Analysis of Circuits with Partially Correlated Multi-Vt Cell Variations using Sensitivity Modeling and Propagation

BY

MOHD ABU UBAIDA

(MT23216)

ELECTRONICS & COMMUNICATION ENGINEERING

Under the Supervision of

Prof. Sneh Saurabh

Indraprastha Institute of Information Technology Delhi

New Delhi

CERTIFICATE

This is to certify that the thesis titled "**Analysis of Circuits with Partially Correlated Multi-Vt Cell Variations using Sensitivity Modeling and Propagation**" submitted by **Mohd Abu Ubaida** for the partial fulfillment of the requirements for the degree of **Master of Technology** in Electronics and Communication Engineering is a record of the bonafide work carried out by him under my guidance and supervision at Indraprastha Institute of Information Technology, Delhi. This work has not been submitted anywhere else for the reward of any other degree/ diploma.

Sneh Saurabh

Prof. Sneh Saurabh
Professor
Department of Electronics & Communication Engineering
Indraprastha Institute of Information Technology Delhi
New Delhi, 110020

Place: New Delhi
Date: 19 May 2025

Acknowledgement

This thesis was undertaken at IIIT-Delhi during the academic year 2024-2025.

I am profoundly grateful to my research advisor, **Prof. Sneh Saurabh**, for his exceptional mentorship, insightful guidance, and unwavering support throughout the course of this work. His expertise and patience have been instrumental in shaping the direction and quality of this research.

I would also like to extend my sincere appreciation to the administrative and technical staff at IIIT-Delhi for their prompt assistance and continued support, which greatly facilitated the progress of this project.

I am thankful to **Mr. Venkatraman Ramakrishnan(Onsemi)** and **Mr. Ajoy Mandal(Texas Instrument)** for their valuable feedback and technical insights, which enriched the scope and depth of this study.

A heartfelt thanks goes to **Ms. Pooja Beniwal** for her constant encouragement, guidance, and thoughtful suggestions.

I am also sincerely thankful to my wonderful parents and supportive friends for their unwavering support, motivation, and unconditional love throughout this rewarding journey.

Mohd Abu ubaida

Abstract

Traditional Static Timing Analysis (STA) tools, commonly used in digital circuit design, usually assume that variations in low-threshold voltage (LVT) and high-threshold voltage (HVT) transistors are fully correlated. This means they treat both types of transistors as if they vary in the same way during the manufacturing process. However, in reality, LVT and HVT transistors are made differently, so their variations are only partly related. This incorrect assumption can lead to errors in estimating the delay and performance of circuits that use both LVT and HVT cells. In this work, we present a new analytical method that accurately calculates the delay variation of logic gates by considering the partial correlation between LVT and HVT transistors. Our method includes a new way to compute the total delay variation of a signal path made up of both LVT and HVT gates. By accounting for the real behavior of process variations, our method improves the accuracy of timing analysis for mixed-Vt circuits. To make this possible, we create a special sensitivity library that helps us measure how changes in transistor properties affect the delay of each cell. This includes not only how LVT cell delays respond to changes in LVT transistors, but also how they respond to changes in HVT transistors and the other way around. This "cross sensitivity" helps us better understand how variations in one type of device affect the other. We test our method on a chain of inverters made from both LVT and HVT gates. The results from our analytical model closely match those from detailed Monte Carlo SPICE simulations, with less than 5% error. Our approach also works well under different input slews, output loads, correlation values between LVT and HVT, and circuit setups. Overall, this framework helps designers more accurately predict timing variations in modern digital circuits and shows how using both LVT and HVT cells can help reduce the negative effects of process variation. It provides useful insights for making more reliable and efficient digital designs.

Contents

List of Tables

List of Figures

Notation

σ	Sigma
ρ	Correlation Coefficient
t_{ox}	Oxide Thickness of transistor
W	Width of transistor
v_{th}	Threshold voltage of transistor
μ	Mobility
C_Y	covariance Matrix
nm	Nano Meter
S_{IN}	Input Slew
C_L	Output Load
D	Delay
ps	Pico Second
$S_{y_i}^D$	Sensitivity of Delsy with respect to device parameter

Chapter 1

Introduction

As semiconductor technology continues to evolve, integrated circuits are becoming smaller and more densely packed. This trend toward miniaturization brings benefits like better performance, lower power usage, and higher integration. However, it also introduces serious challenges, especially during manufacturing. One of the biggest issues is the increased sensitivity of devices to process variations, which can negatively affect both performance and manufacturing yield [1], [2].

When transistors are shrunk to nanoscale sizes, even small errors during fabrication can lead to noticeable differences in how the devices behave. These variations can happen at different stages of the chip-making process. For example, during the chemical mechanical polishing (CMP) step, uneven pressure or inconsistent slurry can cause non-uniform surface smoothing. This leads to variations in layer thickness, which then affects the electrical properties of devices. Similarly, in the photolithography process, problems like lens distortion or incorrect focus can create unintended changes in the printed patterns. These changes affect important physical parameters like gate length, gate width, and oxide thickness—each of which plays a critical role in transistor behavior [3], [4].

Process variations are usually grouped into two categories: systematic and random. Systematic variations are predictable and repeatable, often caused by known issues like layout patterns, proximity effects, or specific equipment. These can often be reduced or corrected using design techniques such as layout optimization or proximity corrections. On the other hand, random (or non-systematic) variations are unpredictable and come from factors like material randomness or atomic-level fluctuations. These are harder to model or control and pose a greater risk to the reliability of chip designs [5], [6].

Random variations can be further broken down into global and local types. Global variations affect the entire wafer or a batch of chips the same way—such as changes in temperature or deposition rates. Local variations, in contrast, happen at a much smaller scale and can cause differences even between neighboring transistors on the same chip. These might result from effects like line-edge roughness (LER), which changes the channel length, or random dopant fluctuations (RDF), which af-

fect threshold voltage differently across devices.
These kinds of variations directly influence key transistor properties such as oxide thickness, channel size, threshold voltage (V_{th}), capacitance, and current drive strength. At the circuit level, this can lead to changes in delay, power consumption, and noise margins. Timing is especially affected since delay depends on how quickly transistors switch. If timing becomes uncertain, it can lead to setup and hold violations, or even complete circuit failure in extreme cases [7], [8].
To address these issues, there is a growing need for better modeling techniques and robust design practices. These tools help predict and reduce the negative effects of process variations, making it possible to maintain performance and yield even as we continue scaling down chip sizes.

1.1 Multi-threshold Transistor

The threshold voltage of a MOSFET is the minimum gate voltage needed for the transistor to turn on and allow current to flow from the source to the drain.
This threshold voltage can be adjusted by changing the amount of doping in the substrate during fabrication. By adding extra dopants in certain areas, manufacturers can create transistors with different threshold voltages on the same chip. Typically, three types of transistors are used based on their threshold voltages:

1. Low Threshold Voltage (LVT) transistors
2. High Threshold Voltage (HVT) transistors
3. Standard or Nominal Threshold Voltage (SVT) transistors

Using a mix of these transistors allows circuit designers to balance power, speed, and leakage in different parts of the design. For example, faster paths might use LVT transistors, while power-sensitive paths might use HVT ones.

1.2 Motivation

As semiconductor technology scales down, the impact of process variations becomes more significant, especially in advanced technology nodes. Traditional static timing analysis (STA) tools typically use corner-based methods that assume full correlation among devices, including low-threshold voltage (LVT) and high-threshold voltage (HVT) transistors[9]. However, this assumption often leads to inaccurate timing

predictions, particularly in mixed-Vt designs where different threshold voltages are used to optimize power and performance.

This work addresses this gap by introducing a new analytical model that incorporates partial correlation between LVT and HVT devices when estimating delay variance in inverter chains. A key innovation of this method is the use of a sensitivity-based library that enables fast and accurate calculation of the standard deviation (σ) of sum of delays across a path. Allowing seamless integration with modern STA flows.

Accurate modeling of delay variance helps designers better anticipate timing failures and improves the robustness of variation-aware designs for advanced nodes. By exposing hidden delays in mixed LVT-HVT configurations that are missed under full-correlation assumptions, the proposed model enables more reliable design closure.

The tables 1.1 – 1.2 below compare the standard deviation (σ) of delay for different inverter pair configurations under full correlation ($\rho = 1$) and partial correlation ($\rho = 0.8$). Results show that partial correlation leads to more **realistic and slightly pessimistic** estimates, especially for mixed LVT-HVT configurations, which helps avoid over-optimistic timing closure.

Inverter Pair Config.	$\boldsymbol{\sigma(D_1 + D_2), \rho = 1}$	$\boldsymbol{\sigma(D_1 + D_2), \rho = 0.8}$	**Remark**
LVT – LVT	0.6	0.6	—
LVT – HVT	0.53	0.69	optimistic Analysis
HVT – LVT	0.83	0.98	optimistic Analysis
HVT – HVT	0.9	0.9	—

Table 1.1: $\sigma(D_1 + D_2)$ for Full Correlation and Partial Correlation between HVT and LVT, Fall Delay D_1 and Rise Delay D_2, in (ps)

Inverter Pair Config.	$\boldsymbol{\sigma(D_1 + D_2), \rho = 1}$	$\boldsymbol{\sigma(D_1 + D_2), \rho = 0.8}$	**Remark**
LVT – LVT	0.39	0.39	—
LVT – HVT	0.31	0.56	optimistic Analysis
HVT – LVT	0.50	0.70	optimistic Analysis
HVT – HVT	0.52	0.52	—

Table 1.2: $\sigma(D_1 + D_2)$ for Full Correlation and Partial Correlation between HVT and LVT, Rise Delay D_1 and Fall Delay D_2, in (ps)

1.3 Contribution

In this work, we propose an analytical model to estimate the standard deviation (sigma) of delay in a chain of low-threshold voltage (LVT) and high-threshold voltage (HVT) inverters, by capturing how delay changes with variations in key device parameters. The model is designed to include the effect of global process variations, and it also accounts for the correlation between different threshold voltage transistors, which is important for accurate timing analysis in realistic circuits.
To support our method, we characterized a sensitivity library in NLDM format. This library provides delay sensitivity values with respect to various process parameters like threshold voltage, mobility, oxide thickness, etc. Using these sensitivities, we calculate the sigma of the total delay in a chain of inverters. We validated our analytical model by comparing it against results from Monte Carlo simulations (with 2500 runs). The maximum error was within 5%, showing good accuracy. This makes it much more efficient and suitable for integration with modern static timing analysis (STA) tools.

1.4 Thesis Organization

This thesis is organized into five chapters: Chapter 2 provides a brief overview of the existing literature relevant to this work. Chapter 3 presents the derivation of the analytical model used to compute the standard deviation (or variance) of delay for a single inverter. It also extends the model to compute the standard deviation of the total delay in a chain of inverters composed of both LVT and HVT devices. Additionally, the concept and method of sensitivity propagation are discussed in detail. Chapter 4 describes the characterization of the sensitivity library and explains how it is used to compute the standard deviation of delay for a chain of inverters using pre-characterized sensitivity data. Chapter 5 validates the proposed analytical model by comparing its results against those obtained from Monte Carlo SPICE simulations. The comparison is presented for single inverters, inverter pairs, and a four-stage inverter chain. Furthermore, we identify the optimized path configuration that yields the minimum standard deviation of total delay, as well as the worst-case configuration that produces the maximum standard deviation for the inverter chain consisting of 4 inverters in the chain.

Chapter 2

Literature Survey

This chapter briefly reviews the existing literature relevant to this work. In [10], the authors have thoroughly explained the genesis of process variations and how the current and delays are impacted. They have also discussed the maximum delay approximation for points where multiple paths converge on a timing graph with various correlation factors ρ. They have also mentioned the complexity of non-normal distributions, which arises when the maximum operation is performed during SSTA on a timing graph. Several approaches such as the numerical integration method, Monte Carlo simulations, probabilistic analysis, and path-based and block-based methods were discussed through which SSTA is performed. The path-based approach deals with the correlation that arises between different paths due to geometric proximity. The block-based approach deals with the spatial correlation problem, in which the gates in closer proximity are tightly correlated, whereas those a bit far away on the path are loosely correlated.

In [11], the authors have mentioned the drawbacks of path-based SSTA and introduced a novel statistical incremental timer which performs breadth-first traversal (path-based approach use depth-first traversal), which helps to overcome the drawbacks by enabling incremental processing and also by identifying the critical path based on their novel concept of tightness probability. The tightness probability measures the probability of variations of a particular path for a particular process corner. It enables us to compute the maximum path delay value between different paths analytically with very little computation. The incremental timer propagates the sensitivities (delay with respect to variations in different device parameters) and also propagates the tightness probabilities. The authors have presented a time-efficient and memory-efficient method to implement path-based SSTA, but the idea of spacial correlation and how will it be integrated with this is missing.

In [12], the authors have presented an analytical approach to compute delay as a function of variations in different device parameters. They have presented a method to approximate the delay function using first-order Taylor series approximation about the nominal values of various device parameters. Also, they have discussed methods to compute the effects of both inter-die and intra-die variations, along with taking

spatial correlation into account. The authors have also mentioned an approach to approximate interconnect delays. However, their approximation model is valid only for Gaussian random variations, thus leaving scope to incorporate non-Gaussian distributions, which generally arise while performing MAX operation on a timing path[10].

In [13], the authors have built over the contributions of [11], [12], [14] and have approximated the delay with respect to variations in various device parameters using Taylor series approximation where the degree of polynomial for individual parameters depends on the magnitude of variation and the desired level of accuracy. They also discussed a method to analytically perform MAX operation during timing graph traversal[15], for which they used regression modeling. They have also discussed that their regression model is applicable to any kind of variations and eliminates the need for random variations/delays to be Gaussian in nature, as in [12]. Though [13] provides a good level of accuracy along with incorporating both local and global variations and also the spatial correlation, the impact of correlation between different threshold voltage transistors is omitted. This work is motivated to address this need.

In [16], the authors have put forward the necessity to take into account threshold voltage variation within the same die (intra-die V_t variation) as its impact on both the delay and leakage current is significant. The delay variation was shown to be highest when only a single input was toggled, whereas the smallest variability was shown when more input toggles were considered. Though this work takes into account single V_t transistors, it established that the intra-die V_t variations are of serious concern.

In work[17], the authors have discussed a simplified SSTA framework named Parametric On-Chip Variations (POCV). They have proposed a framework where each gate is modeled separately as a Gaussian distribution, and only a single POCV coefficient (mean/sigma) will be propagated along the timing graph.

All these works on statistical static timing analysis have successfully dealt with path-to-path correlation, spatial correlation, and both local and global variations using different methodologies. However, the impact of correlated variation amongst different V_t transistors is not incorporated in the existing literature. However, in recent times, most designs use multi-threshold voltage transistors. The commonality in the fabrication processes of multi-V_t transistors results in correlation in their device parameters. Hence, the delays and other timing attributes correlate between the cells

implemented using multi-V_t transistors. In this work, the correlation in the timing attributes of the cells implemented using multi-V_t transistors is investigated in the following chapters.

Chapter 3

Analytical Model for Variations in CMOS Invertor Delay

3.1 MOSFET Variation Model in Model File

A device model file for a specific technology includes SPICE models for various types of MOSFETs. It also describes how different parameters change under various conditions. These conditions are usually grouped into process corners such as Typical-Typical (TT), Fast-Fast (FF), Slow-Slow (SS), Fast-Slow (FS), and Slow-Fast (SF). In addition to these corner-based variations, the model file also supports statistical variations, which are used for running Monte Carlo simulations to analyze how random changes in device parameters affect circuit behavior.

In this work, we employ a 45 nm technology process design kit (PDK) to demonstrate the computational framework and its effectiveness in modeling the impact of global process-induced variations on transistor parameters. Specifically, the transistor parameter y_k is modeled as a linear combination of V random variables, each representing a different source of process variation. The equation governing this relationship is:

$$y_k = \sum_{i=1}^{V} a_{ki} r_i \tag{3.1}$$

where: r_i is the i-th random variable, modeled as a Gaussian random variable with mean zero, standard deviation 1, and independent of other random variables r_j. a_{ki} represents the weight or coefficient of the i-th random variable in the variation of the k-th transistor parameter.

The PDK under consideration identifies four key transistor parameters that are influenced by global variations for each transistor type: - Gate oxide thickness (t_{ox}) - Transistor width (W) - Threshold voltage (V_{th}) - Carrier mobility (μ)

These parameters are listed for each type of transistor in Table 3.1. It is important to note that, for a given transistor, two parameters y_k and y_l are correlated because they depends on the same random variable r_i as shown in Equation (3.1).

	Gate Ox. Thick.	Width	Th. Volt.	Mobility
LVT NMOS	$t_{oxn,L}(y_1)$	$W_{n,L}(y_2)$	$V_{thn,L}(y_3)$	$\mu_{n,L}(y_4)$
LVT PMOS	$t_{oxp,L}(y_5)$	$W_{p,L}(y_6)$	$V_{thp,L}(y_7)$	$\mu_{p,L}(y_8)$
HVT NMOS	$t_{oxn,H}(y_9)$	$W_{n,H}(y_{10})$	$V_{thn,H}(y_{11})$	$\mu_{n,H}(y_{12})$
HVT PMOS	$t_{oxp,H}(y_{13})$	$W_{p,H}(y_{14})$	$V_{thp,H}(y_{15})$	$\mu_{p,H}(y_{16})$

Table 3.1: Transistor parameters impacted by the global process-induced variations and their corresponding symbols.

The variations in these parameters are influenced by global process-induced variations, and the correlation between them can affect the overall performance of integrated circuits. Understanding these correlations and the impact of each parameter is crucial for accurate performance prediction and optimization of designs.

The random variables a_5, a_6, a_7 and a_8 are for HVT devices and a_9, a_{10}, a_{11} and a_{12} are for LVT devices. These random variables are defined in the model file as Gaussian random variables with mean '0' and standard deviation '1'.
There are a total of 16 parameters: 4 for LVT NMOS, 4 for LVT PMOS, 4 for HVT NMOS, and 4 for HVT PMOS.
The coefficients of these linear equations are defined in the following tables:

	t_1	t_2	t_3	t_4
LVT NMOS	-1.60$\times10^{-13}$	-1.08$\times10^{-11}$	-2.55$\times10^{-13}$	7.73$\times10^{-12}$
LVT PMOS	-1.60$\times10^{-13}$	-1.08$\times10^{-11}$	-2.55$\times10^{-13}$	7.73$\times10^{-12}$
HVT NMOS	5.84$\times10^{-13}$	-1.30$\times10^{-11}$	3.04$\times10^{-13}$	2.88$\times10^{-13}$
HVT PMOS	-3.4$\times10^{-13}$	-1.19$\times10^{-11}$	-5.1$\times10^{-13}$	5.77$\times10^{-12}$

Table 3.2: Coefficients corresponding to t_{ox} variations.

	w_1	w_2	w_3	w_4
LVT NMOS	9.39×10^{-11}	3.00×10^{-9}	-7.30×10^{-10}	-2.80×10^{-10}
LVT PMOS	9.39×10^{-11}	3.00×10^{-9}	-7.30×10^{-10}	-2.80×10^{-10}
HVT NMOS	1.88×10^{-10}	2.70×10^{-9}	-9.50×10^{-10}	-2.30×10^{-10}
HVT PMOS	1.88×10^{-10}	2.70×10^{-9}	-9.50×10^{-10}	-2.30×10^{-10}

Table 3.3: Coefficients corresponding to width (dxw) variations.

	v_{t1}	v_{t2}	v_{t3}	v_{t4}
LVT NMOS	-5.90×10^{-3}	-5.05×10^{-3}	7.41×10^{-3}	3.60×10^{-3}
LVT PMOS	-4.73×10^{-3}	-54.07×10^{-3}	5.79×10^{-3}	2.88×10^{-3}
HVT NMOS	-7.31×10^{-3}	-6.09×10^{-3}	3.66×10^{-3}	5.76×10^{-3}
HVT PMOS	-4.41×10^{-3}	-4.66×10^{-3}	5.86×10^{-3}	3.05×10^{-3}

Table 3.4: Coefficients corresponding to threshold voltage V_{t} variations.

	u_1	u_2	u_3	u_4
LVT NMOS	4.80×10^{-5}	3.00×10^{-6}	-9.86×10^{-5}	2.80×10^{-17}
LVT PMOS	-1.50×10^{-4}	-4.70×10^{-4}	-4.66×10^{-6}	3.23×10^{-17}
HVT NMOS	-1.05×10^{-5}	-3.46×10^{-5}	-3.97×10^{-5}	$+1.26\times10^{-17}$
HVT PMOS	-2.70×10^{-5}	-3.72×10^{-5}	-4.41×10^{-5}	7.70×10^{-17}

Table 3.5: Coefficients corresponding to mobility (du0) variations.

3.2 Variations and Their Dispersion

To identify which device parameter exhibits the most variation, we use a normalized statistical metric known as the coefficient of variation (CV). This helps us compare the spread of each parameter relative to its average value, regardless of units.

$$\mathrm{CV} = \frac{\sigma}{\mu}$$

Here, σ is the standard deviation, and μ is the mean value of the parameter. Tables 3.6 and 3.7 summarize the CV values for low and high threshold voltage transistors.

Parameter	**Mean (μ)**	**Std. Dev. (σ)**	**CV (%)**
$t_{ox\_n\_lvt}$ (nm)	2.41	1.295×10^{-2}	0.54
dxw_{n_lvt} (nm)	120	3.0818	2.57
$dvth0_{n\_lvt}$ (V)	0.24925	0.0114	4.57
$du0_{n\_lvt}$ ($\frac{m^2}{Vs}$)	0.0152801	1.1106×10^{-4}	0.73
$t_{ox\_p\_lvt}$ (nm)	2.41	1.295×10^{-2}	0.54
dxw_{p_lvt} (nm)	200	3.0818	1.54
$dvth0_{p\_lvt}$ (V)	-0.202698	0.0090	4.44
$du0_{p\_lvt}$ ($\frac{m^2}{Vs}$)	0.0230515	4.9266×10^{-4}	2.14

Table 3.6: Coefficient of Variation for LVT Transistor Parameters

Parameter	**Mean (μ)**	**Std. Dev. (σ)**	**CV (%)**
$t_{ox\_n\_hvt}$ (nm)	2.41	1.2947×10^{-2}	0.54
dxw_{n_hvt} (nm)	120	2.0883	1.74
$dvth0_{n\_hvt}$ (V)	0.42225	0.0117	2.77
$du0_{n\_hvt}$ ($\frac{m^2}{Vs}$)	0.01528	5.3716×10^{-5}	0.35
$t_{ox\_p\_hvt}$ (nm)	2.41	1.2994×10^{-2}	0.54
dxw_{p_hvt} (nm)	180	2.8659	1.59
$dvth0_{p\_hvt}$ (V)	-0.368698	0.0093	2.52
$du0_{p\_hvt}$ ($\frac{m^2}{Vs}$)	0.0230515	3.7065×10^{-4}	1.61

Table 3.7: Coefficient of Variation for HVT Transistor Parameters

Since the dispersion (CV) for all device parameters is below 5%, we can safely use a Taylor series approximation to estimate the effect of these variations on the transistor's saturation current and gate delay. In the next sections, we develop an analytical method to quantify how these variations influence the performance of inverter chains. We can compactly express the set of equations (as indicated in Eq. (3.1)) governing the variations in the 16 transistor parameters using the following matrix representation:

$$Y = AR \tag{3.2}$$

where Y is a 16×1 column vector representing the transistor parameters listed in Table 3.1, A is a 16×16 coefficient matrix, and R is a 16×1 column vector comprising

the random variables that model process variations. The structure of the coefficient matrix A is given as:

$$A = \begin{bmatrix} P_{ij} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ P'_{ij} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & Q_{ij} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & Q'_{ij} & \mathbf{0} \end{bmatrix}$$

Here, P_{ij} corresponds to LVT NMOS, P'_{ij} to LVT PMOS, Q_{ij} to HVT NMOS, and Q'_{ij} to HVT PMOS devices. In this model, global variations in LVT PMOS and NMOS transistors are considered correlated. Similarly, correlations exist between HVT PMOS and HVT NMOS transistors. Moreover, the model assumes partial correlation between LVT and HVT devices, quantified using a correlation coefficient ρ, where $0 \leq \rho \leq 1$. Traditional analyses typically assume $\rho = 1$; however, in realistic scenarios, $\rho < 1$, and this work accounts for such partial correlation.

The covariance matrix of the random variable vector R, denoted as C_R, is defined as:

$$C_R = \begin{bmatrix} I_8 & \rho I_8 \\ \rho I_8 & I_8 \end{bmatrix}$$

where I_8 is the 8×8 identity matrix. Given C_R and Eq. (3.2), the covariance matrix C_Y of the transistor parameter vector Y is computed as:

$$C_Y = A C_R A^T \tag{3.3}$$

where A^T represents the transpose of the matrix A.

The primary objective of this work is to formulate an analytical framework for computing the variance in path delay for a mixed Low-Vt/High-Vt (LVT-HVT) inverter chain, while incorporating the partial correlation between LVT and HVT transistors. To this end, we first establish a delay variance model for an individual inverter, utilizing sensitivity analysis with respect to the transistor parameters and their correlation. Subsequently, we introduce a methodology for propagating these sensitivities along an inverter chain.

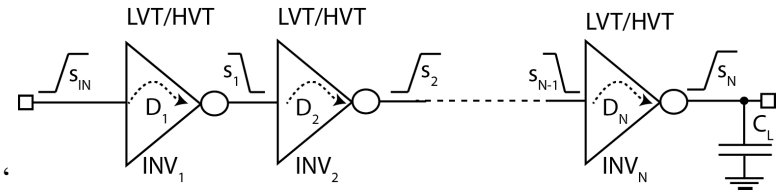


Figure 3.1: Chain of N Inverter.

It is important to note that variations in the driving inverter (e.g., the first stage in Fig. 3.1 influence the output slew, which in turn affects the input slew of the subsequent stage. This cascading alteration in slew impacts the delay and output behavior of all downstream inverters, necessitating accurate modeling of sensitivity propagation. Additionally, in a given standard cell, delay sensitivity is a function of both input slew and output load. This dependency is captured using a pre-characterized sensitivity library, which is described in later chapters.

3.3 Taylor Approximation Model

The Taylor series is a mathematical tool used to approximate a function around a specific point. For a function $f(x)$ expanded around the point $x = a$, the approximation is written as:

$$f(x) \approx f(a) + \frac{f'(x)}{1!}(x-a) + \frac{f''(x)}{2!}(x-a)^2 + \dots \tag{3.4}$$

When the function depends on multiple variables, partial derivatives are used instead of ordinary derivatives. Each partial derivative is multiplied by the corresponding small change (or variation) in that variable.

In our work, we consider variations in four key transistor parameters: oxide thickness (t_{ox}), channel width (w), threshold voltage (v_{th}), and mobility (u). The Taylor series expansion for a function f with respect to these variables can be approximated as:

$$f(x)+\Delta f \approx f(t_{ox_0}, w_0, v_{th_0}, u_0)+\frac{\partial f}{\partial t_{ox}}\Delta t_{ox}+\frac{\partial f}{\partial w}\Delta w+\frac{\partial f}{\partial v_{th}}\Delta v_{th}+\frac{\partial f}{\partial u}\Delta u \tag{3.5}$$

Here, $f(t_{ox_0}, w_0, v_{th_0}, u_0)$ represents the value of the function at the nominal (mean) values of the device parameters. The partial derivatives indicate how sensitive the function is to each individual parameter. These sensitivity values help us estimate how changes in the parameters affect the overall output of the function.

3.4 Variance in Gate Delay D for Single Inverter Case

Let us consider the delay of an inverter, denoted as D_I, to be a function of M underlying device parameters, represented as:

$$D_I = f(y_1, y_2, y_3, \dots, y_M) \tag{3.6}$$

For a single inverter operating under fixed input slew and output load conditions, the delay variation is primarily influenced by $M = 8$ parameters—four corresponding to the PMOS transistor and four to the NMOS transistor, as listed in Table 3.1. During a rising output transition, the PMOS transistor is active (ON), while the NMOS transistor is inactive (OFF), meaning the PMOS parameters predominantly affect the rise delay. Conversely, during a falling transition, the NMOS parameters govern the fall delay.

However, when the input transition time (slew) is large, the inverter operates for a longer period with both the PMOS and NMOS transistors simultaneously in the saturation region. In such cases, both transistors contribute to the delay behavior, making it essential to account for the variation in all device parameters for accurate modeling of delay variation.

Assuming small deviations of the device parameters from their nominal values, we can apply a first-order Taylor series expansion to approximate the incremental change in delay, ΔD_I, as shown in Eq. (3.5):

$$\Delta D_I \approx \sum_{i=1}^{M} \frac{\partial D_I}{\partial y_i} \Delta y_i + \text{higher-order terms} \tag{3.7}$$

Since the variations in parameters are assumed to be small, we neglect the higher-order terms. As a result, the variance in inverter delay can be approximated using the linearized model:

$$\mathrm{Var}(D_I) \approx \sum_{i=1}^{M} \left(S_{y_i}^D\right)^2 \mathrm{Var}(y_i) + 2 \sum_{i=1}^{M} \sum_{j=i+1}^{M} S_{y_i}^D S_{y_j}^D \mathrm{Cov}(y_i, y_j) \tag{3.8}$$

Here, $S_{y_i}^D = \frac{\partial D_I}{\partial y_i}$ denotes the sensitivity of the inverter delay with respect to parameter y_i, $\mathrm{Var}(y_i)$ is the variance of y_i, and $\mathrm{Cov}(y_i, y_j)$ is the covariance between parameters y_i and y_j.

The variances and covariances of these parameters can be derived from the transistor variation model (as defined by Eq. (3.3)). Therefore, with known delay sensitivities, it becomes feasible to compute the variance or standard deviation of the inverter delay. In this work, we propose using a pre-characterized sensitivity library to retrieve these sensitivity values, as detailed in the following sections.

$$
\begin{aligned}
Var(D) = {} & c_1^2 Var(t_{oxnl}) + c_2^2 Var(w_{nl}) + c_3^2 Var(v_{thnl}) + c_4^2 Var(u_{nl}) \\
& + c_5^2 Var(t_{oxpl}) + c_6^2 Var(w_{pl}) + c_7^2 Var(v_{thpl}) + c_8^2 Var(u_{pl}) \\
& + 2\Big(c_1c_2 Cov(t_{oxnl}, w_{nl}) + c_1c_3 Cov(t_{oxnl}, v_{thnl}) + c_1c_4 Cov(t_{oxnl}, u_{nl}) \\
& \quad + c_2c_3 Cov(w_{nl}, v_{thnl}) + c_2c_4 Cov(w_{nl}, u_{nl}) + c_3c_4 Cov(v_{thnl}, u_{nl}) \\
& \quad + c_5c_6 Cov(t_{oxpl}, w_{pl}) + c_5c_7 Cov(t_{oxpl}, v_{thpl}) + c_5c_8 Cov(t_{oxpl}, u_{pl}) \\
& \quad + c_6c_7 Cov(w_{pl}, v_{thpl}) + c_6c_8 Cov(w_{pl}, u_{pl}) + c_7c_8 Cov(v_{thpl}, u_{pl}) \\
& \quad + c_1c_5 Cov(t_{oxnl}, t_{oxpl}) + c_1c_6 Cov(t_{oxnl}, w_{pl}) + c_1c_7 Cov(t_{oxnl}, v_{thpl}) \\
& \quad + c_1c_8 Cov(t_{oxnl}, u_{pl}) + c_2c_5 Cov(w_{nl}, t_{oxpl}) + c_2c_6 Cov(w_{nl}, w_{pl}) \\
& \quad + c_2c_7 Cov(w_{nl}, v_{thpl}) + c_2c_8 Cov(w_{nl}, u_{pl}) + c_3c_5 Cov(v_{thnl}, t_{oxpl}) \\
& \quad + c_3c_6 Cov(v_{thnl}, w_{pl}) + c_3c_7 Cov(v_{thnl}, v_{thpl}) + c_3c_8 Cov(v_{thnl}, u_{pl}) \\
& \quad + c_4c_5 Cov(u_{nl}, t_{oxpl}) + c_4c_6 Cov(u_{nl}, w_{pl}) + c_4c_7 Cov(u_{nl}, v_{thpl}) \\
& \quad + c_4c_8 Cov(u_{nl}, u_{pl}) \Big)
\end{aligned}
\tag{3.9}
$$

The above formula is the detailed formula for computing the Var(D) for single inverter case. Where:

- $\mathrm{c_i} = \mathrm{S}_{y_i}^{D} = \frac{\partial D_I}{\partial y_i}$ is the **sensitivity** of the inverter delay to device parameter y_i,
- $\mathrm{Var}(y_i)$ is the **variance** of device parameter y_i,
- $\mathrm{Cov}(y_i, y_j)$ is the **covariance** between parameters y_i and y_j,
- The summation includes **all 28 unique cross terms** (since $\binom{8}{2} = 28$).

General formula

$$
\mathrm{Var}(D) = \sum_{i=1}^{8} c_i^2 \,\mathrm{Var}(y_i) + 2 \sum_{i=1}^{7} \sum_{j=i+1}^{8} c_i c_j \,\mathrm{Cov}(y_i, y_j) \tag{3.10}
$$

3.5 Variance of Sum of Delay for the Chain of LVT HVT inverters

For a chain of N inverters, as shown in Fig. 3.1, the total path delay D is simply the sum of the delays of each individual inverter:

$$
D = D_1 + D_2 + \cdots + D_i + \cdots + D_N \tag{3.11}
$$

Here, D_i represents the delay of the i-th inverter in the chain.
Since a typical inverter chain can include all four types of transistors listed in Table 3.1 (LVT NMOS, LVT PMOS, HVT NMOS, HVT PMOS), the total path delay D depends on $M = 16$ different transistor parameters. Therefore, the variance of the path delay, Var(D), can be computed using Equation (3.8), which considers sensitivities and correlations between parameters.
The sensitivity of the overall path delay D with respect to any transistor parameter y_k can be expressed as:

$$\frac{\partial D}{\partial y_k} = \sum_{i=1}^{N} \frac{\partial D_i}{\partial y_k} \tag{3.12}$$

This means that the sensitivity of the total delay to a specific parameter y_k is the sum of the sensitivities of each inverter's delay to that parameter.
For inverters built with LVT (Low-Vt) transistors, we can directly obtain the sensitivity of their delays to the 8 LVT parameters from a pre-characterized sensitivity library. However, it becomes more challenging when we want to calculate the sensitivity of an LVT inverter's delay to HVT (High-Vt) transistor parameters—this is known as cross-sensitivity. These cross-sensitivities arise indirectly. For instance, if an upstream HVT inverter's delay changes due to variations in its parameters, it affects the output slew of that stage. This altered output slew becomes the input for the next stage, influencing its delay.
Therefore, for accurate modeling, we must account for how variations propagate from one stage to the next. This effect also applies when computing cross-sensitivities for HVT inverters due to upstream LVT stages. We propose to handle these cross-sensitivities using a technique called **sensitivity propagation**, which will be explained in the following section.

3.5.1 Sensitivity Propagation

Let us consider a simple case where $\texttt{INV}_1$ is a Low-Vt (LVT) inverter and $\texttt{INV}_2$ is a High-Vt (HVT) inverter, as shown in Fig. 3.2. Due to global variation in one of the LVT transistor parameters, say y_1, the output slew of $\texttt{INV}_1$, denoted as s_1, gets affected.

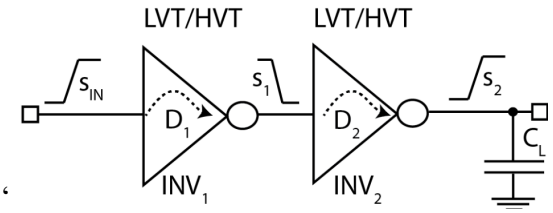


Figure 3.2: Inverter pair configuration.

Since the output of $\texttt{INV}_1$ is connected to the input of $\texttt{INV}_2$, any change in s_1 will also impact the delay of $\texttt{INV}_2$, denoted by D_2. Therefore, the delay of $\texttt{INV}_2$ becomes indirectly sensitive to the parameter y_1 of the preceding stage.
We can express this relationship mathematically using the chain rule as follows:

$$\frac{\partial D_2}{\partial y_1} = \frac{\partial s_1}{\partial y_1} \cdot \frac{\partial D_2}{\partial s_1} \tag{3.13}$$

In the proposed sensitivity library, we store:

- $\frac{\partial s_1}{\partial y_1}$: the sensitivity of the output slew to the inverter's device parameters.
- $\frac{\partial D_2}{\partial s_1}$: the sensitivity of the delay to its input slew.

Since both of these quantities can be retrieved from the sensitivity library, we can easily compute the cross-sensitivity $\frac{\partial D_2}{\partial y_1}$.
This idea can be extended to later stages in the inverter chain. For example, to compute the sensitivity of the i-th inverter's delay D_i with respect to the parameter y_1 from $\texttt{INV}_1$, we apply the following generalized equation:

$$\frac{\partial D_i}{\partial y_1} = \frac{\partial s_1}{\partial y_1} \cdot \prod_{k=2}^{i} \frac{\partial s_k}{\partial s_{k-1}} \cdot \frac{\partial D_i}{\partial s_{i-1}} \tag{3.14}$$

Here:

- $\frac{\partial s_1}{\partial y_1}$ is the sensitivity of $\texttt{INV}_1$'s output slew to its parameter.
- $\frac{\partial s_k}{\partial s_{k-1}}$ represents how the input slew of stage k is affected by the output slew of stage $k-1$.
- $\frac{\partial D_i}{\partial s_{i-1}}$ is the sensitivity of the delay of stage i to its input slew.

All of these sensitivity terms can be extracted from the proposed library. Using this approach, we can accurately compute the cross-sensitivity of the delay at any stage in the chain to a device parameter in an earlier stage.

Chapter 4

Sensitivity Library Characterization

4.1 Sensitivity Library Characterization and Usage

Process variations in advanced semiconductor technologies significantly affect circuit timing, leading to performance degradation and timing failures. As a result, statistical timing analysis has become an essential part of the design flow. However, full Monte Carlo SPICE simulations for large designs are computationally expensive. To enable fast and accurate estimation of delay variations, we propose a method based on a pre-characterized sensitivity library. This library captures how small perturbations in key transistor and environmental parameters impact the delay and slew of standard cells. This section presents a detailed methodology for constructing this sensitivity library and utilizing it to compute delay variance in a circuit.

4.2 Sensitivity Extraction

As discussed in the previous chapter, the variance in Gate delay is influenced by the following factors:

- The intrinsic variance of individual device parameters.
- The covariance between different device parameters.
- The sensitivity of the output parameter with respect to variations in these device parameters.

We define the sensitivity S_Y^X as the rate of change of parameter X with respect to a change in parameter Y. Mathematically, it can be expressed as:

$$S_Y^X = \frac{\Delta X}{\Delta Y} \tag{4.1}$$

In practice, this sensitivity is computed using a finite-difference approach:

$$S_Y^X = \frac{X_{+2.5\%} - X_{-2.5\%}}{Y_{+2.5\%} - Y_{-2.5\%}} \tag{4.2}$$

Here, $Y_{+2.5\%}$ and $Y_{-2.5\%}$ refer to the values of parameter Y when perturbed by +2.5% and -2.5%, respectively. Correspondingly, $X_{+2.5\%}$ and $X_{-2.5\%}$ represent the resulting values of X under those same conditions.
This numerical method allows us to estimate the local sensitivity of performance metrics like of delay to key process parameters such as threshold voltage, oxide thickness, channel width, and carrier mobility.

4.2.1 Characterization of Sensitivity Library

The construction of the sensitivity library begins with the extraction of transistor-level characteristics of each standard cell using SPICE-level simulations. The complete flow is illustrated in Fig. 4.1. The key inputs to the characterization algorithm are:

- **Transistor-Level SPICE Netlist:** This describes the detailed internal circuit of each standard cell, including the topology and connectivity of transistors.
- **MODEL File of the Transistor:** This includes process technology-specific parameters such as threshold voltage (V_{th}), Width (W), oxide thickness (t_{ox}), and mobility (μ) for both low-V_{th} (LVT) and high-V_{th} (HVT) devices.
- **SPICE Simulation Environment:** This is used to run transient simulations for analyzing delay and slew under nominal and perturbed conditions.

The goal is to extract the partial derivatives (sensitivities) of delay and output slew with respect to various sources of variation. These include both device-level parameters (e.g., t_{ox}, W, V_{th}, μ) and environmental parameters (e.g., input slew and output load). The procedure consists of the following steps:

1. For each standard cell, a nominal SPICE simulation is run for a range of input slews (10 ps to 200 ps) and output loads (0.5 fF to 25 fF) to establish baseline delay and output slew metrics.
2. For each parameter of interest (say V_{th}), a small perturbation (typically $\pm 2.5\%$) is applied while keeping other parameters constant.
3. Transient simulations are repeated for each perturbed case to measure the new delay and slew values.

4. Sensitivities are computed using central finite difference approximation:

$$\frac{\partial X}{\partial Y} \approx \frac{X_{+2.5\%} - X_{-2.5\%}}{Y_{+2.5\%} - Y_{-2.5\%}} \tag{4.3}$$

   where X is the delay or output slew and Y is the perturbed parameter.

5. The sensitivities are tabulated for each standard cell as a function of input slew and output load, creating a 2D table (similar in structure to NLDM models). Separate entries are created for rising and falling transitions.

This tabulated data, called the **Sensitivity Library**, forms the basis for quick estimation of delay and slew variations during circuit-level analysis. It eliminates the need for repeated SPICE simulations once the library is characterized.

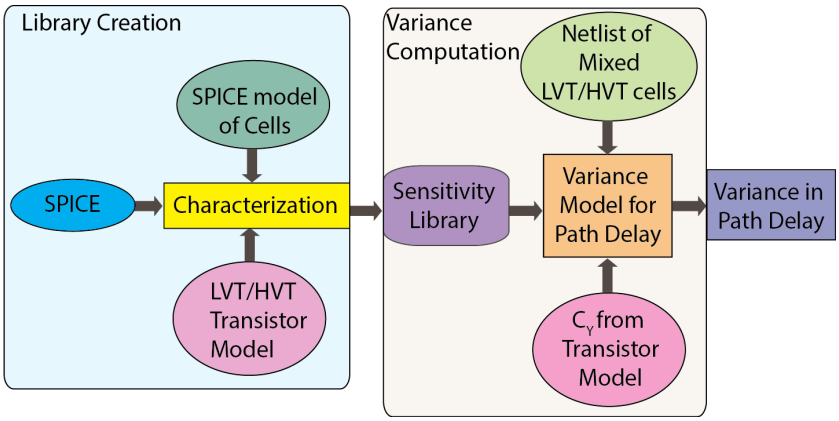


Figure 4.1: Block diagram of sensitivity library characterization and using Library to find the σ of Sum of delay of inverter Chain.

4.2.2 Computation of Variance of Delay Using Sensitivity Library

Once the sensitivity library has been built, it can be used to estimate the delay variance of an entire circuit. Fig. 4.1 shows the flow for using the library in statistical delay analysis. The main goal is to compute the standard deviation of the delay of a circuit path, which depends on variations in device parameters and their correlations.

Required Inputs

To compute delay variance, the following inputs are required:

- **Characterized Sensitivity Library:** This provides sensitivities of delay with respect to various parameters for different cells, output loads, and input slews(NLDM Model).

- **Standard Cell-Level Netlist:** The gate-level representation of the design, which describes the connectivity of various standard cells.

- **Covariance Matrix (C_Y):** This contains the variances and covariances of the transistor parameters. For example, covariances between the threshold voltage (v_{th}) and width(W) Transistors.

Delay Variance Computation Algorithm

The algorithm proceeds as follows:

1. For each cell in the netlist:

    - Determine its local input slew and output load (can be computed through slew propagation).

    - Retrieve or interpolate the corresponding delay sensitivities from the sensitivity library.

2. Extract the sensitivities of the delay and output with respect to all relevant process parameters for that cell (typically 16 parameters that include both NMOS and PMOS devices if both LVT and HVT type cells are present in the design).

3. Using the proposed analytical model to compute the delay variance of each gate

This technique allows for efficient computation of delay variance across the circuit without performing Monte Carlo simulations. The approach is scalable, accurate, and particularly suitable for variation-aware static timing analysis.

4.3 Sensitivity Library Structure

The sensitivity library is a critical component for estimating the variance in delay of digital circuits due to process variations. For each standard cell and transition type (rise/fall), the library stores both sensitivity information and statistical data required for variance computation. All entries are characterized over a grid of input slews and output loads, similar to the Nonlinear Delay Model (NLDM) format.

1. Sensitivity Information

For a given standard cell, the following partial derivatives are stored:

- **Sensitivities of delay (D) and output slew (S_o) w.r.t. device parameters (y_i):**

$$\frac{\partial D}{\partial y_i}, \quad \frac{\partial S_o}{\partial y_i}$$

- **Sensitivities of delay (D) and output slew (S_o) w.r.t. input slew (S_i):**

$$\frac{\partial D}{\partial S_i}, \quad \frac{\partial S_o}{\partial S_i}$$

These sensitivities are stored as a 2D lookup table over a range of:

- **Input Slew:** typically from $10\,ps$ to $200\,ps$
- **Output Load:** typically from $0.5\,fF$ to $25\,fF$

2. Statistical Information for Device Parameters

To account for process variation, the sensitivity library also includes statistical descriptors for all device parameters (e.g., threshold voltage V_{th}, mobility μ, oxide thickness T_{ox}, channel width W).

- **Variances:**

$$\mathrm{Var}(y_i)$$

- **Covariances between device parameters:**

$$\mathrm{Cov}(y_i, y_j)$$

- **Correlation Coefficients**

$$\rho = \frac{\mathrm{Cov}(y_i, y_j)}{\sigma_{y_i}\sigma_{y_j}}$$

These values are usually extracted from the process design kit (PDK) and (ρ) value is provided by foundry.

3. Summary of Stored Quantities

The sensitivity library stores the following key items for each standard cell:

Stored Quantity	**Description**
$\frac{\partial D}{\partial y_i}$	Delay sensitivity to device parameter y_i
$\frac{\partial S_o}{\partial y_i}$	Output slew sensitivity to device parameter y_i
$\frac{\partial D}{\partial S_i}$	Delay sensitivity to input slew
$\frac{\partial S_o}{\partial S_i}$	Output slew sensitivity to input slew

Table 4.1: Sensitivity information stored per standard cell in the sensitivity library

Stored Quantity	**Description**
$\mathrm{Var}(y_i)$	Variance of parameter y_i
$\mathrm{Cov}(y_i, y_j)$	Covariance between y_i and y_j
ρ	Correlation coefficient between LVT and HVT devices

Table 4.2: Variance, covariances and Correlation Coefficient global information for all Cells

All these values are tabulated for a range of input slews and output loads to allow for accurate interpolation and delay variance estimation in arbitrary circuit contexts.

Chapter 5

Experimental Validation of Proposed Analytical Model

5.1 Experimental Setup

In this study, we utilized the 45nm gpdk045 technology model provided by cadence [18]. All circuit simulations were carried out using Cadence Virtuoso [19], with Monte Carlo analysis performed through Cadence ADE-XL [20]. A total of 2500 Monte Carlo runs were executed to extract the statistical characteristics, specifically, the mean and standard deviation of the inverter delays.

5.1.1 Experimental setup for Single Inverter Delay D case

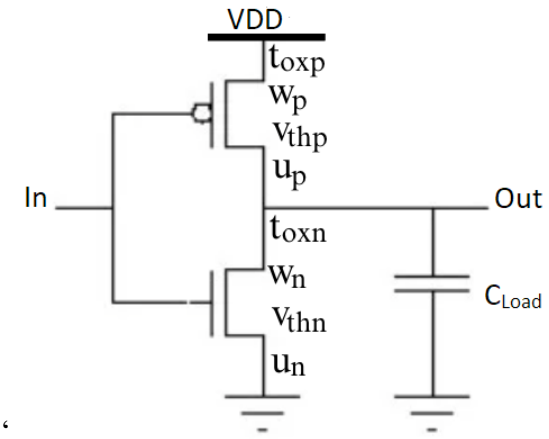


Figure 5.1: Schematic of single inverter with the 4 key Parameters.

We have considered both types of inverters—LVT and HVT—for computing the standard deviation (σ) of the delay, as illustrated earlier. The nominal values of the relevant device parameters, as specified in the technology model file, for all four transistor types (LVT PMOS, LVT NMOS, HVT PMOS, and HVT NMOS) are summarized in Table 5.1.

Devices	t_{ox0} (nm)	w_0 (nm)	v_{th0} (V)	μ_0 (m^2/Vs)
LVT NMOS	2.41	120	0.24925	0.0152801
LVT PMOS	2.41	200	-0.202698	0.0230515
HVT NMOS	2.40	120	0.42225	0.0152801
HVT PMOS	2.40	180	-0.368698	0.0230515

Table 5.1: Nominal values of device parameters for different transistor types given in the model file

Delays were extracted using circuit shown in Fig 5.1. The device parameters and simulation setup values are shown in Tab 5.2

	w_n(nm)	w_p (nm)	Input Slew Range(ps)	C_L (fF)
LVT INVERTER	520	780	10-200	5
HVT INVERTER	520	780	10-200	5

Table 5.2: Simulation setup values of Single Inverter

5.2 Calculation of variance of Delay D for Single Inverter using Analytical Model and Characterized Sensitivity Library

We use the schematic shown in Figure 5.1 to compute the delay variance of a single inverter implemented with either low threshold voltage (LVT) or high threshold voltage (HVT) transistors. This computation is based on Equation (3.10), which models the propagation of process variations into delay variability.
To calculate the delay variance for this configuration, we extract the following sensitivity information from the characterized sensitivity library:

- Sensitivity of the inverter delay D_1 with respect to device parameters such as threshold voltage (V_{th}), mobility (μ), etc.
- Sensitivity of the output slew S_l (at the inverter output) with respect to the same device parameters.

These sensitivities are obtained by performing a lookup in the characterized sensitivity table, indexed by the given values of:

- Input slew S_{IN} applied at the input of the inverter.
- Output load C_L connected at the inverter output.

The sensitivity library contains pre-characterized 2D lookup tables for both LVT and HVT inverter variants, enabling accurate extraction of delay and slew sensitivities for any operating condition of interest. These values are then used to evaluate the delay variance using the analytical expression provided in Equation (3.10), thereby eliminating the need for full Monte Carlo simulation for each case.

$$\text{Var}(D) = \sum_{i=1}^{8} c_i^2 \,\text{Var}(Y_i) + 2 \sum_{i=1}^{7} \sum_{j=i+1}^{8} c_i c_j \,\text{Cov}(Y_i, Y_j)$$

5.2.1 Comparison of Results obtained using the proposed model with the experimental results(MC) Single Inverter Case

We have compared the standard deviation of inverter delays obtained through Monte Carlo simulations with those computed using our proposed analytical model. The analytical approach leverages a pre-characterized sensitivity library to compute the delay variance (σ^2), accounting for process variations. The results of this comparison are summarized in Tables 5.3–5.6, which include both the Monte Carlo-based and

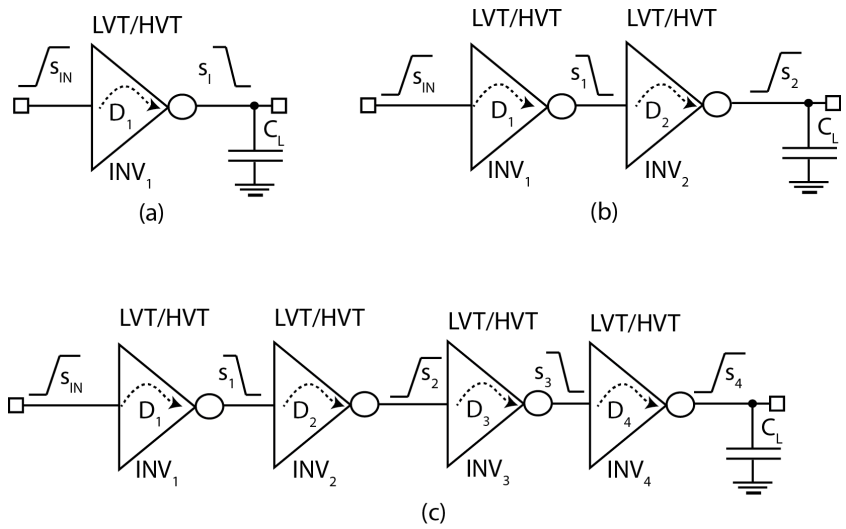


Figure 5.2: Various configurations of inverter chain (a) single inverter (b) inverter pair (c) chain of four inverters.

computed standard deviation (σ) values, as well as the corresponding percentage error between them. This comparison validates the accuracy of the proposed model across a wide range of input slew values.

Input Slew (ps)	**MC (ps)**	**Comp. (ps)**	**% Error**
10	0.36	0.35	2.78
20	0.40	0.41	-2.50
30	0.45	0.44	2.22
40	0.54	0.51	5.56
50	0.64	0.62	3.13
60	0.74	0.73	1.35
75	0.90	0.89	1.11
100	1.17	1.14	2.56
150	1.68	1.65	1.79
200	2.19	2.17	0.91

Table 5.3: Comparison of Standard Deviation of Delay of LVT inverter (Fall Case) between Monte Carlo σ and Computed σ by proposed model

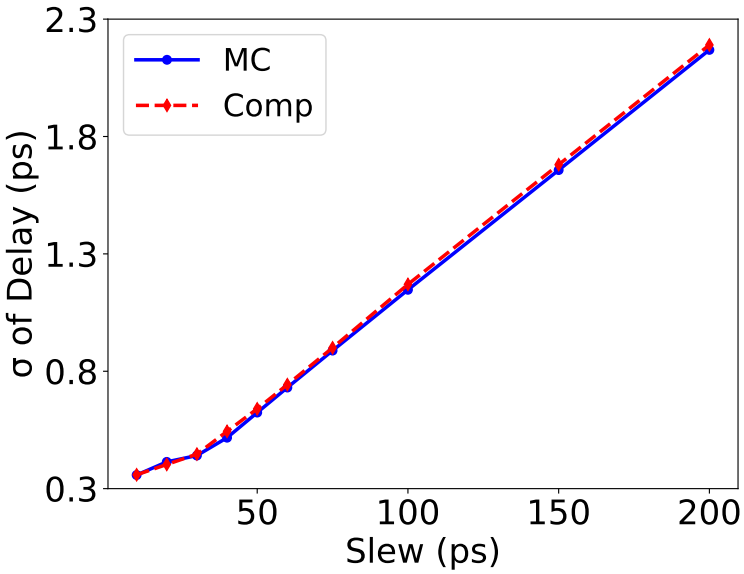


Figure 5.3: Standard deviation of fall delay for different slews of an LVT inverter ($C_L = 5fF$)

Input Slew (ps)	**MC (ps)**	**Comp. (ps)**	**% Error**
10	0.28	0.29	-3.57
20	0.32	0.33	-3.13
30	0.36	0.37	-2.78
40	0.42	0.44	-4.76
50	0.52	0.53	-1.92
60	0.61	0.62	-1.64
75	0.76	0.77	-1.32
100	0.99	1.01	-2.02
150	1.46	1.50	-2.74
200	1.93	1.96	-1.55

Table 5.4: Comparison of Standard Deviation of Delay of LVT inverter (Rise Case) between Monte Carlo σ and Computed σ by proposed model

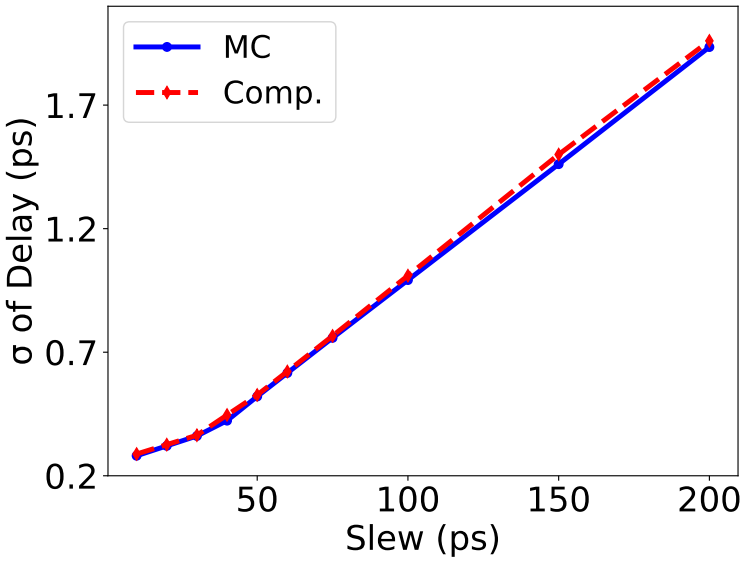


Figure 5.4: Standard deviation of Rise delay for different slews of an LVT inverter ($C_L = 5fF$)

Input Slew (ps)	**MC (ps)**	**Comp. (ps)**	**% Error**
10	0.71	0.70	1.41
20	0.74	0.73	1.35
30	0.78	0.77	1.28
40	0.82	0.81	1.22
50	0.88	0.84	4.55
60	0.91	0.900	1.10
75	0.97	0.98	-1.03
100	1.16	1.15	0.86
150	1.69	1.68	0.59
200	2.23	2.20	1.35

Table 5.5: Comparison of Standard Deviation of Delay of HVT inverter (Fall Case) between Monte Carlo σ and Computed σ by proposed model

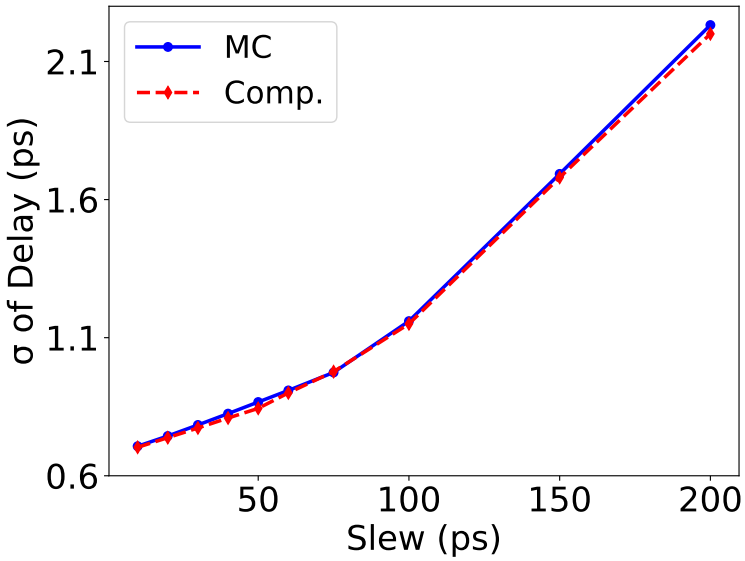


Figure 5.5: Standard deviation of fall delay for different slews of an HVT inverter ($C_L = 5fF$)

Input Slew (ps)	**MC (ps)**	**Comp. (ps)**	**% Error**
10	0.51	0.520	-1.96
20	0.56	0.55	1.79
30	0.59	0.58	1.69
40	0.62	0.61	1.61
50	0.67	0.65	2.99
60	0.70	0.69	1.43
75	0.77	0.81	-5.19
100	0.97	1.01	-4.12
150	1.43	1.44	-0.70
200	1.90	1.89	0.53

Table 5.6: Comparison of Standard Deviation of Delay of HVT inverter (Rise Case) between Monte Carlo σ and Computed σ using the proposed model

The comparison between the proposed model and Monte Carlo (MC) simulations for the standard deviation of inverter delay, as shown in Table 5.3–5.6 and Figure 5.3–5.6, demonstrates strong agreement for the LVT fall delay case. Across all evaluated cases—LVT rise, LVT fall, HVT rise, and HVT fall—the maximum error remains under 6%. Specifically, the maximum errors observed are: 5.22% for LVT fall,

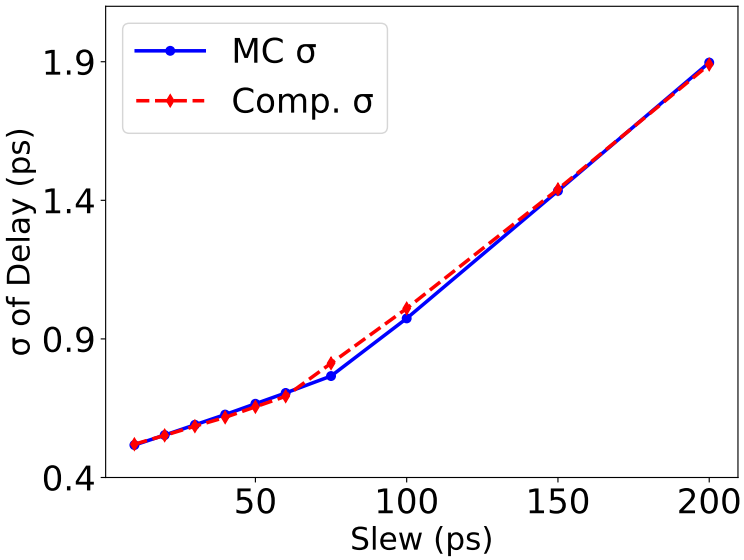


Figure 5.6: Standard deviation of fall delay for different slews of an LVT inverter ($C_L = 5fF$)

5.44% for LVT rise, 2.65% for HVT fall, and 6.01% for HVT rise. All figures show a consistent trend: as the input slew increases, the standard deviation of delay (σ) also increases. Moreover, the proposed model closely follows the MC results across the full range of slews, validating its accuracy and reliability. Therefore, the model can be confidently used for timing analysis in both LVT and HVT scenarios.

5.3 Calculation of variance of Delay D for Chain of Inverter using Analytical Model and Characterized Sensitivity Library

To compute the delay variance of an inverter chain comprising both low threshold voltage (LVT) and high threshold voltage (HVT) transistors, we extend the framework used for single inverters by incorporating sensitivity propagation across the stages. This computation relies on Equation (3.8), which models how variations in device parameters affect the total path delay through both direct and propagated sensitivities.

For each inverter stage in the chain, we extract the following sensitivity information from a pre-characterized sensitivity library:

- Sensitivity of the inverter delay D_i with respect to transistor parameters such as threshold voltage (V_{th}), carrier mobility (μ), oxide thickness, etc.
- Sensitivity of the output slew S_i with respect to the same device parameters.

- Sensitivity of the delay to input slew, i.e., $\frac{\partial D_i}{\partial S_{i-1}}$, and sensitivity of output slew to input slew, i.e., $\frac{\partial S_i}{\partial S_{i-1}}$.

These sensitivity values are obtained from 2D lookup tables in the library, indexed by the input slew and output load of each inverter stage. The sensitivity library supports both LVT and HVT standard cell variants.

In a chain of N inverters, global variation in the parameters of an upstream stage (e.g., y_1 in the first inverter) can indirectly affect the delays of downstream stages due to changes in the propagated slew. This dependency is captured through *sensitivity propagation*, where the delay sensitivity of inverter i with respect to a parameter y_k is computed as:

$$\frac{\partial D_i}{\partial y_k} = \frac{\partial S_k}{\partial y_k} \cdot \prod_{j=k+1}^{i} \frac{\partial S_j}{\partial S_{j-1}} \cdot \frac{\partial D_i}{\partial S_{i-1}}$$

This recursive formulation ensures that the influence of global variation is accurately tracked throughout the chain.

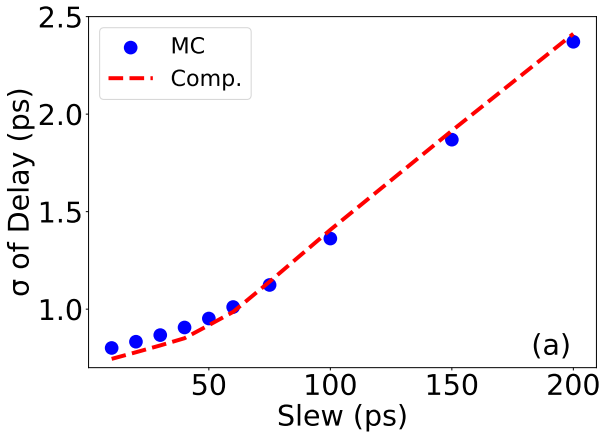
By summing the sensitivities of all inverter stages and applying the analytical expression for delay variance, we can compute the total path delay variability efficiently, without resorting to full-scale Monte Carlo simulations.

5.3.1 Comparison of Results obtained using the proposed model with the experimental results(MC) for Pair of Inverter Case

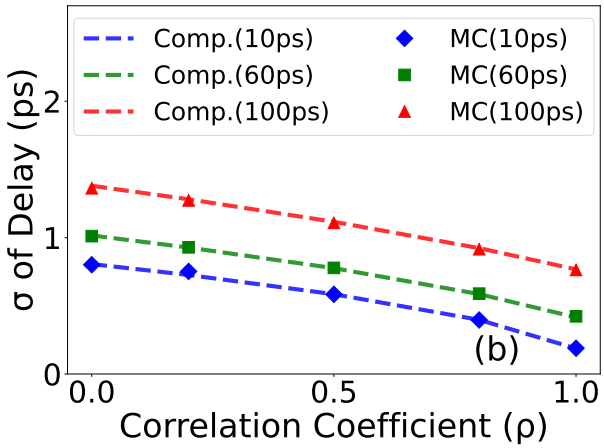
We compare the standard deviation of inverter pair delays obtained from Monte Carlo simulations with those computed using our analytical model, which leverages a pre-characterized sensitivity library to estimate delay variance (σ^2) under process variations and got less error,this validates the model's accuracy across various input slew conditions. In this case, parameter variations in the LVT inverter influence the HVT delay via output slew. Using sensitivity propagation, our model captures this effect accurately, showing close agreement with Monte Carlo results even in multi-stage scenarios.

Here, we analyze the delay variation in an inverter pair, as shown in Fig. 5.2(b), where the first inverter uses LVT devices and the second uses HVT devices. Monte Carlo (MC) simulations are performed to obtain the standard deviation of the total delay across various input slews (S_{IN}), and the results are shown in Fig. 5.7(a). As expected, the standard deviation increases with higher input slew values.

We then calculate the delay standard deviation using the proposed analytical model, which relies on pre-characterized delay sensitivities (from the sensitivity library) and



(a) Standard deviation of $D = D_1 + D_2$ for different input slews S_{IN} at $\rho = 0.8$ ($C_{L1} = 5fF$,$C_{L2} = 2.82fF$)



(b) Standard deviation of $D = D_1 + D_2$ for different ρ and input Slew ($C_{L1} = 5fF$,$C_{L2} = 2.82fF$)

Figure 5.7: Standard deviation of $D = D_1 + D_2$ for varying input slews S_{IN} ,varying ρ

Eq. 3.8, using $M = 16$ samples. The sensitivities are extracted based on slew and load conditions provided by from Characterized Sensitivty Librarary. Additionally, the sensitivity of the HVT inverter's delay (D_2) to variations in the LVT inverter's parameters is computed using Eq. 3.14. From Fig. 5.7(a), we observe that the standard deviations predicted by our model closely match the MC results, with the percentage error remaining below 7% across all input slews. This confirms the accuracy of the proposed model in estimating the total delay variation of an inverter pair.

We also study the effect of parameter correlation between LVT and HVT devices. Fig. 5.7(b) shows how the total delay standard deviation (σ_D) changes with different

correlation coefficients (ρ) for multiple input slews. The results reveal that as ρ increases, the total standard deviation decreases. This is because a higher correlation leads to a stronger cancellation of the delay variations contributed by each stage.
The first inverter's fall delay (D_1) is primarily affected by the NMOS transistor, while the second inverter's rise delay (D_2) is mostly influenced by the PMOS transistor. Since their delay sensitivities respond oppositely to the same variation, a higher correlation between NMOS and PMOS parameters causes these effects to partially cancel out, reducing the total delay variability. Our analytical model incorporates this effect through the covariance terms in Eq. 3.8. The close match between the computed and simulated values, with a maximum error under 7%, confirms that the proposed model effectively captures correlation-driven cancellation, making it suitable for timing analysis of multi-stage circuits. This error mainly arises due to interpolation from the coarse 5×5 NLDM table. Using a finer-grained table(10×10 NLDM table) with more entries would further reduce this error.

5.3.2 Comparison of Results obtained using the proposed model with the experimental results(MC) for chain of Inverter containing 4 inverter in chain

We investigate the impact of device variations in a 4-stage inverter chain, shown in Fig. 5.2(c), where each inverter can be implemented using either LVT or HVT devices. This results in $2^4 = 16$ possible configurations. For convenience, each configuration is represented by a 4-character string using 'L' (or '0') for LVT and 'H' (or '1') for HVT inverters.
We first analyze three representative configurations: LLLL, LHLH, and HLHL. Monte Carlo (MC) simulations are performed across different input slews S_{IN} and a correlation coefficient $\rho = 0.8$. The standard deviations of the total delay $D = D_1 + D_2 + D_3 + D_4$ for these configurations are plotted in Fig. 5.8. As seen, the standard deviation increases with input slew, consistent with earlier observations in single inverter and inverter pair cases. Moreover, the standard deviation differs across configurations, which is expected since the variation in input slew for a given inverter depends on delay variations in its preceding stages. We analyze this configuration-dependent behavior in more detail below.
We also compute the standard deviations using our proposed technique, which involves Eq. 3.8, a characterized sensitivity library, and sensitivity propagation. The computed results (also shown in Fig. 5.10) match closely with MC simulations,

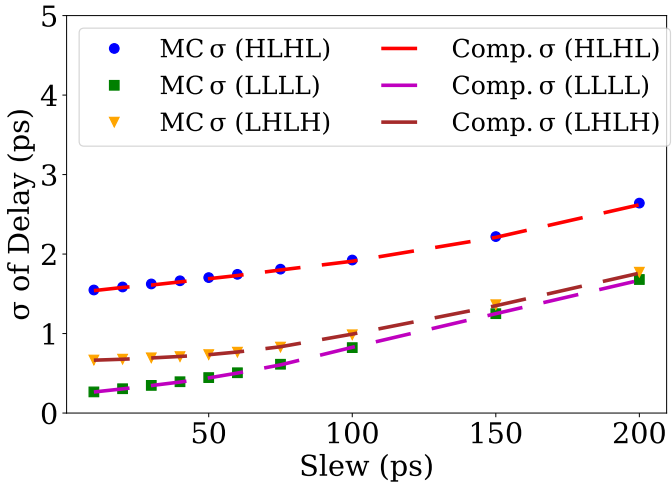


Figure 5.8: Standard deviation of D at different input slews for HLHL, LLLL and LHLH configurations ($\rho = 0.8$).

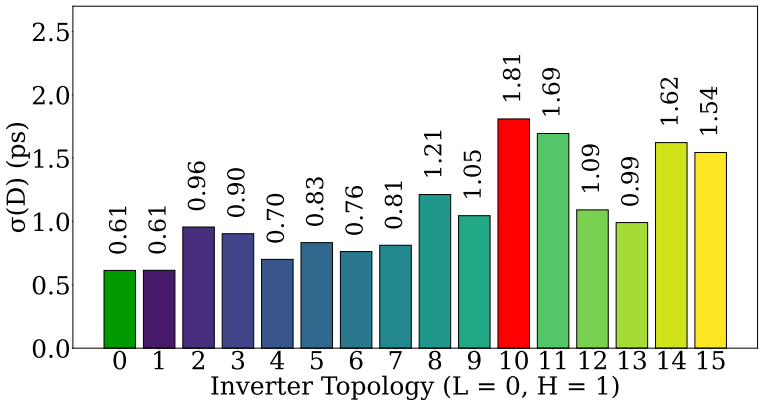


Figure 5.9: Standard deviation of total delay for 4-stage inverter configurations ($\rho = 0.8$, $S_{IN} = 75$ps). Configurations are encoded in decimal (LVT = 0, HVT = 1). For example, LLLL, LHLH, and HHHH are encoded as 0, 5, and 15 respectively.

demonstrating that our method accurately captures the effects of input slew and configuration by leveraging extracted and propagated sensitivities.

Next, we explore how configuration impacts the standard deviation of total delay across all 16 possibilities. Fig. 5.9 shows these results. The LLLL configuration has the lowest standard deviation (≈ 0.61 ps), while HLHL exhibits the highest (≈ 1.81 ps). This trend can be explained by the following:

- **Initial inverter type:** LVT inverters exhibit smaller output slew variations than HVT for the same input slew and output load. Hence, configurations starting with LVT devices (indices 0–7) generally show reduced variation in propagated

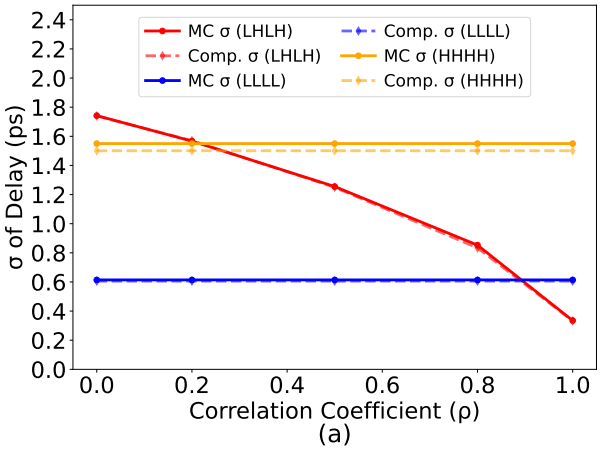
slews and total delay.

- **Majority device type:** LVT devices have lower delay sensitivities than HVT devices under similar conditions. Thus, configurations with a majority of LVT inverters tend to have lower standard deviations. This explains the higher variability seen in configurations dominated by HVT (e.g., 11, 13, 14, 15) compared to those dominated by LVT (e.g., 0, 1, 2, 3).
- **Compensation across stages:** Variation in one stage can be partially compensated by the next, depending on device type, transition direction, input slew, and load—since sensitivity propagation depends on these parameters. In configuration 10 (HLHL), compensation between consecutive stages is weaker, leading to a higher standard deviation than in configurations like HLHH (11) and HHHH (15).

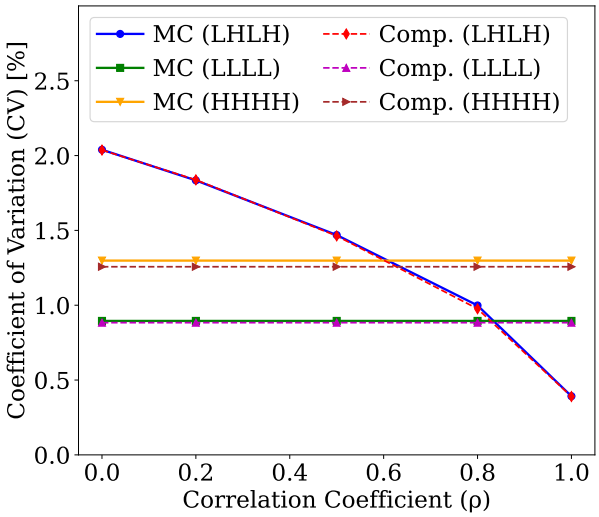
Interestingly, the configuration that results in the minimum standard deviation σ varies with the correlation coefficient ρ. As shown in Fig. 5.10(a), for $S_{IN} = 75\,\mathrm{ps}$, the LLLL configuration exhibits the lowest σ when $\rho = 0.8$. However, as ρ approaches 1, indicating perfect correlation among devices, the LHLH configuration demonstrates the lowest standard deviation.

This behavior can be attributed to the compensatory effect between alternating LVT and HVT inverters. At $\rho = 1$, variations in adjacent stages with differing threshold voltages (LVT-HVT) tend to cancel each other due to their perfect correlation, thereby reducing the overall delay variation.

Figure 5.10(b) shows a similar trend for the coefficient of variation (CV), supporting this conclusion. These observations suggest that circuit designers can exploit correlation-aware configuration selection to mitigate the impact of process-induced variations. Specifically, when ρ is close to unity, alternating LVT-HVT chains can be advantageous. Conversely, when ρ is lower, using an all-LVT configuration is more effective in minimizing delay variations.



(a) Standard deviation vs. ρ ($S_{IN} = 75\,\text{ps}$ for different configurations)



(b) Coefficient of variation (CV) vs. ρ ($S_{IN} = 75\,\text{ps}$ for different configurations)

Figure 5.10: (a) Standard deviation and (b) coefficient of variation (CV) of total delay as a function of correlation coefficient ρ, for LLLL, LHLH, and HHHH configurations.

Chapter 6

Conclusion

In this work, we developed a novel analytical framework to compute the variance in total path delay of digital circuits composed of mixed threshold voltage (LVT-HVT) cells, addressing the limitations of traditional STA methods that assume fully correlated device variations. Our method explicitly models the partial correlation between LVT and HVT devices through a covariance matrix and leverages a pre-characterized sensitivity library combined with a sensitivity propagation technique to accurately capture delay variations across different circuit configurations.
The proposed framework was rigorously validated using SPICE-based Monte Carlo simulations on single inverters, inverter pairs, and multi-stage inverter chains. Results show a close agreement, with error margins within 5%, demonstrating that the method effectively captures the influence of varying input slews, output load conditions, and correlation coefficients (ρ) on delay variability. Importantly, the framework not only provides accurate predictions but also delivers significant computational efficiency by eliminating the need for exhaustive statistical simulations.
Through detailed configuration-based analysis, we showed that:

- LVT-dominant configurations yield lower delay variability due to smaller delay and slew sensitivities.
- Alternating LVT-HVT configurations (e.g., LHLH) exhibit compensatory effects that reduce delay variation significantly when the correlation between device parameters is high ($\rho \approx 1$).
- When the correlation is weaker, uniform LVT configurations are preferable for minimizing variability.

While our study focused on CMOS inverters and inverter chains, the methodology is general and can be extended to other types of combinational logic gates. The use of a 2D sensitivity library indexed by slew and load allows this framework to be scalable and practical for full-chip statistical timing analysis. However, for more complex topologies involving reconvergent paths, further work is required to model and propagate sensitivities accurately in the presence of multiple fan-ins and shared dependencies.

In conclusion, this work provides a robust, accurate, and scalable technique for statistical analysis and optimization of multi-V_t circuits under partial correlation. It offers circuit designers a valuable tool for making correlation-aware configuration decisions, ultimately enabling more variation-tolerant and reliable digital systems.

References

[1] A. S. Petrov and I. V. Kudryavchenko, "Overview on executive report international technology roadmap for semiconductors (itrs) 2015," in *Recent Achievements and Prospects of Innovations and Technologies*, 2017, pp. 73–76.

[2] I. L. Markov, "Limits on fundamental limits to computation," *Nature*, vol. 512, no. 7513, pp. 147–154, 2014.

[3] F.-L. Yang, J.-R. Hwang, and Y. Li, "Electrical characteristic fluctuations in sub-45nm cmos devices," in *IEEE Custom Integrated Circuits Conference*, IEEE, 2006, pp. 691–694.

[4] P. A. Stolk, F. P. Widdershoven, and D. Klaassen, "Modeling statistical dopant fluctuations in mos transistors," *IEEE Transactions on Electron Devices*, vol. 45, no. 9, pp. 1960–1971, 1998.

[5] A. Malinowski, J. Chen, S. K. Mishra, S. Samavedam, and D. K. Sohn, "What is killing moore's law? challenges in advanced finfet technology integration," in *2019 MIXDES—26th International Conference on Mixed Design of Integrated Circuits and Systems*, IEEE, 2019, pp. 46–51.

[6] Y. Ye, F. Liu, M. Chen, S. Nassif, and Y. Cao, "Statistical modeling and simulation of threshold variation under random dopant fluctuations and line-edge roughness," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 19, no. 6, pp. 987–996, 2010.

[7] A. Zjajo, *Stochastic Process Variation in Deep-Submicron CMOS*. Springer, 2016.

[8] A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub 50 nm mosfets: A statistical 3d atomistic simulation study," *Nanotechnology*, vol. 10, no. 2, p. 153, 1999.

[9] S. Saurabh, *Introduction to VLSI Design Flow*. Cambridge University Press, 2023.

[10] D. Blaauw, K. Chopra, A. Srivastava, and L. Scheffer, "Statistical timing analysis: From basic principles to state of the art," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, no. 4, pp. 589–607, 2008.

[11] C. Visweswariah, K. Ravindran, K. Kalafala, S. G. Walker, and S. Narayan, "First-order incremental block-based statistical timing analysis," in *Proceedings of the 41st Annual Design Automation Conference*, 2004, pp. 331–336.

[12] H. Chang and S. S. Sapatnekar, "Statistical timing analysis under spatial correlations," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 9, pp. 1467–1482, 2005.

[13] V. Khandelwal and A. Srivastava, "A general framework for accurate statistical timing analysis considering correlations," in *Proceedings of the 42nd Annual Design Automation Conference*, 2005, pp. 89–94.

[14] J. Le, X. Li, and L. T. Pileggi, "Stac: Statistical timing analysis with correlation," in *Proceedings of the 41st Annual Design Automation Conference*, 2004, pp. 343–348.

[15] M. Sentovich, "Sis: A system for sequential circuit synthesis," *Memorandum No. UCB/ERL M92/41*, 1992.

[16] J. P. de Gyvez and R. Rodríguez-Montañés, "Threshold voltage mismatch (ΔV_T) fault modeling," in *Proceedings of the 21st VLSI Test Symposium*, IEEE, 2003, pp. 145–150.

[17] X. Peng, H. Wang, S. Wang, and J. Du, "A new generation of static timing analysis technology based on n7+ process—pocv," in *2019 IEEE 4th International Conference on Integrated Circuits and Microsystems (ICICM)*, IEEE, 2019, pp. 199–203.

[18] Cadence, *Gpdk045 model file*, Access granted by IIIT-Delhi.

[19] Cadence Design Systems, *Cadence virtuoso*, Accessed: 2025-04-29. [Online]. Available: `https://www.cadence.com/`.

[20] Cadence Design Systems, *Cadence ade-xl*, Accessed: 2025-04-29. [Online]. Available: `https://www.cadence.com/`.

Paper Based on this Work

1. Mohd Abu Ubaida, Prashasti Pandey, Sneh Saurabh, Ajoy Mandal and Ramakrishnan Venkatraman. “Analysis of Circuits with Partially Correlated Multi-Vt Cell Variations using Sensitivity Modeling and Propagation” [Submitted in VDAT Conference]

Brief Introduction of the Author

Mohd Abu Ubaida completed his B.Tech in Electronics and Communication Engineering from Zakir Husain College of Engineering and Technology, Aligarh Muslim University (AMU), Aligarh in 2023. He pursued his M.Tech in Electronics and Communication Engineering from the Indraprastha Institute of Information Technology, Delhi (IIIT-Delhi).

He is particularly interested in digital circuit design, with a strong focus on the VLSI Design Flow from RTL to GDS. His areas of specialization include Physical Design and Static Timing Analysis (STA). He is currently undertaking an internship at **Qualcomm**, where he is gaining hands-on experience in Physical Design.

He can be reached at: ubaida78690@gmail.com or abu2316@iiitd.ac.in