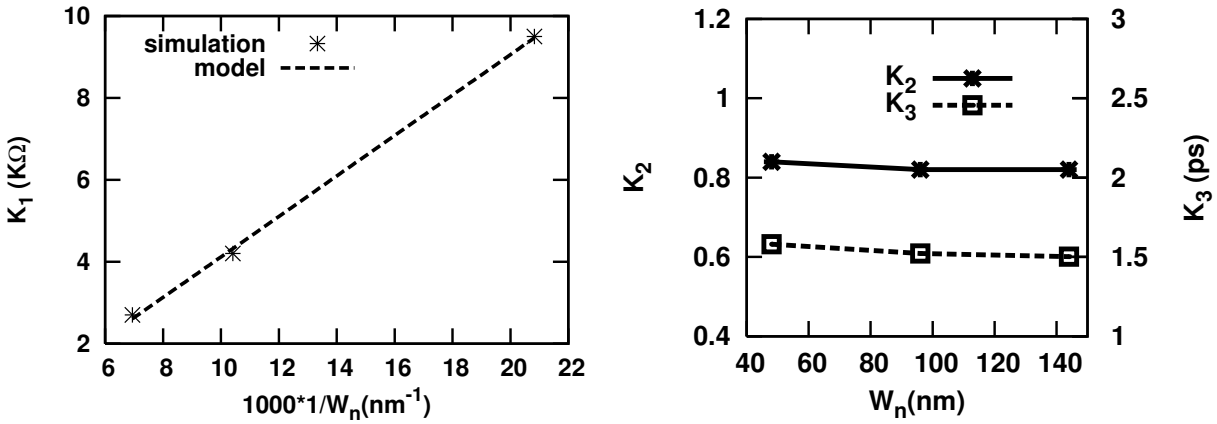


Figure 3.10: Variation of K_1, K_2 and K_3 with W_n at 22nm CMOS technology node.

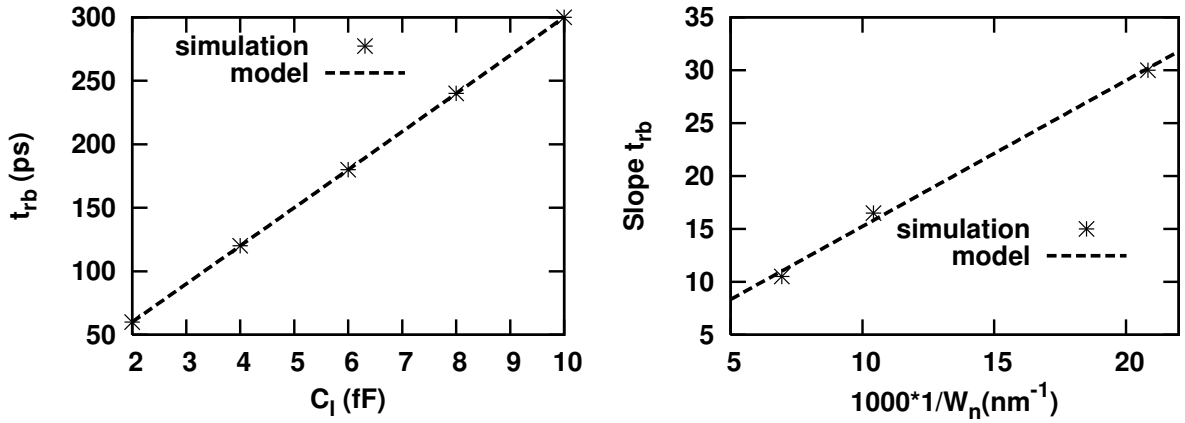


Figure 3.11: S_{trb} variation with W_n at 22nm CMOS technology node.

3.5 Impact of technology scaling on timing models

For a timing model to be used in standard cell characterization, it should be valid as well as accurate with respect to technology scaling. In this section, we show that our

proposed models remain valid with technology scaling while maintaining its accuracy. For validation, we perform our analysis at 22nm CMOS technology node.

We verified Observation 1-4 with HSPICE 22nm BSIM PTM as shown in Fig. 3.10 and 3.11. The t_{TCP} model given by (4.1) for Region I has thus been verified using HSPICE simulated data for $TCP_{60\%}$. We verified Observation 5-7 with HSPICE 22nm BSIM PTM as shown in Fig. 3.12 and 3.13. The t_{TCP} model given by (4.1) for Region II has thus been verified using HSPICE simulated data for $TCP_{90\%}$.

3.6 Summary

In this chapter, we proposed models for $t_{TCP}s$ of output voltage transition for standard cell CMOS inverter. The t_{TCP} values are derived in terms of T_R and C_l. We also derived the region of validity of these models in T_R, C_l space. The relationship between cell size and model coefficients are also derived. Further, we derived the relationship of model coefficients with the technology scaling. The proposed model are in good agreement with HSPICE simulations with a maximum error of 2.5%. We later used these models to reduce the number of HSPICE simulations by about half in ECSM characterization of standard cell CMOS inverter.

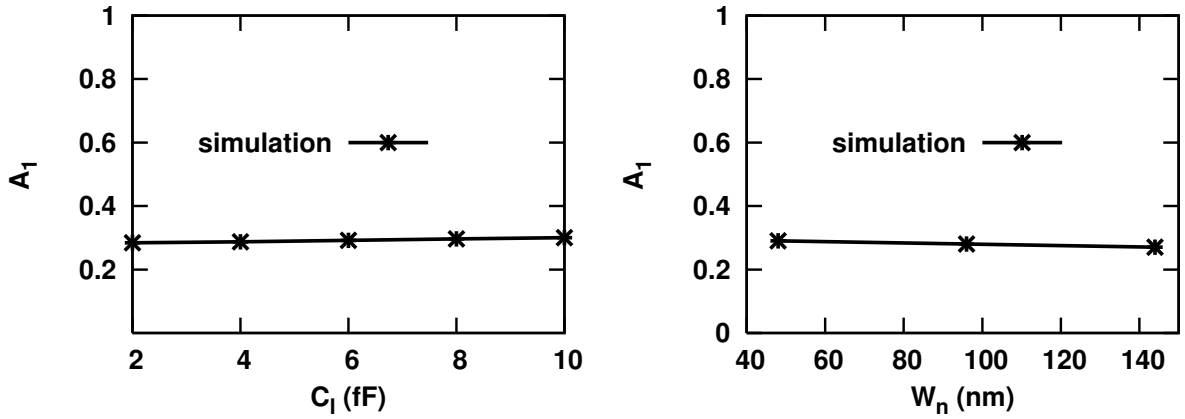


Figure 3.12: Variation of A_1 with C_l and W_n at 22nm CMOS technology node.

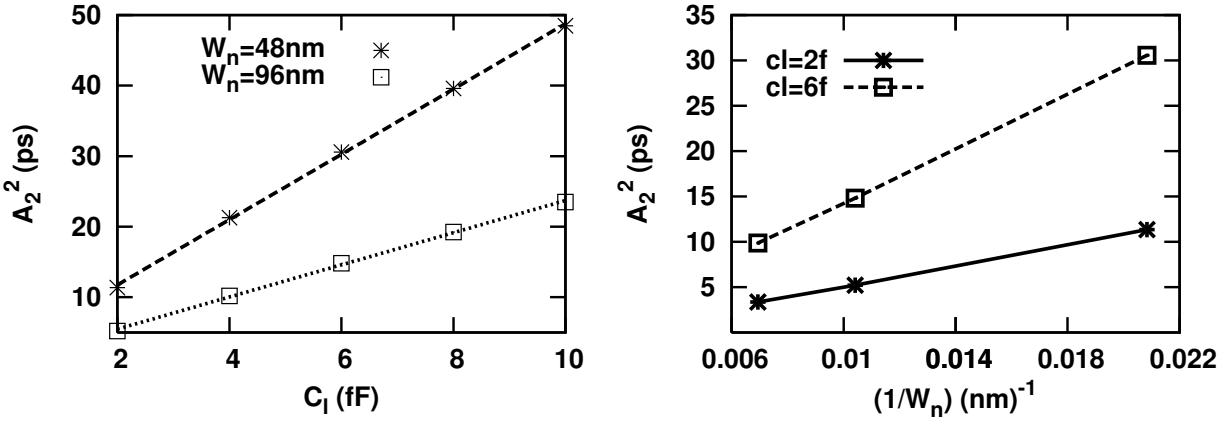


Figure 3.13: Variation of A_2 with C_l and $\frac{1}{W_n}$as predicted by (4.7) at 22nm CMOS technology node.

Chapter 4

Efficient ECSM Characterization of CMOS Inverter Standard Cell Considering PVT Variations

4.1 Overview

In the previous chapter, we discussed our proposed timing models and its region of validity for efficient characterization of CMOS inverter standard cell. In this chapter, we show that our proposed models remain valid with voltage, temperature and stress variability. We derive relationships of variation of our model coefficients and regions of validity with cell size in mechanical stress enabled CMOS technologies, considering cell layout parameters. We also derive relationships of our model coefficients with on-chip supply voltage and temperature variations. We use these relationships in reducing number of HSPICE simulations in ECSM re-characterization significantly. We show the results of $t_{TCP60\%}$ as a representative of t_{TCP} in Region I and $t_{TCP90\%}$ as a representative of t_{TCP} in Region II.

The chapter is organized as follows. In Section 6.2, we describe our simulation setup. In Section 4.3, we derive the relationship of $t_{TCP}s$ with T_R, C_l and NF for two different regions in stress enabled technologies. In Section 4.4, we derive the relationship $t_{TCP}s$ with T_R, C_l and temperature (T) for two different regions. In Section 4.5, we derive the relationship $t_{TCP}s$ with T_R, C_l and supply voltage (V_{dd}) for two different regions.

4.2 Simulation Setup

In this chapter, we use the same simulation setup for standard cell CMOS inverter, as we used in previous chapter. To account for voltage variability, we considered a $\pm 10\%$ variation in nominal supply voltage. For temperature variability, we vary the temperature range from $298K$ to $423K$. For stress variability, we consider the variation of device channel's mechanical stress as a function of the number of fingers (NF) of device sharing an active

area in an inverter's layout. This is because, often the size of a standard cell is increased by increasing NF, the consequent variation in channel mechanical stress is important. Stress induced in the channel of pMOS (nMOS) using various sources like Compressive/Tensile Etch Stop Liner (c/t-ESL), embedded SiGe (eSiGe) and Stress Memorization Technique (SMT) [52] have been considered in this work.

To emulate the effect of stress in our HSPICE simulations, we model the value of PTM parameters MULU0 (mobility multiplier) and DELVT0 (threshold voltage shift) as a function of the nMOS (pMOS) device's NF in layout. The values of MULU0 and DELVT0 are calculated as a function of average channel stress for a given value of NF, as we explained in [89, 90]. The procedure is as follows:

The device structure, doping profile and I-V characteristics of nMOS (pMOS) devices in our Sentaurus TCAD simulation setup are well calibrated to match those corresponding to an equivalent PTM device model. The value of average channel stress is obtained using TCAD process simulations corresponding to the device structure and NF. Values of MULU0 and DELVT0 for a given value of NF in HSPICE simulation are obtained from TCAD average channel stress value. In Section 4.3 (which address stress induced variability), we use HSPICE simulations at 45nm CMOS technology node. This is because we have calibrated our 45nm BSIM PTM model with our Sentaurus TCAD simulation setup, which considers mechanical stress. In Section 4.4 and 4.5, we verify our models for temperature and supply voltage variability with HSPICE simulations at 32nm CMOS technology node. In this way, we have also verified the validity of our t_{TCP} models at 45nm CMOS technology node.

4.3 TCP models considering stress variability for CMOS inverter standard cell

To circumvent scaling related issues such as large-off state leakage current and mobility degradation due to high channel doping, stress engineering techniques are being used in sub $90nm$ CMOS technologies [52, 91]. In standard cells, the size of a cell is typically increased by increasing the number of device fingers in a layout since the distance between power supply and ground line is fixed. However, this increase in NF (which represents cell size) leads to a change in average value of stress in the channel [89]. This leads to a stress induced performance variability of standard cells as their size is changed[1]. Therefore, the drive capability of strain-engineered (MFGSs) does not increase linearly with the NFs.

In this section, we derive the change in coefficients in (4.1) and (4.2) as a function of NF in a stress enabled 45nm CMOS technology. First, we recall these models equation derived in previous chapter as:.

$$t_{TCP_{60\%}} = K_1 C_l + K_2 T_R + K_3 \tag{4.1}$$

[1]For example, an inverter standard cell FO4 delay would change with its size, which is contrary to conventional expectation

$$t_{TCP90\%} = A_1 T_R + A_2 \sqrt{T_R} \tag{4.2}$$

Where,

$$K_1 = \frac{(V_{dd} - V_{dsat})}{I_{ON}} \tag{4.3}$$

$$K_2 = \left[0.8 - \frac{S_T}{I_{ON}}\right] \tag{4.4}$$

$$K_3 = C_p \left[\frac{(V_{dd} - V_{dsat})}{I_{ON}}\right] \tag{4.5}$$

$$A_1 = \left(\frac{V_{th}}{V_{dd}}\right) \tag{4.6}$$

$$A_2 = \left(\sqrt{\frac{0.2(C_l + C_p)}{\beta_s}}\right) \tag{4.7}$$

The values of I_{ON}, I_{lin} and V_{dsat} used in this model are derived from alpha power law model [9] such as:

$$I_{ON} = v_{sat}\, W_n\, P_s\, (V_{gs} - V_{th}) \tag{4.8}$$

$$I_{lin} = \mu\, \frac{W_n}{L_{eff}}\, P_l\, (V_{gs} - V_{th})\, V_{ds} \tag{4.9}$$

$$V_{dsat} = \frac{v_{sat}\, P_s\, (V_{gs} - V_{th})\, L_{eff}}{\mu\, P_l} \tag{4.10}$$

Where, P_s, P_l are technology dependent parameters, v_{sat} is the saturation velocity and μ is the mobility of the nMOS device. We also derive the change in region of validity in T_R, C_l space with inverter standard cell size (*i.e.*, NF). As discussed in Section 6.2, the effect of change in channel stress as a function of NF is captured by PTM parameters MULU0 and DELVT0 in our HSPICE simulations.

Now we derive coefficients of (4.1) and (4.2) as a function of NF. We use a set of empirical equations suggested in [89, 92, 93], to relate device level electrical parameters with stress. These equations are:

$$\mu\,(\sigma) = [P_1\, \sigma(NF) + 1]\, \mu_0 \tag{4.11}$$

$$v_{sat}\ (\sigma) = [P_1\, \sigma(NF) + 1]\, v_{sat} \tag{4.12}$$

$$I_{ON}\,(\sigma) = [P_1\, P_2\, \sigma\,(NF) + 1]\, I_{ON} \tag{4.13}$$

$$V_{th}\ (\sigma) = [V_{th} + P_3\, \sigma\,(NF)] \tag{4.14}$$

Where, $\mu\,(\sigma)$, $v_{sat}\,(\sigma)$, $I_{ON}\,(\sigma)$,$V_{th}\,(\sigma)$ are stress dependent mobility, saturation velocity, drive current and threshold voltage parameters respectively. Whereas, μ_0, v_{sat}, I_{ON}, V_{th} are unstressed parameters and P_1 is the piezoresistive coefficient. P_2 and P_3 are technology dependent coefficients extracted by fitting the above equations into HSPICE simulated I-V data, as discussed in [89, 90, 93]. Here $\sigma\,(NF)$ represents the average stress in the fingers in MFGSs. As discussed in [89], a relation between this average stress and NF is:

$$\frac{\sigma\,(NF)}{\sigma_{ref}\,(NF=1)} = M_1 + \frac{M_2}{NF + M_3} \tag{4.15}$$

In (5.25), M_1, M_2, M_3 are fitting parameters specific to given technology node where, M_1 denotes $\sigma\,(NF \to \infty)\,/\sigma\,(NF=1)$, while M_2 and M_3 control the rate of change of stress as NF is increased (discussed in detail in [89]).

4.3.1 Impact of stress induced variability in Region I

In this subsection, we derive the relationships of $t_{TCP}s$ with T_R, C_l and NF for Region I in stress enabled technologies. In (4.3)-(4.5), μ, v_{sat}, I_{ON}, and V_{th} are now given by (5.21)-(5.24). Again we take $t_{TCP_{60\%}}$ as a representative of all TCPs in Region I and derive a model for $t_{TCP_{60\%}}$ considering the impact of channel stress, as it lies in Region I. In this derivation, we assume that $(V_{dd} - V_{th})$ is independent of NF. This is justified since V_{th} is much smaller than V_{dd}. From (4.3)-(4.5) and (5.21)-(5.25), we obtain :

$$K_1 = (NF + M_3)\left(\frac{D_1}{NF + D_2}\right) \tag{4.16}$$

Where,

$$D_1 = \frac{(V_{dd} - V_{dsat})}{((M_1 P_1 P_2 + 1) I_{ON})} \tag{4.17}$$

$$D_2 = (M_3 + \frac{M_2 P_1 P_2}{(M_1 P_1 P_2 + 1)}) \tag{4.18}$$

We observe that (5.26) fits well with HSPICE simulated data as shown in Fig. 5.19. In these HSPICE simulations, MULU0 and DELVT0 vary with NF in accordance with (5.25) (*i.e.* the PTM model incorporates channel stress variability effects).

Likewise, solving (4.5), we found the relation between K_3 and NF as:

$$K_3 = C_p\, K_1 = C_p \left[(NF + M_3)\left(\frac{D_1}{NF + D_2}\right)\right] \tag{4.19}$$

Where, C_p is parasitic capacitance (due to gate-drain overlap, drain-bulk junction capacitance etc.) which is linearly related to NF in MFGS. This can be seen in the inset of Fig. 5.19. We observe that (5.29) fits well with HSPICE simulated data as shown in Fig.

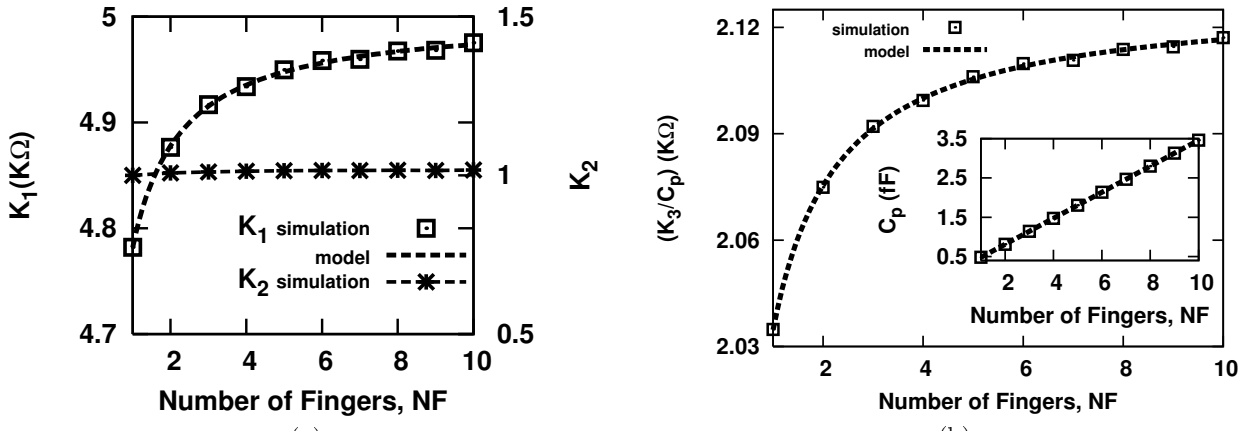


Figure 4.1: K_1, K_2 and K_3 as a function of NF (which also represents to channel stress).

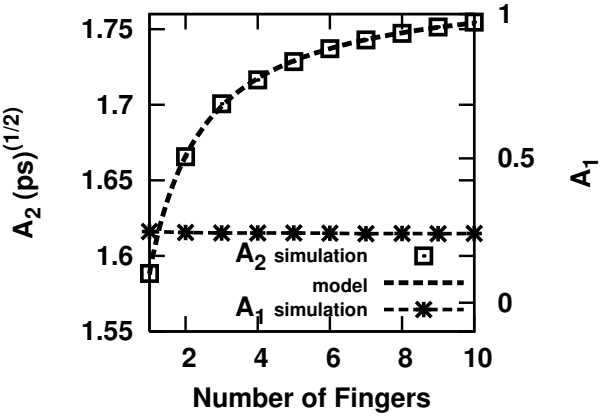


Figure 4.2: A_1 and A_2 as a function of NF.

5.19. Thereafter, we observe that K_2 is independent of NF (because S_T and I_{ON} are both proportional to NF) as shown in Fig. 5.19. In Fig. 5.19, K_2 is normalized with value of K_2 at $NF = 1$. In this subsection, we observe that our model in Region I is valid with respect to stress variability.

4.3.2 Impact of stress induced variability in Region II

In this subsection, we derive the relationships of $t_{TCP}s$ with T_R, C_l and NF for Region II in stress enabled technologies. In (4.6) and (4.7), v_{sat}, V_{th} are now given by (5.22) and (5.24). Again we take $t_{TCP_{90\%}}$ as a representative of all TCPs in Region II and derive a model for $t_{TCP_{90\%}}$ considering the impact of channel stress, as it lies in Region II. From (4.6), (4.7) and (5.22), (5.24), (5.25), we obtain that A_1 is independent of $V_{th}(\sigma)$. We have verified through simulations that variation in $V_{th}(\sigma)$ with NF is very small (in our TCAD calibrated simulations variation in $V_{th}(\sigma)$ with NF is less than 3%). This small variation in $V_{th}(\sigma)$ has also been reported in [94]. Therefore, A_1 is independent of $\sigma(NF)$ as shown in Fig. 4.2.

We find the relationship between A_2 and NF as:

$$A_2 = \sqrt{\left(\frac{S_1(NF + M_3)}{(NF + S_2)}\right)} \tag{4.20}$$

Where,

$$S_1 = \left(\frac{0.2(C_l + C_p)}{W_n P_s \, v_{sat} \, (M_1 P_1 + 1)}\right) \tag{4.21}$$

$$S_2 = \left(\frac{1}{M_3 + \frac{P_1 M_2}{(M_1 P_1 + 1)}}\right) \tag{4.22}$$

We observe that (5.31) fits well on our stress aware HSPICE simulation data as shown in Fig. 4.2. In this subsection, we observe that our model in Region II is valid with respect to stress variability.

4.3.3 Efficient stress aware ECSM characterization

In this subsection, we show that using our t_{TCP} models, ECSM characterization of inverter standard cells would need significantly lesser number of HSPICE simulations. We generated 7×7 LUTs having $t_{TCP60\%}$ values with varying C_l and T_R values for different inverter cell sizes (represented by NF). Conventionally, this would require 343 HSPICE simulations for cell sizes corresponding to NF=1 to 7. However, using our models only 88 HSPICE simulations (including simulations to extract the model coefficients) were needed to generate the same LUT.

For computing the remaining values of $t_{TCP60\%}$, we used our model. In this way, we saved 255 HSPICE simulations to generate the 7×7 LUTs for NF=1 to 7. Therefore, proposed model can be used to save $\backsimeq 74\%$ HSPICE simulations for all $TCPs$ in Region I. Likewise, to generate the 7×7 LUTs for NF=1 to 7 in Region II, the proposed model can be used to save $\backsimeq 45\%$ HSPICE simulations for all $TCPs$ in Region II. We then compare the LUTs for $t_{TCP}s$ generated using our above approach with conventional fully HSPICE generated ECSM LUTs. We observe that the values of our LUT's $t_{TCP}s$ are different from the conventional LUT by a maximum of 2.5%.

4.4 TCP models considering temperature variability for CMOS inverter standard cell

In present day CMOS technologies one of the major sources of circuit performance variability is on-chip temperature variation. It can cause significant change in electron and hole mobility as well as in threshold voltage [95]. An increase in temperature results in lowering the mobility thereby reducing ON current which in turn increases the delay values. Therefore, a re-characterization of standard cells for several values of temperature becomes necessary that requires a huge computational effort and time.

In this section, we derive the change in coefficients in (4.1) and (4.2) as a function

of temperature at 32nm CMOS technology. In this work, we take a realistic range of temperature variation due to on-chip heating from $298K$ (room temperature) to $423K$. In this work, we use an empirical expression suggested in [23, 29], to consider the impact of temperature variability on carrier mobilities. The expression is given as:

$$\mu\left(T\right) = \mu(T_0)\left(\frac{T}{T_0}\right)^{-\theta} \tag{4.23}$$

Where, $\mu(T)$ is the temperature dependence of mobility, T is the temperature, T_0 is the nominal temperature *i.e.* $298\,K$ and θ is technology dependent temperature coefficient. For our PTM CMOS technology, we extract the value of $\theta = 2.3$ by maximum transconductance (g_m) method. We use the empirical relation (given in [29]) between carrier saturation velocity and temperature as:

$$v_{sat}\ \left(T\right) = v_{sat}\left(T_0\right) - \eta\left(T - T_0\right) \tag{4.24}$$

Where, $v_{sat}\left(T\right)$ is the temperature dependence of saturation velocity and η is the temperature coefficient. The threshold voltage of devices also gets affected by an increase in temperature due to change in fermi level location and band gap energy [96]. In [96], the temperature dependence of threshold voltage is given by:

$$V_{th}\ \left(T\right) = V_{th}\left(T_0\right) - \kappa\left(T - T_0\right) \tag{4.25}$$

Where, κ is the temperature dependence coefficient of threshold voltage. The value of η and κ can be extracted from simulated HSPICE I-V data for a given CMOS technology. We verified the validity of (5.34), (5.35) and (5.36) using Sentaurus TCAD device simulations. In TCAD simulations, we use 25nm drawn gate length nMOS and pMOS devices. We now discuss our approach for deriving temperature variation aware t_{TCP} models.

4.4.1 Impact of temperature variability in Region I

In this subsection, we derive the relationships of t_{TCPs} with T_R, C_l and temperature T for Region I of CMOS inverter standard cell. In (4.3)-(4.5), v_{sat}, μ and V_{th} are now given by (5.34), (5.35) and (5.36). Again we take $t_{TCP_{60\%}}$ as a representative of all TCPs in Region I and derive a model for $t_{TCP_{60\%}}$ considering the impact of temperature variability, as it lies in Region I. In this derivation, we assume that $(V_{dd} - V_{th})$ is independent of T. This is justified since V_{th} is much smaller than V_{dd}. From (4.3)-(4.5) and (5.34), (5.35), we obtain :

$$K_1 = \left(R_1\,T^{2.3} + R_2\,T + R_3\right) \tag{4.26}$$

Where,

$$R_1 = \left[\frac{\left(\frac{L_{eff}}{W_n P_l}\right)}{\left(\mu(T_0)\left(\frac{1}{T_0}\right)^{-2.3}\right)} \right] \tag{4.27}$$

$$R_2 = \left[\frac{V_{dd}}{W_n P_s (V_{gs} - V_{th})} \frac{\eta}{v_{sat}(T_0)^2} \right] \tag{4.28}$$

$$R_3 = \left[\frac{V_{dd}}{W_n P_s (V_{gs} - V_{th})} \frac{(1 - \eta T_0)}{v_{sat}(T_0)^2} \right] \tag{4.29}$$

We observe that (5.37) fits well on HSPICE simulated data as shown in Fig. 4.3. Likewise, solving (4.5), we find the relation between K_3 and T as:

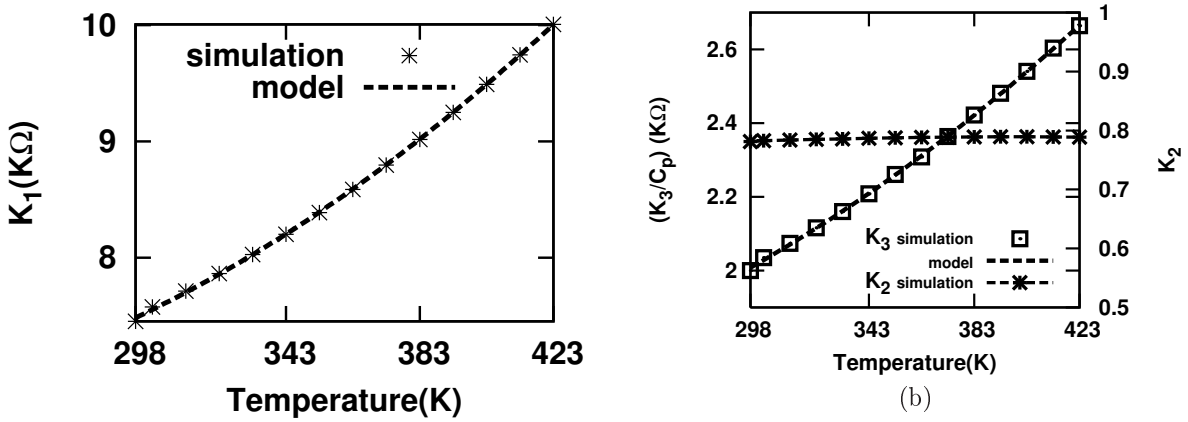


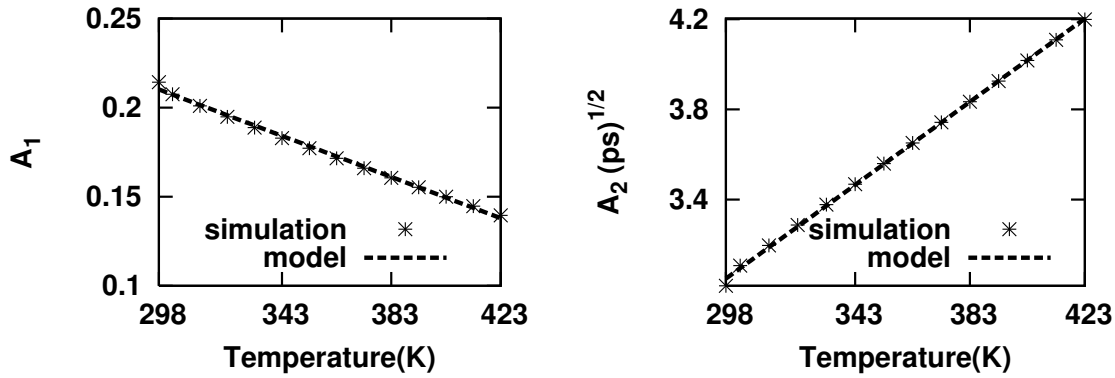
Figure 4.3: Variation of K_1, K_2 and K_3 with temperature.

$$K_3 = C_p K_1 = C_p \left(R_1 T^{2.3} + R_2 T + R_3\right) \tag{4.30}$$

Where, C_p is parasitic capacitance. We observe using HSPICE simulations that change in C_p with T varying from 298K to 423K is negligibly small. We observe that (5.41) fits well with HSPICE simulated data as shown in Fig. 4.3. Thereafter, solving (4.4), we observe that K_2 is independent of T as shown in Fig. 4.3 (because S_T and I_{ON} both are the function of T). In this subsection, we observe that our model in Region I is valid with respect to temperature variability.

4.4.2 Impact of temperature variability in Region II

In this subsection, we derive the relationships of $t_{TCP}s$ with T_R, C_l and T for Region II of CMOS inverter standard cell. In (4.6), (4.7) v_{sat} and V_{th} are now given by (5.35) and (5.36). Again we take $t_{TCP_{90\%}}$ as a representative of all TCPs in Region II and derive a model for $t_{TCP_{90\%}}$ considering the impact of temperature variability as it lies in Region II.



(a) (b)

Figure 4.4: Variation of A_1 and A_2 with temperature.

From (4.6), we observe that A_1 is dependent on V_{th} only. From (5.36), we expect A_1 to reduce linearly with T, which we verify through HSPICE simulations in Fig. 4.4.

We obtain the relationship between A_2 and T as:

$$A_2 = \left[\left(\sqrt{\frac{0.2(C_l + C_p)}{W_n P_s \, v_{sat}(T_0)}}\right)\left(1 + \frac{\eta \, (T - T_0)}{2 \, v_{sat}(T_0)}\right)\right] \tag{4.31}$$

We observe that (5.42) fits well on our HSPICE simulation data with temperature variability as shown in Fig. 4.4. In this subsection, we observe that our model in Region II is valid with respect to temperature variability.

4.4.3 Efficient temperature variation aware ECSM characterization

In this subsection, we show that using our t_{TCP} models, ECSM characterization of inverter standard cells would need significantly lesser number of HSPICE simulations. We generated 7×7 LUTs having $t_{TCP60\%}$ values with varying C_l and T_R values for different temperature values . Conventionally, this would require 343 HSPICE simulations for 7 different values of temperature T from $298K$ to $423K$. However, using our models only 77 HSPICE simulations (including simulations to extract the model coefficients) were required to generate the same LUT. For computing the remaining values of $t_{TCP60\%}$, we use our models. In this way, we saved 266 HSPICE simulations to generate the 7×7 LUTs for $T = 298K$ to $423K$. Therefore, proposed model can be used to save $\backsimeq 78\%$ HSPICE simulations for all TCPs in Region I.

Likewise, to generate the 7×7 LUT for 7 different values of temperature T from $298K$ to $423K$ in Region II, the proposed model can be used to save $\backsimeq 50\%$ HSPICE simulations for all TCPs in Region II. We then compare the LUTs for $t_{TCP}s$ generated using our above approach with conventional fully HSPICE generated ECSM LUTs. We observe that the values of our LUT's $t_{TCP}s$ are different from the conventional LUT by a maximum of 2.5%.

4.5 TCP models considering supply voltage variability for CMOS inverter standard cell

At nanometer range technologies, power supply noise dominates due to a significant increase in the ratio of peak noise voltage to the ideal supply voltage [97]. This effect becomes more pronounced with technology scaling. Problems with power supply voltage level drop in the on-chip power distribution network also becomes pronounced at these technology nodes [98]. As a result, voltage fluctuation of ±10% from the nominal power supply levels is considered acceptable [98]. This results in re-characterization of standard cell libraries at several values of power supply voltages.

In this section, we derive a power supply voltage variation aware t_{TCP} models which we use to reduce the re-characterization effort significantly. We consider the ±10% change in power supply voltage (V_{dd}) from the its nominal value of $V_{dd} = 0.9V$.

4.5.1 Impact of supply voltage variability in Region I

In this subsection, we derive the relationships of $t_{TCP}s$ with T_R, C_l and V_{dd} for Region I of CMOS inverter standard cell. Again we take $t_{TCP_{60\%}}$ as a representative of all TCPs in Region I and derive a model for $t_{TCP_{60\%}}$ considering the impact of supply voltage variability, as it lies in Region I. From (4.3), we obtain the expression for K_1 as :

$$K_1 = \frac{V_{dd}}{v_{sat}\, W_n\, P_s\, (V_{dd} - V_{th})} - \frac{L_{eff}}{W_n\, P_l\, \mu} \tag{4.32}$$

We observe that (4.32) fits well with HSPICE simulated data as shown in Fig. 5.23. We now discuss the variation of K_3 with V_{dd}. In (4.5), C_p includes the voltage dependent junction capacitance which is given by [94] as:

$$C_j(V) = \frac{A.C_{j0}}{\sqrt{\left(1 + \frac{V}{\phi_0}\right)}} \tag{4.33}$$

Where, A indicates the junction area, C_{j0} is zero-bias junction capacitance per unit area, V is the reverse bias voltage and ϕ_0 is built-in potential. Using (4.5) and 4.33, we obtain the expression for K_3 as :

$$K_3 = \left(\frac{A.C_{j0}}{\sqrt{\left(1 + \frac{V}{\phi_0}\right)}}\right)\left[\frac{V_{dd}}{v_{sat}\, W_n\, P_s\, (V_{dd} - V_{th})} - \frac{L_{eff}}{W_n\, P_l\, \mu}\right] \tag{4.34}$$

We observe that (4.34) fits well with HSPICE simulated data as shown in Fig. 5.23. Using (4.4) and (4.8), the expression for K_2 will get reduced to :

$$K_2 = \frac{(a_7 * V_{dd} + a_8)}{(V_{dd} + a_9)} \tag{4.35}$$

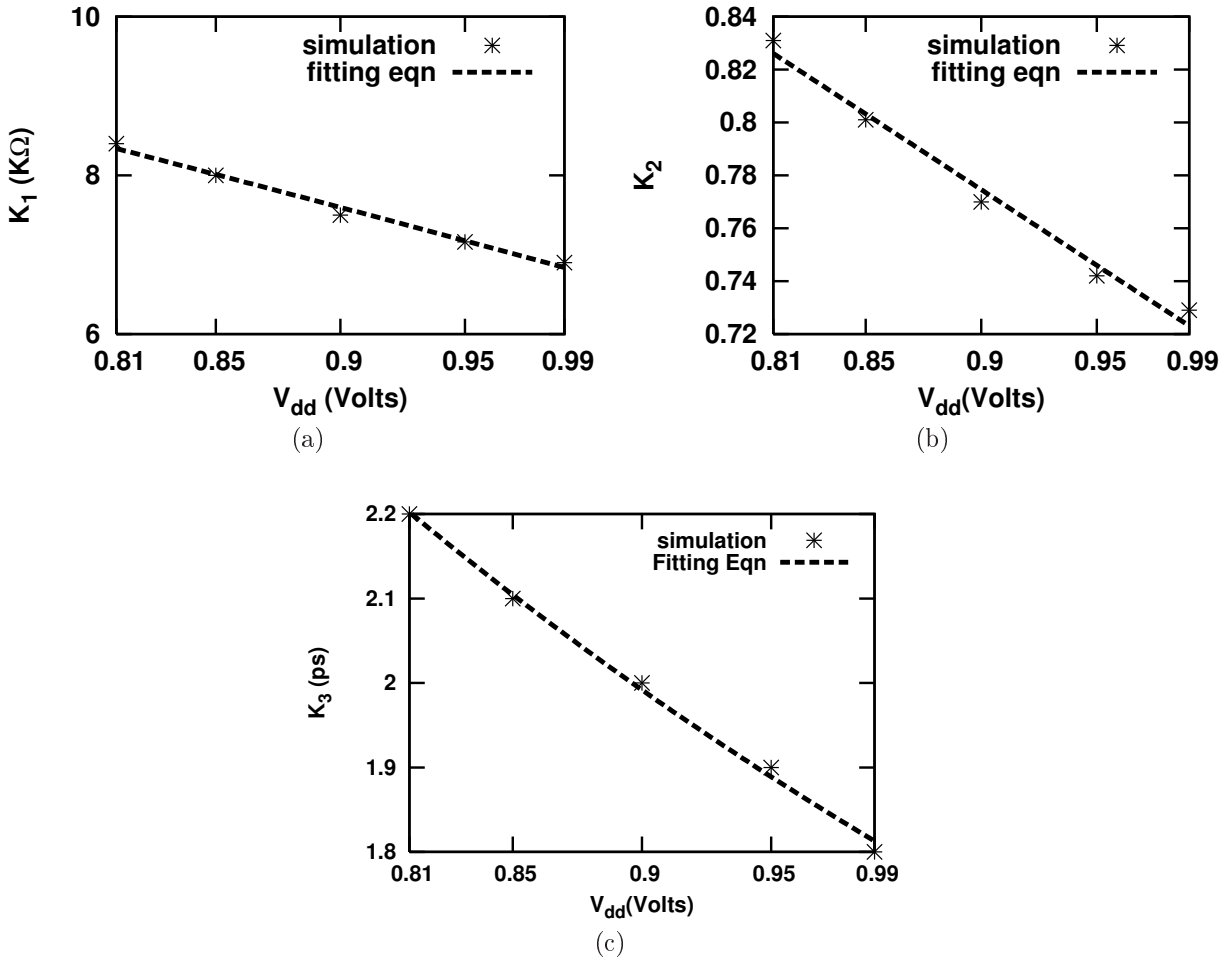


Figure 4.5: Variation of K_1, K_2 and K_3 with supply voltage.

Where, $a_7 = 0.3$, $a_8 = 0.2V_{th}$ and $a_9 = -V_{th}$. We observe that (4.35) fits well with HSPICE simulated data as shown in Fig. 5.23. In this subsection, we observe that our model in Region I is valid with respect to supply voltage variability.

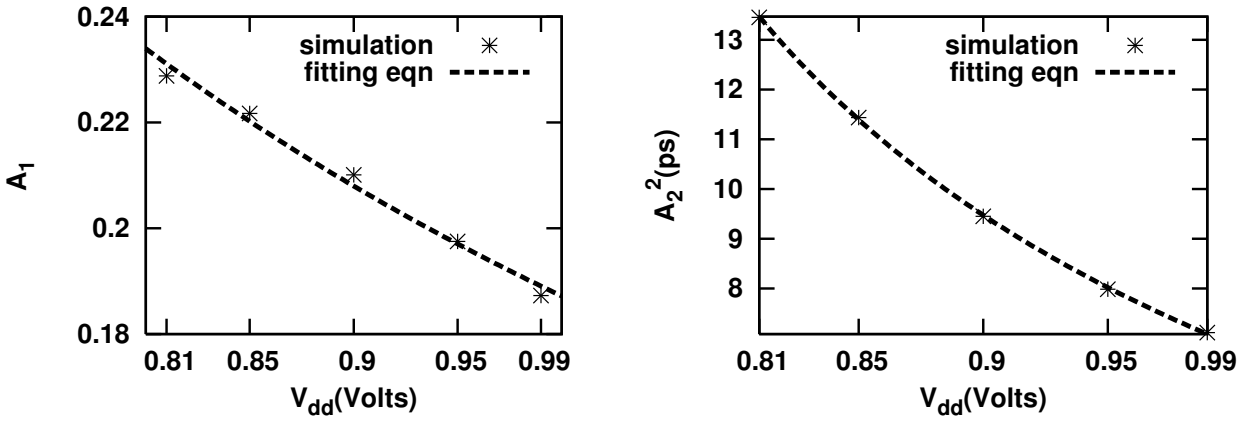
4.5.2 Impact of supply voltage variability in Region II

In this subsection, we derive the relationships of $t_{TCP}s$ with T_R, C_l and V_{dd} for Region II of CMOS inverter standard cell. Again we take $t_{TCP_{90\%}}$ as a representative of all TCPs in Region II and derive a model for $t_{TCP_{90\%}}$ considering the impact of supply voltage variability, as it lies in Region II. In (4.6), A_1 is inversely proportional to supply voltage which is also verified through HSPICE simulations as shown in Fig. 4.6.

In (4.7), we can see that only C_p is supply voltage dependent term which decreases with supply voltage. Therefore, expression for A_2 will get reduced to:

$$A_2^2 = b_1 + \frac{b_2}{\sqrt{(b_3 + V_{dd})}} \tag{4.36}$$

Where,



(a) (b)

Figure 4.6: Variation of A_1 and A_2 with supply voltage.

$$b_1 = \frac{0.2C_l}{W_n P_s \nu_{sat}} \tag{4.37}$$

$$b_2 = \frac{0.2}{W_n P_s \nu_{sat}} \left(A.C_{j0} \sqrt{\phi_0} \right) \tag{4.38}$$

and

$$b_3 = \phi_0 \tag{4.39}$$

We observe that (4.36) fits well with HSPICE simulated data as shown in Fig. 4.6. In this subsection, we observe that our model in Region II is valid with respect to voltage variability.

4.5.3 Efficient supply voltage variation aware ECSM characterization

In this subsection, we show that using our t_{TCP} models, ECSM characterization of inverter standard cells would need significantly lesser number of HSPICE simulations. We generated 7×7 LUTs having $t_{TCP60\%}$ values with varying C_l and T_R values for different supply voltages. Conventionally, this would require 343 HSPICE simulations for 7 different values of supply voltage V_{dd} from $0.81V$ to $0.99V$. However, using our models only 84 HSPICE simulations (including simulations to extract the model coefficients) were required to generate the same LUT. For computing the remaining values of $t_{TCP60\%}$, we used our model. In this way, we saved 259 HSPICE simulations to generate the 7×7 LUTs for $V_{dd} = 0.81V$ to $0.99V$. Therefore, the proposed model can be used to save $\backsimeq 76\%$ HSPICE simulations for all TCPs in Region I.

Likewise, to generate the 7×7 LUTs for 7 different values of supply voltage V_{dd} from $0.81V$ to $0.99V$ in Region II, the proposed model can be used to save $\backsimeq 50\%$ HSPICE simulations for all TCPs in Region II. We then compare the LUTs for $t_{TCP}s$ generated using our above

approach with conventional fully HSPICE generated ECSM LUTs. We observe that the values of our LUT's $t_{TCP}s$ are different from the conventional LUT by a maximum of 2.5%.

4.6 Summary

In this chapter, we derived the relationship of model coefficients with voltage, temperature and stress variability. We derived the relationships of model coefficients and their regions of validity with cell size in mechanical stress enabled CMOS technologies, considering cell layout parameters. We also derived the relationships of model coefficients with on-chip supply voltage and temperature variations. We later used these relationships to reduce the number of HSPICE simulations by nearly half for standard cell CMOS inverter in ECSM re-characterization significantly. We observed that the proposed models are in good agreement with HSPICE simulations with a maximum error of 2.5%.

Chapter 5

Timing model for 2-input NAND gate standard cell considering PVT variation

5.1 Overview

This chapter focuses on modeling of $t_{TCP}s$ as a function of T_R and C_l for a 2-input CMOS NAND gate. Further, the region of validity of the model in T_R, C_l space is derived. We then derive the relationships of the model coefficients with the NAND gate size, V_{dd} and temperature. We also consider layout dependent effects due to mechanical stress in deriving these relationships.

The chapter is organized as follows. In section 6.2, we describe our simulation setup. In Section 5.3, based on the switching of the input of series stack of 2-input NAND gate, we categorize the timing models into two cases. In this chapter, Case 1 corresponds to the switching of the upper transistor of the series stack, whereas, Case 2 corresponds to the switching of the lower transistor of the series stack. In Section 5.4 and 5.6, we derive the t_{TCP} models and their region of validity as a function of T_R and C_l for Case 1 and 2, respectively. Later, we use these models to reduce the number of HSPICE simulations for ECSM characterization. In Section 5.5 and 5.7, we derive the relationships of the model coefficients with mechanical stress, temperature and V_{dd} variability for Case 1 and 2, respectively.

5.2 Simulation Setup

In this work, we use $32nm$ Predictive Technology device Model (PTM)[1] for HSPICE simulations. The widths of nMOS and pMOS devices are chosen to obtain equal output rising and falling transition using the procedure discussed in [22]($\frac{W_p}{W_n} = \frac{120nm}{96nm}$ [2]). The channel lengths are kept at their minimum allowed value. Since width ratios of all transistors for standard cell of a given type (say, 2-input NAND gate) is fixed, the value of W_n also represents the cell size. Throughout this work, we consider the case of rising input transition as

[1] Obtained from http://ptm.asu.edu/

[2] W_p: pMOS device width; W_n: nMOS device width

shown in Fig. 5.1; The case of falling input transition can be handled in similar manner. In Fig. 5.1, C_{out} represents the sum of the parasitic capacitance between node 'Out' and 'gnd' and the external load capacitance (C_l) appearing at node 'out'. The symbol C'_X represents the parasitic capacitance between node 'X' and 'gnd'.

To emulate the effect of stress in our HSPICE simulations, we model the value of PTM parameters MULU0 (mobility multiplier) and DELVT0 (threshold voltage shift) as a function of the nMOS (pMOS) device's NF in layout. The values of MULU0 and DELVT0 are calculated as a function of average channel stress for a given value of NF, as we explained in [89, 90]. To account for voltage variability, we considered a $\pm 10\%$ variation in nominal supply voltage. For temperature variability, we vary the temperature range from $298K$ to $423K$. For stress variability, we consider the variation of device channel mechanical stress as a function of NF in inverter layout (discussed in detail in the simulation setup of Chapter 4).

5.3 Timing model for 2-input NAND gate standard cell

This chapter focuses on modeling of timing values of $TCPs$ as a function of T_R and C_l for 2-input CMOS NAND gate. The NAND gate ECSM characterization done for the following two cases:

- Case 1: When upper nMOS transistor in series-stack switches
- Case 2: When lower nMOS transistor in series-stack switches

In this work, we derive physics based models for TCPs as a function of T_R and C_l. We consider the intermediate node voltage transition of the series stack of the 2-input NAND gate. The region of validity of the model in both cases (Case 1 and 2) is derived. The relationship between cell size and model coefficient is also derived. Further, we derive the relationships of model coefficients with mechanical induced stress as a function of the NF, temperature and supply voltage variability, for both the cases.

5.4 Case 1: Derivation and Validation of t_{TCP} model

In this section, we derive the t_{TCP} model for the switching of the upper nMOS transistor in the series stack of 2-input NAND gate. We derive the relationship of $t_{TCP}s$ with the T_R, C_l and cell size. We also derive the region of validity of these models in T_R, C_l plane. The model remains valid for the case when $V_{out}(T_R) \geq V_{dsat}$, where V_{dsat} is the saturation drain-source voltage (V_{ds}) at which velocity saturation occurs for the NAND gate's upper nMOS device. In this derivation, we assume the t_{TCP} models for 2-input NAND gate remains same as for CMOS inverter (discussed in Chapter 3), because the lower nMOS transistor

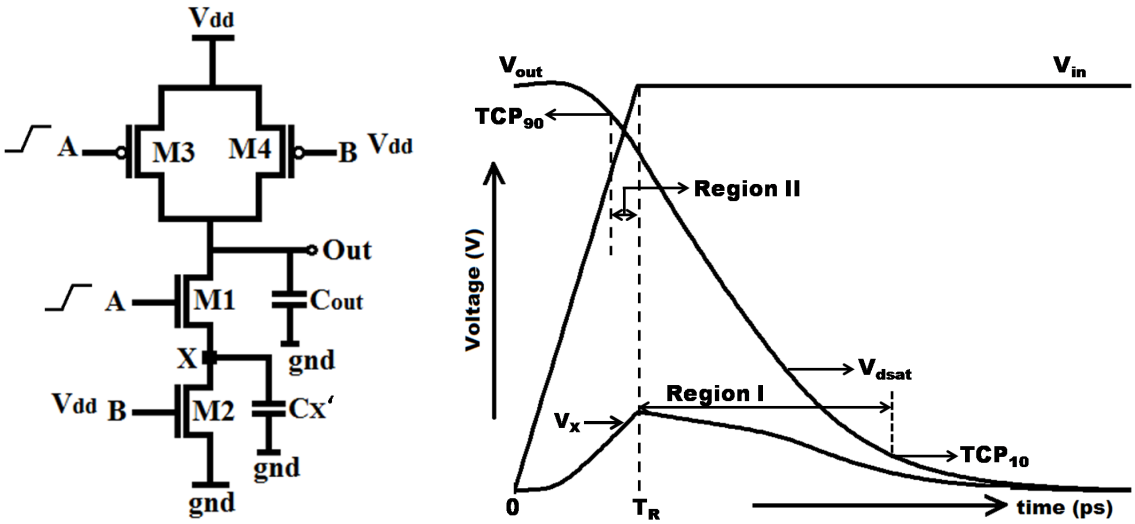


Figure 5.1: Case 1: (a) 2-input NAND gate schematic (b) its I/O waveform.

(M_2) in the series stack adds some resistance to the source of upper nMOS transistor (M_1). In this work, we ignore the negligibly small current flowing through pMOS transistor (max. current$\approx 3\mu A$). We classify the TCPs into two regions:

- Region I: When $V_{in} = V_{dd}$ (for $t_{TCP}s > T_R$)
- Region II: When $V_{in} < V_{dd}$ (for $t_{TCP}s < T_R$)

In this section, we derive the timing model of $TCP_{10\%}$ in Region I and timing model of $TCP_{90\%}$ in Region II (as shown in Fig. 5.1b).

5.4.1 Derivation of the model in Region I

In this subsection, we derive the model for the NAND gate shown in Fig. 5.1a. Region I contains all the $TCPs$ having values V_{TCP} smaller than $V_{out}(T_R)$, as we depict in Fig. 5.1b. In this derivation, we assume that the upper nMOS (discharging) device operates either in saturation or in linear region and lower nMOS (discharging) device always operates in linear region when $V_{TCP} < V_{out}(T_R)$.

To find the $t_{TCP10\%}$ model, we first integrate the saturation current through M_1 during the input transition $V_{in-A}(t) = V_{dd}(\frac{t}{T_R})$ for $0 \leq t \leq T_R$. Then integrate the discharging current $I_{ON}\,(V_{in-A} = Vdd)$ through M_1, from $t = T_R$ to $t = t_{TCP}$. We equate the sum of these integrals to $(C_l + C_p)\,(V_{out}(T_R) - V_{TCP})$ to obtain t_{TCP} . Here C_l is the load capacitance, C_p is the parasitic capacitance between node 'Out' and 'gnd'. If the $V_{TCP} < V_{dsat}$, we integrate I_{ON} through M_1 from $t = T_R$ to $t = t_{sat}$ (where t_{sat} represents the time when $V_{out} = V_{dsat}$). We equate the sum of these integrals to$(C_l + C_p)\,(V_{out}(T_R) - V_{dsat})$ to obtain t_{sat}. From $t = t_{sat}$ to $t = t_{TCP}$, both the series stacked nMOS devices operate

in linear region. These devices act as a resistance where the value of the resistance R_1 and R_2 can be obtained by equating $V_{gs1} = V_{gs2} = V_{dd}$. Therefore, the time duration from t_{sat} to t_{TCP} is proportional to time constant $R_{eq}(C_l + C_p)$. Where, R_{eq} represents an equivalent resistance of series stacked nMOS devices. As we discuss later in this section, for $V_{TCP} < V_{dsat}$ the value of V_X is so low that the change in V_X during this RC discharge need not to be considered.

To determine the integral of I_{M1}, we need V_X, which is the value of the voltage at node 'X'. This is the source voltage of M_1 and having a non-zero value. To determine V_X, we apply KCL at node 'X' which gives an expression as follows:

$$I_{M2} = I_{M1} + I_{C_{C1}} + I_{C'_X} \tag{5.1}$$

Where, I_{C_X} is the current flowing out of the parasitic capacitance (of M_1 and M_2) between node 'X' and 'gnd'. $I_{C_{C1}}$ is the current flowing through gate-to-source coupling capacitance of M_1, it consists of the gate-to-source overlap capacitance and a part of the gate-to-channel capacitance. I_{M1} is saturation current flowing through M_1 and I_{M2} is linear current flowing through M_2, which is given by alpha power law model as follows [9]:

$$I_{M1} = I_{sat} = \nu_{sat} W_n P_s (V_{gs1} - V_{th1})^{\alpha} \tag{5.2}$$

$$I_{M2} = I_{lin} = \mu_n \frac{W_n}{L_{eff}} P_l (V_{gs2} - V_{th})^m V_{ds} \tag{5.3}$$

Where, ν_{sat} is saturation velocity, W_n is width of nMOS device, μ_n is the mobility of nMOS device and P_s, P_l are the technology dependent parameters. In our case, we have used m and α values to be 1 and verified this through HSPICE simulations. In this chapter, we use $\beta_s = \nu_{sat} W_n P_s$ and $\beta_l = \mu_n \frac{W_n}{L_{eff}} P_l$. Now, we rewrite (5.1) as:

$$I_{M2} = I_{M1} - (C'_X + C_{C1}) \frac{dV_X}{dt} + C_{C1} \frac{dV_{in-A}}{dt} \tag{5.4}$$

Where,

$V_{in-A} = V_{dd} \left(\frac{t}{T_R} \right)$ for $0 \leq t \leq T_R$

In this derivation, to simplify the expression, we are taking $(C'_X + C_{C1}) = C_X$. Now, we write (5.4) as:

$$\beta_l (V_{dd} - V_{th}) V_X(t) = \beta_s (V_{in-A} - V_X(t) - V_{th1}) - C_X \frac{dV_X}{dt} + C_{C1} \frac{V_{dd}}{T_R} \tag{5.5}$$

This equation holds true only if $V_X(t)$ is a linear function of time. As M_1 is in saturation region, therefore I_{M1} is a function of V_{gs1} only (from (5.2)). From (5.2) and (5.5), we observe that V_X follows the change in $V_{in-A} = V_{dd} \left(\frac{t}{T_R} \right)$. Therefore, $C_X \frac{dV_X}{dt}$ would be a constant, we represent this constant term as I_X which is proportional to W_n.

After solving (5.5), we find the expression for V_X as:

$$V_X = \frac{(\beta_s \frac{V_{dd}}{T_R})\, t}{[\beta_l(V_{dd} - V_{th}) + \beta_s]} + \frac{(C_{C1}\frac{V_{dd}}{T_R} - I_X - \beta_s V_{th1})}{[\beta_l(V_{dd} - V_{th}) + \beta_s]} \tag{5.6}$$

As we discussed earlier, to find the $t_{TCP10\%}$ we first integrate the saturation current through M_1 during the input transition $V_{in-A}(t) = V_{dd}(\frac{t}{T_R})$ for $0 \leq t \leq T_R$. Then, we integrate the discharging current I_{ON} $(V_{in-A} = Vdd)$ through M_1, from $t = T_R$ to $t = t_{TCP}$. We equate the sum of these integrals to $(C_l + C_p)\,(V_{out}(T_R) - V_{TCP})$ to obtain t_{TCP}. From $t = T_R$ to t_{sat}, the lower nMOS transistor (M_2) operates in linear region, therefore we consider M_2 as a linear resistor (represented as R_2) which gives the value of intermediate node voltage $V_X = I_{M1}R_2$. This R_2 results in the the lowering of saturation current of M_1 by a constant value that is technology independent (since, $I_{M1} = \frac{I_{ON}}{(1+\beta_s R_2)}$). Therefore, in our derivation we represent the same current as I_{ON} from $t = T_R$ to t_{sat}. If $V_{TCP} < V_{dsat}$, then we find $t_{TCP10\%}$ such as $t_{TCP10\%} = t_{sat} + \triangle t$. Where $\triangle t$ represents the time during which both nMOS transistors operate in linear region. Therefore, we add a time constant term '$R_{eq}(C_l + C_p)$' to resultant term obtained till t_{sat}. It gives the expression for $t_{TCP10\%}$ as:

$$t_{TCP10\%} = K_1 C_l + K_2 T_R + K_3 \tag{5.7}$$

$$\text{Where, } K_1 = \frac{(V_{dd} - V_{dsat})}{I_{ON}} + R_{eq} \tag{5.8}$$

$$K_2 = -\frac{a_1}{I_{ON}} \tag{5.9}$$

$$K_3 = \left(\frac{(V_{dd} - V_{dsat})}{I_{ON}} + R_{eq}\right) C_p + \frac{a_2}{I_{ON}} \tag{5.10}$$

$$a_1 = \beta_s \frac{Vdd}{2}\left(1 - \frac{\beta_s}{[\beta_l(V_{dd} - V_{th}) + \beta_s]}\right) + \beta_s\left(\frac{I_X + \beta_s V_{th1}}{[\beta_l(V_{dd} - V_{th}) + \beta_s]} - V_{th1}\right) \tag{5.11}$$

$$a_2 = \frac{C_{C1} V_{dd} \beta_s}{[\beta_l(V_{dd} - V_{th}) + \beta_s]} \tag{5.12}$$

Where, a_1, a_2 are constants proportional to W_n. For all the $t_{TCP}s$ which fall under Region I, (5.7) remain valid. The model of (5.7) in Region I has thus been verified using HSPICE simulated data.

As a representative, we show the simulation results of $t_{TCP_{60\%}}$. This is because $t_{TCP_{60\%}} > T_R$ in the whole T_R, C_l space used in characterization LUTs. For $t_{TCP60\%}$, the form of model remains same. The coefficient values also remains same, the only difference is that the term R_{eq} will not be present. For $t_{TCP60\%}$, model coefficients K_1 and K_3 will be:

$$K_1 = \frac{(V_{dd} - V_{dsat})}{I_{ON}} \tag{5.13}$$

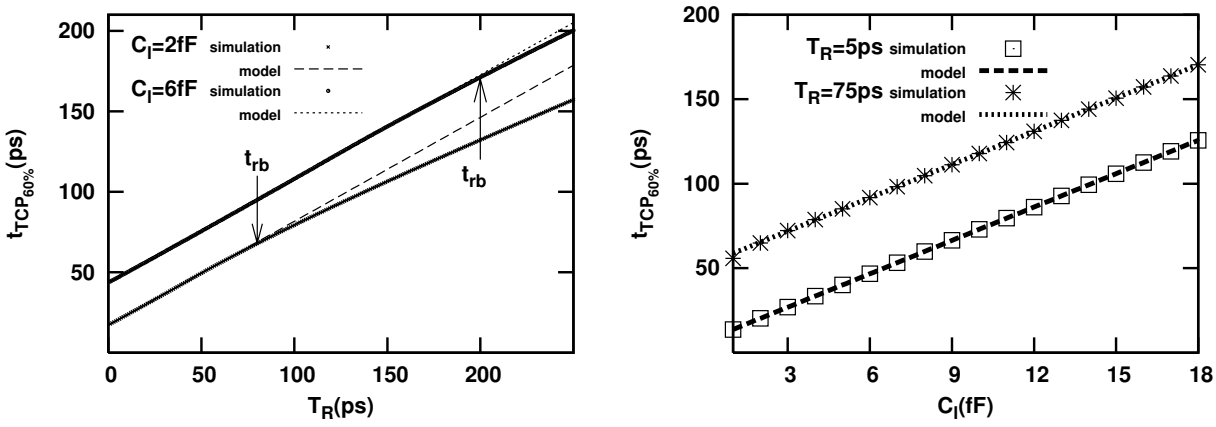


Figure 5.2: Case 1: Variation of $t_{TCP60\%}$ with respect to T_R and C_l values.

$$K_3 = \left(\frac{V_{dd} - V_{dsat}}{I_{ON}}\right) C_p + \frac{a_2}{I_{ON}} \tag{5.14}$$

So, t_{sat} would also have same model as a $t_{TCP60\%}$. Using curve fitting (as shown in Fig. 5.2) on the simulated values of $t_{TCP60\%}$, we extracted the coefficients K_1, K_2, K_3 of (5.7). From (5.7), we observe:

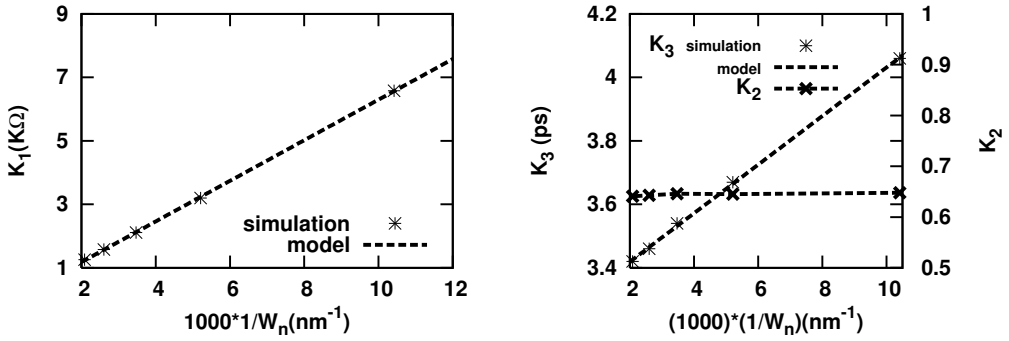
- Observation 1: K_1 is a linear function of $1/W_n$
- Observation 2: K_2 is independent of W_n
- Observation 3: K_3 is a linear function of $1/W_n$

As explained earlier, (5.7) is valid if $V_{out}(T_R) \geq V_{dsat}$, this imposes the following constraint on the region of validity:

$$\triangle Q(T_R) = (a_1 T_R - a_2) \leq (C_l + C_p)(V_{dd} - V_{dsat}) \tag{5.15}$$

Where, $\triangle Q(T_R)$ is the output discharge from 0 to T_R and V_{dd} is the power supply voltage. C_p is NAND gate's parasitic capacitance (due to gate-drain over-lap, drain-bulk junction capacitance etc.) between node 'Out' and 'gnd', and it is linearly dependent on W_n. The maximum value of T_R which satisfies (5.15), will be represented as t_{rb}.

Further, we verify Observation $1-3$ with HSPICE 32nm PTM. We observe that (5.7) fits well on HSPICE simulation data as shown in Fig. 5.2. We verify the coefficients (K_1, K_2 and K_3) behavior with W_n (shown in Fig. 5.3a and 5.3b). In this sub-subsection, we observe that our approach in Region I remains valid with variation in cell size.



(a) (b)

Figure 5.3: Case 1: Variation of K_1, K_2 and K_3 with cell size (W_n).

5.4.2 Derivation of the model in Region II

In this subsection, we derive the relationship of $t_{TCP}s$ with T_R, C_l and cell size for Region II. Region II contains all the $TCPs$ having value of $V_{TCP} > V_{out}(T_R)$, as we depict in Fig. 5.1b. In this derivation, we consider that the discharging nMOS transistor (M_1) always operates in velocity saturation (since $V_{ds}(M_1) > V_{dsat}$) and M_2 operates in linear region only, when $V_{out}(T_R) \geq V_{dsat}$ constraint is followed. As a representative, we show the simulation results of $t_{TCP_{90\%}}$ (shown in Fig. 5.1b). This is because $t_{TCP_{90\%}} < T_R$ in the whole T_R, C_l space used for characterization of LUTs.

To find the $t_{TCP90\%}$ model, we integrate the saturation current flowing through M_1, from $t = \left(\frac{V_{th}}{V_{dd}}\right) T_R$ to $t = t_{TCP}$. We equate this integral to $(C_l + C_p)\,(V_{dd} - V_{TCP})$ to obtain the t_{TCP}. Solving this integral using (5.2), (5.3) and (5.6), we get the expression for $t_{TCP90\%}$ as:

$$t_{TCP90\%} = A_1 T_R + A_2 \sqrt{T_R} \tag{5.16}$$

$$\text{Where, } A_1 = \left(\frac{V_1}{V_{dd}\left(1 - \frac{\beta_s}{[\beta_l(V_{dd} - V_{th}) + \beta_s]}\right)} \right) \tag{5.17}$$

$$V_1 = \left[V_{th1} - \frac{(\beta_s V_{th1} - C_{C1}\frac{V_{dd}}{T_R} + I_X)}{[\beta_l(V_{dd} - V_{th}) + \beta_s]} \right] \tag{5.18}$$

$$A_2 = \sqrt{\frac{0.2\,(C_l + C_p)}{\beta_s\left(1 - \frac{\beta_s}{[\beta_l(V_{dd} - V_{th}) + \beta_s]}\right)}} \tag{5.19}$$

For all the $t_{TCP}s$ which fall under Region II, (5.16) remains valid. The model of (5.16) in Region II has thus been verified using HSPICE simulated data. Using curve fitting (as shown in Fig. 5.4) on the simulated values of $t_{TCP90\%}$, we extracted the coefficients A_1 and A_2 of (5.16). Please note that we have considered the impact of V_X on the discharge of

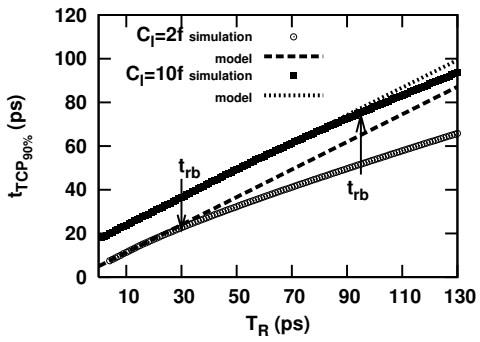


Figure 5.4: Case 1: Variation of $t_{TCP90\%}$ with respect to T_R.

node 'Out'.

From (5.16), we observe :

- Observation 4: A_1 is independent of W_n and C_l
- Observation 5: A_2^2 is a linear function of C_l
- Observation 6: A_2^2 is proportional to $1/W_n$

As explained earlier, (5.16) is valid if $V_{out}(T_R) \geq V_{dsat}$, this imposes the following constraint on the region of validity:

$$\triangle Q(T_R) = (a_1 T_R - a_2) > (C_l + C_p)(V_{dd} - V_{TCP}) \tag{5.20}$$

For a given set of values of C_l , T_R in the LUT, we first verify through (5.20) whether a given $\alpha\%$ TCP (represented as $TCP - \alpha\%$) falls in Region I. We then use the model of (5.7) to obtain $t_{TCP-\alpha\%}$. If $TCP - \alpha\%$ falls in Region II for this C_l , T_R values, we use the model of (5.16) to obtain $t_{TCP-\alpha\%}$. This saves a large number of HSPICE simulations in ECSM (or CCSM) characterization. Next, we verify observations 4-6 against HSPICE simulation data using 32nm PTM model. Using curve fitting (as shown in Fig. 5.4) on the simulated values of $t_{TCP90\%}$, we extract the coefficients (A_1 and A_2) of (5.16). In Fig. 5.5a and 5.5b, we verify that the model coefficient's (A_1 and A_2) behavior with W_n and C_l is in accordance with our prediction.

5.4.3 Efficient ECSM Characterization

In this subsection, we demonstrate the effectiveness of our models (derived in Subsection 5.4.1 and 5.4.2) in saving the number of HSPICE simulations for ECSM characterization of 2-input NAND gate. Using (5.7) and (5.16), within the regions of their validity, we get the values of all TCPs directly (without any HSPICE simulation). Hence, it saves characterization effort for the standard cell. On the other hand, the TCP values which

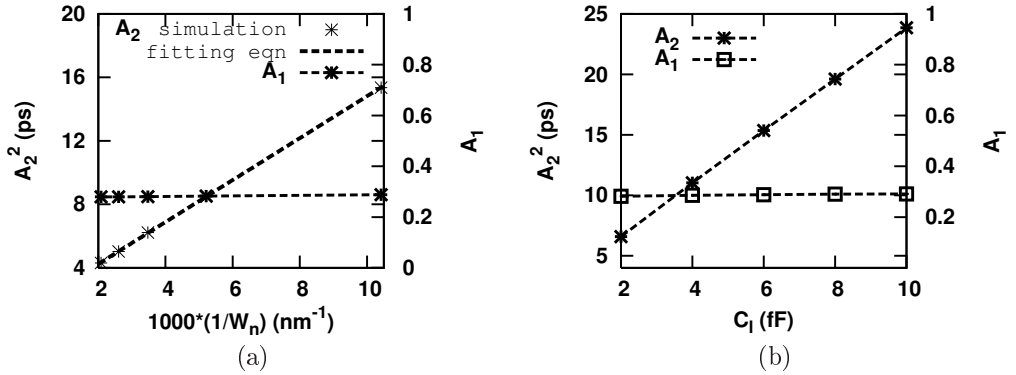


Figure 5.5: Case 1: Variation of A_1 and A_2 with cell size (W_n) and load capacitance (C_l).

are out of the region of validity, would be obtained from HSPICE simulations. For a 2-input NAND gate standard cell, we first extract the values of K_1, K_2 and K_3 for TCPs in Region I using 7 HSPICE simulations and the values of A_1, A_2 for TCPs in Region II using 4 HSPICE simulations. We then calculate the values of $t_{TCP}s$ (entries shown by numeric values in Table 5.1) for C_l , T_R values lying within the region of validity of (5.7) and (5.16). For 2-input NAND gate standard cells of other sizes, all TCPs (within the region of validity of (5.7) and (5.16)) are obtained using Observation $4-6$. In Table 5.1, $t_{TCP60\%}$ values calculated using our model are shown. In Table 5.2, $t_{TCP60\%}$ values generated using conventional method (*i.e.* fully HSPICE generated) are shown. In Table 5.3, the percentage error in $t_{TCP60\%}$ values calculated using our model compared to conventional method is shown. In Table 5.1, $t_{TCP60\%}$ values calculated using our model show some entries as 'HSPICE', it corresponds to the data points (*i.e.* $t_{TCP60\%}$ values) of C_l , T_R values which are out of region of validity for (5.7). And these values would be generated using HSPICE simulations.

We observe that the percentage saving in HSPICE simulations using our method of generating LUTs explained above is minimum 67.35% and 77.55% for $TCP_{60\%}$ and $TCP_{90\%}$, respectively (including the simulations required to obtain the model coefficients). We find that the values of $t_{TCP}s$ in our LUTs differ by a maximum of 1.28% (for 7×7matrix size of $t_{TCP60\%}$) and 3.08% (for 7×7matrix size of $t_{TCP90\%}$) from those in fully HSPICE generated conventional LUTs. Therefore using proposed models, standard cell characterization can be done with a significantly lesser number of HSPICE simulations (approximately 67% reduction in HSPICE simulations). For the $t_{TCP} < T_R$, both the Region I and II models are used .

5.5 Case 1: Variation aware TCP models

In this section, we consider layout dependent effects due to mechanical stress in deriving the model coefficients. We also derive the relationships of the model coefficients with the temperature (T) and supply voltage (V_{dd}). We have already discussed the motivation towards the analysis of variation aware (considering process induced mechanical stress,

Table 5.1: Case 1: LUT of $TCP_{60\%}$ for 2-input CMOS NAND gate using our model.

$C_l(fF)$	$T_R(ps)$						
	2.20	4.84	10.66	23.46	51.62	113.60	250.00
1.51	15.45	17.17	20.95	29.27	HSPICE	HSPICE	HSPICE
2.28	20.54	22.26	26.04	34.36	52.67	HSPICE	HSPICE
3.45	28.24	29.95	33.73	42.05	60.36	HSPICE	HSPICE
5.22	39.86	41.58	45.36	53.68	71.99	112.27	HSPICE
7.88	57.43	59.15	62.93	71.25	89.56	129.84	HSPICE
11.91	83.99	85.70	89.48	97.80	116.11	156.40	245.06
18.00	124.11	125.83	129.61	137.93	156.24	196.52	285.18

Note: Entries shown by 'HSPICE' can be obtained by HSPICE simulation whereas entries shown by numeric values are obtained using our model

Table 5.2: Case 1: ECSM LUT of $TCP_{60\%}$ for 2-input CMOS NAND gate obtained using HSPICE simulations.

$C_l(fF)$	$T_R(ps)$						
	2.20	4.84	10.66	23.46	51.62	113.60	250.00
1.51	15.35	17.01	20.69	28.9	47.00	81.89	150.21
2.28	20.45	22.11	25.80	33.98	52.33	89.56	161.03
3.45	28.14	29.81	33.50	41.67	59.88	99.41	175.10
5.22	39.77	41.43	45.12	53.28	71.40	111.95	193.11
7.88	57.32	58.99	62.68	70.84	88.89	129.11	215.92
11.91	83.86	85.53	89.22	97.37	115.37	155.34	244.72
18.00	123.95	125.62	129.31	137.46	155.43	195.22	283.85

Table 5.3: Case 1: Percentage error in proposed model's LUT with respect to fully HSPICE generated ECSM LUT of $TCP_{60\%}$ for 2-input CMOS NAND gate.

$C_l(fF)$	$T_R(ps)$						
	2.20	4.84	10.66	23.46	51.62	113.60	250.00
1.51	0.65	0.94	1.26	1.28	×	×	×
2.28	0.44	0.68	0.93	1.12	0.65	×	×
3.45	0.36	0.47	0.69	0.91	0.80	×	×
5.22	0.23	0.36	0.53	0.75	0.83	0.29	×
7.88	0.19	0.27	0.40	0.58	0.75	0.57	×
11.91	0.16	0.20	0.29	0.44	0.64	0.68	0.14
18.00	0.13	0.17	0.23	0.34	0.52	0.67	0.47

Note: Entries shown by '×' correspond to the values we obtain using HSPICE simulations in Table 5.1 (not through our models)

supply voltage and temperature variability) TCP models in Chapter 4 of the thesis. Here, we derive and validate the behavior of model coefficients with PVT variations. Later, we use these models and their derived relationships to reduce the number of HSPICE simulations significantly.

5.5.1 TCP models considering stress variability

In this subsection, we derive the change in coefficients in (5.7) and (5.16) as a function of NF in a stress enabled 45nm CMOS technology. The novelty of this work is due to the prediction of entire output waveform of the standard cell as a function of channel stress (or NF) for a wide range of T_R and C_l. As discussed in Section 6.2, the effect of change in channel stress as a function of NF is captured by PTM parameters MULU0 and DELVT0 in our HSPICE simulations.

Now we derive the model coefficients of (5.7) and (5.16) as a function of NF. We use a set of empirical equations suggested in [89, 92, 93], to relate device level electrical parameters with stress. These equations are:

$$\mu\left(\sigma\right) = \left[P_1\,\sigma(NF) + 1\right]\mu_0 \tag{5.21}$$

$$v_{sat}\ \left(\sigma\right) = \left[P_1\,\sigma(NF) + 1\right]v_{sat} \tag{5.22}$$

$$I_{ON}\left(\sigma\right) = \left[P_1\,P_2\,\sigma\left(NF\right) + 1\right]I_{ON} \tag{5.23}$$

$$V_{th}\ \left(\sigma\right) = \left[V_{th} + P_3\,\sigma\left(NF\right)\right] \tag{5.24}$$

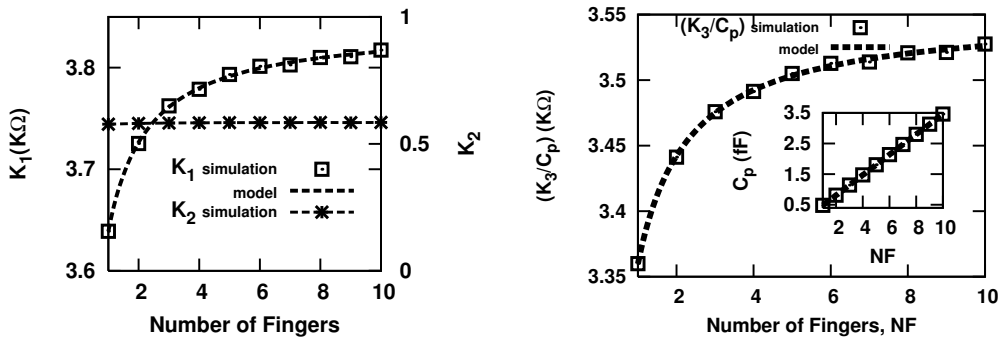
Where $\mu\left(\sigma\right)$, $v_{sat}\left(\sigma\right)$, $I_{ON}\left(\sigma\right)$, $V_{th}\left(\sigma\right)$ are stress dependent mobility, saturation velocity, drive current and threshold voltage parameters, respectively. Whereas, μ_0, v_{sat}, I_{ON}, V_{th} are unstressed parameters and P_1 is the piezoresistive coefficient. P_2 and P_3 are technology dependent coefficients extracted by fitting the above equations into HSPICE simulated I-V data, as discussed in [89, 93]. Here, $\sigma\left(NF\right)$ represents the average stress in the fingers in MFGSs. As we discussed in [89], a relation between this average stress and NF is:

$$\frac{\sigma\left(NF\right)}{\sigma_{ref}\left(NF=1\right)} = M_1 + \frac{M_2}{NF + M_3} \tag{5.25}$$

In (5.25), M_1, M_2, M_3 are fitting parameters specific to given technology node, where M_1 denotes $\sigma\left(NF \to \infty\right)/\sigma\left(NF=1\right)$, while M_2 and M_3 control the rate of change of stress as NF is increased (as discussed in detail in [89]).

5.5.1.1 Impact of stress induced variability in Region I

In this sub-subsection, we derive the relationship of $t_{TCP}s$ with T_R, C_l and NF for Region I in stress enabled technologies. In (5.9), (5.13) and (5.14), μ, v_{sat}, I_{ON}, and V_{th} are now given by (5.21)-(5.24). We derive the model for $t_{TCP60\%}$ considering the impact of



(a) (b)

Figure 5.6: Case 1: K_1, K_2 and K_3 as a function of NF.

channel stress in Region I. In this derivation, we assume that $(V_{dd} - V_{th})$ is independent of NF. This is justified since V_{th} is much smaller than V_{dd}. From (5.13) and (5.21)-(5.25), we obtain :

$$K_1 = (NF + M_3)\left(\frac{D_1}{NF + D_2}\right) \tag{5.26}$$

Where,

$$D_1 = \frac{(V_{dd} - V_{dsat})}{((M_1 P_1 P_2 + 1)\, I_{ON})} \tag{5.27}$$

$$D_2 = M_3 + \frac{M_2 P_1 P_2}{(M_1 P_1 P_2 + 1)} \tag{5.28}$$

We observe that (5.26) fits well with HSPICE simulated data as shown in Fig. 5.6a. In these HSPICE simulation MULU0 and DELVT0 vary with NF in accordance with (5.25) (*i.e.* the PTM model incorporates channel stress variability effects). Likewise, solving (5.14), we found the relation between K_3 and NF as:

$$(K_3/C_p) = K_1 + \frac{a_2\,(\sigma)}{C_p\, I_{ON}\,(\sigma)} = \left[(NF + M_3)\left(\frac{D_3}{NF + D_2}\right)\right] \tag{5.29}$$

Where,

$$D_3 = [(V_{dd} - V_{dsat}) + a_2]\,\frac{1}{((M_1 P_1 P_2 + 1)\, I_{ON})} \tag{5.30}$$

Where, C_p is parasitic capacitance (due to gate-drain overlap, drain-bulk junction capacitance etc.) which is linearly related to NF in MFGS. This can be seen in the inset of Fig. 5.6b. We observe that (5.29) fits well with HSPICE simulated data as shown in Fig. 5.6b. Thereafter, from (5.9), we observe that K_2 is independent of NF (because a_1 and I_{ON} are both proportional to NF) as shown in Fig. 5.6a.

5.5.1.2 Impact of stress induced variability in Region II

In this sub-subsection, we derive the relationships of $t_{TCP}s$ with T_R, C_l and NF for

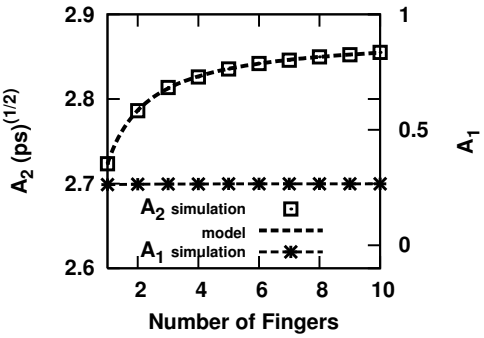


Figure 5.7: Case 1: A_1 and A_2 as a function of NF.

Region II in stress enabled technologies. In (5.17) and (5.19), μ, v_{sat}, I_{ON} and V_{th} are now given by (5.21)-(5.24). We derive a model for $t_{TCP90\%}$ considering the impact of channel stress in Region II. From (5.17) and (5.21)-(5.25), we obtain that A_1 is independent of $V_{th}(\sigma)$. We have verified through simulations that variation in $V_{th}(\sigma)$ with NF is very small (in our TCAD calibrated simulations variation in $V_{th}(\sigma)$ with NF is less than 3%). This small variation in $V_{th}(\sigma)$ has also been reported in [94]. Therefore, A_1 is independent of $\sigma(NF)$ as shown in Fig. 5.7. Using (5.19), (5.22) and (5.25), We find the relationship between A_2 and NF as:

$$A_2 = \sqrt{\left(\frac{S_1(NF + M_3)}{(NF + S_2)}\right)} \tag{5.31}$$

Where,

$$S_1 = \left(\frac{0.2\,(C_l + C_p)}{\beta_s\left(1 - \frac{\beta_s}{[\beta_l(V_{dd} - V_{th}) + \beta_s]}\right)}\right) \tag{5.32}$$

$$S_2 = \left(\frac{1}{M_3 + \frac{P_1 M_2}{(M_1 P_1 + 1)}}\right) \tag{5.33}$$

We observe that (5.31) fits well on our stress aware HSPICE simulation data as shown in Fig. 5.7.

5.5.1.3 Efficient stress aware ECSM characterization

In this sub-subsection, we show that using our t_{TCP} models, ECSM characterization of 2-input NAND gate standard cell would need significantly lesser number of HSPICE simulations. We generated 7×7 LUTs having $t_{TCP60\%}$ values with varying C_l and T_R values for different cell sizes (represented by NF). Conventionally, this would require 343 HSPICE simulations for cell sizes corresponding to NF=1 to 7. However using our models only 68 HSPICE simulations (including simulations to extract the model coefficients) were needed to generate the same LUT. For computing the remaining values of $t_{TCP60\%}$, we used our model. In this way, we saved 275 HSPICE simulations to generate the 7×7 LUTs for NF=1 to 7. Therefore, proposed model can be used to save $\backsimeq$ 80.18% HSPICE

simulations for all $TCPs$ in Region I if $V_{TCP} < V_{out}(T_R)$. For all the TCPs having value $V_{TCP} > V_{out}(T_R)$, the proposed models (Region I and II models) can be used to save 97.38% HSPICE simulations.

5.5.2 TCP models considering temperature variability

In this subsection, we analytically model the coefficients of (5.7) and (5.16) as a function of temperature. We also model the region of validity of (5.7) and (5.16) in C_l, T_R space as a function of temperature. In this work, we take a realistic range of temperature variation due to on-chip heating from $298K$ (room temperature) to $423K$. We use an empirical expression suggested in [23, 29], to consider the impact of temperature variability on carrier mobilities. This expression is:

$$\mu\left(T\right) = \mu(T_0)\left(\frac{T}{T_0}\right)^{-\theta} \tag{5.34}$$

Where, $\mu(T)$ is the temperature dependence of mobility, T is the temperature, T_0 is the nominal temperature *i.e.* $298\,K$ and θ is technology dependent temperature coefficient. For our PTM CMOS technology we extract the value of $\theta = 2.3$ by maximum transconductance (g_m) method. We use the empirical relation (given in [29]) between carrier saturation velocity and temperature as:

$$v_{sat}\ \left(T\right) = v_{sat}\left(T_0\right) - \eta\left(T - T_0\right) \tag{5.35}$$

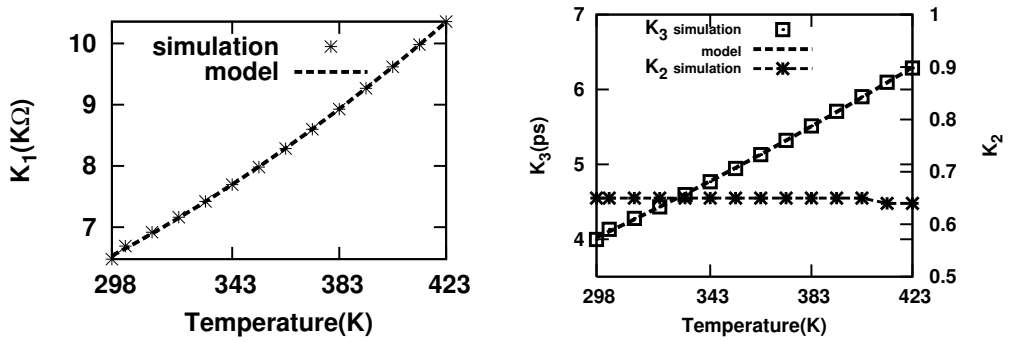
Where, $v_{sat}\left(T\right)$ is the temperature dependence of saturation velocity and η is the temperature coefficient. The threshold voltage of devices also gets affected by an increase in temperature due to change in fermi level location and band gap energy [96]. In [96], the temperature dependence of threshold voltage is given by:

$$V_{th}\ \left(T\right) = V_{th}\left(T_0\right) - \kappa\left(T - T_0\right) \tag{5.36}$$

Where, κ is the temperature dependence coefficient of threshold voltage. The value of η and κ can be extracted from simulated HSPICE I-V data for a given CMOS technology. We also verified the validity of (5.34), (5.35) and (5.36) relations using Sentaurus TCAD device simulations. In TCAD simulations, we use 25nm drawn gate length nMOS and pMOS devices. We now discuss our approach for deriving temperature variation aware t_{TCP} models.

5.5.2.1 Impact of temperature variability in Region I

In this sub-subsection, we derive the relationship of $t_{TCP}s$ with T_R, C_l and temperature (T) for Region I. In (5.9), (5.13) and 5.14, v_{sat}, μ and V_{th} are now given by (5.34)-(5.36). We derive a model for $t_{TCP60\%}$ considering the impact of temperature variability in Region I. In this derivation, we again assume that $(V_{dd} - V_{th})$ is independent of T, as in the previous section. From (5.13) and (5.34), (5.35), we obtain :



(a) (b)

Figure 5.8: Case 1: Variation of K_1, K_2 and K_3 with temperature.

$$K_1 = \left(R_1\, T^{2.3} + R_2\, T + R_3\right) \tag{5.37}$$

Where,

$$R_1 = \left[\frac{\left(\frac{L_{eff}}{W_n\, P_l}\right)}{\left(\mu(T_0)\left(\frac{1}{T_0}\right)^{-2.3}\right)}\right] \tag{5.38}$$

$$R_2 = \left[\frac{V_{dd}}{W_n\, P_s\,(V_{gs} - V_{th})}\frac{\eta}{v_{sat}\,(T_0)^2}\right] \tag{5.39}$$

$$R_3 = \left[\frac{V_{dd}}{W_n\, P_s\,(V_{gs} - V_{th})}\frac{(1 - \eta\, T_0)}{v_{sat}\,(T_0)^2}\right] \tag{5.40}$$

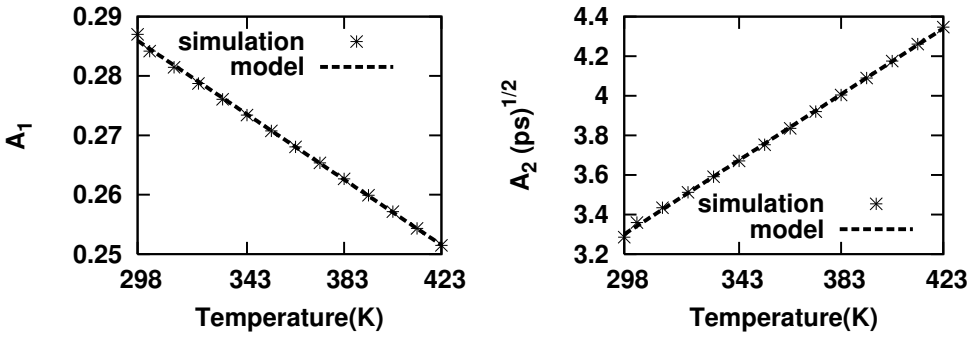
R_1, R_2 and R_3 are the technology dependent constants obtained from the derivation of (5.13) using (5.34) and (5.35). We observe that (5.37) fits well on HSPICE simulated data as shown in Fig. 5.8a. Likewise solving (5.14), we find the relation between K_3 and T as:

$$K_3 = C_p\, K_1 + \frac{a_2}{I_{ON}} = \left(R_4\, T^{2.3} + R_5\, T + R_6\right) \tag{5.41}$$

Where, R_4, R_5 and R_6 are the technology dependent constants obtained after deriving the (5.14) using (5.34)-(5.36). Using HSPICE simulations, we find that the change in C_p with T varying from 298K to 423K is negligibly small. We observe that (5.41) fits well with HSPICE simulated data as shown in Fig. 5.8b. Thereafter, solving (5.9), we observe that K_2 is independent of T as shown in Fig. 5.8b (because a_1 and I_{ON} both are proportional to T).

5.5.2.2 Impact of temperature variability in Region II

In this sub-subsection, we derive the relationship of $t_{TCP}s$ with T_R, C_l and T for Region II. In (5.17) and (5.19), v_{sat}, μ and V_{th} are now given by (5.34)-(5.36). We derive a model for $t_{TCP90\%}$ considering the impact of temperature variability in Region II. From (5.17), we observe that A_1 is dependent on V_{th} only. From (5.36), we expect A_1 to reduce linearly with



(a) (b)

Figure 5.9: Case 1: Variation of A_1 and A_2 with temperature.

T, which we verify through HSPICE simulations in Fig. 5.9a. We obtain the relationship between A_2 and T from (5.19) and (5.35):

$$A_2 = (R_7 T + R_8) \tag{5.42}$$

Where, R_7 and R_8 are technology dependent constants obtained from the derivation of (5.19) using (5.35). We observe that (5.42) fits well on our HSPICE simulation data with temperature variability as shown in Fig. 5.9b.

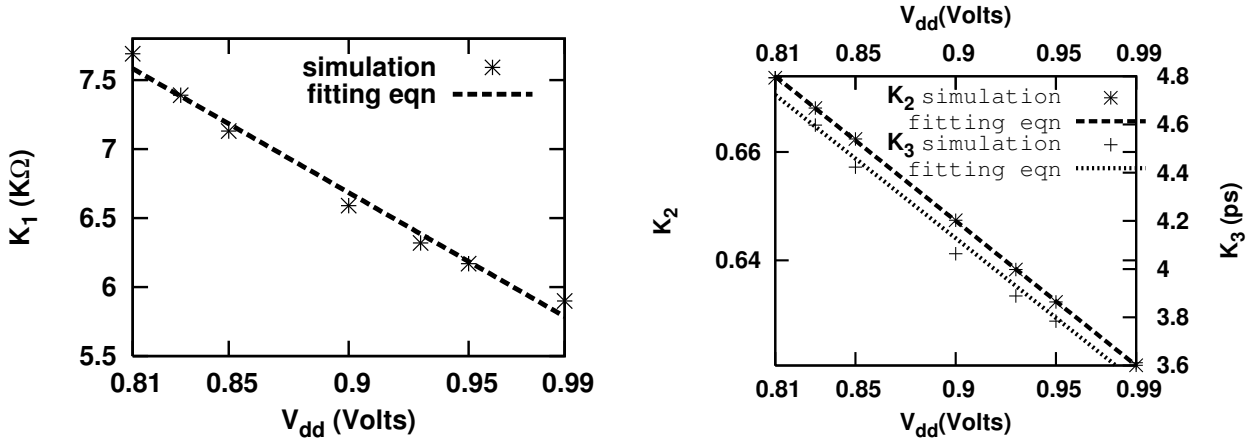
Therefore, we observe that our models in Region I and II are valid with temperature variability.

5.5.2.3 Efficient temperature variation aware ECSM characterization

In this sub-subsection, we show that using our t_{TCP} models, ECSM characterization of 2-input NAND gate standard cell would need significantly lesser number of HSPICE simulations. We generated 7×7 LUTs having $t_{TCP60\%}$ values with varying C_l and T_R values for different temperature values. Conventionally, this would require 343 HSPICE simulations for 7 different values of temperature T from $298K$ to $423K$. However, using our models only 69 HSPICE simulations (including simulations to extract the model coefficients) were needed to generate the same LUT. For computing the remaining values of $t_{TCP60\%}$, we used our model. In this way, we saved 274 HSPICE simulations to generate the 7×7 LUTs for $T = 298K$ to $423K$. Therefore, proposed model can be used to save $\backsimeq 79.88\%$ HSPICE simulations for all TCPs in Region I if $V_{TCP} < V_{out}(T_R)$. For all the TCPs having value $V_{TCP} > V_{out}(T_R)$, our models (Region I and II models) can be used to save 97.09% HSPICE simulations.

5.5.3 TCP models considering Supply Voltage Variability

In this subsection, we derive the power supply voltage variation aware t_{TCP} models which we use to reduce the re-characterization effort significantly. We consider the $\pm 10\%$ change in power supply voltage (V_{dd}) from the its nominal value of $V_{dd} = 0.9V$ [99].



(a) (b)

Figure 5.10: Case 1: Variation of K_1, K_2 and K_3 with supply voltage.

5.5.3.1 Impact of supply voltage variability in Region I

In this sub-subsection, we derive the relationship of $t_{TCP}s$ with T_R, C_l and V_{dd} for Region I. We derive a model for $t_{TCP60\%}$ considering the impact of supply voltage variability in Region I. In (5.13), we observe that K_1 decreases with an increase in supply voltage. We verify this observation in Fig. 5.10a.

We now discuss the variation of K_3 with V_{dd}. In (5.14), C_p contains the voltage dependent junction capacitance which is given by [94] as $C_j(V) = \frac{A.C_{j0}}{\sqrt{\left(1+\frac{V}{\phi_0}\right)}}$. Where, A indicates the junction area, C_{j0} is zero-bias junction capacitance per unit area, V is the reverse bias voltage and ϕ_0 is built-in potential. In (5.14), we observe that the behavior of K_3 with supply voltage remains same as of K_1. From (5.9), we observe that K_2 is inversely proportional to V_{dd}. We verify our observation related to K_2 and K_3 with V_{dd} in Fig. 5.10b.

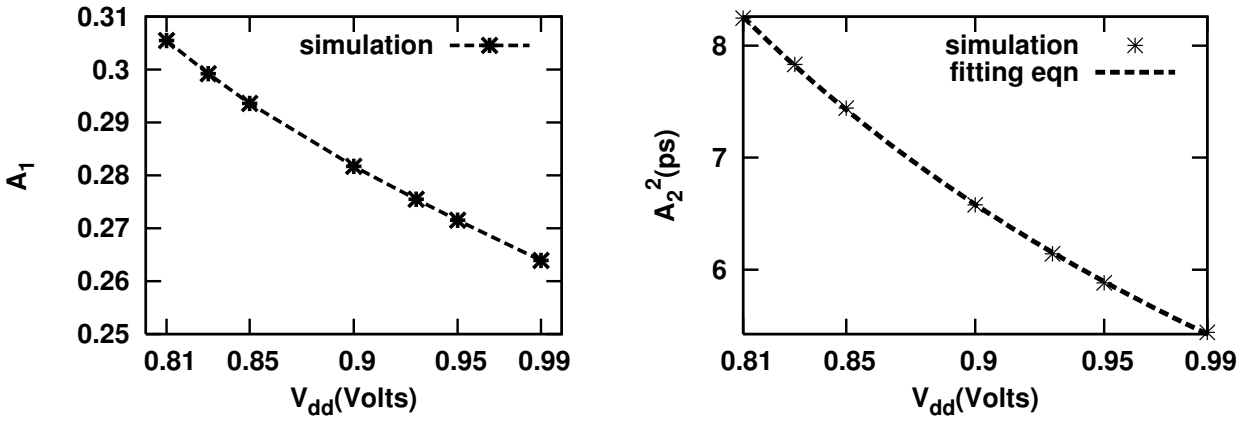
5.5.3.2 Impact of supply voltage variability in Region II

In this sub-subsection, we derive the relationship of $t_{TCP}s$ with T_R, C_l and V_{dd} for Region II. We derive a model for $t_{TCP90\%}$ considering the impact of supply voltage variability in Region II. In (5.17), A_1 is inversely proportional to supply voltage and it is verified through HSPICE simulations as shown in Fig. 5.11a. In (5.19), we can see that only C_p is supply voltage dependent term which decreases with supply voltage. Using this analysis, we observe that A_2^2 is inversely proportional to V_{dd}. We verify this observation in Fig. 5.11b.

Thus, we observe that our model in Region I and Region II incorporates the effect of supply voltage variability quite satisfactorily.

5.5.3.3 Efficient supply voltage variation aware ECSM characterization

In this sub-subsection, we show that using our t_{TCP} models, ECSM characterization of 2-input NAND gate standard cell would need significantly lesser number of HSPICE simulations. We generated 7×7 LUTs having $t_{TCP60\%}$ values with varying C_l and T_R values



(a) (b)

Figure 5.11: Case 1: Variation of A_1 and A_2 with supply voltage.

for different supply voltages. Conventionally, this would require 343 HSPICE simulations for 7 different values of supply voltage V_{dd} from 0.81V to 0.99V. However using our models only 69 HSPICE simulations (including simulations to extract the model coefficients) were needed to generate the same LUT. For computing the remaining values of $t_{TCP60\%}$, we used our model. In this way, we saved 274 HSPICE simulations to generate the 7×7 LUTs for $V_{dd} = 0.81V$ to $0.99V$. Therefore, proposed model can be used to save $\backsimeq 79.88\%$ HSPICE simulations for all TCPs in Region I if $V_{TCP} < V_{out}(T_R)$. For all the TCPs having value $V_{TCP} > V_{out}(T_R)$, our models (Region I and II models) are used to save 95.92% HSPICE simulations.

5.6 Case 2: Derivation and Validation of t_{TCP} model

In this section, we derive the t_{TCP} model for the switching of the lower nMOS transistor in the series stack of 2-input NAND gate. We derive the relationship of $t_{TCP}s$ with the T_R, C_l and cell size. We also derive the region of validity of these relationships in T_R, C_l space. We derive the model for the case when $V_{out}(T_R) \geq V_{dsat}$. In this work, we ignore the negligibly small current flowing through pMOS transistor (max. current$\approx 3\mu A$). We classify the TCPs into two regions:

- Region I: When $V_{in} = V_{dd}$ (for $t_{TCP}s > T_R$)
- Region II: When $V_{in} < V_{dd}$ (for $t_{TCP}s < T_R$)

In this section, we derive the timing model of $TCP_{10\%}$ in Region I and timing model of $TCP_{90\%}$ in Region II (as shown in Fig. 5.12b).

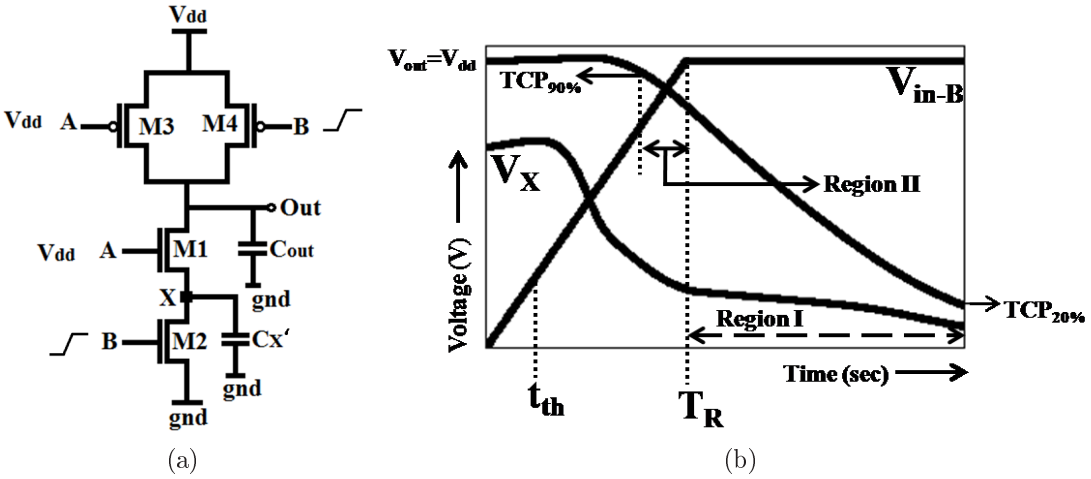


Figure 5.12: Case 2: (a) 2-input NAND gate schematic (b) its I/O waveform.

5.6.1 Derivation of the model in Region I

In this subsection, we derive the model for $t_{TCP}s$ in Region I for the 2-input NAND gate shown in Fig. 5.12a. We first model the voltage transition at the intermediate node 'X' considering different operating regions of lower nMOS transistor. As a representative, we show the simulation results of $t_{TCP_{60\%}}$. This is because $t_{TCP_{60\%}} > T_R$ is true in the whole T_R, C_l space used in characterization LUTs, as we discuss later.

To derive the model for $t_{TCP10\%}$, we proceed as follows: The transistor M_1 starts conducting at $t = t_{th} = \frac{V_{th}}{V_{dd}} T_R$. The transistor M_2 operates in saturation region till V_X reaches V_{dsat} (corresponding time is represented as t_{xsat}). We observe from simulation that within the whole T_R, C_l space used in characterization LUTs, $t_{xsat} < T_R$ (which we observe in our simulation). From t_{xsat} to t_{TCP}, M_2 operates in linear region. We first integrate the current through M_1 from t_{th} to T_R and equate the integral to $(C_l + C_p)\,(V_{dd} - V_{out}(T_R))$ to obtain the value of $V_{out}(T_R)$. If the $V_{TCP} \geq V_{dsat}$, we integrate the saturation current through M_1 from $t = T_R$ to $t = t_{TCP}$. We equate the sum of these integrals to $(C_l + C_p)\,(V_{out}(T_R) - V_{TCP})$ to obtain t_{TCP} . Here C_l is the load capacitance, C_p is the parasitic capacitance. If the $V_{TCP} < V_{dsat}$, we integrate I_{ON} through M_1 from $t = T_R$ to $t = t_{sat}$ (where $t = t_{sat}$ represents the time at which $V_{out} = V_{dsat}$). From $t = T_R$ to t_{sat}, the lower nMOS transistor (M_2) operates in linear region, therefore we consider M_2 as a linear resistor (represented as R_2) which gives the value of intermediate node voltage $V_X = I_{M1} R_2$. This R_2 results in the the lowering of saturation current of M_1 by a constant value which is technology independent (since, $I_{M1} = \frac{I_{ON}}{(1+\beta_s R_2)}$). Therefore, in our derivation we represent this saturation current I_{M1} by the symbol I_{ON} from $t = T_R$ to t_{sat}. We equate the sum of these integrals to $(C_l + C_p)\,(V_{out}(T_R) - V_{dsat})$ to obtain t_{sat}. From$t = t_{sat}$ to $t = t_{TCP}$, both the series stacked nMOS devices operate in linear region. These devices act as a resistances R_1 and R_2 whose values can be obtained by equating $V_{gs1} = V_{gs2} = V_{dd}$. Therefore, the time duration from t_{sat} to t_{TCP} is proportional to time constant $R_{eq}(C_l + C_p)$. Where, R_{eq} represents an equivalent resistance of series stacked nMOS devices. As we discuss later in this section, for $V_{TCP} < V_{dsat}$ the value of V_X is so low that the change in V_X during this

RC discharge need not to be considered.

Now we discuss in detail the derivation of $t_{TCP10\%}$ model. Following the procedure as explained in the previous paragraph, we write an expression for the output voltage discharge as:

$$\int_{t_{th}}^{T_R} I_{M1} dt = (C_l + C_p)\,(V_{dd} - V_{out}(T_R)) \tag{5.43}$$

Taking L.H.S. and solving it :

$$\int_{t_{th}}^{T_R} I_{M1} dt = \beta_{s-M1} \int_{t_{th}}^{T_R} (V_{dd} - V_X(t) - V_{th1})\, dt \tag{5.44}$$

In this derivation, we represent the threshold voltage of M_1 as V_{th1} (threshold voltage of M_2 as V_{th}).

$$= \beta_{s-M1} \int_{t_{th}}^{T_R} (V_{dd} - V_{th1})\, dt - \beta_{s-M1} \left[\int_{t_{th}}^{t_{xsat}} V_X(t)\, dt + \int_{t_{xsat}}^{T_R} V_X(t)\, dt \right] \tag{5.45}$$

From $t = t_{th}$ to t_{xsat}, both nMOS transistors operate in saturation region. To determine V_X, we apply KCL at node 'X' which gives an expression as follows:

$$I_{M2} = I_{M1} + I_{C_{CX}} + I_{C_X} \tag{5.46}$$

$$I_{M2} = I_{M1} - (C'_X + C_{CX}) \frac{dV_X}{dt} + C_{CX} \frac{dV_{in-B}}{dt} \tag{5.47}$$

Where,

$$V_{in-B} = V_{dd} \left(\frac{t}{T_R} \right) \text{ for } 0 < t < T_R \tag{5.48}$$

Where, $I_{C_{CX}}$ is the current flowing through gate-to-drain coupling capacitance of M_2, it consists of the gate-to-drain overlap capacitance and a part of the gate-to-channel capacitance. The other symbols used in (5.46), have the same meaning as in Subsection 5.4.1 of this chapter. Taking $(C'_X + C_{C1}) = C_X$, we rewrite (5.46) as:

$$\beta_{s-M2}(V_{in-B} - V_{th}) = \beta_{s-M1}(V_{dd} - V_X(t) - V_{th1}) - C_X \frac{dV_X}{dt} + C_{CX} \frac{V_{dd}}{T_R} \tag{5.49}$$

This equation is true only if $V_X(t)$ is a linear function of time since V_{in-B} is proportional to t. Therefore,

$$\int_{t_{th}}^{t_{xsat}} V_X(t)\, dt = \frac{1}{2}(t_{xsat} - t_{th})(V_{dd} - V_{th1} - V_{dsat}) \tag{5.50}$$

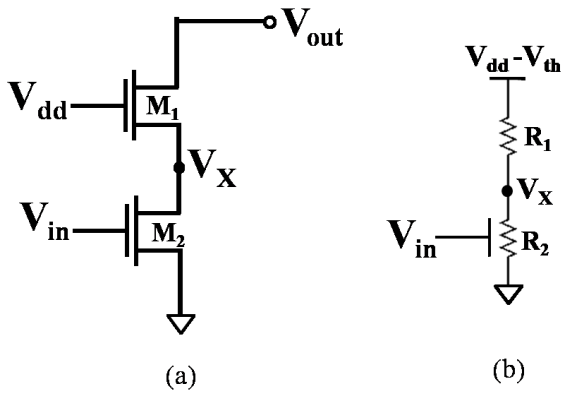


Figure 5.13: (a) Pull Down part of 2-input NAND gate (b) Its equivalent circuit looking at node 'X' when M_1 operates in saturation region and M_2 operates in linear region.

Where,

$$t_{xsat} = \left(\frac{V_{in,sat}}{V_{dd}}\right) T_R \tag{5.51}$$

$$t_{th} = \left(\frac{V_{th}}{V_{dd}}\right) T_R \tag{5.52}$$

Here, $V_{in,sat}$ represents the rising input voltage at time $t = t_{xsat}$. From $t = t_{xsat}$ to T_R, M_1 and M_2 operates in saturation and linear region, respectively.

To find out $\int_{t_{xsat}}^{T_R} V_X(t)\,dt$ in (5.45), we show with the help of figure (see Fig. 5.13) that the pull down part of NAND gate in Fig. 5.13 (a) and (b) are equivalent looking from the intermediate node 'X'. Here, $R_1 = \frac{1}{\beta_{s-M1}}$. We observe that the current through M_1 is equal in Fig. 5.13 (a) and (b).Therefore, we obtain the value of V_X from Fig. 5.13 (b) as:

$$V_X = \frac{(V_{dd} - V_{th1})}{1 + \frac{\beta_{l,\,M2}}{\beta_{s,\,M1}}(V_{in-B} - V_{th})} \tag{5.53}$$

Using (5.53), we write the intermediate node voltage as:

$$\int_{t_{xsat}}^{T_R} V_X(t)\,dt = V_1 T_R \tag{5.54}$$

Where,

$$V_1 = \left[\frac{\beta_{s,\,M1}}{\beta_{l,\,M2}} \frac{(V_{dd} - V_{th1})}{V_{dd}} ln\left[\frac{V_{dd}}{V_{in,sat}}\right]\right] \tag{5.55}$$

Where, $V_{in,sat}$ is the input voltage, V_{in-B} at $t = t_{xsat}$. In deriving the value of V_1, we use the following two points:

1. The value of V_{in-B} at t_{xsat} is a technology based constant and is independent of values of C_l, T_R and cell size. We now explain the reason for this. The discharge of node 'X' from $V_{dd} - V_{th1}$ to V_{xsat} (=V_{dsat} of M_2) happens linearly with time with the slope

$1/T_R$ as can be seen from (5.49).

2. Therefore, with an increase in T_R, the time duration of fall in voltage of node 'X' from $V_{dd} - V_{th1}$ to V_{xsat} is proportional to T_R. Since this discharge happens for input V_{in-B}'s value increasing from $V_{in-B} = V_{th}$ to $V_{in-B} = V_{insat}$, V_{insat} is independent of T_R.

3. From (5.53), we extracted the value of $\left(\frac{\beta_{l,\,M2}}{\beta_{s,\,M1}}\right)$and using this value, we find that the term $\left(1 - \frac{\beta_{l,\,M2}}{\beta_{s,\,M1}} V_{th}\right) << \frac{\beta_{l,\,M2}}{\beta_{s,\,M1}} V_{insat}$ (and therefore $\frac{\beta_{l,\,M2}}{\beta_{s,\,M1}} V_{dd}$).

The value of $\frac{V_{dd}}{V_{insat}}$ in 32nm PTM technology used by us is 1.28 ($\frac{V_{dd}}{V_{insat}}$=x=1.28). Therefore, the higher order terms in $ln(x) = \left[\left(\frac{x-1}{x}\right) + \frac{1}{2}\left(\frac{x-1}{x}\right)^2 + \ldots\right]$ (valid for $\left(x > \frac{1}{2}\right)$) can be neglected. Using (5.45), (5.50) and (5.54), we obtain the expression for $V_{out}(T_R)$ as:

$$V_{out}(T_R) = V_{dd} - \left(\frac{A_{11} T_R}{C_l + C_p}\right) \tag{5.56}$$

Where, $A_{11}$is the constant independent of W_n:

$$A_{11} = \frac{\beta_{s,\,M1}}{\beta_{l,\,M2}} \frac{(V_{dd} - V_{th1})}{V_{dd}} \left(\frac{V_{in,sat}}{V_{dd}} - 1\right) \tag{5.57}$$

Using the above discussed approach and (5.43), (5.45), (5.50) and (5.54), we get an expression for $t_{TCP10\%}$ as:

$$t_{TCP10\%} = K_1 C_l + K_2 T_R + K_3 \tag{5.58}$$

Where,

$$K_1 = \frac{(V_{dd} - V_{dsat})}{I_{ON}} + R_{eq} \tag{5.59}$$

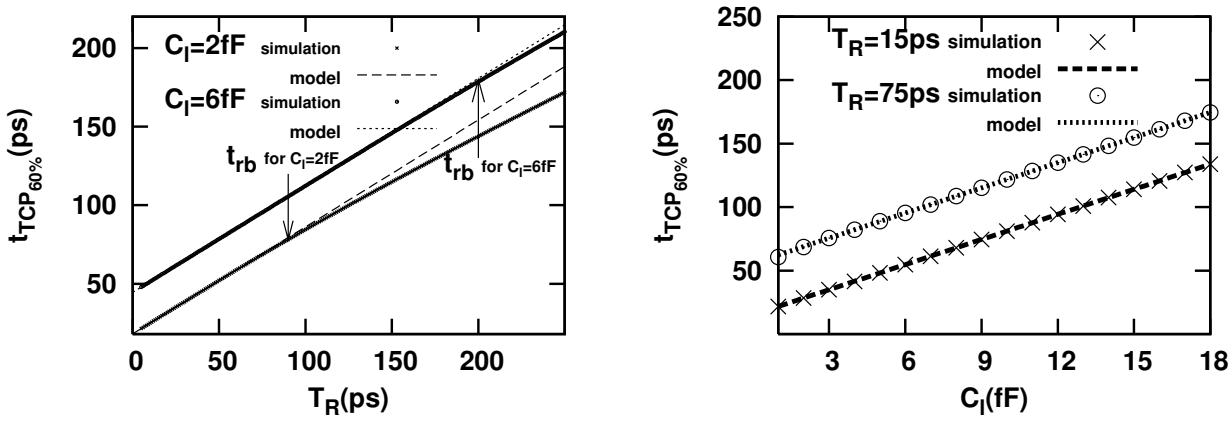
$$K_2 = -\frac{A_{11}}{I_{ON}} \tag{5.60}$$

$$K_3 = \left[\frac{(V_{dd} - V_{dsat})}{I_{ON}} + R_{eq}\right] C_p \tag{5.61}$$

The behavior of model coefficients of (5.58) remain same as of (5.7) with the cell size. As explained earlier, (5.58) is valid if $V_{out}(T_R) \geq V_{dsat}$, this imposes the following constraint on the region of validity:

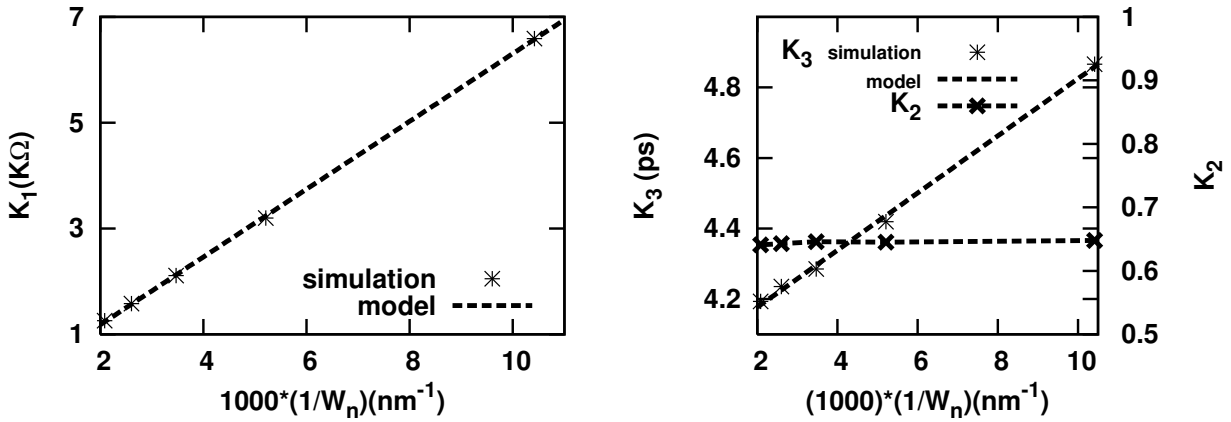
$$\triangle Q(T_R) = (A_{11} T_R) \leq (C_l + C_p)(V_{dd} - V_{dsat}) \tag{5.62}$$

Where, $\triangle Q(T_R)$ is the output discharge from 0 to T_R and V_{dd} is the power supply voltage. C_p is NAND gate's parasitic capacitance (due to gate-drain over-lap, drain-bulk



(a) (b)

Figure 5.14: Case 2: Variation of $t_{TCP60\%}$ with respect to T_R and C_l values.



(a) (b)

Figure 5.15: Case 2: Variation of K_1, K_2 and K_3 with cell size (W_n).

junction capacitance etc.) present at the out node, which is a linear function of W_n. The maximum value of T_R which satisfies (5.62), will be represented as t_{rb}.

For all the $t_{TCP}s$ which fall under Region I, (5.58) remain valid. The model of (5.58) in Region I has thus been verified using HSPICE simulated data. Using curve fitting (as shown in Fig. 5.14) on the simulated values of $t_{TCP60\%}$, we extracted the coefficients K_1, K_2, K_3 of (5.58). We verify the coefficients (K_1, K_2 and K_3) behavior with W_n (shown in Fig. 5.15). In this subsection, we observe that our approach in Region I remains valid with variation in cell size.

5.6.2 Derivation of the model in Region II

In this subsection, we derive the relationship of $t_{TCP}s$ with T_R, C_l and cell size for Region II. Region II contains all the $TCPs$ having values $V_{TCP} > V_{out}(T_R)$, as we depict in Fig. 5.12b. In this derivation, we consider that the discharging nMOS transistor (M_1) always operates in velocity saturation (since $V_{ds}(M_1) > V_{dsat}$) and M_2 operates in saturation

and linear region depending on the intermediate node voltage (V_X), when $V_{out}(T_R) \geq V_{dsat}$ constraint is followed. We consider that the nMOS transistor M_2 remains in saturation region till $t = t_{xsat}$, beyond this M_2 operates in linear region from $t > t_{xsat}$ to t_{TCP}. As a representative, we show the simulation results of $t_{TCP_{90\%}}$. This is true because $t_{TCP_{90\%}} < T_R$ in the whole T_R, C_l space used in characterization LUTs. To find the $t_{TCP90\%}$ which lies in Region II, we proceed as follows:

First, we integrate the saturation current through M_1 from $t = t_{th}$ to t_{TCP}. We equate the sum of these integrals to $(C_l + C_p)(V_{dd}{-}V_{TCP})$ to obtain the expression for t_{TCP}. Following the procedure as explained above, we write an expression for the output voltage discharge as:

$$\int_{t_{th}}^{t_{TCP}} I_{M1} dt = (C_l + C_p)(V_{dd}{-}0.9V_{dd}) \tag{5.63}$$

Taking L.H.S. and solving it :

$$\int_{t_{th}}^{t_{TCP}} I_{M1} dt = \beta_{s-M1} \int_{t_{th}}^{t_{TCP}} (V_{dd} - V_X(t) - V_{th1})\, dt \tag{5.64}$$

$$= \beta_{s-M1} \int_{t_{th}}^{t_{TCP}} (V_{dd} - V_{th1})\, dt - \beta_{s-M1} \left[\int_{t_{th}}^{t_{xsat}} V_X(t)\, dt + \int_{t_{xsat}}^{t_{TCP}} V_X(t)\, dt \right] \tag{5.65}$$

The intermediate node voltage expression remains same as given in (5.50), when lower transistor operates in saturation region (since, $t_{xsat} < t_{TCP}$). From $t = t_{xsat}$ to t_{TCP}, we find the expression for V_X as:

$$\int_{t_{xsat}}^{t_{TCP}} V_X(t)\, dt = V_1 T_R \tag{5.66}$$

Where,

$$V_1 = \left[\frac{\beta_{s,M1}}{\beta_{l,M2}} \frac{(V_{dd} - V_{th1})}{V_{dd}} ln \left[\frac{t_{TCP}}{t_{xsat}} \right] \right] \tag{5.67}$$

The value of $\frac{t_{TCP}}{t_{xsat}}$ in 32nm PTM technology used by us is 1.33 ($\frac{t_{TCP}}{t_{xsat}}{=}x{=}1.33$). Therefore, the higher order terms in $ln(x) = \left[\left(\frac{x-1}{x}\right) + \frac{1}{2}\left(\frac{x-1}{x}\right)^2 + \ldots \right]$ (valid for $\left(x > \frac{1}{2}\right)$) can be neglected. Using (5.65), (5.50) and (5.66), we find the expression for $t_{TCP90\%}$ as follows:

$$t_{TCP90\%} = A_1 T_R + A_2 + \sqrt{(A_3 T_R + A_2^2)} \tag{5.68}$$

Where,

$$A_1 = \left[\frac{1}{2\, V_{dd}} \left(\frac{V_{th1}}{2} + \frac{V_{in,sat}}{V_{dd}} \frac{\beta_{s,M1}}{\beta_{l,M2}} \right) \right] \tag{5.69}$$

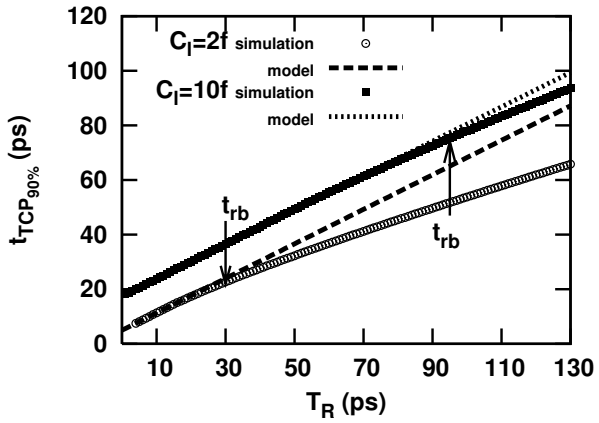
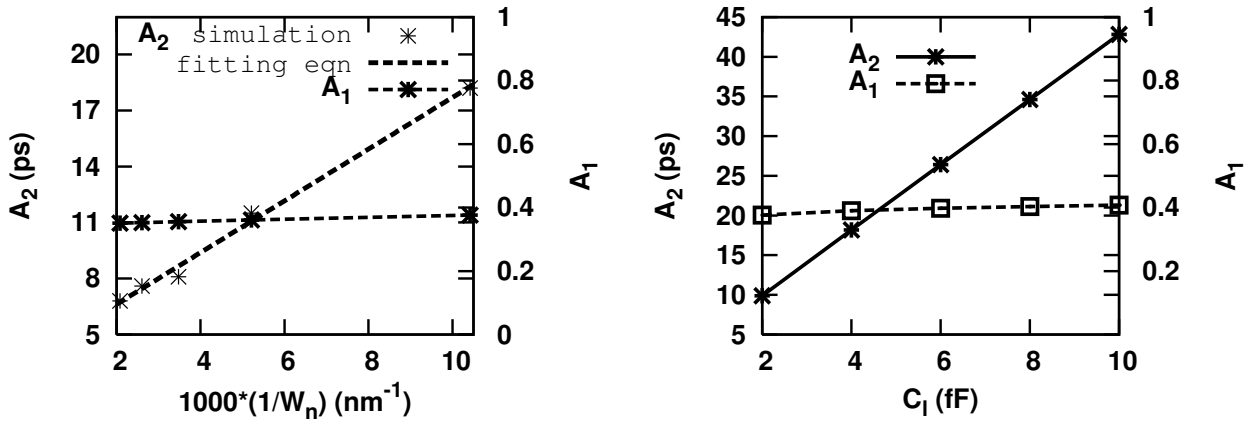


Figure 5.16: Case 2: Variation of $t_{TCP90\%}$ with respect to T_R.



(a) (b)

Figure 5.17: Case 2: Variation of A_1 and A_2 with cell size (W_n) and load capacitance (C_l).

$$A_2 = \frac{0.05\, V_{dd}(C_l + C_p)}{\beta_{s,\,M1}(V_{dd} - V_{th1})} \tag{5.70}$$

$$A_3 = \left(\frac{0.05\,(C_l + C_p)}{\beta_{s,\,M1}(V_{dd} - V_{th1})}\right)\left[\frac{1}{V_{dd}}\left(\frac{V_{th1}}{2} + \frac{V_{in,sat}}{V_{dd}}\frac{\beta_{s,\,M1}}{\beta_{l,\,M2}}\right)\right] \tag{5.71}$$

For all the $t_{TCP}s$ which fall under Region II, (5.68) remains valid. The model of (5.68) in Region II has thus been verified using HSPICE simulated data. Using curve fitting (as shown in Fig. 5.16) on the simulated values of $t_{TCP90\%}$, we extracted the coefficients A_1, A_2, A_3 of (5.68).

The following observations have been made from the derivation of (5.68):

- Observation 4: A_1 is independent of $1/W_n$
- Observation 5: A_2 and A_3^2 are linear function of $1/W_n$

The model of (5.68) remain valid, if $V_{out}(T_R) \geq V_{dsat}$, this imposes the following constraint on the region of validity:

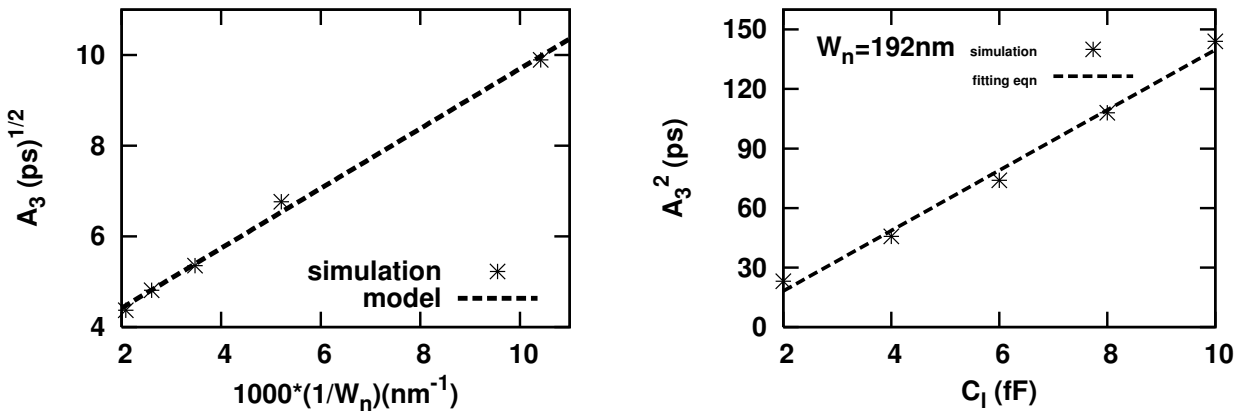


Figure 5.18: Case 2: Variation of A_3 with cell size (W_n) and load capacitance (C_l).

$$\triangle Q(T_R) = (A_{11}T_R) > (C_l + C_p)(V_{dd} - V_{TCP}) \tag{5.72}$$

Please note that the Region II Model can be used to find $t_{TCP}s$ only if $T_R > t_{rb}$, but if $T_R \leq t_{rb}$ then Region I Model can be used depending on the C_l , T_R values used in characterization of LUT's. Therefore, it saves the maximum number of HSPICE simulations in ECSM (or CCSM) characterization.

In Fig. 5.17-5.18, we verify the coefficient's behavior with W_n and C_l.

5.6.3 Efficient ECSM Characterization

In this subsection, we use the models derived in Subsection 5.6.1 and 5.6.2 to save the number of HSPICE simulations in ECSM (or CCSM) characterization of 2-input NAND gate standard cell. Using (5.58) and (5.68), within the region of validity, we get the values of all TCPs directly directly from our model, thus, saving upon the characterization effort. On the other hand, the t_{TCP} values which are out of validity bound in Region I and II, would be obtained from HSPICE simulations. For a 2-input NAND standard cell, we first extract the values of K_1, K_2 and K_3 for $t_{TCP}s$ in Region 1 using 7 HSPICE simulations and A_1, A_2, A_3 using 5 HSPICE simulations. We then calculate the values of $t_{TCP60\%}$ (entries shown by numeric values in Table 5.4) for C_l , T_R values lying within the region of validity of (5.58). For 2-input NAND gate standard cells of other sizes, all TCPs (within the region of validity of (5.58) and (5.68)) are obtained using Observation 4−6. In Table 5.4, t_{TCP} values shown by 'HSPICE', it corresponds to the data points (*i.e.* $t_{TCP60\%}$ values) of C_l , T_R values which are out of region of validity of (5.58). In Table 5.5, t_{TCP} values are shown which are fully generated through HSPICE simulations. Table 5.6 shows the percentage error in t_{TCP} values, generated using our approach (described above) with respect to the t_{TCP} values generated using conventional approach (*i.e.* fully HSPICE generated) for $TCP_{60\%}$ of 2-input CMOS NAND gate.

We observe that the minimum percentage saving in HSPICE simulations using our

Table 5.4: Case 2: LUT of $TCP_{60\%}$ for 2-input CMOS NAND gate using our model.

$C_l(fF)$	$T_R(ps)$						
	2.20	4.84	10.66	23.46	51.62	113.60	250.00
1.51	16.32	18.12	22.07	30.77	49.92	HSPICE	HSPICE
2.28	21.41	23.20	27.16	35.86	55.01	HSPICE	HSPICE
3.45	29.10	30.90	34.85	43.55	62.70	104.85	HSPICE
5.22	40.72	42.52	46.47	55.18	74.33	116.47	HSPICE
7.88	58.29	60.08	64.04	72.74	91.89	134.04	226.79
11.91	84.83	86.63	90.58	99.29	118.44	160.58	253.34
18.00	124.95	126.74	130.70	139.40	158.55	200.70	293.45

Note: Entries shown by 'HSPICE' can be obtained by HSPICE simulation whereas entries shown by numeric values are obtained using our model

Table 5.5: Case 2: ECSM LUT of $TCP_{60\%}$ for 2-input CMOS NAND gate obtained using HSPICE simulations.

$C_l(fF)$	$T_R(ps)$						
	2.20	4.84	10.66	23.46	51.62	113.60	250.00
1.51	16.34	18.19	22.18	30.88	49.86	88.59	165.32
2.28	21.44	23.29	27.29	35.99	55.05	95.27	175.34
3.45	29.14	30.99	34.99	43.69	62.78	104.15	187.97
5.22	40.76	42.62	46.62	55.33	74.43	116.27	203.93
7.88	58.33	60.19	64.19	72.89	92.01	133.93	224.61
11.91	84.87	86.73	90.73	99.44	118.56	160.54	252.51
18.00	124.98	126.84	130.8	139.55	158.68	200.69	292.87

Table 5.6: Case 2: Percentage error in proposed model's LUT with respect to fully HSPICE generated ECSM LUT of $TCP_{60\%}$ for 2-input CMOS NAND gate.

$C_l(fF)$	$T_R(ps)$						
	2.20	4.84	10.66	23.46	51.62	113.60	250.00
1.51	0.12	0.38	0.50	0.36	0.12	×	×
2.28	0.14	0.39	0.48	0.36	0.07	×	×
3.45	0.14	0.29	0.40	0.32	0.13	0.67	×
5.22	0.10	0.23	0.32	0.27	0.13	0.17	×
7.88	0.07	0.18	0.23	0.21	0.13	0.08	0.97
11.91	0.05	0.12	0.17	0.15	0.10	0.02	0.33
18.00	0.02	0.08	0.11	0.11	0.08	0.00	0.20

Note: Entries shown by '×' correspond to the values we obtain using HSPICE simulations in Table 5.4 (not through our models)

method of generating LUTs (as explained above) is 67.35% and 75.51% (for 7×7 matrix size) for $TCP_{60\%}$ and $TCP_{90\%}$, respectively. We find that the values of $t_{TCP}s$ in our LUTs differ by a maximum of 0.97% (for 7×7 matrix size) from those in fully HSPICE generated conventional LUTs. Therefore, we observe that standard cell characterization can be done with a significantly lesser number of HSPICE simulations (approximately 67% reduction in HSPICE simulations). For the $t_{TCP} < T_R$, both the Region I and II models are used.

5.7 Case 2: Variation aware TCP models

In this section, we consider layout dependent effects due to mechanical stress in deriving the model coefficients. We also derive the relationships of the model coefficients with the temperature (T) and supply voltage (V_{dd}). We have already discussed the motivation towards the analysis of variation aware (considering process induced mechanical stress, supply voltage and temperature variability) TCP models in Chapter 4 of the thesis. Here, we derive and validate the behavior of model coefficients with PVT variations. Later, we use the model and the derived relationships to reduce the number of HSPICE simulations significantly.

5.7.1 TCP models considering stress variability

In this subsection, we derive the change in model coefficients of (5.58) and (5.68) as a function of NF in a stress enabled 45nm CMOS technology. We follow the same procedure, as we did in Section 5.5.1.

5.7.1.1 Impact of stress induced variability in Region I

In this sub-subsection, we derive the relationship of $t_{TCP}s$ with T_R, C_l and NF for Region I and II in stress enabled technologies. In (5.59)-(5.61) and (5.69)-(5.71), μ, v_{sat}, I_{ON}, and V_{th} are now stress dependent values, given as (5.21)-(5.24). We derive the model for $t_{TCP60\%}$ considering the impact of channel stress in Region I. We assume that $(V_{dd} - V_{th})$ is independent of NF (as explained in Subsection 5.5.1). Using the stress dependent parameters in (5.59)-(5.61) (as we told earlier), we obtain the relation between model coefficients and NF as:

$$K_1 = (NF + M_3)\left(\frac{D_1}{NF + D_2}\right) \tag{5.73}$$

Where, D_1 and D_2 are the technology dependent constants, given in Subsection 5.5.1 as:

$$D_1 = \frac{(V_{dd} - V_{dsat})}{((M_1 P_1 P_2 + 1)\, I_{ON})} \tag{5.74}$$

$$D_2 = M_3 + \frac{M_2 P_1 P_2}{(M_1 P_1 P_2 + 1)} \tag{5.75}$$

We observe that (5.73) fits well with HSPICE simulated data as shown in Fig. 5.19a. In

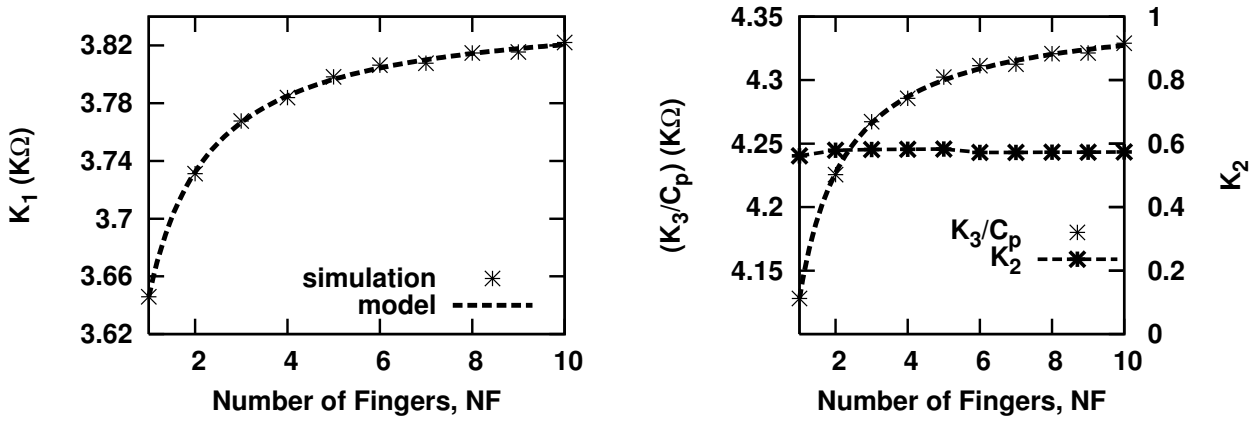


Figure 5.19: Case 2: K_1, K_2 and K_3 as a function of NF (which also represents to channel stress).

these HSPICE simulation MULU0 and DELVT0 vary with NF in accordance with (5.25) (*i.e.* the PTM model incorporates channel stress variability effects). From (5.61), the relation between K_3 and NF is:

$$K_3 = C_p(K_1) = C_p\left[(NF + M_3)\left(\frac{D_1}{NF + D_2}\right)\right] \tag{5.76}$$

We observe that (5.76) fits well with HSPICE simulated data as shown in Fig. 5.19b. Thereafter, we observe that K_2 is independent of NF (because A_{11} and I_{ON} are both proportional to NF) as shown in Fig. 5.19b.

5.7.1.2 Impact of stress induced variability in Region II

In this sub-subsection, we derive a model for $t_{TCP90\%}$, considering the impact of channel stress in Region II. Using stress dependent parameters in (5.69) and (5.25), we obtain that A_1 is independent of $V_{th}(\sigma)$. Where, $V_{th}(\sigma)$ with NF is negligibly small [94]. Therefore, A_1 is independent of $\sigma(NF)$ as shown in Fig. 5.20a. Using stress dependent parameters in (5.70) and (5.25), We find the relationship between A_2 and NF as:

$$A_2 = \left(\frac{S_1(NF + M_3)}{(NF + S_2)}\right) \tag{5.77}$$

Where,

$$S_1 = \frac{1}{(M_1P_1 + 1)}\left(\frac{0.05\,V_{dd}(C_l + C_p)}{\beta_{s,\,M1}(V_{dd} - V_{th1})}\right) \tag{5.78}$$

$$S_2 = \left(M_3 + \frac{P_1M_2}{(M_1P_1 + 1)}\right) \tag{5.79}$$

We observe that (5.77) fits well on our stress aware HSPICE simulation data as shown in Fig. 5.20a. Likewise, from (5.71) and (5.25), We find the relationship between A_3 and NF as:

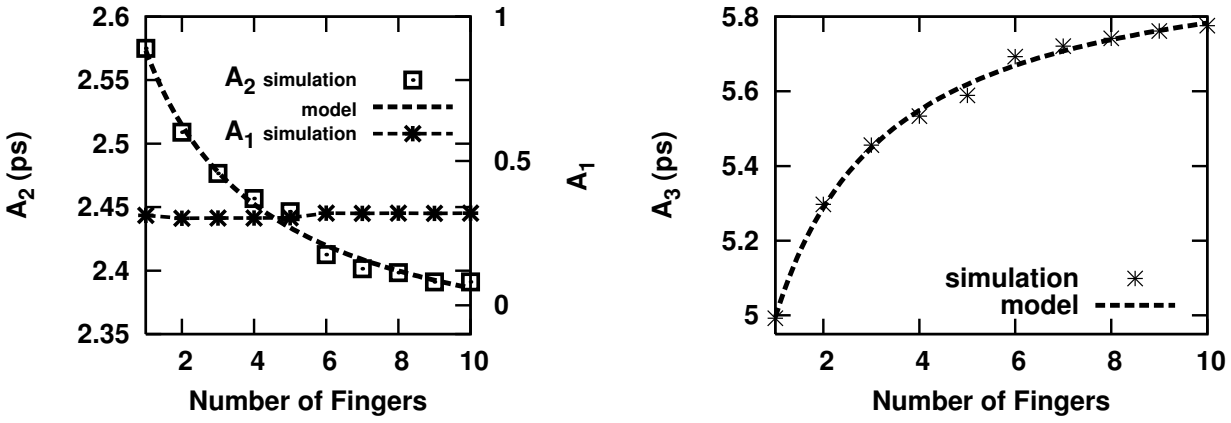


Figure 5.20: Case 2: A_1, A_2 and A_3 as a function of NF.

$$A_3 = \left(\frac{S_3(NF + M_3)}{(NF + S_2)}\right) \tag{5.80}$$

Where,

$$S_3 = \frac{1}{(M_1P_1 + 1)} \left(\frac{0.05\, V_{dd}(C_l + C_p)}{\beta_{s,\,M1}(V_{dd} - V_{th1})}\right) \left[\frac{1}{V_{dd}} \left(\frac{V_{th1}}{2} + \frac{V_{in,sat}}{V_{dd}} \frac{\beta_{s,\,M1}}{\beta_{l,\,M2}}\right)\right] \tag{5.81}$$

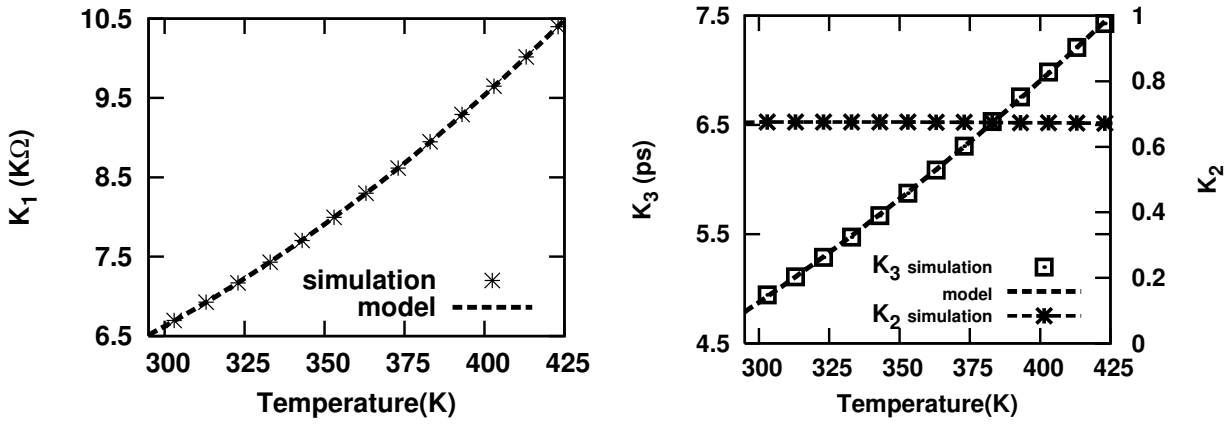
We observe that (5.80) fits well on our stress aware HSPICE simulation data as shown in Fig. 5.20b.

5.7.1.3 Efficient stress aware ECSM characterization

In this sub-subsection, we make use of relationships of model coefficients with stress as a function of NF, derived in previous sub-subsection (as described in Subsection 5.5.1) to reduce the number of HSPICE simulations required for ECSM characterization when the lower input of the series nMOS stack of 2-input NAND gate standard cell switches. The LUT size is 7×7 matrix. Using our model (5.58) for $V_{TCP} < V_{out}(T_R)$, 47 HSPICE simulations and using model (5.58) and (5.68) for $V_{TCP} > V_{out}(T_R)$ only 14 HSPICE simulations, were needed to generate the LUT, whereas a conventional procedure would require 343 HSPICE simulations. Saving in number of HSPICE simulations using Region I model (for $V_{TCP} < V_{out}(T_R)$) is $\backsimeq 86\%$ and using both the models corresponding to Region I and II (for $V_{TCP} > V_{out}(T_R)$) is 96% (compared to conventional approach) .

5.7.2 TCP models considering temperature variability

In this subsection, we derive the temperature variation aware t_{TCP} models, which we later use to reduce the re-characterization effort significantly. In this work, we consider a realistic range of temperature variation due to on-chip heating, from 298K (room temperature) to 423K.



(a) (b)

Figure 5.21: Case 2: Variation of K_1, K_2 and K_3 with temperature.

5.7.2.1 Impact of temperature variability in Region I

In this sub-subsection, we derive the relationships of t_{TCPs} with T_R, C_l and temperature (T) for Region I and II. In (5.59)-(5.61) and (5.69)-(5.71), v_{sat}, μ and V_{th} are now given by (5.34)-(5.36). We derive a model for $t_{TCP60\%}$, considering the impact of temperature variability in Region I. In this derivation, we assume that $(V_{dd} - V_{th})$ is independent of T. This is justified since V_{th} is much smaller than V_{dd}. From (5.59)-(5.61), we obtain :

$$K_1 = \left(R_1\, T^{2.3} + R_2\, T + R_3\right) \tag{5.82}$$

Where R_1, R_2 and R_3 are technology dependent constants obtained after deriving the (5.59) using temperature dependent parameters. We observe that (5.82) fits well on HSPICE simulated data as shown in Fig. 5.21a. Likewise solving (5.61), we find the relation between K_3 and T as:

$$K_3 = C_p\, K_1 = C_p\, \left(R_1\, T^{2.3} + R_2\, T + R_3\right) \tag{5.83}$$

We observed using HSPICE simulations that change in C_p with T varying from 298K to 423K is negligibly small. We observe that (5.83) fits well with HSPICE simulated data as shown in Fig. 5.21b. Thereafter, solving (5.60), we find that K_2 is independent of T as shown in Fig. 5.21b (because A_{11} and I_{ON} both are the function of T).

5.7.2.2 Impact of temperature variability in Region II

We derive a model for $t_{TCP90\%}$, considering the impact of temperature variability in Region II. From (5.69), we observe that A_1 is independent of T. From (5.70), we obtain that A_2 is proportional to T, which we verify through HSPICE simulations in Fig. 5.22a. We obtain the relationship between A_2 and T as:

$$A_2 = (R_4\, T + R_5) \tag{5.84}$$

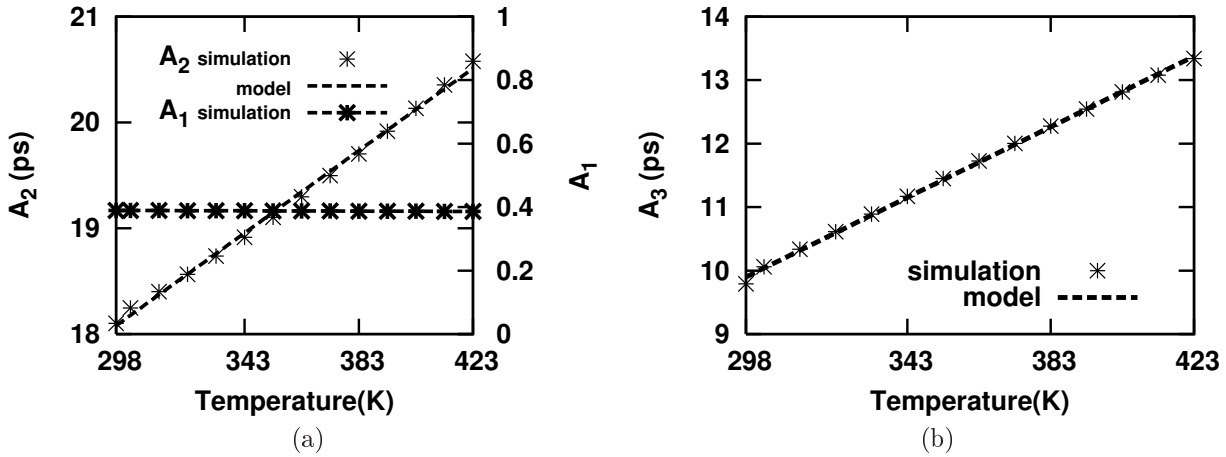


Figure 5.22: Case 2: Variation of A_1, A_2 and A_3 with temperature.

Where, R_4 and R_5 are technology dependent constants obtained after deriving the (5.70). We observe that (5.84) fits well on our HSPICE simulation data with temperature variability as shown in Fig. 5.22a. Likewise, solving (5.71), we find that A_3 increases linearly with T as shown in Fig. 5.22b.

5.7.2.3 Efficient temperature variation aware ECSM characterization

In this sub-subsection, we use the relationships of model coefficients with temperature, to reduce the number of HSPICE simulations required for ECSM characterization, when the lower input of the series nMOS stack of 2-input NAND gate standard cell switches. The LUT size is 7×7 matrix. Using our model (5.58) for $V_{TCP} < V_{out}(T_R)$, 50 HSPICE simulations and using model (5.58) and (5.68) for $V_{TCP} > V_{out}(T_R)$ only 13 HSPICE simulations, were needed to generate the LUT, whereas a conventional procedure would require 343 HSPICE simulations. Saving in number of HSPICE simulations using Region I model (for $V_{TCP} < V_{out}(T_R)$) is $\backsimeq 85\%$ and using both the models corresponding to Region I and II (for $V_{TCP} > V_{out}(T_R)$) is 96% (compared to conventional approach) .

5.7.3 TCP models considering Supply Voltage Variability

In this subsection, we derive a power supply voltage variation aware t_{TCP} models, which we later use to reduce the re-characterization effort significantly. We consider the $\pm 10\%$ change in power supply voltage (V_{dd}) from the its nominal value of $V_{dd} = 0.9V$.

5.7.3.1 Impact of supply voltage variability in Region I

In this sub-subsection, we derive the relationships of $t_{TCP}s$ with T_R, C_l and V_{dd} for Region I. We derive a model for $t_{TCP60\%}$, considering the impact of supply voltage variability in Region I. In (5.59), we observe and later verify that K_1 decreases with an increase in supply voltage as shown in Fig. 5.23a.

We now discuss the variation of K_3 with V_{dd}. From (5.61), we observe that the behavior of K_3 with supply voltage remains same as of K_1. We verify this observation in Fig. 5.23b.

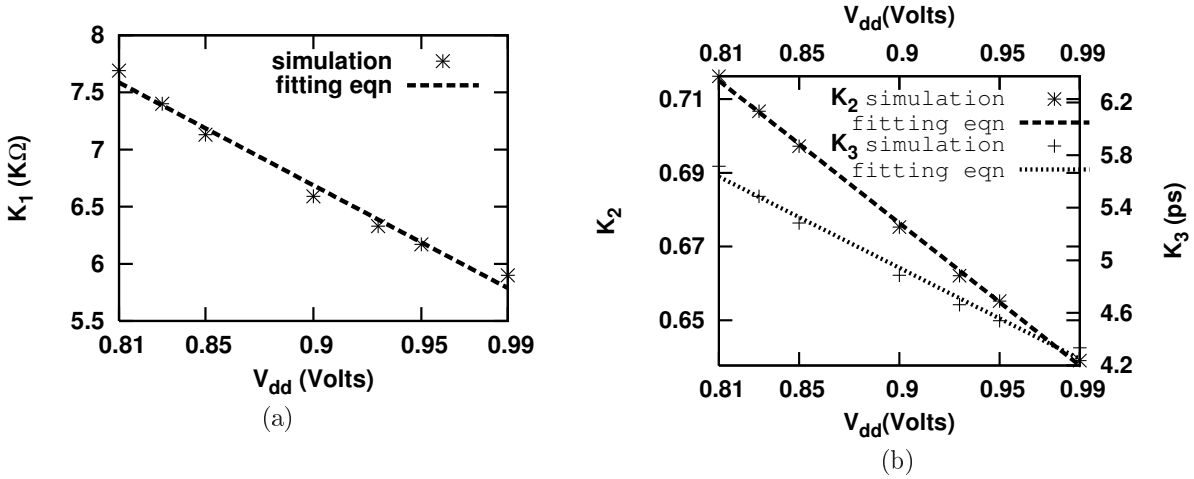


Figure 5.23: Case 2: Variation of K_1, K_2 and K_3 with supply voltage.

In (5.60), we observe that K_2 is inversely proportional to V_{dd}. We verify this observation in Fig. 5.23b.

5.7.3.2 Impact of supply voltage variability in Region II

In this sub-subsection, we derive a model for $t_{TCP90\%}$, considering the impact of supply voltage variability in Region II. In (5.69), A_1 is inversely proportional to supply voltage and it is verified through HSPICE simulations as shown in Fig. 5.24a. In (5.70) and (5.71), we observe that A_2 and A_3 both decreases with supply voltage. We verify this observation in Fig. 5.24b and 5.24c.

5.7.3.3 Efficient supply voltage variation aware ECSM characterization

In this sub-subsection, we use the relationships of model coefficients with supply voltage, to reduce the number of HSPICE simulations required for ECSM characterization, when the lower input of the series nMOS stack of 2-input NAND gate standard cell switches. The LUT size is 7×7 matrix. Using our model (5.58) for $V_{TCP} < V_{out}(T_R)$, 48 HSPICE simulations and using model (5.58) and (5.68) for $V_{TCP} > V_{out}(T_R)$ only 14 HSPICE simulations, were needed to generate the LUT, whereas a conventional procedure would require 343 HSPICE simulations. Saving in number of HSPICE simulations using Region I model (for $V_{TCP} < V_{out}(T_R)$) is $\backsimeq 86\%$ and using both the models corresponding to Region I and II (for $V_{TCP} > V_{out}(T_R)$) is 96% (compared to conventional approach) .

5.8 Summary

In this chapter, we proposed models for t_{TCPs} of output voltage transition for 2-input NAND gate standard cell, for Case 1 and 2. First, we derived the model for the switching of upper nMOS transistor in the series stack of the NAND gate (represented as Case 1), second, we derived the model for the switching of lower nMOS transistor in the series stack of the NAND gate (represented as Case 2). In this work, we considered the impact of the voltage transition at the intermediate node in the series stack of nMOS transistors in the

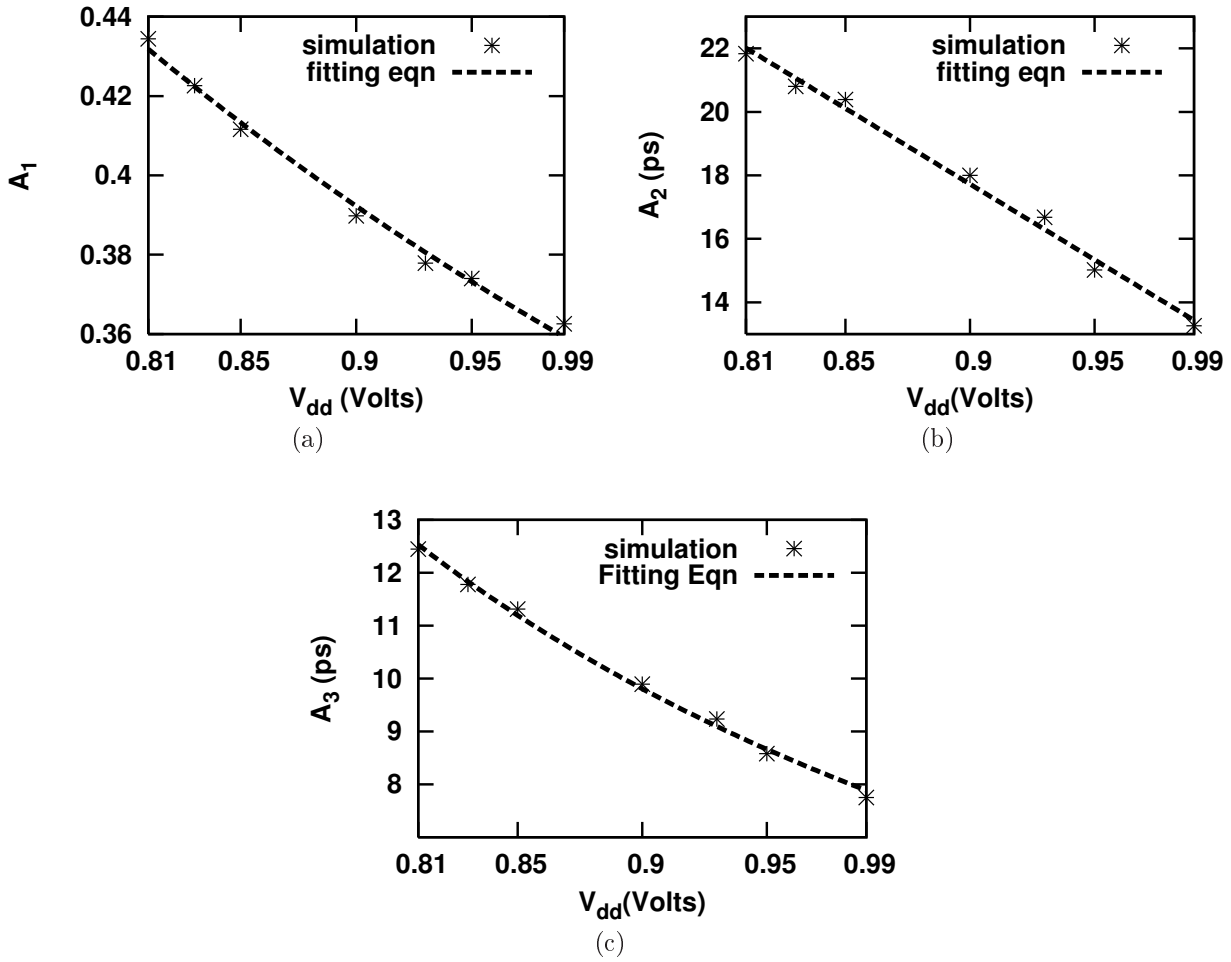


Figure 5.24: Case 2: Variation of A_1, A_2 and A_3 with supply voltage.

NAND gate. The t_{TCP} values are derived in terms of T_R and C_l. We also derived the region of validity of these models in C_l, T_R space. Later, we used these models in reducing the number of HSPICE simulations in ECSM characterization of 2-input NAND gate standard cell by $\approxeq 67\%$. We also derived the relationship of the model coefficients with the NAND gate size, V_{dd}, carrier mobility, threshold voltage and temperature. While considering layout dependent effects due to mechanical stress, we derived the relationship of the model coefficients with stress as a function of NF. We later used these relationships to reduce the number of HSPICE simulations by about 80.18% and 86% in ECSM characterization of a 2-input NAND gate standard cell having a different value of NF for Case 1 and 2, respectively. Further, we included the effect of temperature variation in our t_{TCP} models to reduce the number of HSPICE simulations by $\simeq 79.88\%$ for Case 1 and 85% for Case 2. We also included the effect of supply voltage variation in our t_{TCP} models to reduce the number of HSPICE simulations by $\simeq 79.88\%$ for Case 1 and 86% for Case 2.

We thus proposed the comprehensive models capturing all the timing information regarding the output transition of a 2-input NAND gate for the switching of one of its inputs. These models relate all the $t_{TCP}s$ with circuit/layout level parameters such as T_R, C_l, W_n, NF, T and V_{dd}. We also derived the region of validity of these models with respect to each

of these circuit/layout level parameters. Further, we apply all this work in reducing the number of HSPICE simulations required for ECSM characterization of 2-input NAND gate significantly (about 67-96%).

Chapter 6

Overshoot Timing Model for CMOS Inverter and NAND Gate Standard Cells

6.1 Overview

In this chapter, we propose an analytical model to estimate overshoot time for CMOS inverter and NAND gate standard cells. We separately model the over-shoot time period for switching of each of the inputs of a 2-input NAND gate standard cell. We observe that the duration of the overshoot is significantly (about 40%) larger for NAND gates compared to inverter for a given fan-out for the case when the lower nMOS device in the series stack switches. To model the overshoot time period for this case, we show that considering the impact of the intermediate node of nMOS series stack is critical. Therefore, we first model the voltage transition at this node in our work, which we further use to derive overshoot time period for the lower transistor switching case. We observe and explain that the overshoot timing model of a CMOS inverter remains valid for the case when the upper nMOS device in the series stack of the NAND gate switches. This is because the lower nMOS device in the series stack is ON and only adds a resistance to the source of the upper nMOS device. The model also considers the influence of T_R, C_l and cell size on overshoot time which makes it useful in standard cell characterization. The proposed model is in good agreement with HSPICE simulations with a maximum error of 2.5%.

The chapter is organized as follows. In Section 6.2, we describe our simulation setup. In Section 6.3, we derive the relationship of overshoot time (t_{crit}) with T_R, C_l for CMOS inverter standard cell. We then verify the model with respect to T_R, C_l values and cell size of CMOS inverter. In Section 6.4, we derive the relationship of t_{crit} with T_R, C_l for 2-input NAND gate standard cell. We then verify the model with respect to T_R, C_l values and cell size of NAND gate.

6.2 Simulation Setup

In this Chapter, we use HSPICE simulations at $32nm$ CMOS technology node. In these simulations, we use BSIM 4.0 Predictive Technology device Model (PTM) [87]. We use a CMOS inverter (shown in Fig. 6.1a) and a 2-input NAND gate standard cells (shown in Fig. 6.2a) for the overshoot timing model derivation. The widths of pMOS transistor and those of nMOS transistors are kept in such a way that the output rising and falling transition times are equal for both inverter and NAND gate standard cells as described in [22]. The minimum transistor widths for inverter and NAND gate standard cells are $^{W_P}/_{W_n} = \frac{120nm}{48nm}$ and $^{W_P}/_{W_n} = \frac{120nm}{96nm}$, respectively. The channel lengths are kept at their minimum allowed value for all the transistors (*i.e.* 32nm). Since the $^{W_P}/_{W_n}$ ratio for standard cell of a given type (say, 2-input NAND gate are fixed), the value of W_n also represents the cell size. Throughout this work, we consider the case of rising input transition: The case of falling input transition can be handled in similar manner. Please note that we have not considered

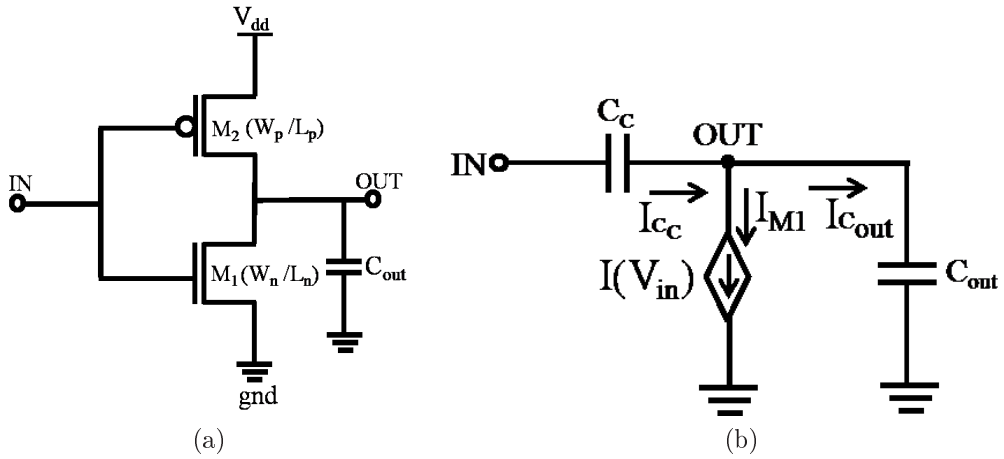


Figure 6.1: (a) CMOS inverter schematic (b) its equivalent circuit during overshoot period.

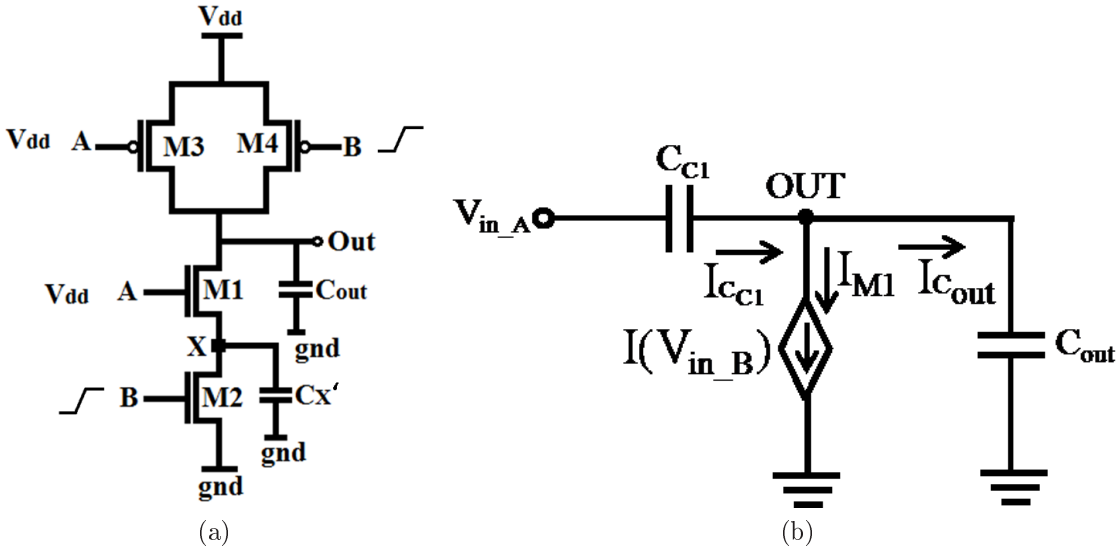


Figure 6.2: (a) 2-input NAND gate Schematic (b) its equivalent circuit during overshoot period.

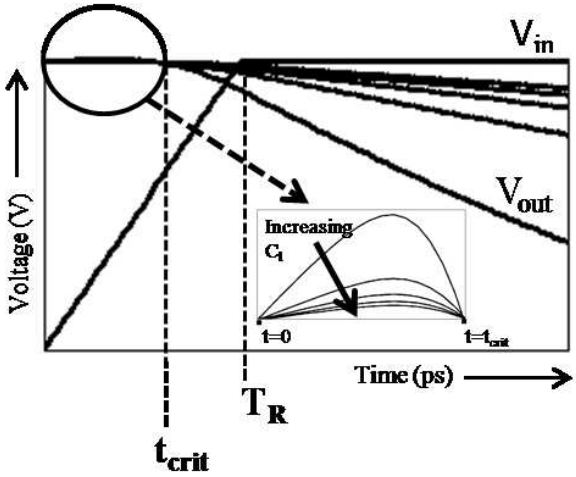


Figure 6.3: I/O waveform of standard cell CMOS inverter.

the inverse narrow width effects in the proposed work. Since, the threshold voltage V_{th} is the parameter considered in the derivation, narrow width effects can be considered in the model.

6.3 Our approach towards CMOS inverter overshoot modeling

To derive the overshoot time for CMOS inverter standard cell (shown in Fig. 6.1) , we use the following assumptions and observations :

- The pMOS transistor (M_2) is operating in deep triode region and negligibly small current is flowing through it (*i.e.* $\approxeq 5\mu A$). Therefore, we ignore this negligible pMOS current in our derivation.
- The time at which the overshoot reaches its peak value V_p, is t_p. We assume that $t_p = \frac{V_{th}}{V_{dd}} T_R$, because the nMOS device (M_1) is in OFF state for $V_{in} < V_{th}$. The increase in output voltage till this time is only because of capacitive coupling between nodes IN and OUT. Therefore, t_p is also independent of C_l and W_n. This assumption is validated by our results that we discuss later in the paper.
- We observe from HSPICE simulations that the overshoot period t_{crit} is independent of C_l (as shown in Fig. 6.3). We explain this as follows:

For a given value of T_R, the charging current through the coupling capacitance C_C during the input transition $t = 0$ to T_R has a value $C_C \frac{V_{dd}}{T_R}$, therefore the total charge that contributes to the overshoot at the output is independent of C_l and T_R. This charge is now discharged through the nMOS such that V_{out} reduces to a value $V_{out} = V_{dd}$. During the

discharge of the overshoot the nMOS operates in saturation region and therefore, its drain current is independent of the value of V_{out} (this is also because, the V_{ds} of nMOS almost doesn't vary and has a value $V_{out} \approxeq V_{dd}$ during the overshoot period).

6.3.1 Model for Overshoot Time t_{crit}

In this subsection, we derive the relationship between t_{crit} and T_R, C_l values. We consider the rising input transition for the derivation of overshoot time (shown in Fig. 6.3). As we explained earlier that t_p is the time after which transistor M_1 starts to conduct. Till t_p, there will be only capacitive coupling at 'OUT' node. To find the t_p, we apply KCL at node 'OUT'. Therefore, we have the expression:

$$I_{C_C} = I_{C_{out}} + I_{M_1} \tag{6.1}$$

I_{C_C} is current flowing through gate-to-drain capacitance of M_1 and M_2, since M_1 is operating either in cut-off (for $t < t_p$) or in saturation for ($t \geq t_p$), it consists of the overlap capacitance component only. Whereas, since M_2 operates in linear region during the overshoot period so its gate-to-drain overlap capacitance consists of both channel and overlap components. $I_{C_{out}}$ is current flowing through the load capacitance and parasitic capacitance (other than C_C) at the output node. I_{M_1} is the saturation current flowing through M_1 which is non-zero for $t \geq t_p$. Therefore, we have an equivalent circuit as shown in Fig. 6.1b. At $t = t_p$ solving KCL at node 'OUT' , we rewrite (6.1) as:

$$C_C \frac{dV_{in}}{dt} = I_{M_1} \text{where } V_{in} = V_{dd}\frac{t}{T_R} \text{for } 0 \leq t \leq T_R \tag{6.2}$$

And, I_{M_1} is given by alpha power law model as [9]:

$$I_{M1} = I_{ON} = \nu_{sat} W_n P_s (V_{gs} - V_{th})^m \tag{6.3}$$

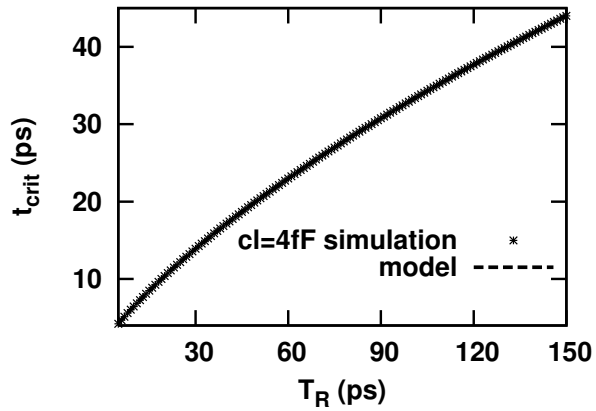
We assume $\beta_s = \nu_{sat} W_n P_s$, where ν_{sat} is saturation velocity, W_n is width of nMOS device, P_s is technology dependent parameter and exponent m is velocity saturation index. In our case, we have used $m = 1$ and it is verified through HSPICE simulations. Solving (6.2), we get an expression for t_p as follows:

$$t_p = \frac{C_C}{\beta_s} + \frac{V_{th}}{V_{dd}} T_R \tag{6.4}$$

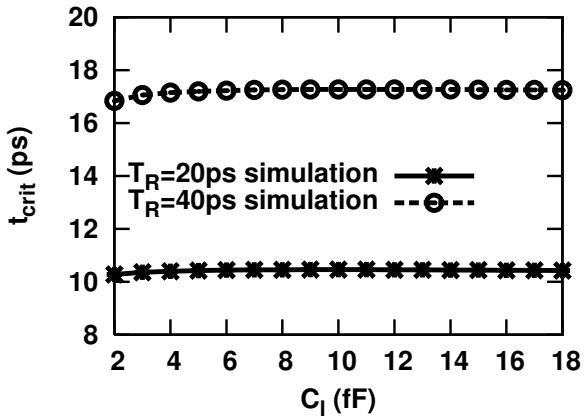
To obtain the t_{crit}, we first integrate (6.1) from 0 to t_p and find an expression as follows:

$$C_C \frac{V_{dd}}{T_R}(t_p - 0) = (C_{out} + C_C)(V_p - V_{dd}) + \int_0^{t_p} I_{M1} dt \tag{6.5}$$

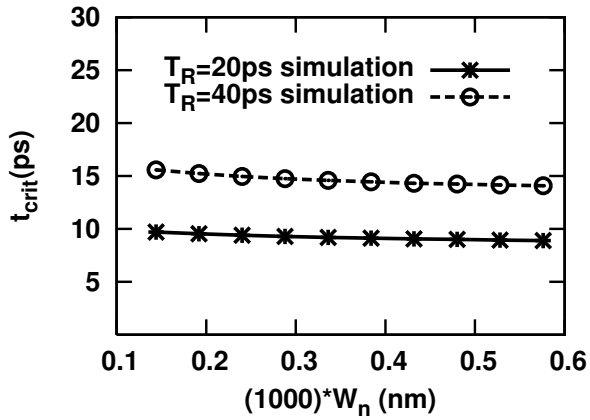
Where, $\int_0^{t_p} I_{M1} dt \simeq 0$ as we explained earlier that from 0 to t_p negligibly small current



(a) Variation of t_{crit} with T_R



(b) Variation of t_{crit} with C_l



(c) Variation of t_{crit} with W_n

Figure 6.4: Variation of t_{crit} with T_R, C_l and W_n.

is flowing. Now, we integrate (6.1) from t_p to t_{crit} and find an expression as follows:

$$C_C \frac{V_{dd}}{T_R}(t_{crit} - t_p) = -(C_{out} + C_C)(V_p - V_{dd}) + \int_{t_p}^{t_{crit}} I_{M1} dt \tag{6.6}$$

Using (6.3), (6.4), (6.5) and (6.6), we get final expression for t_{crit} as follows:

$$t_{crit} = \left(\frac{C_C}{2\beta_s} + \frac{V_{th}}{V_{dd}} T_R\right) + \sqrt{\left(\frac{5C_C^2}{4\beta_s^2} + \frac{C_C}{\beta_s}\frac{V_{th}}{V_{dd}} T_R\right)} \tag{6.7}$$

We observe that the t_{crit} is independent of C_l and W_n from the derivation of (6.7). The later is because both C_C and β_s are proportional to W_n.

The model of (6.7) has been verified using HSPICE simulation data as shown in Fig. 6.4. As observed in (6.7), the coefficients of the model are the constants for a given CMOS technology and can be extracted from HSPICE simulations/measurements. Once these coefficients are extracted, (6.7) can be used to obtain t_{crit} for any value of T_R and inverter cell size. The (6.7) can therefore be used in standard cell library characterization to obtain overshoot time period.

6.4 NAND gate overshoot modeling

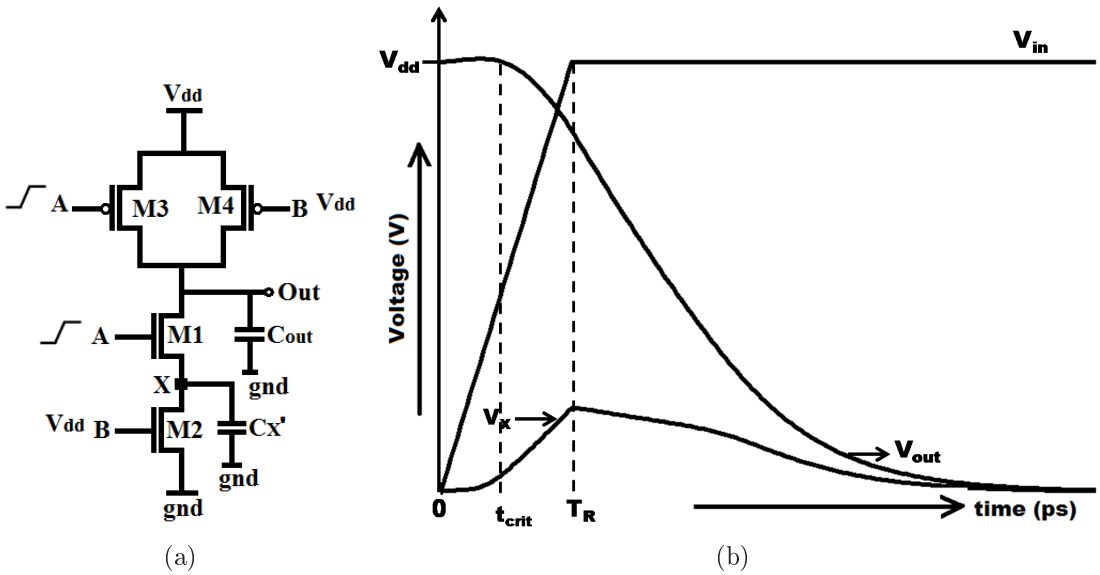
In this section, we derive the relationships of t_{crit} with T_R , C_l values and W_n for a 2-input NAND gate standard cell. We derive the models for the case where input is rising, a similar approach can be used for falling input transition case too. Since typically single input switching would be more frequent compared to multiple input switching, we consider single input switching for our derivation. There are two possibilities for 2-input NAND gate single input switching.

- Case 1: When$V_{in-B} = V_{dd}$ and$V_{in-A} = V_{dd}\frac{t}{T_R}$ (for $0 \leq t \leq T_R$ *i.e.* when upper nMOS device (M_1) in the series stack switches) (refer to Fig. 6.5a)
- Case 2: When $V_{in-A} = V_{dd}$ and $V_{in-B} = V_{dd}\frac{t}{T_R}$ (for $0 \leq t \leq T_R$ *i.e.* when lower nMOS device (M_2) in the series stack switches) (refer to Fig. 6.7a)

We derive the model of t_{crit} for the above 2 cases of 2-input NAND gate in the following subsections.

6.4.1 NAND gate overshoot modeling for Case 1

To derive the model of t_{crit} for 2-input NAND gate (for Case 1) standard cell, shown in Fig. 6.5, we use the following assumptions and observations :



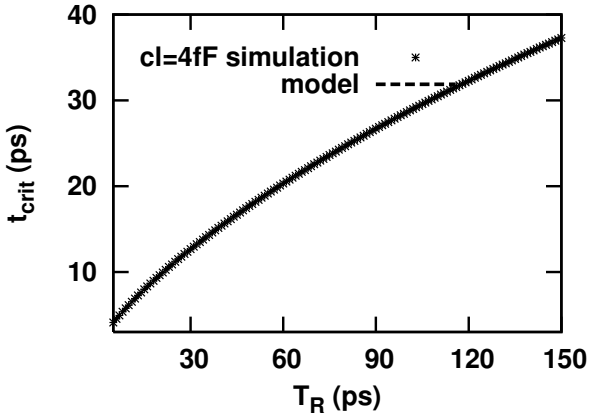
(a) (b)

Figure 6.5: Case 1: (a) NAND gate schematic (b) its I/O waveform.

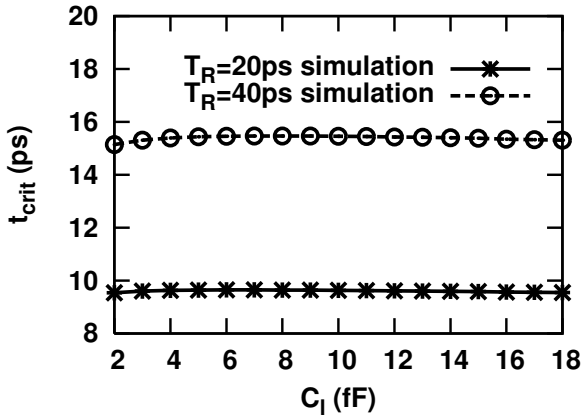
- The pMOS transistor (M_4) is in OFF state whereas M_3 is in deep triode region and negligibly small current is flowing through it. Therefore, we ignore this negligible pMOS current in our derivation.

- The nMOS transistor M_1 will be ON at $\frac{V_{th1}}{V_{dd}}T_R$ whereas M_2 is in deep triode region during overshoot period.

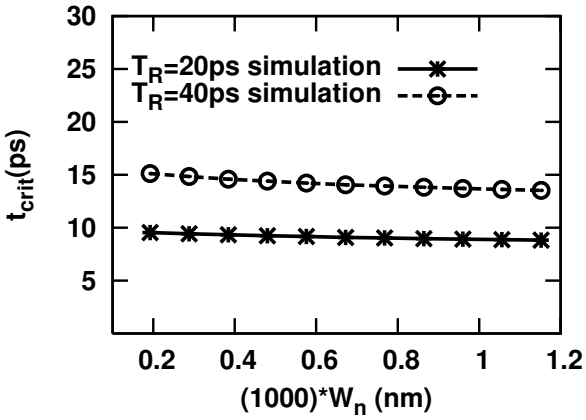
- Expression for t_{crit} remains same as for CMOS inverter (given in (6.7)), because in Case 1 behavior of NAND gate remains same as of CMOS inverter except that the lower nMOS transistor in the series stack adds some resistance to the source of upper nMOS transistor.



(a) Variation of t_{crit} with T_R



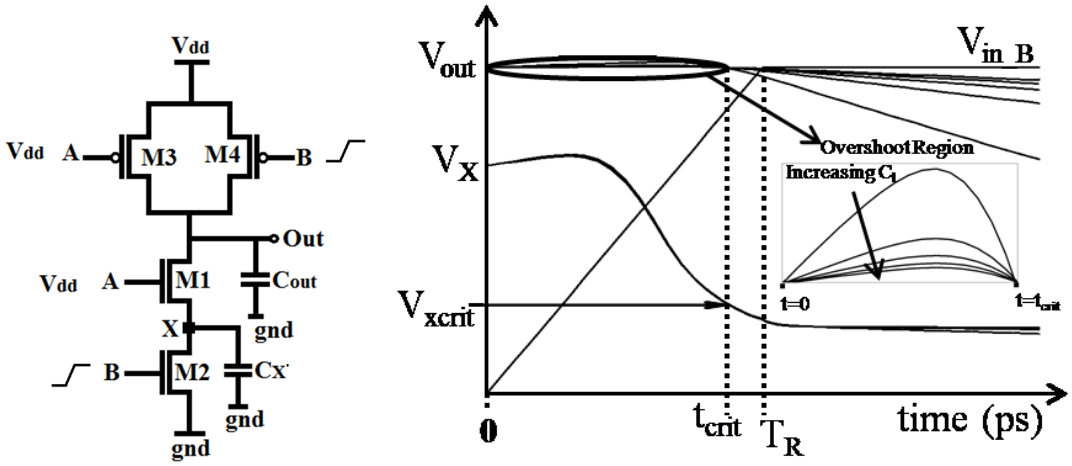
(b) Variation of t_{crit} with C_l



(c) Variation of t_{crit} with W_n

Figure 6.6: Case 1: Variation of t_{crit} with T_R, C_l and W_n.

The proposed model of (6.7) for t_{crit} has been verified using HSPICE simulated data for Case 1. Figure 6.6 shows the verification of t_{crit} model with respect to T_R, C_l values and cell size for Case 1 of NAND gate.



(a) (b)

Figure 6.7: Case 2: NAND gate schematic and its I/O waveform.

6.4.2 NAND gate overshoot modeling for Case 2

To derive the model of t_{crit} for 2-input NAND gate (for Case 2), shown in Fig. 6.5, we use the following assumptions and observations :

- The pMOS transistor (M_3) is in OFF state whereas M_4 is in linear region and negligibly small current is flowing through it. Therefore, we ignore that negligible pMOS current throughout overshoot modeling.
- Threshold voltage of M_1 is always somewhat larger than M_2 , because of potential difference between source and bulk of M_1 (*i.e.* $V_{sb(M1)} \neq 0$).
- The nMOS transistor M_2 will be ON at $t = \frac{V_{th}}{V_{dd}} T_R$ whereas M_1 starts to conduct later for a technology dependent value of $V_X = V_{Xa}$. V_{Xa} is the value of the source voltage of an nMOS of the the given CMOS technology having its $V_g = V_d = V_{dd}$.
- In this work, $V_X = V_{Xa}$ at time $t = t_{Xa}$. Till t_{Xa}, current through M_1 is almost zero. And output voltage will be increased by capacitive coupling only. When $t \geq t_{Xa}$, current starts to flows through M_1 and then equivalent circuit of NAND can be represented as shown in Fig. 6.2b.
- Overshoot time (denoted as t_{crit} in our case) is independent of C_l as shown in Fig.6.8. We explain this as follows:

For a given value of T_R, the charging current through the coupling capacitance C_{C1} during the input transition $t = 0$ to T_R has a value $C_{C1}\frac{V_{dd}}{T_R}$, therefore the total charge that contributes to the overshoot at the output is independent of C_l and T_R. This charge is now discharged through the nMOS such that V_{out} reduces to a value $V_{out} = V_{dd}$. During the discharge of the overshoot the nMOS operates in saturation region and therefore its drain

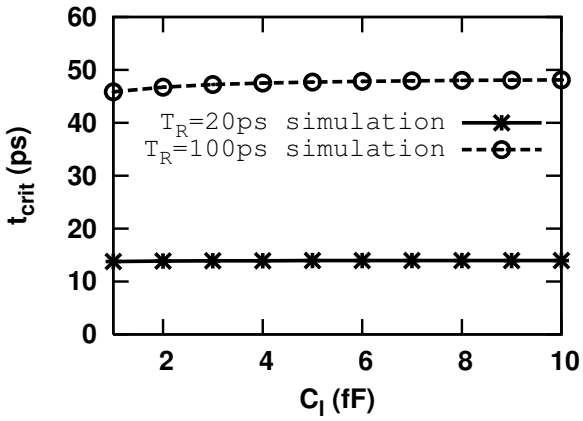


Figure 6.8: Case 2: t_{crit} is independent of C_l.

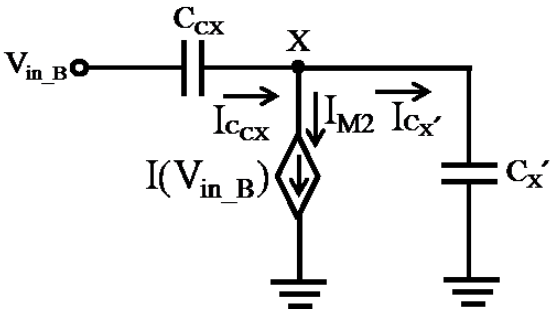


Figure 6.9: Case 2: NAND gate equivalent circuit at node 'X'.

current is independent of the value of V_{out} (this is also because of the value V_{ds} of nMOS almost doesn't vary and has a value $V_{out} \cong V_{dd}$ during the overshoot period). To derive the t_{crit} model, first we need to model the t_{Xa}.

6.4.2.1 Proposed Model for t_{Xa}

In this sub-subsection, we derive the relationship between t_{Xa} and T_R, C_l values. To find the t_{Xa}, we apply KCL at 'X' node for $t < t_{Xa}$ (refer to Fig. 6.9):

$$I_{C_{CX}} + I_{M2} = I_{C_X} \tag{6.8}$$

$I_{C_{CX}}$ is current flowing through gate-to-drain coupling capacitance of M_2, it consists of the gate-to-drain overlap capacitance and may also consist a part of the gate-to-channel capacitance (as we elaborate later in this paper).We use the symbol C'_X for the parasitic capacitance (of M_1 and M_2) between node 'X' and 'gnd'. $I_{C'_X}$ is current flowing out of this parasitic capacitance. And I_{M2} is saturation current flowing through M_2 given in (6.3).

Now, we rearrange (6.8) as:

$$C_{CX}\frac{dV_{in-B}}{dt} + I_{M2}(t) = (C'_X + C_{CX})\frac{dV_X}{dt} \tag{6.9}$$

Where,

$$V_{in-B} = V_{dd}\left(\frac{t}{T_R}\right) \text{ for } 0 \leq t \leq T_R \tag{6.10}$$

In this derivation, we are taking $(C'_X + C_{CX}) = C_X$, to simplify the expression. To find the t_{Xa}, integrate (6.9) from $t = t_{th}$ to t_{Xa}. Where t_{th} is time at which lower nMOS transistor M_2 starts to conduct (*i.e.* $t_{th} = \frac{V_{th}}{V_{dd}}T_R$, when V_{in-B}=V_{th}). Before t_{th}, node 'X' gets charged only through capacitive coupling as M_1 and M_2 are both OFF. At t =t_{th}, value of V_X is given as :

$$V_{Xth} - V_{X0} = \left(\frac{C_{CX}}{C_X}\right)V_{th} \tag{6.11}$$

Where, V_{Xth} is voltage at node 'X' when V_{in-B} reaches threshold voltage of M_2 and V_{X0} is the voltage at node 'X' when $t = 0$ (since, $V_{X0} \approx V_{dd} - V_{th}$). Using (6.3), (6.10), (6.11) and solve for t_{Xa} by integrating (6.9), we have the expression as:

$$\left(\frac{C_{CX}}{C_X}\frac{V_{dd}}{T_R} + \frac{\beta_{M2}V_{th}}{C_X}\right)(t_{Xa}-t_{th}) - \left(\frac{\beta_{M2}V_{dd}}{2\,C_X T_R}\right)(t_{Xa}^2-t_{th}^2) + \left(\frac{C_{CX}}{C_X}V_{th} + \triangle V_{Xa}\right) = 0 \tag{6.12}$$

Where,

$$(V_{Xa} - V_{Xth}) = \{(V_{Xa} - V_{X0}) - (V_{Xth} - V_{X0})\} = \left(\triangle V_{Xa} - \frac{C_{CX}}{C_X}V_{th}\right) \tag{6.13}$$

Where, $\triangle V_{Xa} = -(V_{Xa} - V_{X0})$ is a positive value since $V_{Xa} < V_{X0}$. After solving (6.12), we get the final equation of t_{Xa} as:

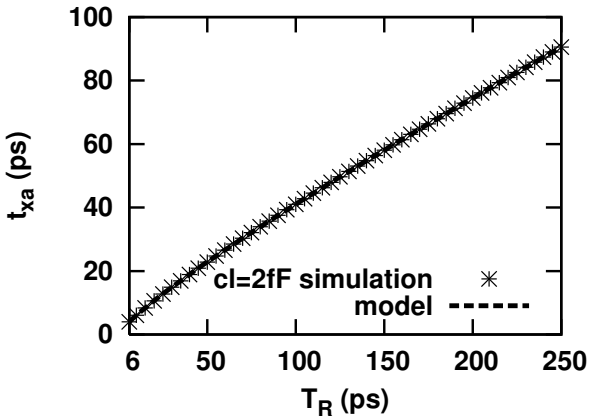
$$t_{Xa} = \frac{C_{CX}}{\beta_{M2}} + \frac{V_{th}}{V_{dd}}T_R + \sqrt{\left[\frac{C_{CX}}{\beta_{M2}}\left(\frac{C_{CX}}{\beta_{M2}} + \left(\frac{6\,V_{th}}{V_{dd}} + \frac{2\,\triangle V_{Xa}}{V_{dd}}\right)T_R\right)\right]} \tag{6.14}$$

It can be observed from (6.14) that t_{Xa} is independent of C_l and W_n. The independence of t_{Xa} with W_n is because C_{CX} and β_{M2} are both proportional to W_n.

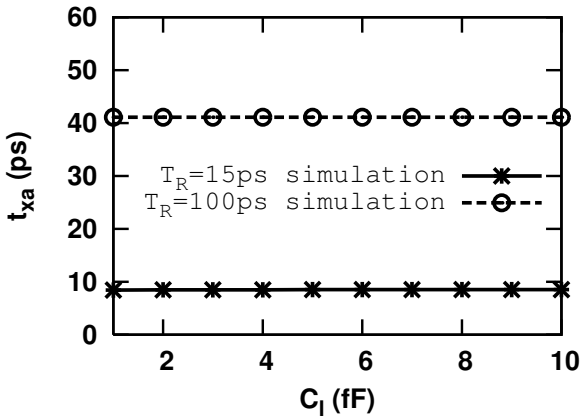
The proposed model of (6.14) has been verified using HSPICE simulated data for t_{Xa}. Figure 6.10, shows the validation of t_{Xa} model with respect to T_R, C_l values and cell size.

6.4.2.2 Proposed Model for t_{crit}

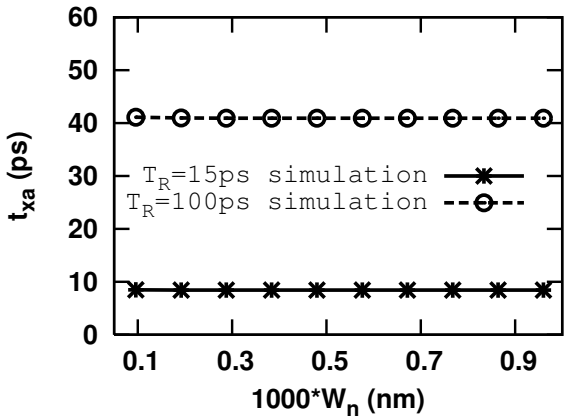
In this sub-subsection, we derive the relationship between t_{crit} and T_R, C_l values for 2-input NAND gate standard cell. Till t_{Xa}, the transistor M_1 is OFF and node 'OUT' charges through capacitive coupling (through the C_{gd} of pMOS transistor (M_4) in Fig. 6.7). For $t > t_{xa}$, the node 'OUT' also discharges through the drain current of M_1 (I_{M1}). To find I_{M1}, we apply KCL at node 'X'. We integrate I_{M1} from t_{Xa} to t_{crit} to account for the overshoot. We explain this approach to derive t_{crit} in detail in the following paragraph. We model the t_{crit} for the large values of T_R (from 15ps to 250ps).



(a) Variation of t_{Xa} with T_R



(b) Variation of t_{Xa} with C_l



(c) Variation of t_{crit} with W_n

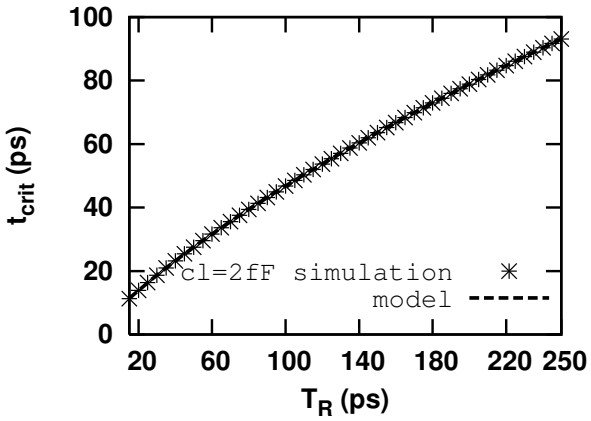
Figure 6.10: Case 2: Variation of t_{Xa} with T_R, C_l and W_n.

As the minimum $FO1$ delay of CMOS inverter and NAND gate is 11ps and 14ps respectively. In general, it is very rare to get the real input signal having transition time equal to or less than the $FO1$ delay of CMOS inverter. Therefore, we consider the range of T_R (from 15ps to 250ps) which is very practical in standard cell characterization. Within this range of T_R, $V_{x,crit} \geq V_{dsat}$ and, at $t = t_{crit}$, both nMOS transistors operate in velocity saturation region. Here, $V_{x,crit}$ is the voltage at node 'X' when $t = t_{crit}$ and V_{dsat} is the saturation drain source voltage of M_2 transistor. We assume that the value of V_{dsat} is weakly dependent on the values of V_{ds}, as in (given in [9]). For our PTM CMOS technology, we extract the value of $V_{dsat} = 0.28V$ from $I_d - V_{ds}$ characteristics (as done in [9]).

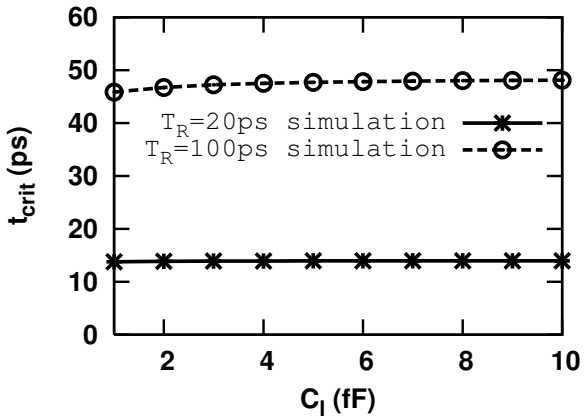
During overshoot period, NAND gate equivalent circuit consists of coupling capacitance (C_{C1}), total output capacitance (C_{out}) and a dependent current source (I_{M1}) having a non-zero current from t_{Xa} to t_{crit} (as shown in Fig. 6.2b). To obtain the value of I_{M1}, We apply KCL at node 'X':

$$I_{M2} = I_{C_{CX}} + I_{C'_X} + I_{M1} \tag{6.15}$$

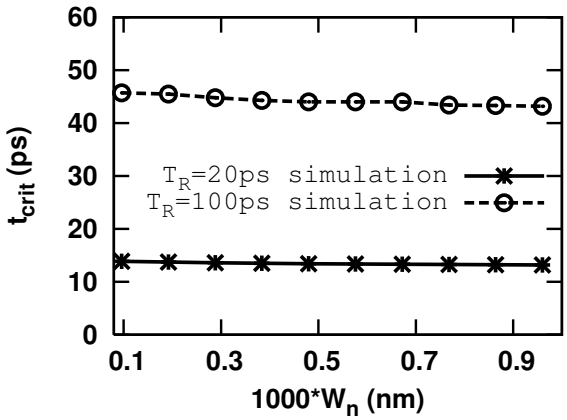
We can write (6.15) as:



(a) Variation of t_{crit} with T_R



(b) Variation of t_{crit} with C_l



(c) Variation of t_{crit} with W_n

Figure 6.11: Case 2: Variation of t_{crit} with T_R, C_l and W_n.

$$\beta_{M2}\left(\frac{V_{dd}}{T_R}t - V_{th}\right) = C_{CX}\frac{dV_{in-B}}{dt} - C'_X\frac{dV_X}{dt} + \beta_{M1}\left(V_{dd} - V_X - V_{th1}\right) \tag{6.16}$$

At $t = t_{crit}$, we can write (6.16) as:

$$\beta_{M2}\left(\frac{V_{dd}}{T_R}t_{crit} - V_{th}\right) = C_{CX}\frac{V_{dd}}{T_R} + C'_X\frac{dV_{Xcrit}}{dt} + \beta_{M1}\left(V_{dd} - V_{Xcrit} - V_{th1}\right) \tag{6.17}$$

At $t = t_{Xa}$, we can write (6.16) as:

$$\beta_{M2}\left(\frac{V_{dd}}{T_R}t_{Xa} - V_{th}\right) = C_{CX}\frac{V_{dd}}{T_R} + C'_X\frac{dV_{Xa}}{dt} + \beta_{M1}\left(V_{dd} - V_{Xa} - V_{th1}\right) \tag{6.18}$$

Subtracting (6.18) from (6.17) to get the value of t_{crit}. As M_1 and M_2 transistors are of same size and both are operating in saturation region, therefore $\beta_{M1} = \beta_{M2}$. Now, we have the expression of t_{crit} as:

$$t_{crit} = \left(\frac{V_{Xa} - V_{Xcrit}}{V_{dd}}\right)T_R + t_{Xa} \tag{6.19}$$

In this derivation, we ignore the current $I_{C'_X}$ (refer to (6.15)), as its value is very small compared to the ON current of M_1. It can be observed from (6.19) that t_{crit} is independent of C_l and W_n.

The proposed model of (6.19) has been verified using HSPICE simulated data for t_{crit}. Figure 6.11 shows the validation of t_{crit} model with respect to T_R, C_l values and cell size.

6.5 Summary

In this chapter, we proposed t_{crit} model for CMOS inverter and 2-input NAND gate. We then verified the relationships of proposed model with T_R, C_l and cell size (in terms of W_n). The model equation as well as simulation results shows that t_{crit} is independent of C_l and W_n. The model covers a wide range of T_R values (from 15ps to 150ps). The proposed models are in good agreement with HSPICE simulation with maximum error of 2.5%.

Chapter 7

Conclusion and Future Scope

In this chapter, a summary of the research work carried along with the conclusions. The future scope in this area of study is also pointed out.

The problem addressed in this thesis is briefly defined here. Standard cell library characterization in nanometer range CMOS technologies, consumes huge time because of numerous cells of different sizes, layout dependent effects, temperature and supply voltage variation and frequent device model updates. To address this issue, delay/timing models with clearly defined regions of validity in transition time (T_R), load capacitance (C_l) space and which also consider cell size, layout parameters, temperature and supply voltage variation are needed.

7.1 Conclusions

The major contributions and conclusions from this work are summarized below.

In the first stage of the work, we proposed an analytical model for the Timing values of Threshold Crossing Point ($t_{TCP}s$) of output voltage transition as a function of T_R and C_l for a standard cell CMOS inverter. Subsequently, the region of validity of the model in T_R, C_l space used in characterization Lookup Tables (LUTs) is derived. We developed the relationships of the model coefficients with the cell size. Further, the impact of technology scaling on these model coefficients is investigated. The results show that the proposed model is in good agreement with HSPICE simulations with a maximum error of 2.5%. Later, we use this model to reduce the number of HSPICE simulations in ECSM characterization by nearly half. We then obtained the relationships of the model coefficients with process induced mechanical stress as a function of Number of Fingers (NF), temperature and power supply voltage (V_{dd}) variability. We later expand this model for different values of NF and include the impact of temperature and supply voltage variations, to reduce the number of SPICE simulations. We observed that the model helps to reduce the number of HSPICE simulations by about 50% in ECSM characterization of standard cell CMOS inverter.

In the next phase of the work, we proposed the t_{TCP} model for 2-input NAND gate standard cell, for Case 1 and 2, which we now discuss. First, we developed the model for the case (Case 1) when the upper nMOS transistor in the series stack of the NAND gate

switches. After this, we developed the model for the case (Case 2) when the lower nMOS transistor in the series stack of the NAND gate switches. In this work, we considered the impact of the voltage transition at the intermediate node in the series stack of nMOS transistors in the NAND gate. For this, we considered the input to intermediate node capacitive coupling effect, parasitic capacitances at the intermediate node and the regions of operation of the two nMOS devices placed in series stack. The t_{TCP} models are derived as a function of T_R and C_l. We also derived the region of validity of these models in C_l, T_R space. Later, we used these models in reducing the number of HSPICE simulations in ECSM characterization of 2-input NAND gate standard cell by $\approxeq 67\%$. We then developed the relationship of the model coefficients with the cell size, power supply voltage (V_{dd}) and temperature. While considering layout dependent effects due to mechanical stress, we developed the relationship of the model coefficients with stress as a function of NF. We later used these relationships to reduce the number of HSPICE simulations by about 80.18% and 86% in ECSM characterization of a 2-input NAND gate standard cell having a different value of NF for Case 1 and 2, respectively. Further, we included the effect of temperature variation in our t_{TCP} models to reduce the number of HSPICE simulations by $\simeq 79.88\%$ for Case 1 and 85% for Case 2. We also included the effect of supply voltage variation in our t_{TCP} models to reduce the number of HSPICE simulations by $\simeq 79.88\%$ for Case 1 and 86% for Case 2.

Further, we developed an analytical model to estimate overshoot time considering the influence of T_R, C_l and cell size, for CMOS inverter and NAND gate standard cells. In nanometer range technologies, the parasitic capacitances would increase thereby increasing the importance on the overshoot time period. However, in 32nm CMOS technology node we considered, the overshoot time period is important only for the TCP coming immediately after the overshoot. We separately model the overshoot time for switching of each of the inputs of a 2-input NAND gate standard cell. In this regard, we first model the voltage transition at this node in our work, which we later use to derive overshoot time period for the lower transistor's switching case. We observed that the overshoot timing model of a CMOS inverter remains valid for the case when the upper nMOS device in the series stack of the NAND gate switches. This is because the lower nMOS device in the series stack is ON and only adds a resistance to the source of the upper nMOS device. We observed that the proposed models are in good agreement with HSPICE simulations with a maximum error of 2.5%.

Hence, we conclude that the standard cell characterization effort can be reduced significantly (nearly 50% for CMOS inverter and 67% for 2-input NAND gate) using our models while maintaining the accuracy close to HSPICE. The validation of these model's coefficients with the cell size, process parameters, V_{dd} and temperature variations, minimizes the re-characterization effort significantly in standard cell libraries.

7.2 Scope for Future Research

In this section, we concisely discuss some future/prospective directions for further research in the same area:

1. The work can be generalized by developing standard cell library and hence, timing models for other cells like 2-stage buffer, AND and OR gate can be derived. To reduce the re-characterization effort in standard cell library, the relationships of model coefficients with the cell size, process parameters, V_{dd} and temperature variations can be derived.

2. Timing model for inverter followed by transmission gate can be derived on the lines of our NAND gate timing model.

3. For accurate timing model, an overshoot timing model can be used in ECSM characterization of standard cell libraries. This will ensure that even the $\simeq 5\%$ error (for NAND gate) seen in t_{TCP} estimation for the TCP closest to the overshoot can also be reduced significantly.

4. This work can be used in improving the efficiency of standard cell characterization EDA tools significantly.

LIST OF PUBLICATIONS

Based upon the research work carried out, following papers are published or communicated for publication :

1. B. Kaur, N. Alam, S. K. Manhas, and B. Anand, "Efficient ECSM Characterization Considering Voltage, Temperature and Mechanical Stress Variability", *IEEE Transactions on Circuits and Systems-I*, (in Press).

2. B. Kaur, S. Vundavalli, S. K. Manhas, S. Dasgupta, and B. Anand, "An Accurate Current Source Model for CMOS Based Combinational Logic Cell", in *13th ISQED'12*, pp. 561-565, 2012.

3. B. Kaur, S. Miryala, S. K. Manhas, and B. Anand, "An Efficient Method for ECSM Characterization of CMOS Inverter in Nanometer Range Technologies", in *14th ISQED'13*, pp. 665-669, 2013.

4. S. Miryala, B. Kaur, B. Anand, and S. K. Manhas, "Efficient Nanoscale VLSI Standard Cell Library Characterization Using a Novel Delay Model", in *12th ISQED'11*, pp. 458-463, 2011.

5. B. Kaur, A. Sharma, N. Alam, S. K. Manhas, and B. Anand, "A Variation Aware Timing Model for a 2-Input NAND Gate and Its Use in Sub-65nm CMOS Standard Cell Characterization-Part I", *IEEE Transactions on Circuits and Systems-I*, (Communicated).

6. B. Kaur, A. Sharma, N. Alam, S. K. Manhas, and B. Anand, "A Variation Aware Timing Model for a 2-Input NAND Gate and Its Use in Sub-65nm CMOS Standard Cell Characterization-Part II", *IEEE Transactions on Circuits and Systems-I*, (Communicated).

Bibliography

[1] I. Synopsys, "Library compiler: Modeling timing and power," Chapter-II, 2003.

[2] S. Gummalla, A. R. Subramaniam, Y. Cao, and C. Chakrabarti, "An analytical approach to efficient circuit variability analysis in scaled CMOS design," *13th Int. Symp. Qual. Electron. Des.*, pp. 641–647, 2012.

[3] H. Fatemi, S. Nazarian, and M. Pedram, "Statistical logic cell delay analysis using a current-based model," in *Proc. Des. Autom. Conf.*, 2006, pp. 253–256.

[4] S. Gupta and S. S. Sapatnekar, "Current source modeling in the presence of body bias," *Proc. IEEE 15 th Asia and South Pacific Design Automation Conference (ASP-DAC 2010), Taipei*, pp. 199–204, Jan 2010.

[5] Silvaco, "Introduction to Cell Characterization," May 2008. [Online]. Available: http://www.silvaco.com

[6] Y. B. Kim, "Challenges for nanoscale MOSFETs and emerging nanoelectronics," *Trans. Electr. Electron. Mater*, vol. 3, pp. 93–105, 2010.

[7] K. Peng, Y. Huang, P. Mallick, W. T. Cheng, and M. Tehranipoor, "Full-Circuit SPICE Simulation Based Validation of Dynamic Delay Estimation," *Proc. IEEE Eur. Test Symp.*, pp. 101–106, 2010.

[8] S. Louis, L. Luciano, and M. Grant, *EDA for IC Implementation, Circuit Design, and Process Technology.* Taylor & Francis Group, 2006.

[9] T. Sakurai and A. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," *IEEE J. Solid-State Circuits*, vol. 25, no. 2, pp. 584–594, 1990.

[10] Synopsys, "Scalable polynomial delay and power model," October 2002.

[11] G. Mekhtarian. (2005, November) Composite current source (ccs) modeling technology backgrounder. Retrieved from http://www.synopsys.com.

[12] Cadence, "Encounter Library Characterizer," Dec. 2012. [Online]. Available: http://www.cadence.com

[13] S. Nazarian, H. Fatemi, and M. Pedram, "Accurate timing and noise analysis of combinational and sequential logic cells using current source modeling," *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 19, pp. 92–103, 2011.

[14] J. F. Croix and D.F.Wong, "Blade and Razor: Cell and interconnect delay analysis using current-based models," in *Proc. Des. Autom. Conf.*, 2003, pp. 386–389.

[15] I. Keller, K. Tseng, and N. Verghese, "A robust cell-level crosstalk delay change analysis," *IEEE/ACM Int. Conf. Comput. Des. Dig. Tech. Pap. ICCAD*, pp. 147–154, 2004.

[16] P. Li and E. Acar, "A waveform independent gate model for accurate timing analysis," in *Proc. IEEE Int. Conf. Comput. Des. VLSI Comput. Process.*, 2005, pp. 363–365.

[17] C. Knoth, V. B. Kleeberger, P. Nordholz, and U. Schlichtmann, "Fast and waveform independent characterization of current source models," *IEEE Int. Behav. Model. Simul. Work.*, pp. 90–95, 2009.

[18] J. K. Ousterhout, "Switch level delay models for digital mos vlsi," *IEEE 21st Design Automation Conference*, pp. 542–548, 1984.

[19] N. Hedenstierna and K. Jeppson, "CMOS Circuit Speed and Buffer Optimization," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 6, 1987.

[20] W. Shockley, "A unipolar field effect transistor," *IRE*, vol. 40, pp. 1365–1376, November 1952.

[21] S. Dutta, S. S. Shetti, and S. L. Lushy, "A comprehensive delay model for CMOS inverters," *IEEE J. Solid-State Circuits*, vol. 30, no. 8, pp. 864–871, 1995.

[22] I. E. Sutherland, B. F. Sproull, and D. L. Harris, *Logical Effort: Designing Fast CMOS Circuits*. Morgan Kaufmann Publishers, 1998.

[23] K. Kanda, K. Nose, H. Kawaguchi, and T. Sakurai, "Design impact of positive temperature dependence on drain current in sub-1-v CMOS VLSIs," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 10, pp. 1559–1564, 2001.

[24] K. Okada, K. Yamaoka, and H. Onodera, "A statistical gate delay model for intra-chip and inter-chip variabilities," in *Proc. ASP-DAC Asia South Pacific Des. Autom. Conf. 2003.*, 2003.

[25] S. Sengupta, K. Saurabh, and P. E. Allen, "A process, voltage, and temperature compensated CMOS constant current reference," *IEEE Int. Symp. Circuits and Syst. (ISCAS)*, vol. 1, pp. 325–328, 2004.

[26] S. P. S. Pant and D. Blaauw, "Static timing analysis considering power supply variations," *ICCAD-2005. IEEE/ACM Int. Conf. Comput. Des. 2005.*, 2005.

[27] O. Unsal, J. Tschanz, K. Bowman, V. De, X. Vera, A. Gonzalez, and O. Ergin, "Impact of Parameter Variations on Circuits and Microarchitecture," *IEEE Micro*, vol. 26, 2006.

[28] R. Kumar and V. Kursun, "Reversed Temperature-Dependent Propagation Delay Characteristics in Nanometer CMOS Circuits," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 53, 2006.

[29] J. C. Ku and Y. Ismail, "On the scaling of temperature-dependent effects," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 26, no. 10, pp. 1882–1888, 2007.

[30] S. Basu, P. Thakore, and R. Vemuri, "Process Variation Tolerant Standard Cell Library Development Using Reduced Dimension Statistical Modeling and Optimization Techniques," *8th Int. Symp. Qual. Electron. Des.*, 2007.

[31] K. Agarwal and S. Nassif, "Characterizing Process Variation in Nanometer CMOS," *2007 44th ACM/IEEE Des. Autom. Conf.*, 2007.

[32] M. Abu Rahma and M. Anis, "A Statistical Design-Oriented Delay Variation Model Accounting for Within-Die Variations," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 27, 2008.

[33] S. Aftabjahani and L. Milor, "Compact Variation-Aware Standard Cell Models for Timing Analysis - Complexity and Accuracy Analysis," *9th Int. Symp. Qual. Electron. Des. (isqed 2008)*, 2008.

[34] C. Agostino, P. Flatresse, E. Beigne, and M. Belleville, "Statistical leakage modeling in CMOS logic gates considering process variations," *IEEE International Conference on Integrated Circuit Design and Technology and Tutorial (ICICDT)*, pp. 301–304, 2008.

[35] B. Das, V. Janakiraman, B. Amrutur, H. Jamadagni, and N. Arvind, "Voltage and Temperature Scalable Gate Delay and Slew Models Including Intra-Gate Variations," *21st Int. Conf. VLSI Des. (VLSID 2008)*, 2008.

[36] J. Viraraghavan, B. P. Das, and B. Amrutur, "Voltage and Temperature Scalable Standard Cell Leakage Models Based on Stacks for Statistical Leakage Characterization," *21st Int. Conf. VLSI Des. (VLSID 2008)*, 2008.

[37] S. A. Aftabjahani and L. Milor, "Timing analysis with compact variation-aware standard cell models," *Integr. VLSI J.*, vol. 42, pp. 312–320, 2009.

[38] S. Aftabjahani and L. Milor, "Fast Variation-Aware Statistical Dynamic Timing Analysis," *WRI World Congr. Comput. Sci. Inf. Eng.*, vol. 3, pp. 488–492, 2009.

[39] E. Y. Chin, C. S. Levy, and A. R. Neureuther, "Variability aware timing models at the standard cell level," in *Proc. SPIE - Int. Soc. Opt. Eng.*, vol. 7641, 2010.

[40] V. Janakiraman, A. Bharadwaj, and V. Visvanathan, "Voltage and Temperature Aware Statistical Leakage Analysis Framework Using Artificial Neural Networks," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 29, 2010.

[41] S. Wu, S. Chakravarty, and L. C. Wang, "Impact of multiple input switching on delay test under process variation," *VLSI Test Symp. (VTS), 2010 28th*, 2010.

[42] M. Chen, Y. Yi, W. Zhao, and D. Ma, "Variation-aware deep nanometer gate performance modeling: An analytical approach," in *Proc. 2011 Int. Symp. VLSI Des. Autom. Test*, 2011, pp. 1–4.

[43] B. P. Das, B. Amrutur, H. S. Jamadagni, N. V. Arvind, and V. Visvanathan, "Voltage and Temperature-Aware SSTA Using Neural Network Delay Model," *IEEE Trans. Semicond. Manuf.*, vol. 24, pp. 533–544, 2011.

[44] J. Freijedo, J. Semiao, A. J. Rodriguez, F. Vargas, I. C. Teixeira, and J. P. Teixeira, "Modeling the effect of process variations on the timing response of nanometer digital circuits," *12th Latin American Test Workshop (LATW)*, pp. 1–5, 27-30 March 2011.

[45] M. Dave, M. Jain, M. Shojaei Baghini, and D. K. Sharma, "A Variation Tolerant Current-Mode Signaling Scheme for On-chip Interconnects," *IEEE Trans. VLSI Syst.*, vol. 21, no. 2, pp. 342–353, Feb 2013.

[46] L. Ding, Z. Huang, M. Jiang, A. Kurokawa, and Y. Inoue, "Modeling the overshooting effect of multi-input gate in nanometer technologies," *Proc. IEEE 54th Int. Midwest Symp. Circuits Syst.*, pp. 1–4, Aug 2011.

[47] L. Ding, J. Wang, Z. Huang, A. Kurokawa, and Y. Inoue, "An analytical model of the overshooting effect for multiple-input gates in nanometer technologies," *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1712–1715, May 2013.

[48] M. Mehri, M. H. M. Kouhani, N. Masoumi, and R. Sarvari, "New approach to VLSI buffer modeling, considering overshooting effect," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 21, no. 8, pp. 1568–1572, Aug 2013.

[49] Z. Huang, A. Kurokawa, M. Hashimoto, T. Sato, M. Jiang, and Y. Inoue, "Modeling the Overshooting Effect for CMOS Inverter Delay Analysis in Nanometer Technologies," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 29, 2010.

[50] L. M. Brocco, S. P. McCormick, and J. Allen, "Macromodeling CMOS circuits for timing simulation," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 7, no. 12, pp. 1237–1249, 1988.

[51] D. Auvergne, N. Azemard, D. Deschacht, and M. Robert, "Input waveform slope effects in CMOS delays," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, pp. 1588–1590, 1990.

[52] B. Yang and M. Cai, "Advanced strain engineering for state-of-the-art nanoscale CMOS technology," *Sci. China Inf. Sci.*, vol. 54, no. 5, pp. 946–958, 2011.

[53] J. R. Burns, "Switching response of complementary symmetry mos transistor logic circuits," *RCA*, vol. 25, pp. 627–661, Dec 1964.

[54] J. K. Ousterhout, "A switch-level timing verifier for digital MOS VLSI," *IEEE Trans. Computer-Aided Design*, vol. 4, no. 3, pp. 336–349, 1985.

[55] K. Jeppson, "Modeling the influence of the transistor gain ratio and the input-to-output coupling capacitance on the CMOS inverter delay," *IEEE J. Solid-State Circuits*, vol. 29, 1994.

[56] A. Nabavi Lishi and N. Rumin, "Inverter models of CMOS gates for supply current and delay evaluation," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 13, 1994.

[57] S. Embabi and R. Damodaran, "Delay models for CMOS, BiCMOS and binmos circuits and their applications for timing simulations," *IEEE Trans. Computer-Aided Design*, vol. 13, no. 9, pp. 1132–1142, 1994.

[58] T. Choi, W. Cho, and D. Kim, "A simple CMOS delay model for wide applications," in *Proc. APCCAS'96 - Asia Pacific Conf. Circuits Syst.*, 1996.

[59] J. B. Sulistyo and S. H. Dong, "A new characterization method for delay and power dissipation of standard library cells," *VLSI Design*, vol. 15 (3), pp. 667–678, 2002.

[60] D. Patel, "Characterization and modeling system for accurate delay prediction of asic design," *IEEE Proc. of Custom Integrated Circuits Conference*, pp. 9.5.1–9.5.6, 1990.

[61] J. Y. Jou, J. Y. Lin, and W. Z. Shen, "A power modeling and characterization method for the cmos standard cell library," *Digest of Technical Papers, IEEE International conference on computer aided design*, pp. 400–404, 1990.

[62] M. Cirit, "Characterizing a VLSI standard cell library," in *Proc. IEEE 1991 Cust. Integr. Circuits Conf.*, 1991.

[63] J. M. Daga and D. Auvergne, "A comprehensive delay macro modeling for submicrometer CMOS logics," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 1, pp. 42–55, Jan 1999.

[64] D. Auvergne, J. M. Daga, and M. Rezzoug, "Signal transition time effect on CMOS delay evaluation," *IEEE Transactions on Circuits and Systems: I*, vol. 47, no. 9, pp. 1362–1369, Sep 2000.

[65] G. Bai, S. Bobba, and I. Hjj, "Static timing analysis including power supply noise effect on propagation delay in VLSI circuits," in *Proc. 38th Des. Autom. Conf. (IEEE Cat. No.01CH37232)*, 2001.

[66] L. Bisdounis, S. Nikolaidis, and O. Loufopavlou, "Propagation delay and short-circuit power dissipation modeling of the CMOS inverter," *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.*, vol. 45, 1998.

[67] L. Bisdounis, S. Nikolaidis, O. Koufopavlou, and C. Goutis, "Switching response modeling of the CMOS inverter for sub-micron devices," in *Proc. Des. Autom. Test Eur.*, 1998.

[68] L. Bisdounis, S. Nikolaidis, and O. Koufopavlou, "Analytical transient response and propagation delay evaluation of the CMOS inverter for short-channel devices," *IEEE J. Solid-State Circuits*, vol. 33, 1998.

[69] A. Chatzigeorgiou, S. Nikolaidis, I. Tsoukalas, and O. Koufopavlou, "CMOS gate modeling based on equivalent inverter," *Proc. 1999 IEEE Int. Symp. Circuits Syst. VLSI (ISCAS'99.)*, vol. 6, 1999.

[70] H. C. Chow and W. S. Feng, "An analytical CMOS inverter delay model including channel-length modulations," *IEEE J. Solid-State Circuits*, vol. 27, no. 9, pp. 1303–1306, 1992.

[71] F. S. Marranghello, A. I. Reis, and R. P. Ribas, "CMOS inverter delay model based on DC transfer curve for slow input," *14th Int. Symp. Qual. Electron. Des.*, pp. 651–657, 2013.

[72] J. Rossello and J. Segura, "An analytical charge-based compact delay model for submicrometer CMOS inverters," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 51, 2004.

[73] Y. Wang and M. Zwolinski, "Analytical transient response and propagation delay model for nanoscale CMOS inverter," *2009 IEEE Int. Symp. Circuits Syst.*, 2009.

[74] G. Michael. (2006) PrimeTime static timing analysis. Retrieved from http://www.synopsys.com.

[75] B. Mullen, "Ccs technology," November 2005.

[76] B. Tutuianu, R. Baldick, and M. S. Johnstone, "Nonlinear device models for timing and noise analysis," *IEEE trans. on Computer Aided Design*, vol. 23, no. 11, pp. 1510–1521, Nov. 2004.

[77] Synopsys, "CCS timing white paper," 2005.

[78] Cadence, "Delay calculation meets the nanometer era," Cadence Technical Paper, 2005.

[79] F. Dartu, N. Menezes, and L. Pileggi, "Performance computation for precharacterized CMOS gates with RC loads," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 15, no. 5, pp. 544–553, 1996.

[80] S. Y. Kim and S. S. Wong, "Closed-form rc and rlc delay models considering input rise time," *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 54, no. 9, pp. 2001–2010, September 2007.

[81] C. Kashyap, C. Amin, N. Menezes, and E. Chiprout, "A nonlinear cell macro model for digital applications," *Proc. International Conference on Computer-Aided Design, 2007 (ICCAD 2007)*, pp. 678–685, 4-8 Nov. 2007.

[82] V. Veetil, D. Sylvester, and D. Blaauw, "Fast and Accurate Waveform Analysis with Current Source Models," *9th Int. Symp. Qual. Electron. Des. (isqed 2008)*, 2008.

[83] I. Keller, H. King, and V. Kariat, "Challenges in gate level modeling for delay and si at 65nm and below," *Proc. 45th annual Design Automation Conference*, pp. 19–24, June 2008.

[84] Cadence, "Effective current source model (ECSM)," 2007.

[85] Synopsys, "Composite current source model (ccsm)," 2009.

[86] S. Turgis and D. Auvergne, "A novel macromodel for power estimation in CMOS structures," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17, no. 11, pp. 1090–1098, Nov 1998.

[87] Predective technology model. Retrieved from http://www.eas.asu.edu/ ptm/.

[88] L. A. Akers, "The inverse-narrow-width effect," *Electron Device Lett. IEEE*, vol. 7, no. 7, pp. 419–421, 1986.

[89] N. Alam, B. Anand, and S. Dasgupta, "The Impact of Process-Induced Mechanical Stress on CMOS Buffer Design using Multi-Fingered Devices," *Microelectronics Reliability (Elsevier)*, vol. 53, no. 3, pp. 379–385, 2012.

[90] N. Alam, S. Dasgupta, and B. Anand, "Gate-pitch optimization for circuit design using strain-engineered multifinger gate structures," *IEEE Transactions on Electron Devices*, vol. 59, no. 11, pp. 3120–3123, 2012.

[91] C. Wee, S. Maikop, and C. Y. Yu, "Mobility-enhancement technologies," *IEEE Circuits Devices Mag.*, vol. 21, pp. 21–36, 2005.

[92] J. S. Lim, S. E. Thompson, and J. G. Fossum, "Comparison of threshold-voltage shifts for uniaxial and biaxial tensile-stressed n-MOSFETs," *IEEE Electron Device Lett.*, vol. 25, no. 11, pp. 731–733, 2004.

[93] C. Wang, W. Zhao, F. Liu, M. Chen, and Y. Cao, "Modeling of layout-dependent stress effect in cmos design," in *Computer-Aided Design (ICCAD), 2009 IEEE/ACM International Conference on*, 2009, pp. 513–520.

[94] M. Kang and I. Yun, "Modeling electrical characteristics for multi-finger mosfets based on drain voltage variation," *Transactions on Electrical and Electronic Materials*, vol. 12, no. 6, pp. 245–248, 2011.

[95] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits," in *Proc. IEEE*, vol. 91, 2003, pp. 305–327.

[96] S.Sze, *Physics of Semiconductor Devices.* Newyork: Wiley, 1981.

[97] A. Dharchoudhury, R. Panda, D. Blaauw, R. Vaidyanathan, B. Tutuianu, and D. Bearden, "Design and analysis of power distribution networks in powerpc microprocessors," in *Proceedings of the 35th annual Design Automation Conference.* ACM, 1998, pp. 738–743.

[98] M. Iwabuchi, N. Sakamoto, Y. Sekine, and T. Omachi, "A methodology to analyze power, voltage drop and their effects on clock skew/delay in early stages of design," in *Proceedings of the 1999 international symposium on Physical design.* ACM, 1999, pp. 9–15.

[99] International technology roadmap for semiconductors 2011. Retrieved from http://www.itrs.net.

[100] A. Agarwal, D. Blaauw, V. Zolotov, S. Sundareswaran, K. M. Z. Gala, and R. Panda, "Statistical delay computation considering spatial correlations," *Proc. ASP-DAC 2003 Asia South Pacific*, pp. 271–276, 2003.

[101] A. Agarwal, K. Chopra, and D. Blaauw, "Statistical timing based optimization using gate sizing," *Des. Autom. Test Eur.*, 2005.

[102] A. Agarwal, K. Chopra, D. Blaauw, and V. Zolotov, "Circuit optimization using statistical static timing analysis," in *Proc. 42nd Des. Autom. Conf. 2005.*, 2005.

[103] N. Alam, B. Anand, and S. Dasgupta, "An analytical delay model for mechanical stress induced systematic variability analysis in nanoscale circuit design," *IEEE Transactions on Circuits and Systems-I*, vol. 61, no. 6, pp. 1714–1726, 2014.

[104] M. Alioto, M. Poli, and G. Palumbo, "Compact and simple output transition time model in nanometer CMOS gates," *2008 Int. Conf. Microelectron.*, 2008.

[105] B. Amelifard, S. Hatami, H. Fatemi, and M. Pedram, "A Current Source Model for CMOS Logic Cells Considering Multiple Input Switching and Stack Effect," *2008 Des. Autom. Test Eur.*, 2008.

[106] C. Amin, C. Kashyap, N. Menezes, K. Killpack, and E. Chiprout, "A multi-port current source model for multiple-input switching effects in CMOS library cells," *2006 43rd ACM/IEEE Des. Autom. Conf.*, 2006.

[107] S. Bhardwaj, P. Ghanta, and S. Vrudhula, "A Framework for Statistical Timing Analysis using Non-Linear Delay and Slew Models," *2006 IEEE/ACM Int. Conf. Comput. Aided Des.*, 2006.

[108] J. Bhasker and R. Chadha, "Static Timing Analysis for Nanometer Designs," in *Business*, 2009, pp. 179–225.

[109] D. Blaauw, K. Chopra, A. Srivastava, and L. Scheffer, "Statistical Timing Analysis: From Basic Principles to State of the Art," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 27, 2008.

[110] D. Blaauw, V. Zolotov, S. Sundareswaran, C. Oh, and R. Panda, "Slope propagation in static timing analysis," *IEEE/ACM Int. Conf. Comput. Aided Des. ICCAD - 2000. IEEE/ACM Dig. Tech. Pap.*, 2000.

[111] S. Borkar, "Design challenges for 22nm CMOS and beyond," *Electron Devices Meet. (IEDM), 2009 IEEE Int.*, 2009.

[112] L. Brusamarello, G. I. Wirth, P. Roussel, and M. Miranda, "Fast and accurate statistical characterization of standard cell libraries," *Microelectron. Reliab.*, vol. 51, pp. 2341–2350, 2011.

[113] J. Chang and L. Johnson, "A novel delay model of CMOS VLSI circuits," *49th IEEE International Midwest Symposium on Circuits and Systems (MWSCAS '06)*, vol. 2, pp. 481–485, 2006.

[114] A. Chatzigeorgiou, S. Nikolaidis, and I. Tsoukalas, "Single transistor primitive for modeling CMOS gates," in *Proc. ICECS '99. 6th IEEE Int. Conf. Electron. Circuits Syst.*, vol. 3, 1999.

[115] H. Chen and S. Dutta, "A timing model for static CMOS gates," *1989 IEEE Int. Conf. Comput. Des. Dig. Tech. Pap.*, 1989.

[116] M. Chu, Y. Sun, U. Aghoram, and S. E. Thompson, "Strain: A solution for higher carrier mobility in nanoscale mosfets," *Annual Review of Materials Research*, vol. 39, no. 1, pp. 203–229, 2009.

[117] B. Cline, K. Chopra, D. Blaauw, A. Torres, and S. Sundareswaran, "Transistor-Specific Delay Modeling for SSTA," *2008 Des. Autom. Test Eur.*, 2008.

[118] B. Cline, V. Joshi, D. Sylvester, and D. Blaauw, "STEEL: A technique for stress-enhanced standard cell library design," *2008 IEEE/ACM Int. Conf. Comput. Des.*, 2008.

[119] F. Dartu, N. Menezes, J. Qian, and L. Pillage, "A Gate-Delay Model for High-Speed CMOS Circuits," *31st Des. Autom. Conf.*, 1994.

[120] M. Dave, M. Shojaei Baghini, and D. Kumar Sharma, "A process and temperature compensated current reference circuit in CMOS process," *Microelectronics J.*, vol. 43, pp. 89–97, 2012.

[121] U. Doddannagari, S. Hu, and W. Shi, "Fast characterization of parameterized cell library," *2009 10th Int. Symp. Qual. Electron. Des.*, 2009.

[122] F. Driussi, D. Esseni, L. Selmi, P. E. Hellström, G. Malm, J. Hå llstedt, M. Östling, T. J. Grasby, D. R. Leadley, and X. Mescot, "On the electron mobility enhancement in biaxially strained Si MOSFETs," *Solid. State. Electron.*, vol. 52, pp. 498–505, 2008.

[123] M. V. Dunga, "Nanoscale CMOS Modeling," *Spring*, pp. 1–218, 2008.

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

[125] C. Forzan and D. Pandini, "Statistical static timing analysis: A survey," pp. 409–435, 2009.

[126] F. Frustaci, P. Corsonello, and S. Perri, "Analytical Delay Model Considering Variability Effects in Subthreshold Domain," *IEEE Trans. Circuits Syst.*, vol. 59, no. 3, pp. 168–172, 2012.

[127] T. Fukuoka, A. Tsuchiya, and H. Onodera, "Statistical gate delay model for Multiple Input Switching," *2008 Asia South Pacific Des. Autom. Conf.*, 2008.

[128] A. Goel and S. Vrudhula, "Statistical waveform and current source based standard cell models for accurate timing analysis," *2008 45th ACM/IEEE Des. Autom. Conf.*, 2008.

[129] J. Herbert, "An integrated design and characterization environment for the development of a standard cell library," in *Proc. IEEE 1991 Cust. Integr. Circuits Conf.*, 1991.

[130] J. Hilder, J. Walker, and A. Tyrrell, "Optimising variability tolerant standard cell libraries," *2009 IEEE Congr. Evol. Comput.*, 2009.

[131] A. Hirata, H. Onodera, and K. Tamura, "Estimation of propagation delay considering short-circuit current for static CMOS gates," *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.*, vol. 45, 1998.

[132] R. Huang, H. Wu, J. Kang, D. Xiao, X. Shi, X. An, Y. Tian, R. Wang, L. Zhang, X. Zhang, and Y. Wang, "Challenges of 22 nm and beyond CMOS technology," pp. 1491–1533, 2009.

[133] M. E. Hwang, M. Eun, S. O. Jung, and K. Roy, "Slope Interconnect Effort: Gate-Interconnect Interdependent Delay Modeling for Early CMOS Circuit Simulation," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 56, 2009.

[134] E. Jacobs and M. Berkelaar, "Gate sizing using a statistical delay model," in *Proc. Des. Autom. Test Eur. Conf. Exhib. 2000*, 2000.

[135] A. Jambek, A. NoorBeg, and M. Ahmad, "Standard cell library development," in *Proceedings. Elev. Int. Conf. Microelectron.*, 2000.

[136] L. Jing-Jia, L. C. Wang, K. T. Cheng, and A. Krstic, "False-path-aware statistical timing analysis and efficient path selection for delay testing and timing validation," *Proc .Design Autom. Conf.*, pp. 566–569, 2002.

[137] A. Kabbani, D. Al-Khalili, and A. Al-Khalili, "Delay macro modeling of CMOS gates using modified logical effort technique," *2004 IEEE Int. Conf. Semicond. Electron.*, 2004.

[138] A. Kahng and S. Muddu, "Gate load delay computation using analytical models," *Proc. APCCAS'96 - Asia Pacific Conf. Circuits Syst.*, 1996.

[139] S. M. Kang and Y. Leblebici, *Cmos Digital Integrated Circuits, 3/E.* Tata McGraw-Hill Education, 2003.

[140] S. Kannan, N. Sreeram, and B. Amrutur, "Unified Vdd Vth Optimization Based DVFM Controller for a Logic Block," *21st Int. Conf. VLSI Des. (VLSID 2008)*, 2008.

[141] B. Kaur, N. Alam, S. Manhas, and B. Anand, "Efficient ecsm characterization considering voltage, temperature and mechanical stress variability," *IEEE Transactions on Circuits and Systems:I*, in Press.

[142] B. Kaur, S. Miryala, S. K. Manhas, and B.Anand, "An Efficient Method for ECSM Characterization of CMOS Inverter in Nanometer Range Technologies," in *Proc. 14th Int. Symp. Qual. Electron. Des. ISQED 2013*, 2013, pp. 665–669.

[143] B. Kaur, S. Vundavalli, S. K. Manhas, S.Dasgupta, and B.Anand, "An accurate current source model for cmos based combinational logic cell," *Proceedings of the 13th International Symposium on Quality Electronic Design, ISQED 2012*, pp. 561–565, 2012.

[144] A. Kayssi, K. Sakallah, and T. Burks, "Analytical transient response of CMOS inverters," *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.*, vol. 39, 1992.

[145] C. Kim, S. Goodwin Johannson, and D. Sharma, "Constant current contours plots for the description of short channel effects in MOS transistors," *IEEE Trans. Electron Devices*, vol. ED-33, p. 1619, 1986.

[146] K. Kuhn, "CMOS scaling beyond 32nm: Challenges and opportunities," *2009 46th ACM/IEEE Des. Autom. Conf.*, 2009.

[147] A. S. Kumar, M. P. Kumar, S. Murali, V. Kamakoti, L. Benini, and G. D. Micheli, "A Simulation Based Buffer Sizing Algorithm for Network on Chips," *2011 IEEE Comput. Soc. Annu. Symp. VLSI*, pp. 206–211, 2011.

[148] P. Lakshmikanthan, K. Sahni, and A. Nunez, "Design of Ultra-Low Power Combinational Standard Library Cells Using A Novel Leakage Reduction Methodology," *2006 IEEE Int. SOC Conf.*, 2006.

[149] B. Li, N. Chen, M. Schmidt, W. Schneider, and U. Schlichtmann, "On hierarchical statistical static timing analysis," *2009 Des. Autom. Test Eur. Conf. Exhib.*, 2009.

[150] M. Z. Li, C. I. Ieong, M. K. Law, P. I. Mak, M. I. Vai, and R. P. Martins, "Subthreshold standard cell library design for ultra-low power biomedical applications." in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2013, 2013, pp. 1454–7.

[151] Y. Lim and M. Soma, "Statistical estimation of delay-dependent switching activities in embedded CMOS combinational circuits," *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 5, pp. 309–319, 1997.

[152] S. Lin and M. Marek-Sadowska, "An accurate and efficient delay model for CMOS gates in switch-level timing analysis," *IEEE Int. Symp. Circuits Syst.*, 1990.

[153] M. Lundstrom and Z. Ren, "Essential physics of carrier transport in nanoscale MOSFETs," *IEEE Trans. Electron Devices*, vol. 49, 2002.

[154] P. Maurine, M. Rezzoug, N. Azemard, and D. Auvergne, "Transition time modeling in deep submicron CMOS," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 21, 2002.

[155] S. Miryala, B. Kaur, B.Anand, and S. K. Manhas, "Efficient nanoscale VLSI standard cell library characterization using a novel delay model," in *Proc. 12th Int. Symp. Qual. Electron. Des. ISQED'11*, 2011, pp. 458–463.

[156] N. Mohta and S. Thompson, "Mobility enhancement," *IEEE Circuits Devices Mag.*, vol. 21, 2005.

[157] S. Nazarian, M. Pedram, E. Tuncer, T. L. T. Lin, and A. Ajami, "Modeling and propagation of noisy waveforms in static timing analysis," *Des. Autom. Test Eur.*, 2005.

[158] S. M. Nooshabadi, G. S. Visweswaran, and D. Nagchoudhuri, "A MOS transistor thermal sub-circuit for the SPICE circuit simulator," *Microelectronics J.*, vol. 29 (4-5), no. Issues 4-5, pp. 355–366, May 1998.

[159] K. Okada, K. Yamaoka, and H. Onodera, "A statistical gate-delay model considering intra-gate variability," *ICCAD-2003. Int. Conf. Comput. Aided Des.*, 2003.

[160] S. Pavan, "Efficient Simulation of Weak Nonlinearities in Continuous-Time Oversampling Converters," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 57, 2010.

[161] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits a Design Perspective*. Pearson Education, 2003.

[162] I. Rachit and M. Bhat, "AutoLibGen: An open source tool for standard cell library characterization at 65nm technology," *2008 Int. Conf. Electron. Des.*, 2008.

[163] S. Raja, F. Varadi, M. Becer, and J. Geada, "Transistor level gate modeling for accurate and fast timing, noise, and power analysis," *2008 45th ACM/IEEE Des. Autom. Conf.*, 2008.

[164] R. S. Rajan and S. Pavan, "Device noise in continuous-time oversampling converters," pp. 1829–1840, 2012.

[165] N. Rajesh and S. Pavan, "A Lumped Component Programmable Delay Element for Ultra-Wideband Beamforming," in *Proc. 2013 IEEE Cust. Integr. Circuits Conf.*

[166] A. Ricci, I. D. Munari, and P. Ciampolini, "Performance-Effective Compaction of Standard-Cell Libraries for Digital Design," *2009 12th Euromicro Conf. Digit. Syst. Des. Archit. Methods Tools*, 2009.

[167] J. Rossello and J. Segura, "A compact propagation delay model for deep-submicron CMOS gates including crosstalk," *Proc. Des. Autom. Test Eur. Conf. Exhib.*, vol. 2, 2004.

[168] S. Saha, "Modeling Process Variability in Scaled CMOS Technology," *IEEE Des. Test Comput.*, vol. 27, 2010.

[169] J. Samanta and B. Prasad, "Comprehensive analysis of delay in UDSM CMOS circuits," *Int. Conf. Electron. Commun. Comput. Technol.*, pp. 29–32, 2011.

[170] S. Shah, P. Gupta, and A. Kahng, "Standard cell library optimization for leakage reduction," *2006 43rd ACM/IEEE Des. Autom. Conf.*, 2006.

[171] G. Shahidi, "Evolution of CMOS Technology at 32 nm and Beyond," *2007 IEEE Cust. Integr. Circuits Conf.*, 2007.

[172] D. Sharma, K. Narasimhan, N. Periaswamy, and D. Bapat, "Temperature dependence of the electron drift mobility in doped and undoped amorphous silicon," *Phys. Rev.*, vol. B-44, p. 12806, 1991.

[173] D. Sharma and K. Ramanathan, "Modeling thermal effects on MOS I-V characteristics," *IEEE Electron Device Lett.*, vol. 4, 1983.

[174] T. S. Shelar and G. S. Visweswaran, "Inclusion of thermal effects in the simulation of bipolar circuits using circuit level behavioral modeling," in *Proceedings. 17th Int. Conf. VLSI Des.*, 2004.

[175] T. Skotnicki, J. A. Hutchby, T. J. King, H. S. P. Wong, and F. Boeuf, "The end of CMOS scaling: toward the introduction of new materials and structural changes to improve MOSFET performance," *IEEE Circuits Devices Mag.*, vol. 21, 2005.

[176] Y. Song, H. Zhou, Q. Xu, J. Luo, H. Yin, J. Yan, and H. Zhong, "Mobility Enhancement Technology for Scaling of CMOS Devices: Overview and Status," pp. 1584–1612, 2011.

[177] J. Sridharan and T. Chen, "Gate delay modeling with multiple input switching for static (statistical) timing analysis," *19th Int. Conf. VLSI Des. held jointly with 5th Int. Conf. Embed. Syst. Des.*, 2006.

[178] D. Sylvester, *Statistical Analysis and Optimization for VLSI : Timing and Power*, 2005.

[179] Q. Tang, A. Zjajo, M. Berkelaar, and N. van der Meijs, "Statistical delay calculation with Multiple Input Simultaneous Switching," *2011 IEEE Int. Conf. IC Des. Technol.*, pp. 1–4, 2011.

[180] Q. Tang, A. Zjajo, M. Berkelaar, and N. V. D. Meijs, "Transistor-Level Gate Model Based Statistical Timing Analysis Considering Correlations," in *PROC_ DATE*, 2012, pp. 917 – 922.

[181] Y. Taur, D. A. Buchanan, W. Chen, D. J. Frank, K. E. Ismail, S. H. Lo, G. A. Sai Halasz, R. G. Viswanathan, H. J. C. Wann, S. J. Wind, and H. S. Wong, "CMOS scaling into the nanometer regime," in *Proc. IEEE*, vol. 85, 1997.

[182] R. Thakker, C. Sathe, M. Baghini, and M. Patil, "A Table-Based Approach to Study the Impact of Process Variations on FinFET Circuit Performance," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 29, 2010.

[183] R. Thakker, C. Sathe, A. Sachid, M. S. Baghini, V. R. Rao, and M. Patil, "A Novel Table-Based Approach for Design of FinFET Circuits," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 28, 2009.

[184] Y. L. Theng, E. Duncker, N. Mohd Nasir, G. Buchanan, and H. W. Thimbleby, "Design Guidelines and User-Centred Digital Libraries," *Res. Adv. Technol. Digit. Libr.*, pp. 852–853, 2010.

[185] S. E. Thompson, S. Suthram, Y. Sun, G. Sun, S. Parthasarathy, M. Chu, and T. Nishida, "Future of Strained Si / Semiconductors in Nanoscale MOSFETs," in *Proc. IEDM*, 2006, pp. 8–11.

[186] S. Tsukiyama, "Toward stochastic design for digital circuits - statistical static timing analysis," *Proc. ASP-DAC Des. Autom. Conf.*, pp. 762–767.

[187] G. S. Visweswaran, N. K. Jain, and A. B. Bhattacharyya, "Quasistatic Analysis of a Dynamic Sense Amplifier," *Electron. Lett.*, vol. 21, no. (8), pp. 331–332, 1985.

[188] Q. Wang and S. Vrudhula, "A new short circuit power model for complex CMOS gates," in *Proc. IEEE Alessandro Volta Meml. Work. Low-Power Des.*, 1999.

[189] C. Wu, J. Hwang, and C. Chang, "An Efficient Timing Model for CMOS Combinational Logic Gates," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 4, 1985.

[190] H. Xie, "Evaluation of digital libraries: Criteria and problems from users' perspectives," *Libr. Inf. Sci. Res.*, vol. 28, pp. 433–452, 2006.

[191] N. Xu, L. Wang, A. Neureuther, and T. J. K. Liu, "Physically Based Modeling of Stress-Induced Variation in Nanoscale Transistor Performance," *IEEE Trans. Device Mater. Reliab.*, vol. 11, pp. 378–386, 2011.

[192] N. Zang, E. Park, and J. Kim, "Efficient cell characterization for SSTA," *APCCAS 2008 - IEEE Asia Pacific Conf. Circuits Syst.*, 2008.

[193] X. Zhang, L. Yuelin, J. Liu, and Y. Zhang, "Effects of interaction design in digital libraries on user interactions," pp. 438–463, 2008.

[194] J. Zhou, S. Jayapal, B. Busze, L. Huang, and J. Stuyt, "A 40 nm inverse-narrow-width-effect-aware sub-threshold standard cell library," *2011 48th ACM/EDAC/IEEE Des. Autom. Conf.*, pp. 441–446, 2011.

[195] *CMOS Digital Integrated Circuits.* Tata McGraw-Hill.

[196] (2012, December) Encounter library characterizer. Retrieved from http://www.cadence.com.