UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROELETRÔNICA

FELIPE TODESCHINI BORTOLON

Static Noise Margin Analysis for CMOS Logic Cells in Near-Threshold

Thesis presented in partial fulfillment
of the requirements for the degree of
Master of Microeletronics

Advisor: Prof. Dr. Sergio Bampi
Coadvisor: Prof. Dr. Fernando Gehm Moraes

Porto Alegre
April 2018

"It's the questions we can't answer that teach us the most.

They teach us how to think.

If you give a man an answer, all he gains is a little fact.

But give him a question and he'll look for his own answers. "

— PATRICK ROTHFUSS, THE WISE MAN'S FEAR

ACKNOWLEDGMENTS

ABSTRACT

The advancement of semiconductor technology enabled the fabrication of devices with faster switching activity and chips with higher integration density. However, these advances are facing new impediments related to energy and power dissipation. Besides, the increasing demand for portable devices leads the circuit design paradigm to prioritize energy efficiency instead of performance. Altogether, this scenario motivates engineers towards reducing the supply voltage to the near and subthreshold regime to increase the lifespan of battery-powered devices. Even though operating in these regime offer interesting energy-frequency trade-offs, it brings challenges concerning noise tolerance. As the supply voltage reduces, the available noise margins decrease, and circuits become more prone to functional failures. In addition, near and subthreshold circuits are more susceptible to manufacturing variability, hence further aggravating noise issues. Other issues, such as wire minimization and gate fan-out, also contribute to the relevance of evaluating the noise margin of circuits early in the design. Accordingly, this work investigates how to improve the static noise margin of digital synchronous circuits that will operate at the near/subthreshold regime. This investigation produces a set of three original contributions. The first is an automated tool to estimate the static noise margin of CMOS combinational cells. The second contribution is a realistic static noise margin estimation methodology that considers process-voltage-temperature variations. Results show that the proposed methodology allows to reduce up to 70% of the static noise margin pessimism. Finally, the third contribution is the noise-aware cell design methodology and the inclusion of a noise evaluation of complex circuits during the logic synthesis. The resulting library achieved higher static noise margin (up to 24%) and less spread among different cells (up to 62%).

Keywords: Subthreshold. digital circuit. SNM. noise tolerance. digital cell design.

RESUMO

Os avanços na tecnologia de semicondutores possibilitou que se fabricasse dispositivos com atividade de chaveamento mais rápida e com maior capacidade de integração de transistores. Estes avanços, todavia, impuseram novos empecilhos relacionados com a dissipação de potência e energia. Além disso, a crescente demanda por dispositivos portáteis levaram à uma mudança no paradigma de projeto de circuitos para que se priorize energia ao invés de desempenho. Este cenário motivou à reduzir a tensão de alimentação com qual os dispositivos operam para um regime próximo ou abaixo da tensão de limiar, com o objetivo de aumentar sua duração de bateria. Apesar desta abordagem balancear características de performance e energia, ela traz novos desafios com relação a tolerância à ruído. Ao reduzirmos a tensão de alimentação, também reduz-se a margem de ruído disponível e, assim, os circuitos tornam-se mais suscetíveis à falhas funcionais. Somado à este efeito, circuitos com tensões de alimentação nestes regimes são mais sensíveis à variações do processo de fabricação, logo agravando problemas com ruído. Existem também outros aspectos, tais como a miniaturização das interconexões e a relação de *fan-out* de uma célula digital, que incentivam a avaliação de ruído nas fases iniciais do projeto de circuitos integrados. Por estes motivos, este trabalho investiga como aprimorar a margem de ruído estática de circuitos síncronos digitais que irão operar em tensões no regime de tensão próximo ou abaixo do limiar. Esta investigação produz um conjunto de três contribuições originais. A primeira é uma ferramenta capaz de avaliar automaticamente a margem de ruído estática de células CMOS combinacionais. A segunda contribuição é uma metodologia realista para estimar a margem de ruído estática considerando variações de processo, tensão e temperatura. Os resultados obtidos mostram que a metodologia proposta permitiu reduzir até 70% do pessimismo das margens de ruído estática, Por último, a terceira contribuição é um fluxo de projeto de células combinacionais digitais considerando ruído, e uma abordagem para avaliar a margem de ruído estática de circuitos complexos durante a etapa de síntese lógica. A biblioteca de células resultante deste fluxo obteve maior margem de ruído (até 24%) e menor variação entre diferentes células (até 62%).

Palavras-chave: Operação em regime sub-limiar, circuitos digitais, SNM, tolerância à ruído, projeto de células digitais.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

CAD   Computer-Aided Design

CMOS   Complementary Metal Oxide Semiconductor

COC   Current over Capacitance

CS   Calhoun Standard

DIBL   Drain-Induced Barrier Lowering

DNM   Dynamic Noise Margin

EDA   Electronic Design Automation

FF   Fast nMOS -Fast pMOS

HDL   Hardware Description Language

IoT   Internet-of-Things

KS   Keane Standard

KSE   Keane Standard Extended

MAM   Multiple Analysis Mode

MC   Monte Carlo

MEC   Maxmum Equals Criteria

MOS   Metal Oxide Semiconductor

MOSFET   Metal Oxide Semiconductor Field Effect Transistor

MPC   Maximum Product Criteria

NO   Noise Optimized

NS   Nabavi Standard

NSC   Negative Slope Criteria

PVT   Process-Voltage-Temperature

RSCE   Reverse Short-Channel Effect

RTL   Register Transfer Level

SAM	Single Analysis Mode
SET	Static noise margin Estimation Tool
SNM	Static Noise Margin
SRAM	Static Random-Access Memory
SS	Slow nMOS -Slow pMOS
TT	Typical nMOS -Typical pMOS
VTC	Voltage Transfer Characteristic
VS	Voltage Scaling

LIST OF SYMBOLS

α	Transistor stack factor
β	W_p/W_n design factor
γ	Body effect coefficient
λ_d	Drain-induced barrier lowering coefficient
μ	Monte Carlo mean
σ	Monte Carlo standard deviation
σ^2	Monte Carlo variance
I_{DS1}	Drain-source drift current
I_{DS2}	Drain-source diffusion current
I_{DS}	Drain-source current
L_n	nMOS transistor width
L_p	pMOS transistor channel length
NM_H	High noise margin
NM_L	Low noise margin
R_{I_T}	The resistance of transistor T, connected to input I
T_D	Average cell delay
V_t	Thermal Voltage
V_{DD}	Maximum supply voltage value
V_{IH}	Lowest acceptable input voltage to represents logical 1
V_{IL}	Highest acceptable input voltage to represents logical 0
V_{in}	Input Voltage
V_{OH}	Lowest output voltage that represents logical 1
V_{OL}	Highest output voltage that represents logical 0
V_{out}	Output Voltage

V_{SS}	Minimum supply voltage value, e.g., zero
V_{th}	Transistor threshold voltage
VBS	Body-source voltage
VDS	Drain-source voltage
VGS	Gate-source voltage
VLT	Logic threshold voltage
W_n	nMOS transistor width
W_p	pMOS transistor channel length

CONTENTS

1 INTRODUCTION

The number of transistors on a chip has exponentially increased, following Moore's law, over the past decades (MOORE, 2003). This increase led to the adoption of integrated circuits (IC) to perform several computing applications such as in the field of health-care, security, manufacturing, communication and many others. The widespread use of ICs in our daily life and the desire to interconnect those objects, therefore, culminated on the Internet of Things (IoT). The IoT enables these physical objects to see, hear, think, and execute jobs by "talking" together to share information and coordinate decisions (AL-FUQAHA et al., 2015). This network transforms traditional objects into smart devices by exploiting underlying technologies such as ubiquitous computing, embedded devices, communication technologies, sensor networks, Internet protocols, and so on. It is expected that the IoT will increasingly contribute to the quality of life and the world's economic growth.

The IoT environment created a high demand for portable devices and autonomous systems. Hence, this scenario motivated shifting traditional performance oriented IC design to prioritize energy consumption instead (HANSON et al., 2009a). Since IoT encompasses a wide range of applications, there are also many different power requirements. While some devices can be recharged on a daily basis, e.g., smart-phones, other battery-powered devices have strict power budgets, e.g., heart monitor. The later type relies on severe battery-efficiency, i.e., autonomy, and, in some cases, may need energy harvesting mechanisms to enhance its battery life time. Besides these constraints, device integration density is facing new impediments related to energy and power dissipation (DRESLINSKI et al., 2010). Formerly, the supply voltage and threshold voltage scaled about the same factor as the feature size. Thus, designers were able to obtain the corresponding decrease in switching power. Consequently, the power consumption per chip area, i.e., power density, remained approximately constant as technology advanced. This phenomenon is referred as the Dennard Scaling model (DENNARD et al., 1974). Nonetheless, in deep sub-micron technologies dominated by leakage, reducing the threshold voltage results in an exponential increase in leakage power (SHAFIQUE et al., 2014). Since the threshold voltage is no longer scaling, the supply voltage cannot be scaled further without impacting performance. Altogether, the power density has surpassed the amount of power that chips can safely dissipate creating the so-called Dark Silicon era (SHAFIQUE et al., 2014; ESMAEILZADEH et al., 2011).

On the other hand, not all systems require the inherited performance (and problems) of most advanced technology nodes. There are many IoT applications, such as sensors, that can explore more mature technology nodes, e.g., 65nm - 180nm, to minimize power density issues. Moreover, using mature technologies have a profound impact on the budget as their fabrication process involves fewer steps and processes. Despite that, power dissipation is still a problem as these technologies have higher supply voltages and thus higher power consumption. At the circuit design level, a widely adopted approach to decrease energy consumption is to reduce the supply voltage to lower values, i.e., near- and subthreshold regimes. This approach is commonly known as supply voltage scaling or simply voltage scaling (VS). This technique is very effective because the power dissipation in an integrated circuit is quadratically proportional to the supply voltage. Hence, small voltage reductions lead to significant energy savings. Operating at near-threshold or subthreshold offer different trade-offs for energy consumption and delay. In the subthreshold regime occurs the minimum energy operating point, i.e., maximum energy efficiency, but circuits witness severe delay degradation, i.e., ~50-100X more than superthreshold. Alternatively, the first offers moderate delay penalties, i.e. ~10x, and significant energy savings, i.e., ~10X less than super-threshold and ~2X more than subthreshold (DRESLINSKI et al., 2010). Choosing the proper region of operation depends on the target circuit and its requirements. Although reducing the supply voltage has a profound impact on the system performance, many applications are not concerned with this penalty. Some monitoring systems (HANSON et al., 2009b), for example, do not need to execute many instructions and thus can remain on standby for long periods. Such systems are the primary target devices for using mature technologies in IoT.

1.1 Motivation and Objectives

Even though operating at near- and subthreshold regimes offer interesting energy-frequency trade-offs, it brings a set of complications. The three key barriers that must be overcome are the performance loss, increased susceptibility to variations, and functional failures (AL-FUQAHA et al., 2015; HIMANSHU et al., 2012). Among these complications, this work focuses on the last, which is also influenced by variation susceptibility. As voltage reduces, the noise margins degrade and circuits become prone to functional failures. Also, circuits with large fan-in are more vulnerable to failure at low voltages due to worse on-current to off-current ratio compared to low fan-in circuits (DE; VANGAL;

KRISHNAMURTHY, 2017). Wire interconnect minimization also plays an important role in noise margin. This is because the delay of wires in recent technologies, i.e., below 180nm, is comparable to the delay of gates (EDENFELD et al., 2004). Therefore, the interconnect noise impact on signal characteristics and system performance also became significant (KAESLIN, 2008). Hence both voltage scaling and technology minimization affected the robustness of digital integrated circuits, making noise an essential metric in circuit design.

A common approach to assess a circuit tolerance to noise is through the concept of noise margins (KAESLIN, 2008). These margins define how much noise the circuits can endure without affecting its behavior. Accordingly, it is desirable to design circuits that have high noise margins, meaning they are more robust to the noise available in the IC. In this context, this work investigates noise in digital integrated circuits to enhance the available noise margins, i.e., the tolerance to noise. Moreover, it intends to investigate existing metrics and evaluate how to evaluate noise at near- and subthreshold regions properly. The results of this investigation are used to insert a noise analysis within the design project space and thus evaluate noise early in the design.

1.2 Contributions

The contributions of this work are three-fold: *(i)* the static noise margin (SNM) estimation tool (SET); *(ii)* the SNM estimation methodology; and *(iii)* the SNM-aware digital cell design methodology. These contributions were applied to the design of cells which are designed to operate at near-threshold or subthreshold. The contributions can be summarized as follows:

SNM estimation tool

The SET provides means to analyze the static noise margins of combinational cells automatically. Moreover, the tool was explicitly envisioned to assess circuits operating in the sub and near-threshold regime, which are more sensitive to process, voltage and temperature variations.

SNM estimation methodology

The SNM estimation methodology guarantees a consistent approach to estimate SNM considering process, voltage and temperature (PVT) variations. The suggested ap-

proach was capable of reducing inherited pessimism of the static estimations and thus avoid unnecessary circuit optimization.

SNM-aware digital cell design methodology

The main contribution of this work is a systematic approach to design combinational cells considering their static noise margin and traditional power and delay metrics. The resulting standard cell library from this methodology has higher SNM and thus are more resilient to functional failures, when compared to other state-of-art based subthreshold libraries. In addition, this work presents a simple but new methodology to evaluate the SNM of complex circuits after the logic synthesis.

1.3 Document Structure

The remaining of the document is organized as follows. Chapter 2 gives an introduction about noise in digital integrated circuits. This Chapter discusses the different types of noises, their effects (particularly for digital synchronous IC), the existing noise mitigation strategies, and contextualize this work. Chapter 3 presents the static noise margin estimation tool, and explains available features and how to interpret the output data. The discussion about the estimation methodology and its impact on SNM is explored in Chapter 4. The suggested methodology and the tool are explored in chapter 5 to investigate the SNM of existing subthreshold design strategies. This investigation examines the design choices impact on SNM for a wide voltage range, and opens a discussion for the main contribution in chapter 6. Based on the physical effects of transistor dimensions on the noise margins, this chapter presents the SNM-aware digital cell design methodology. Moreover, the proposed methodology is compared against the existing subthreshold design strategies with regard to delay, power, area and, naturally, SNM. Finally, Chapter 7 summarizes the dissertation contributions and discusses possible future development branches.

2 NOISE IN DIGITAL INTEGRATED CIRCUITS

In electrical circuits, noise is defined as undesired fluctuations in current or voltage that can interfere with the functional signal or indirectly degrade the system performance. This definition broadly expands to two primary categories that are device noise and switching noise, as shown in Figure 2.1. The first represents the intrinsic noise in the device while the second refers to noise-induced due to the switching activity of a digital circuit. The latter type of noise is typically two to three orders greater than the device noise. Hence this dissertation shall hold it as its primary focus. Nevertheless, this Chapter briefly explains the intrinsic noise to depict a complete picture of noise in integrated circuits. Accordingly, the remaining of this Chapter is organized as follows. Section 2.1 and 2.2 explain the sub categories of device noise and switching noise respectively. The effects of switching noise in digital IC are explored in Section 2.3, while Section 2.4 introduces the concept of noise margins. Finally, Section 2.5 reviews some methodologies to reduce switching noise, and Section 2.6 summarizes the Chapter.

Figure 2.1: Classification of noise in electrical circuits (SALMAN, 2009).
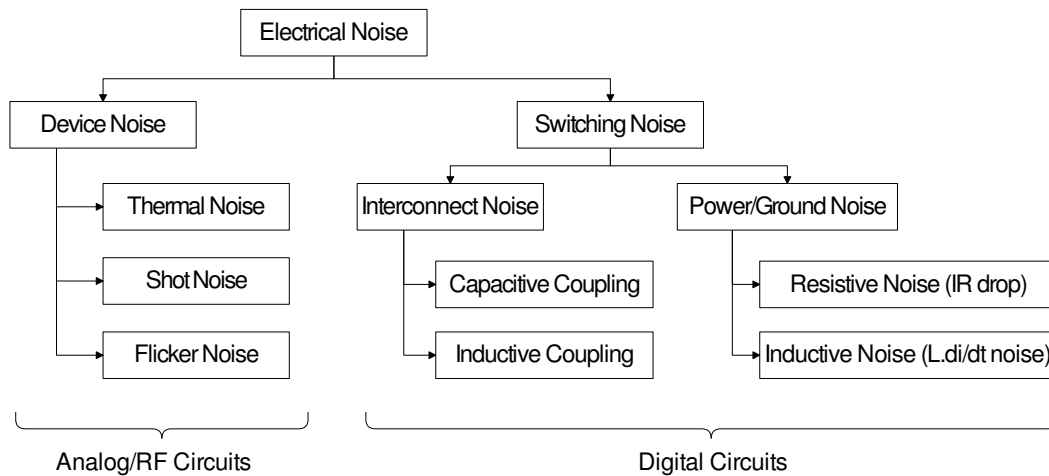
2.1 Device Noise

The device noise arises from intrinsic properties of the device and can be represented as a random signal. In this category there exists three main types: *(i)* shot noise, *(ii)* thermal noise, and *(iii)* flicker noise.

Shot Noise (GRAY, 2010) is associated with fluctuations in the direct-current flow present in diodes, MOS transistors and bipolar transistors. To understand the origin of this

noise, consider the p-n junction diode. In this device, the forward current is composed of holes from the *p* region and electrons from the *n* region that have sufficient energy to overcome the potential barrier at the junction. After the carriers have crossed the junction, they will diffuse as minority carriers. The passage of each carrier depends on it having enough energy and a velocity directed towards the junction, which is a random event. Therefore, even though external current appears to be steady, it actually is composed of random independent current pulses.

Thermal Noise, also known as Nyquist noise or Johnson noise, is generated by an entirely different mechanism from shot noise. It occurs in a conductor due to the random motion of its charge carriers. This random motion is generated by the thermal agitation of electrons at equilibrium, which happens regardless of any applied voltage. Since charge in motion constitutes an electrical current, there must appear some manifestation of this irregular motion in the conductor's terminals (BENNETT, 1960).

Finally, the *Flicker Noise* has varied sources, but is mainly caused by traps associated with contamination and crystal defects. These traps capture and release carriers in a random fashion and the time constants associated with the process give rise to a noise signal with energy concentrated at low frequencies (GRAY, 2010).

2.2 Switching Noise

Distinctly from device noise, this disturbance emerges from the high-to-low and low-to-high logic transitions of a digital circuit. Its two primary manifestation forms are the interconnect noise, or crosstalk noise, and the power/ground noise.

Crosstalk is an interference provoked by unwanted coupling from a neighbor signal wire to a network node. This coupling can be both capacitive and inductive. The capacitive crosstalk is the dominant effect at current switching speeds, although inductive coupling forms a major concern in the design of the input-output circuitry of mixed-signal circuits (RABAEY; CHANDRAKASAN; NIKOLIć, 2007). Figure 2.2 depicts a net susceptible to noise, i.e., the victim, by two neighbors, i.e., aggressors (*a1* and *a2*), through capacitive coupling.

Power/Ground Noise (KAESLIN, 2008) is a type of interference that arises from the fact that, in current technologies, thousands of gates share the same VSS and VDD interconnect lines. The simultaneous switching of these gates requires a significant amount of current drawn from the power supply. This current, in turn, flows through the parasitic

Figure 2.2: Victim net with two coupling aggressors (CADENCE, 2001).
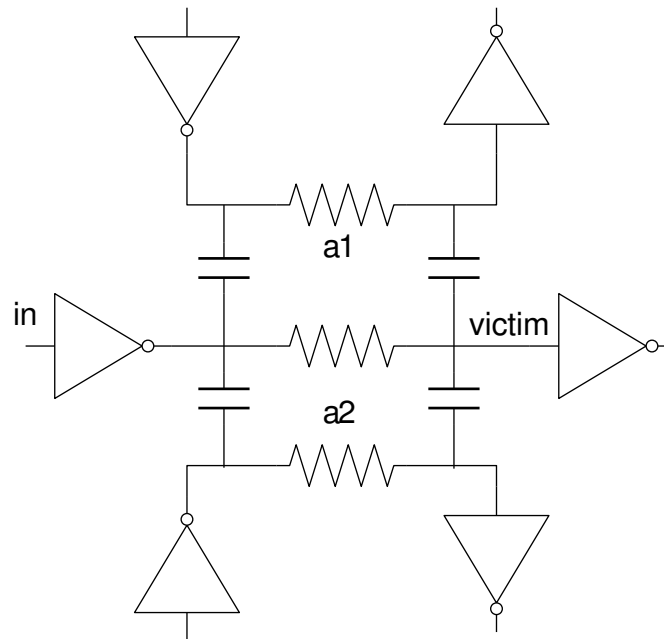
impedance of the power distribution network, causing both static and dynamic voltage fluctuations. This interference on the supply distribution network is also denominated as voltage droop. The ground distribution network suffers from a similar effect, which is referred as ground bounce.

2.3 Switching Noise Effects

Switching noise affects the circuit operation in various ways. One of its effects is the increase in power consumption due to glitch signals. A glitch is a spurious transition in the input or output of a logic gate due to capacitive or inductive coupling, or the circuit switching activity. This transition, i.e., voltage spike, causes the circuit to dissipate unnecessary power in two forms. If the peak voltage is higher than the threshold voltage, the transistor turns on momentarily, dissipating dynamic power. Conversely, if the peak is smaller than the threshold voltage, the glitch contributes to the static power consumption due to leakage current.

Switching noise also generates delay uncertainty in the circuit logic paths or even functional failure. The delay uncertainty originates from a neighbor-net switching activity, i.e., the aggressor, on a victim net during a transition. In this case, the noise can modify the time of flight and slew-rate of the useful signal, and it can cause delay (timing) errors. Aggressors *a1* and *a2* in Figure 2.2, for example, can increase the signal delay

on the victim net, as illustrated in Figure 2.3. The yellow waveform represents the victim without crosstalk interference, while the green when it is attacked by the aggressors switching activity. If the affected signal is part of a critical maximum delay path, then the extra delay can cause the signal to arrive too late at a flip-flop resulting in a timing failure (CADENCE, 2001). Conversely, the noise can also speed up an on-going transition. Noise-induced *jitter*, i.e. uncertainty on signal transition delay, is most vulnerable in (KAESLIN, 2008):

- Clock signals,
- Signals with a small setup margin, and
- Signals with a small hold margin.

Figure 2.3: Crosstalk increasing the signal delay of the victim net (CADENCE, 2001).
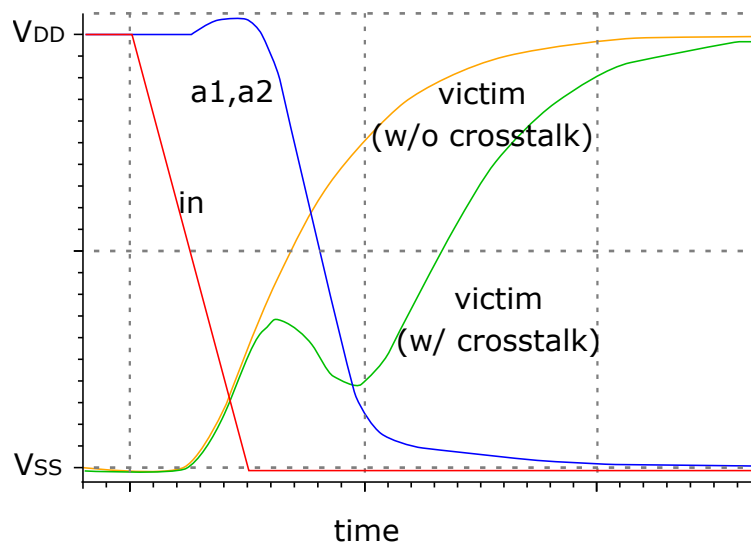
Functional failure occurs when the noise is induced in quiet victim nets by the switching activity of its neighbors, or by the supply voltage. For high levels of induced currents, it can cause unwanted logic activity and even alter the signal value, i.e., the logical state, causing a functional failure. The maximum noise pulse that can be accepted by a gate when used in a system while still giving correct operation is called noise tolerance. A form to represent this parameter is through the noise pulse-width and its amplitude, as in Figure 2.4. Formerly, noise tolerance was also referred as noise margin, and noise immunity was defined as the likelihood of a spurious signal being generated in the circuit (HILL, 1968). Conversely, other authors define noise immunity as a special case of noise tolerance where the noise source is only applied to the input of the gate (KATOPIS, 1985).

Moreover, noise tolerance is called as noise margin only when the noise source is applied both at the input and the power supply of the gate. Given that there is no de facto standard definition, this work adopts Hill's definition (HILL, 1968), since it suffices a single definition of noise tolerance for all cases. Therefore, noise margin and noise tolerance are used interchangeably without loss of its meaning.

Figure 2.4: Typical noise tolerance curve of a gate based on the noise amplitude and width (KATOPIS, 1985).
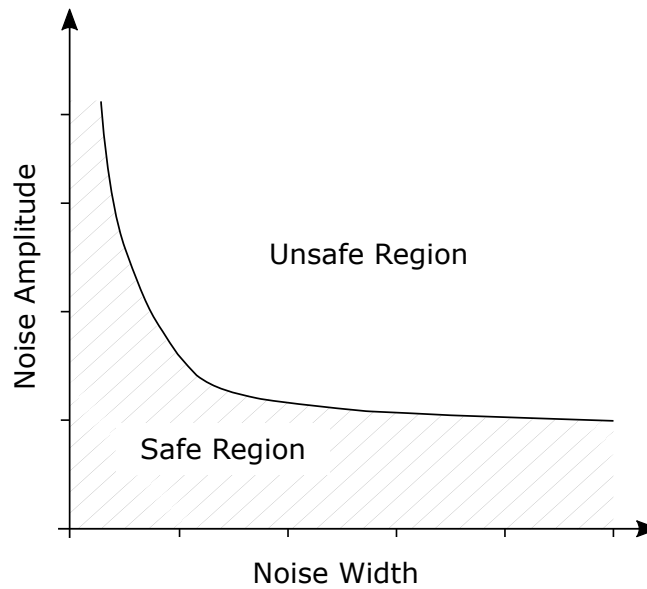
2.4 Noise Estimation Methods

It is desirable that the noise margin of a gate is sufficiently high to ensure that noise at any node does not disrupt the behavior of the current and the next gates. For that purpose, there exists two main metrics to evaluate the noise margin of a circuit. They are the static noise margin (SNM), which is based on a DC voltage transfer characteristic curve (VTC), and the dynamic noise margin (DNM), which is based on the time domain characteristics of the noise. This Section investigates these metrics and assesses their trade-offs for performance and accuracy.

2.4.1 Static Noise Margin

Digital circuits rely on a two-value scheme where a node either holds a logical one or zero. These two states are electrically represented by separate, and non-adjacent, voltage ranges. Gates must generate output values that fall into these ranges to be correctly interpreted by their next gates. To better understand this concept, consider two cascade inverters depicted in Figure 2.5. Let V_{OH} indicate the lowest output voltage produced by a circuit when driving a logical 1 and, analogously, V_{OL} the uppermost voltage when at logical 0. Further let V_{IH} denote the lowest input voltage that gets safely interpreted as a 1, and V_{IL} the highest voltage that gets recognized as 0. Voltages between V_{IH} and V_{IL} could be interpreted as logical 1 by some circuits and 0 by others. This interval is said to form an invalid region, and it represents how much noise the output signal can tolerate to be correctly interpreted at the next gate input. From this definition two possible noise margin arise: high (2.1) and low (2.2) noise margins. The lesser of those defines the maximum noise that can be safely admitted without compromising the circuit's correct behavior and is, therefore, known as **Static Noise Margin** (SNM) (2.3) (KAESLIN, 2008).

$$NM_H = V_{OH} - V_{IH} \tag{2.1}$$

$$NM_L = V_{IL} - V_{OL} \tag{2.2}$$

$$NM = \min(NM_H, NM_L) \tag{2.3}$$

Figure 2.5: Definition of noise margin for cascaded inverter gates.
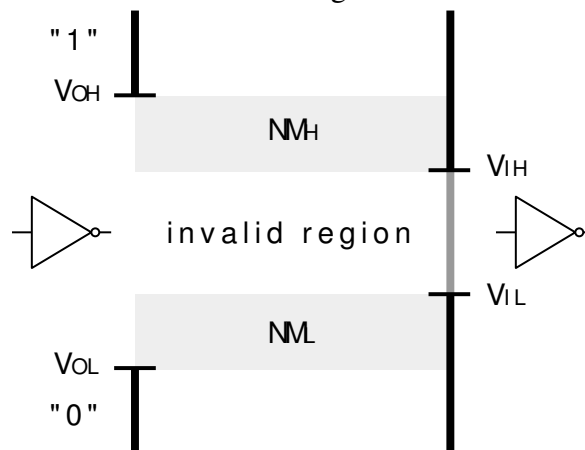
2.4.1.1 Negative Slope Criteria

The definition of V_{IL}, V_{IH}, V_{OL}, and V_{OH} depend on which criteria is adopted. One of the most simplistic approaches, taught in every Computer Major degree, is the negative slope criteria (NSC). The NSC determines these values from two critical points of a gate VTC, where its gain is unity, i.e. $\partial V_{out}/\partial V_{in} = -1$ (RABAEY; CHANDRAKASAN; NIKOLIĆ, 2007). Figure 2.6 illustrates how this criteria extracts these parameters for the Equations (2.1), (2.2), and (2.3).

Figure 2.6: Parameter definition for the negative slope criteria.
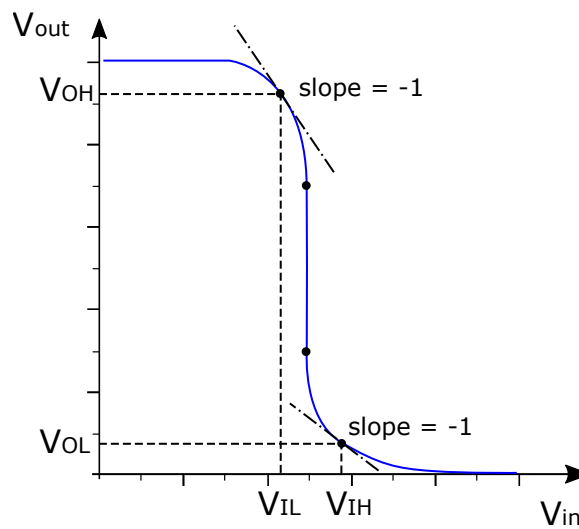
2.4.1.2 Maximum-Equals Criteria

Instead of using these equations, in 1968, Hill (HILL, 1968) created a methodology to graphically estimate the SNM of two back-to-back cells through butterfly plots. The butterfly plot consists of two VTC curves plotted with their axis mirrored, as illustrated by the blue and green curves in Figure 2.7(a). The resulting shape of both curves resembles the two wings of a butterfly, giving origin to its name. Accordingly, Hill defines that the SNM is the side of the largest square that fits in the smaller of those wings. This criteria forces equal high and low noise margins, i.e., $SNM = NM_H = NM_L$, and thus it is referred as the maximum equals criteria (MEC).

The simulation methodology of this approach was later proposed by Seevinck (SEEVINCK; LIST; LOHSTROH, 1987), in 1987, to automatically calculate the SNM. Although this approach methodology is usually used on memory cells, e.g., SRAM, there are no restrictions on using it with combinational cells. As later Chapters demonstrate,

some authors have adopted this strategy.

2.4.1.3 Maximum-Product Criteria

In 1993, Hauser (HAUSER, 1993) introduced a variant to the MEC, which also uses the butterfly plot, called the maximum product criteria (MPC). Instead of imposing equal low and high noise margins, Hauser proposes maximizing the area of a rectangle, i.e., $max(NM_H * NM_L)$, as depicted in 2.7. According to the Author, this approach is preferable to MEC since enforcing equal high and low noise margin appears to be too restrictive. Despite that, Chapter 4 demonstrates through a case study that this might not always be the case for subthreshold circuits.

Figure 2.7: Butterfly plot of two cross-coupled inverters for (a) MEC and (b) MPC

(a) Maximum equals criteria.   (b) Maximum product criteria
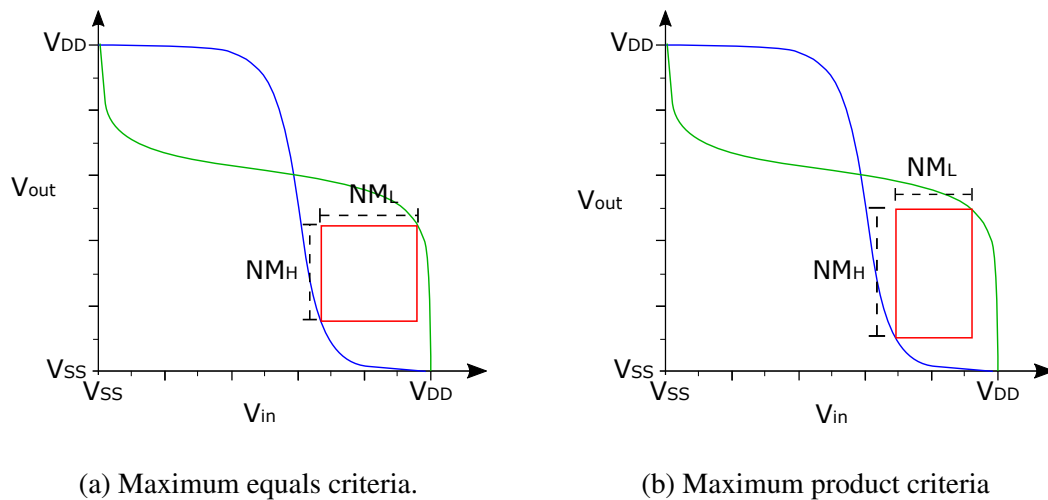
2.4.2 Dynamic Noise Margin

The static noise margin indicates the maximum DC noise that may be withstood by a gate during an infinitely long time without bringing it to the wrong state. If the disturbance is present in a pulse-form the noise amplitudes are allowed to be higher than the static margins without affecting the proper logic states (LOHSTROH, 1979) (Figure 2.4). The concept of Dynamic Noise Margin (DNM), therefore, introduces time-domain characteristics, such as noise amplitude, width, and duration.

Zurada et al. (ZURADA; JOO; BELL, 1989) demonstrates that the noise margins depend on the input rise time and output load capacitance. According to the author, the static VTC only approximates the dynamic behavior if the input transitions are very slow.

Depending on the input transition, DNM may have higher high or low noise margins than SNM due to the movement on V_{IL} and V_{IH} points, respectively. Hence the SNM is a conservative, i.e., pessimistic, approach to estimate noise margin in logic gates.

On the other hand, exact calculations of dynamic margins are very difficult due to the number of parameters involved, hence, promoting computer simulations as a suitable approach. Sephard et al. (SHEPARD; CHOU, 2000) propose a time domain DC noise sensitivity as a failure criterion, where the transient noise characteristics have been considered. His model uses four parameters (V_T, β, G_r and C_{int}) for each cell input for each noise type. Note the complexity involved in the computation where, as demonstrated for an AOI gate, fitting the equation required 250 cases for each input and each noise type. Even though computation complexity is not thoroughly discussed, through its equations is possible to assume that it is not trivial. Ding et al. (DING; MAZUMDER, 2004) proposes a different model from Shepard's which calculates the maximum square on AC transfer curves, instead of traditional the DC approach. The maximum square considerably varies depending on the noise duration w as well as the output capacitance of C_L. Thus several scenarios are considered. Another contribution from this article is a discussion indicating that solely using DNM values as a noise tolerance metric may lead to incorrect conclusions and, hence, proposes an analytic solution. Overall, the pessimism associated with SNM is reduced with dynamic noise margins at the expense of increased computational complexity due to the requirement to calculate the time domain sensitivities (SALMAN, 2009).

2.5 Techniques for Switching Noise Reduction

This Section briefly describes some existing methodologies to alleviate the effects of switching noise in digital IC. Section 2.5.1 summarizes the ones to reduce power/ground noise, while Section 2.5.2, to reduce crosstalk noise. A complete review of this subject is available in (SALMAN, 2009).

2.5.1 Power/Ground Noise Reduction

The logic gates of an IC usually share a common path to their power supplies. Whenever a subcircuit draws current from this path, other elements perceive a voltage

change in the power supply, i.e., noise. If this change is large enough, it may affect the operation of these other elements. To decouple other subcircuits from the effect of the sudden current demand, a decoupling capacitor (DOWNING; GEBLER; KATOPIS, 1993) can be placed between the supply voltage line and its reference, i.e., ground, next to the switched load. The idea behind this strategy is that the capacitor will initially supply the current demand. Ideally, by the time it runs out of charge, the power supply line inductance is saturated and the circuit can draw full current at the normal voltage from the power supply. While the power supply provides the necessary current, the capacitor can recharge. Another approach to further minimize the peak current drawn by the simultaneous switching activity of several gates is to delay the switching time of signals or clock paths (HEYDARI; PEDRAM, 2003). Authors show that by inserting a chain of buffers in the signal path to the output drivers, the ground bounce was attenuated by 65%.

In addition, electromagnetic interference (EMI) is also a concern for high speed synchronous digital systems. These circuits are driven by a clock signal and, due to its period nature, have a very narrow frequency spectrum. A perfect clock, in fact, would have its energy concentrated at a single frequency and its harmonics. Therefore, because of these characteristics, the resulting radiated electromagnetic energy, at specific frequencies, can exceed the regulatory limits for EMI. Clock dithering, or spread-spectrum clock generation (OTT; OTT, 2009), avoids this problem by spreading this energy over a wider bandwidth and thus reducing power/ground noise. This methodology, however, exhibits strong trade-offs between the noise attenuation and the system speed. Finally, the package characteristics influence in the noise of integrated circuits. The package material type, the number of pads, the spacing and interconnects, and others, all influence the electrical characteristics of the circuit. A review on this subject, nonetheless, is outside the scope of this dissertation.

2.5.2 Crosstalk Noise Reduction

A common way to enhance signal integrity and minimize delay uncertainty is to place shields around victim signal lines. There are two types of shields: the passive and the active. Passive shields (Figure 2.8(a)) are power or ground lines inserted between the aggressor and its victim to reduce the capacitive and inductive coupling (ZHANG; FRIEDMAN, 2004). Alternatively, active shields (Figure 2.8(b)) insert adjacent wires that will switch simultaneously with the protected line. This shielding scheme improves per-

Figure 2.8: Shield insertion technique to reduce crosstalk noise (KAUL; SYLVESTER; BLAAUW, 2002).

(a) Passive shield.
             (b) Active shield.
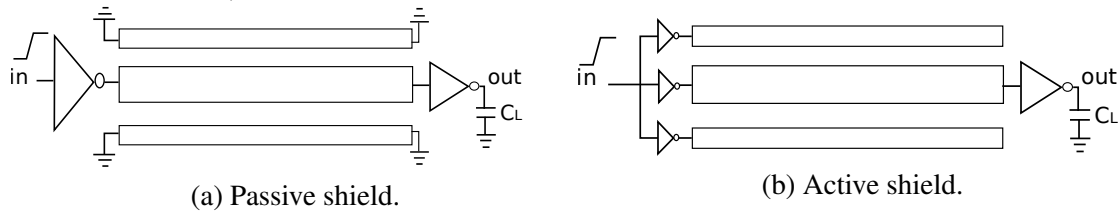
formance by up to 16% compared to passive shields (KAUL; SYLVESTER; BLAAUW, 2002). Nonetheless, the additional switching activity of this shield also increases the circuit power consumption. Both approaches impose additional area.

Breaking long interconnects into smaller subsections with repeaters also lessens the crosstalk noise among interconnects. A repeater, which is an inverter/noninverting buffer, is usually used to restore signal properties and speeding the line delay. Besides offering traditional power and delay advantages, splitting long lines reduces the coupling capacitance, thereby reducing coupling noise. Another approach is to use noise-based interconnect routing to increase the spacing and length of the overlap between victim and aggressor lines (GAO; LIU, 1993). However, inserting a shield line between two coupled interconnects is shown to be more efficient in reducing crosstalk noise than increasing their physical separation (KAUL; SYLVESTER; BLAAUW, 2002).

Finally, it is possible to exploit gate sizing of the aggressor and victim drivers to reduce coupling noise (JIANG; CHANG; JOU, 2000; BECER et al., 2003). Sizing up the victim gate increases its effective capacitance and more efficiently holds a net at a steady voltage value, i.e., V_{DD} or V_{SS}. Increasing the victim size, however, impacts overall area, and can potentially introduce new noise violations, i.e., the victim becomes the aggressor of another net. Alternatively, sizing down the aggressor gate decreases its effective conductance and as a result, noise induced on the victim decreases. Adjusting the aggressor size, however, is bound to slow down the signal path.

2.6 Chapter Summary

This Chapter presented a brief overview of switching noise in synchronous digital integrated circuits. The discussion comprehends from noise types and their origin to available methodologies to attenuate their impact on the system behavior. Accordingly, this work explores gate sizing strategies, but instead of analyzing the circuit after the

routing phase, it proposes to enhance gate noise margins during the cell design phase. The approach presented in the following Chapters intend to automate, study and understand noise margin and hence increase gate robustness to functional failures.

3 THE SNM ESTIMATION TOOL

Electronic Design Automation (EDA) is essential to the progress of semiconductor industry. Over the last quarter of a century, tools raised designer's productivity allowing them to develop, debug, and test complex chips, even with decreasing time-to-market. Without EDA advances, it would not be possible to sustain Moore's law and achieve current chip integration densities. Nevertheless, the integrated circuit evolution lead to new challenges, such as noise robustness. Chapter 1 contributes to this picture, demonstrating the relevance of this parameter in IoT devices. In addition, Chapter 2 shows that there are several methods to estimate noise with different trade-offs. In general, those are classified either as dynamic or static noise margins, respectively, DNM and SNM. The DNM methods produce more accurate results than SNM because it applies a more comprehensive analysis, i.e., considers the noise's temporal characteristics. Nonetheless, its simulation complexity is prohibitively impairing even for small cell subsets. Therefore, this work adopts SNM since it implies on fewer calculations, i.e., computation time, which is an essential aspect for evaluating thousands of digital cells and still provides a qualitative measurement of noise.

Current solutions evaluate noise interference, e.g., crosstalk, late in the design flow, i.e., at the post-routing stage, and thus may imply on re-execution of previous steps or the addition of overhead margins at preliminary design stages. Therefore, it is necessary to develop tools that can assess noise early in the design phase. To address this issue, this Chapter presents the SNM Estimation Tool (SET), which is the **first contribution** of this work. SET allows an engineer to evaluate the SNM of digital CMOS cells early in the design and take actions to optimize it. Section 3.1 presents a general overview of this tool and its functionality, while Section 3.2 explain how to operate the tool. Then, Section 3.3 delves into the details of its implementation for those who intend to continue this project.

3.1 The Tool

SET is a tool that can analyze the SNM of digital CMOS cells early in the design phase. It allows a digital designer to assess cell libraries or designs w.r.t. SNM, and implement techniques to enhance cell robustness. For that purpose, the tool determines the voltage transfer characteristics (VTC) of cells through detailed electrical simulations and, afterward, applies the SNM methods to estimate noise. In its current version, it is possi-

ble to employ either the maximum-equals criteria (2.4.1.2) or maximum product criteria (2.4.1.3) on combinational cells. The classical textbook method, i.e., the negative slope criteria (2.4.1.1), nonetheless, is not implemented because it is not a reliable estimation metric as demonstrated by Hauser et al. (HAUSER, 1993). Chapter 4 also shows that this criterion is more pessimistic than other approaches.

Besides the two SNM criteria, the tool supports temperature, voltage and process variations, which are important attributes to evaluate circuits operating near and below the threshold voltage. Process variations are usually categorized into global and local variations. For global variations on device parameters, such as oxide thickness or dopant concentration varies equally for all transistors. Wafer-to-wafer or lot-to-lot fall into that category. In contrast, for local variations - also known as mismatches, each transistor is affected differently. In other words, variations are distinguished by their spatial correlation distance. For local variations, there is no correlation, while for global variations the correlation is very large. SET supports both variations according to each technology provider specifications.

Figure 3.1: SNM Estimation Tool execution flow.
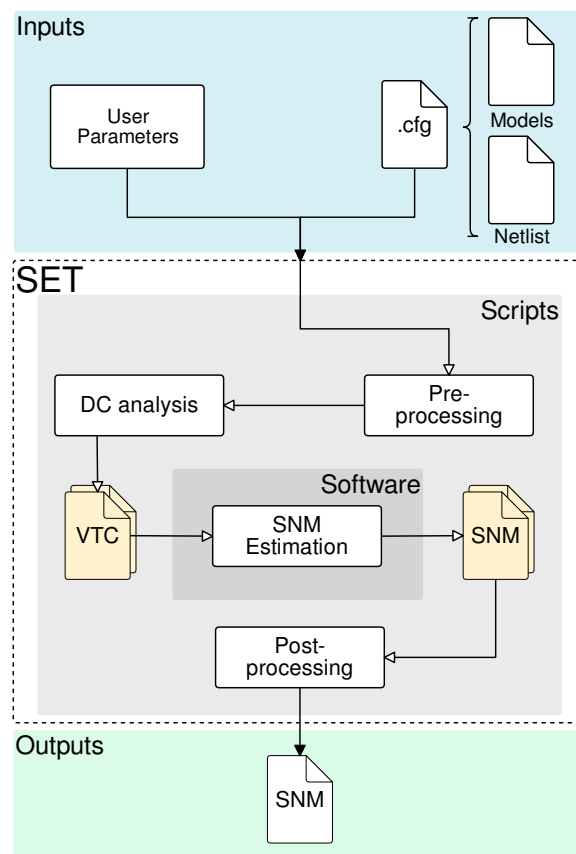
The SET tool flow, depicted in Figure 3.1, comprises a set of scripts that encapsulate a C-software. While the software implements the SNM estimation criteria, the scripts automate electrical simulations and data processing. This combination creates a user-friendly environment that enables a user to use SET with a single configuration file and a few parameters.

The **configuration file** (.cfg), contains attributes that specify the basic aspects of the simulation environment. These aspects are the technology model file, a cell library, the name and terminals of two cells of the library to analyze, the operating voltage, the temperature, and the simulation precision. An example of such file is depicted in Figure 3.2, which contains all the attributes in the same order as they were listed. In the depicted file it is specified that the tool should compute the SNM of two inverters operating at 250 mV and 27°C using the specified model/library and precision. While the .cfg file determines the basic electric characteristics, the **user parameters** define the analysis type. More specifically, it allows the user to execute different experiments regarding SNM. Current available options are (*i*) the technique selection; (*ii*) a list of cells to analyze; (*iii*) enable Monte Carlo simulation; and (*iv*) activate the parametric analysis for either the temperature or the supply voltage. In addition, it is possible to vary transistor dimensions as it will be explored in Chapter 6.

Given a complying configuration file, the **Pre-Processing** block parses its information and any other given options, sets up the environment and loads the electrical simulator. This work uses Cadence® Spectre® SPICE simulator, but any other may be used with minor adjusts. Afterwards, the **DC Analysis** invokes the SPICE simulator, generates the VTC curves, and adjusts the output to a readable format for the SNM estimation software. In addition, this block handles the Monte Carlo and parametric analysis, varying the attributes according to the user parameters. For instance, it may execute the electrical simulations for a set of voltage and temperature conditions as determined by the user. The outcome of this step is a set of VTC files containing the DC response for a particular process, voltage, and temperature (PVT) condition. The **SNM Estimation** iterate over these files to assess the SNM and writes the output for the **Post-Processing** block, which analyzes the data before forwarding the results to the user. This step is crucial for the Monte Carlo analysis where the software produces multiple SNM estimations. Therefore, this block gathers all results and computes the mean (μ), the standard deviation (σ) and the variance (σ^2). Additionally, it generates ready-to-plot logs for a visual inspection of the samples spread and other parameters.

Figure 3.2: Example of a configuration file (.cfg) for SET.

```
1  _PATH_TO_MODELS_ cfg/ibm/130nm/include.scs
2  _PATH_TO_INCLUDE_ ~/research/ibm/130nm/lib/CORE130GP.spi
3  _GATE1_ IS130_IVX1 in out gnd gnd vdd vdd
4  _GATE2_ IS130_IVX1 in out gnd gnd vdd vdd
5  _VOLTAGE_ 0.250
6  _TEMPERATURE_ 27
7  _DCSIZE_ .000250
```

3.2 Modes of Operation

The SET has two operating modes. The Single Analysis Mode (SAM) and the Multiple Analysis Mode (MAM). Each mode provides different parameters that enable different types of analysis and understanding these modes is essential. Therefore, the next Sections explain each separately.

3.2.1 Single Analysis Mode

The Single Analysis Mode (SAM) comprises the SET core functionality and set of scripts. In this operating mode, it is possible to compute the SNM for only a single cell pair with a given temperature and voltage. In the configuration file from Figure 3.2, for example, the tool computes the SNM for two inverters operating at 250 mV and 27°C. This file is the only mandatory argument required by SET and uses the following syntax:

```
set -c config.cfg
```

Besides this file, there are two other **optional** parameters that define *(i)* the technique and *(ii)* if process variations should be taken into account. The first, denoted by the -t flag, accepts the mec and mpc arguments, which stands for maximum-equals criteria, and maximum-product criteria respectively. When omitting this flag, the tool uses the default method: the MEC. The second, denoted by the -m flag, activate the Monte Carlo simulation (mc). Additionally to the mc option, the user has to set the models file (Figure 3.2, first line) as devised by the technology provider. To deactivate process variations, simply omit this flag or set it to (default). In summary, the full SAM mode command and its available options are the following:

```
set -c config.cfg -t [bf|mpc] -m [mc|default]
```

Figure 3.3: Example of a list of cells for Multiple Analysis Mode.

```
1   IS130_IVX1 in out gnd gnd vdd vdd
2   IS130_IVX1 in out gnd gnd vdd vdd
3
4   IS130_IVX1 in out gnd gnd vdd vdd
5   IS130_NAND2X1 in vdd out gnd gnd vdd vdd
6
7           ...
```

3.2.2 Multiple Analysis Mode

The Multiple Analysis Mode (MAM) allows iterative simulations over SAM and, consequently, on the attributes defined in the configuration file. More specifically, the user may choose to dynamically vary the cell pair (_GATE_), the supply voltage (_VOLTAGE_) and the temperature (_TEMPERATURE_). For the first, represented by the -g flag, it is necessary to provide a file containing a list of cell pairs, formatted according to Figure 3.3. In this file, the pattern consists of two cells (lines 1 and 2), with their terminals indicated, followed by a blank line (line 3). The syntax to iterate on this list is the following:

```
set -c config.cfg -g list.cfg
```

Note that the -c option still is mandatory for the MAM since it references the models and other relevant information. For voltage and temperature variations the user has to use a parametric analysis, denoted through the -p flag and the options vs and temp respectively. To specify the parameter interval use the following notation:

```
set -c config.cfg -p [vs|temp] <begin> <end> <step>
```

where begin is the first point, that is incremented by step until it reaches the end value. If the range is not given, the tool uses a default value that depends on the selected parameter. Process variation (-m) and technique selection (-t) are also available in the iterative mode. In summary, the full MAM mode command and its available options are the following:

```
set -c config.cfg -t [bf|mpc] -m [mc|default]
     -g list.cfg -p [vs|temp] <begin> <end> <step>
```

3.3 Noise Estimation Output

The SET tool saves the output in textual and graphical representations. Both contain the description of the VTC curves, which builds the butterfly plot, and the static noise estimation as in Figure 3.4. The tool represents the estimation according to the selected criteria. It draws a square for the MEC, and a rectangle for the MPC. In the text, this information is written as the square's side for MEC and both sides of the rectangle for the MPC. Although it is more common to use MPC lower value, the tool saves both for completeness.

Figure 3.4: SET output format for (a) Maximum Equals Criteria (MEC) and (b) Maximum Product Criteria (MPC).

(a) MEC

(b) MPC
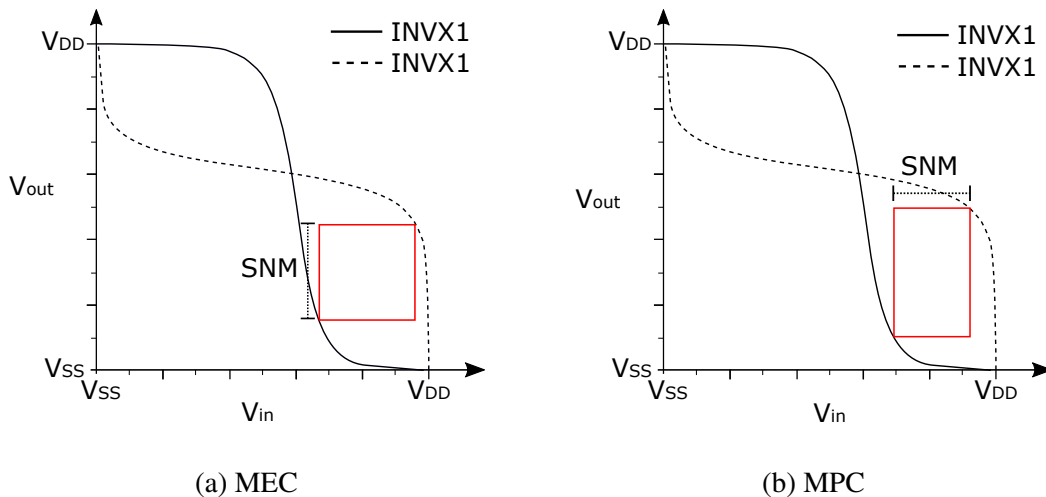
When PVT variations are considered, the SET produces this graph several times. Individually analyzing each graph would be cumbersome; hence the output is converted to express a more meaningful information. For Monte Carlo simulation, for instance, the output produces mean (μ), variance (σ^2), and standard deviation (σ) values, which represent a general behavior of all SNM estimations. Moreover, in the output of the tool in text format, the SNM is written for all PVT conditions, which are then easily converted to a histogram chart, such as the one depicted in Figure 3.5. This Figure demonstrates the Monte Carlo distribution for a pair of inverters for four PVT conditions. The (a) nominal, i.e., Typical nMOS and Typical pMOS (TT corner) at 27 °C, the nominal corner with (b and c) temperature variations, and (d) Fast nMOS and Slow pMOS (FS corner) at 27 °C. Given this Figure, it is easier to interpret and make a conclusion about SET results. In this case, the Figure shows the effect of temperature on SNM, i.e., inversely proportional, and the effect of process variations, i.e., transistors with opposite characteristics (fast/slow)

results in lower SNM. This analysis is considered in more detail in later chapters. The objective here is to demonstrate the SET applicability solely.

Figure 3.5: Example of a Monte Carlo analysis considering mismatch, temperature and corner for an inverter pair.

(a) Corner TT @ 27 °C

(b) Corner TT @ 125 °C

(c) Corner TT @ -25 °C

(d) Corner FS @ 27 °C
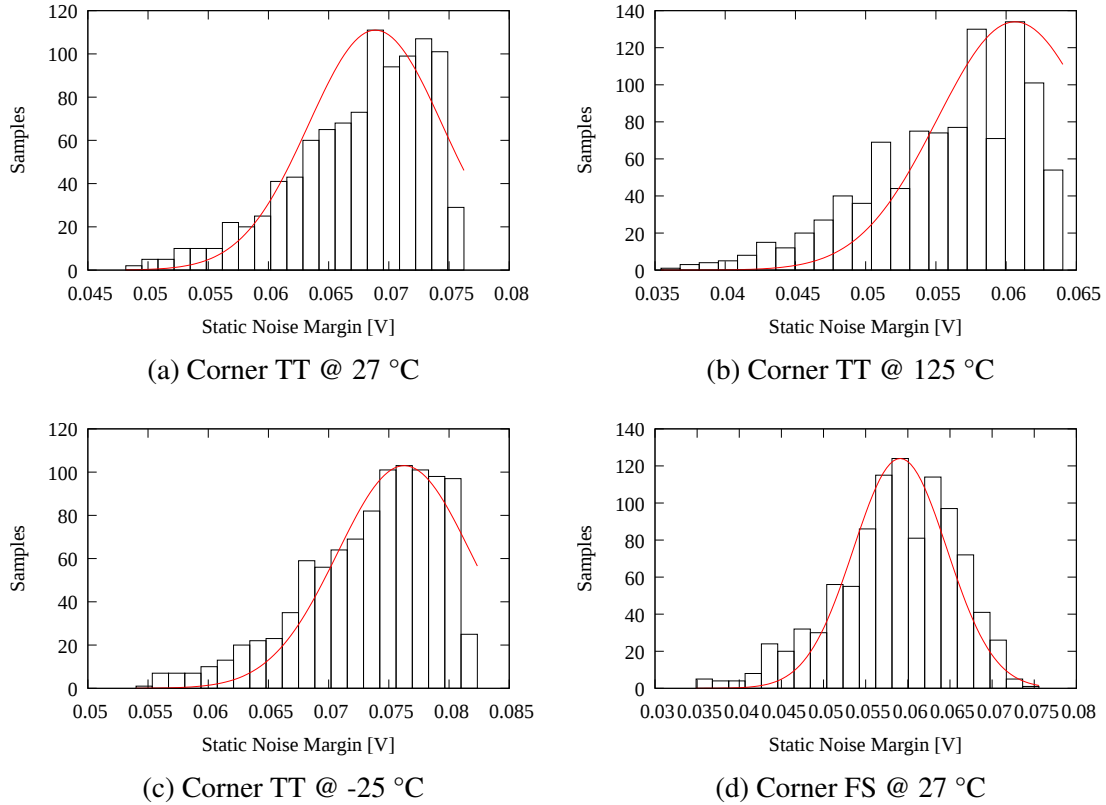
For temperature and voltage variations SET similarly produces adequate figures that are used in later Chapters (Figure 4.8 and 5.2). With SET, there is also the possibility to automatically analyze multiple cell pairs which are formatted as a two column table relating the pair and the resulting SNM. A more complex derivation of this output is used in Chapter 4 (Table 4.3).

3.4 Chapter Summary

This Chapter described the behavior and a basic user guide for SET and demonstrated its versatility and usefulness to analyze SNM data easily. In its current version, SET offers two estimation methods (MEC and MPC), PVT variation and the automated estimation of multiple cells for combinational cells only. However, the tool can be easily extended to support sequential cells. For those who intend to use SET in their work, a more detailed discussion about its development and structure is available in Appendix A.

4 SNM ESTIMATION METHODOLOGY AND CRITERIA

SNM is a simple metric to evaluate the noise tolerance of digital circuits. Although it has been primarily used to design SRAM memories, some authors have recently demonstrated that it is also possible to employ the same techniques to combinational cells (KWONG; CHANDRAKASAN, 2006; BEIU et al., 2013b; BEIU et al., 2013a). The drawback of this approach, however, is the inherited pessimism of SNM estimations (ZURADA; JOO; BELL, 1989) when compared to DNM.

The accuracy of these estimations can significantly affect the noise-aware design cell flow, which is the main contribution of this work (Chapter 6). More specifically, improperly estimating noise can aggravate the SNM's pessimism and generate over-conservative circuit optimization, e.g., increasing cell area more than necessary. Therefore, it is necessary to understand the simulation methodology that extracts the VTC curves and how they imply on SNM values. Moreover, it is necessary to appropriately select an SNM estimation criteria that cope with PVT variations without inflating this issue. This study is the **second** contribution of this work, and it suggests a simulation setup and criteria that significantly improves noise estimation and avoids this pitfall.

The remaining of this Chapter is organized as follows. Section 4.1 studies the influence of the number of inputs and its position on the DC curves and, consequently, on the static noise margin results. The conclusions of this analysis allow establishing a consistent simulation setup, which is used to compare different criteria and select one in Section 4.2. Finally, Section 4.3 explores the Monte Carlo simulations to understand how the SNM metric behaves with corner, mismatch, and temperature variations. Experiments are validated through extensive SET tool simulations with cells in a 65-nm CMOS bulk technology, operating in the subthreshold region ($V_{th} = 373\ mV$).

4.1 The DC Simulation

This Section analyzes the influence of the DC simulation on the SNM and defines a setup that avoids unrealistic estimations. Section 4.1.1 evaluates how the number of inputs affects the SNM, and Section 4.1.2 the input position.

4.1.1 Number of Inputs

The techniques, i.e., criteria, presented in Chapter 2, calculate the SNM through the voltage transfer characteristic (VTC) curve properties, which represents the relationship between the input and output of a logic cell. This curve is directly extracted from a DC simulation by increasing the input voltage from V_{SS} to V_{DD}, and, afterward, it is handed over to a particular SNM criteria to calculate the noise margin. Although this is a straightforward method, there are no studies available on how to simulate logic cells with multiple inputs. As this Section will demonstrate, this has a severe effect on overall noise estimation.

To generate a VTC curve with those cells, there are two common approaches. The first assumes that all inputs switch simultaneously to the same voltage level. The second assumes that one or more inputs switch, while the remaining ones are fixed at a voltage level that is not logic masking, causing the output to switch. A NAND, for example, cannot have inputs set to 0 V; otherwise, the output will remain at V_{DD} regardless of any switching at the other inputs. This condition guarantees that a logic cell behaves as a simple inverter but with added series and parallel resistances. Following this observation, it is possible to evaluate the impact of the number of switching inputs on the VTC and consequently on the SNM.

Figure 4.1: Influence of the number of simultaneously switched inputs on the DC curve of a NAND4 (left) and nor4 (right) gates.
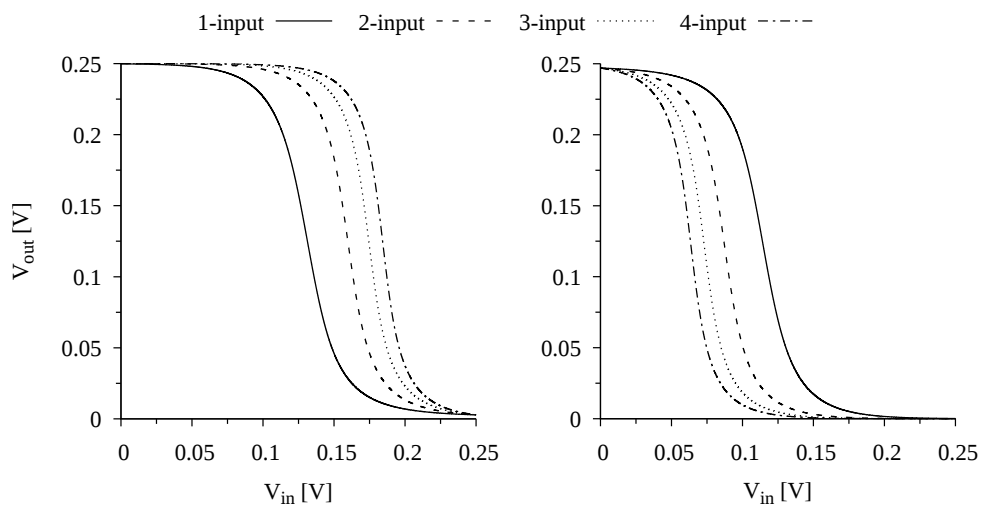
Figure 4.1 demonstrates the results of this analysis for a NAND and a NOR with four inputs, *NAND4* and *NOR4* respectively. As the number of switched inputs increases, observe that each cell deteriorates the logic level controlled by the stacked transistors, which represents the low (high) for the NAND4 (nor4). Since capacitive or inductively

Figure 4.2: Schematic of CMOS NAND cell with four inputs.
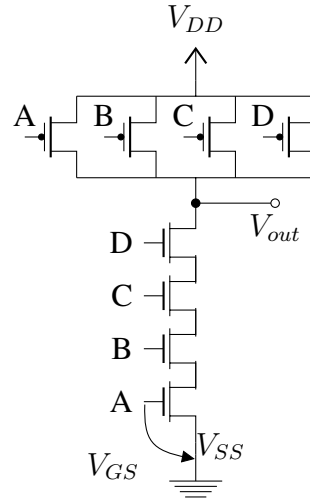
gate loadings are neglected in DC simulations, this variation must be a function of the channel/gate resistance. To understand how this figure scales, consider the NAND4 schematic depicted in Figure 4.2 and the general expressions for the pull-up (R_p) and pull-down (R_n) resistances:

$$R_p = R_{A_p} \parallel R_{B_p} \parallel R_{C_p} \parallel R_{D_p} \tag{4.1}$$

$$R_n = R_{A_n} + R_{B_n} + R_{C_n} + R_{D_n} \tag{4.2}$$

where R_{G_X} represents the resistance of transistor type X (pMOS or nMOS) with gate input G. First, let's examine the case where only input D varies from V_{SS} to V_{DD} while A, B, and C are set to V_{DD}. With this configuration, pMOS transistors A, B and C are in cut-off (high resistivity, R_{max}) and the equivalent nMOS transistors are in the linear region (low resistivity, R_{min}). Both D controlled transistors have their resistance dependent on V_{GS}, but the pMOS will transition from R_{min} to R_{max}, i.e., transistor channel begins open and ends closed, while the nMOS will behave inversely. Given that $R_{max} >> R_{min}$ for digital circuits, Eqs. (4.1) and (4.2) simplify to:

$$R_{p_1} = \begin{cases} \left(\frac{R_{max}}{3} \parallel R_{min} \right) \approx R_{min} & \text{for } V_{in} = V_{SS} \\ \frac{R_{max}}{4} & \text{for } V_{in} = V_{DD} \end{cases} \tag{4.3}$$

$$R_{n_1} = \begin{cases} 3R_{min} + R_{max} \approx R_{max} & \text{for } V_{in} = V_{SS} \\ 4R_{min} & \text{for } V_{in} = V_{DD} \end{cases} \tag{4.4}$$

On the other hand, if all four inputs switch simultaneously, pMOS transistors begin with minimum resistance while nMOS with maximum resistance. Accordingly, Eqs. (4.1) and (4.2) simplify to:

$$R_{p_4} = \begin{cases} \frac{R_{min}}{4} & \text{for } V_{in} = V_{SS} \\ \frac{R_{max}}{4} & \text{for } V_{in} = V_{DD} \end{cases} \tag{4.5}$$

$$R_{n_4} = \begin{cases} 4R_{max} & \text{for } V_{in} = V_{SS} \\ 4R_{min} & \text{for } V_{in} = V_{DD} \end{cases} \tag{4.6}$$

From these equations it is possible to observe that both scenarios have the same resistance at the last DC point, i.e., $V_{in} = V_{DD}$, which is expected. Nevertheless, the starting point, i.e., $V_{in} = V_{SS}$, when switching four inputs has much lower (higher) pull-up (pull-down) resistance. Since transistor resistance is proportional to V_{GS}, R_{p_4} and R_{n_4} will have overall lower and higher resistance than R_{p_1} and R_{n_1}, respectively, making it necessary to apply significant more input voltage to transition from logical one to zero. As V_{in} increases, both systems will converge.

Table 4.1: ΔSNM for different cell pairs, varying the number of inputs switching simultaneously (1 and 4).

Cell Pair	ΔSNM
NAND4-NOR4	68.53%
NAND3-NOR3	53.10%
NAND2-NOR2	33.06%
NAND4-NAND4	20.50%
NOR4-NOR4	22.75%

This shift on the VTC curve produces a significant impact on the static noise margin analysis. With the *maximum-square* method (HILL, 1968), for example, the noise margin estimate for a NAND4 and nor4 gates decreases 68% when comparing presented scenarios, as can be seen through the square reduction depicted in Figure 4.3. As the number of stacked transistor reduces (Table 4.1), the variation between using all (S_a) and only one (S_o) input(s) decreases ($\Delta SNM = (S_o - S_a)/S_o$), but it still is significant. Therefore, it is not recommend to switch more than one input when using SNM methods. Moreover, multiple inputs will rarely transition simultaneously, and such analysis would implement a conservative approach to evaluate noise tolerance.

Figure 4.3: Wing size variation for a NAND4 and nor4 butterfly as the number of inputs switched simultaneously increases.
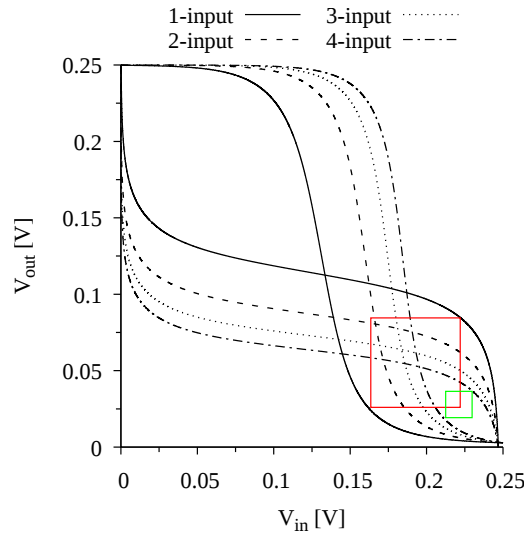
4.1.2 Input Location

The SNM is pessimistic by nature (ZURADA; JOO; BELL, 1989), and to avoid further degrading the estimates, the previous discussion suggested using only one input. Nonetheless, an aspect that was not addressed is how the switching input location, related to the stack of transistors, affects this metric. Using the same NAND4 and NOR4 cells, Figure 4.4 illustrates that despite the input location the produced VTC is very similar. Consequently, the noise estimate only exhibits a small variation of 2% for the NAND4 and NOR4, as depicted in Figure 4.5. To scrutinize this result consider once more the NAND4 schematic in Figure 4.2 and the curves D and A, which represent the corner cases.

Assuming that only input D, the closest one to the output, switches, the V_{GS} from nMOS transistors A, B and C are equal to V_{DD} and thus they are in the on-state. On the opposite case, if only input A switches, nMOS transistors B, C, and D do not possess any connection to the ground rail and, for that reason, are in cut-off. Given this difference, it is possible to perceive that the latter case will slightly transition afterward because its fixed-input nMOS transistors rely on transistor A to start conducting while in the other case, fixed input transistors are always conducting.

To further illustrate this statement, consider Alioto et al. (ALIOTO, 2010) transistor modeling for subthreshold operation. In his analysis the MOS transistor is either represented by a current-source when $V_{DS} \gg v_t$ (4.7) or by an equivalent resistor when

Figure 4.4: Influence of the location of the switched input on the DC curve of a NAND (left) and NOR (right) gates.
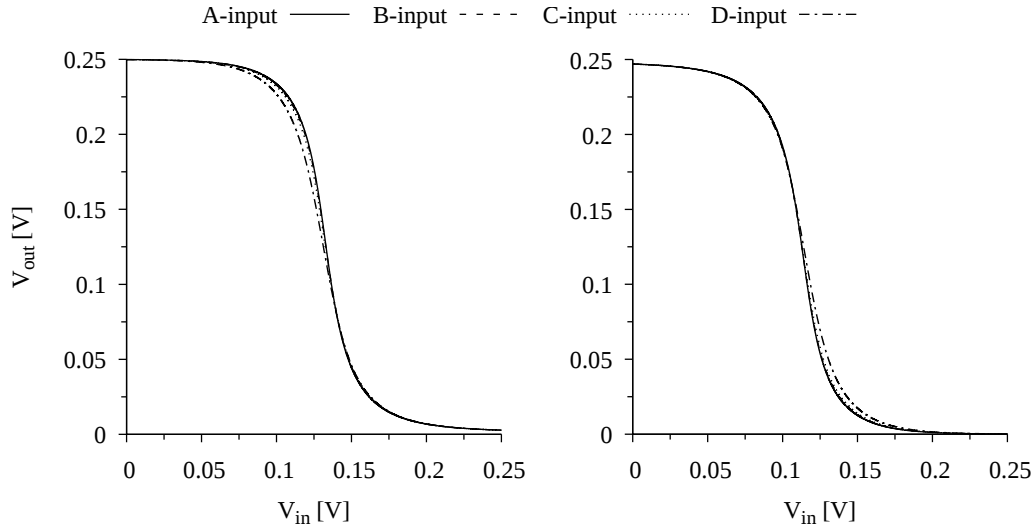
Figure 4.5: Relationship of NAND4 - NOR4 butterfly wing size with the selected input in the DC analysis.
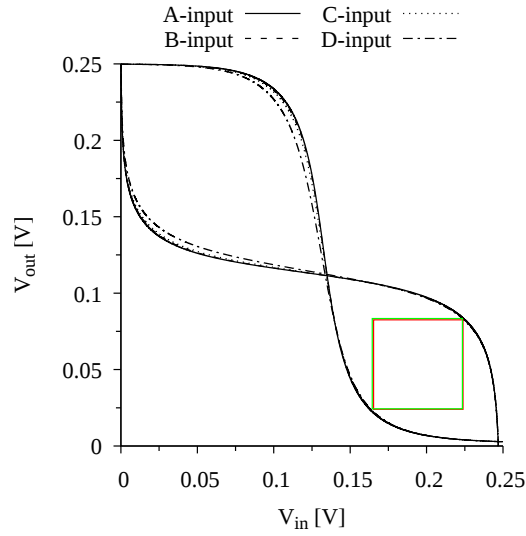
$V_{DS} \ll v_t$ (4.8).

$$I \approx \beta \cdot e^{\frac{V_{GS}}{nv_t}} \qquad \text{for} \quad V_{DS} \gg v_t \tag{4.7}$$

$$R_{eq} = \frac{v_t}{\beta e^{\frac{V_{GS}}{nv_t}}} \qquad \text{for} \quad V_{DS} \ll v_t \tag{4.8}$$

$$\beta = I_0 \frac{W}{L} e^{-\frac{V_{TH0} - \lambda_{BS} V_{BS}}{nv_t}} \tag{4.9}$$

where I_0 and n are technology-dependent parameters, $v_t = kT/q$ is the thermal voltage, W/L is the transistor aspect ratio, V_{TH0} is the threshold voltage, and V_{GS} (V_{DS}) is the gate-source (drain-source) voltage.

Based on those equations, for both presented cases, nMOS transistors are initially represented by Eq. (4.8), which states that V_{GS} is inversely proportional to the resistance. When input D is used, other transistors exhibit small resistance because $V_{GS} = V_{DD}$, while, when input A is used, resistance is high since there is no connection to V_{SS}, i.e., $V_{GS} \approx 0\,V$. Accordingly, resistance in the first case will be lower than in the second until it approximately reaches V_{LT} (4.10) (ALIOTO, 2010)

$$V_{LT} \approx \frac{V_{DD}}{2} + \frac{1}{2}nv_t \ln\left(\frac{\beta_p}{\beta_n}\right) \qquad (4.10)$$

where V_{LT} is the logic threshold voltage. Observe that V_{LT} does not depend on V_{GS} and thus should be approximately the same for all cases, as it can be seen in the DC simulations of Figure 4.4.

In conclusion, the SNM DC simulation should consider switching only the input closest to the output terminal because: (i) it represents the worst single-input case; and (ii) using multiple inputs switching at the same DC value is not realistic for the actual operation of multiple-input logic gates.

The circuit design, however, is not always available to identify each input location. Therefore, the engineer should use SET to simulate every single-input case and thus use the worst case. Nonetheless, only one input should be used to avoid unrealistic results.

4.2 Criteria Comparisons

Likewise the method, the technique also impacts on the static noise margin estimates. Even though Hauser (HAUSER, 1993) provides a valuable discussion about the SNM approaches presented in Chapter 2, his paper did not quantitatively express their differences. Accordingly, this section compares the techniques examined by Hauser and determine which is the most suitable for circuits operating in the sub-threshold region. To achieve this goal, the experiments adopts several pairs of back to back cells described in Table 4.2, and Monte Carlo simulations to account for variation. The back-to-back cells analysis is a standard approach in SRAM memory cells, and it has also been used to evaluate simple NAND/NOR CMOS logic gates.

Table 4.2: Cell set for experiments.

N#	Pair	N#	Pair	N#	Pair
1	INV-INV	6	NOR2-NOR2	11	NOR3-NOR3
2	INV-NAND2	7	INV-NAND3	12	NAND2-NAND3
3	INV-NOR2	8	INV-NOR3	13	NOR2-NAND3
4	NAND2-NOR2	9	NAND3-NOR3	14	NAND2-NOR3
5	NAND2-NAND2	10	NAND3-NAND3	15	NOR2-NOR3

4.2.1 Nominal Evaluation

The three techniques evaluated in Hauser et al. (HAUSER, 1993) are the Negative Slope Criteria (NSC); the Maximum-Product Criteria (MPC); and the Maximum-Equals Criteria (MEC). To assess their difference before Monte Carlo simulations, Figure 4.6 depicts the SNM for the cells described in Table 4.2. As both MPC and NSC estimates separate the high and low noise margins, i.e., NM_H and NM_L, these criteria have two estimates. On the contrary, the MEC has only one value since it forces equal NM_H and NM_L.

By analyzing this figure, it is possible to observe that the MEC gives a midpoint SNM value when compared to the other two criteria. The NSC, on the other hand, provides the highest and lowest values. Comparing both, the NSC has values, on average, 3.5% higher (with a maximum of 5.9%) and 3.7% lower (with a maximum of 6.2%). Given SNM already is a pessimistic metric and that designers usually use lowest values to optimize cells, using NSC would impose extra 6.2% stricter constraints for the worst case. Moreover, using NSC highest estimates might provide too optimistic results since it is 11% larger than its lowest values.

Differently, from the NSC, the MPC results are close to those of the MEC. The maximum difference from both methods is only 2%, and in most cases, the disparity is even lower. This behavior occurs because the wing format makes the MPC achieve maximum product when sides are approximately equal, i.e., a square. Given those results, it is necessary to apply Monte Carlo analysis on MEC and MPC to further compare them. The NSC, however, is not further considered due to its stricter constraints. This conclusion agrees with Hauser's work (HAUSER, 1993) which states that NSC is not a reliable approach to evaluate noise margin.

Figure 4.6: Comparison between SNM criteria at 250 mV supply voltage. X-axis corresponds to the "N#" column of Table 4.2.
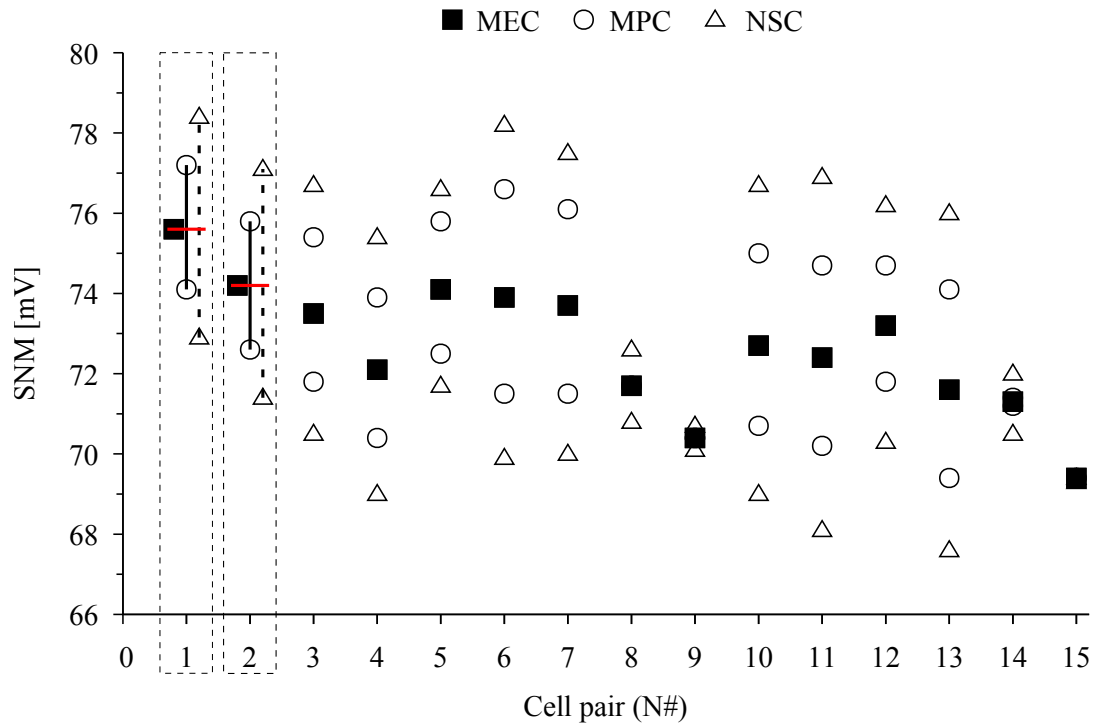
4.2.2 Variability Evaluation

Monte Carlo simulations are necessary for circuits operating at extremely low voltages due to its increased sensitivity to variation. In this type of analysis, it can be considered either or both global and local process variations (3.1). Since the proposed method extracts the SNM of back to back cells, i.e., they are in the same location within the chip, it sensible to apply only local variations (mismatch). However, to also account for global effects, e.g., wafer-to-wafer, the experiment is repeated for each technology corner (e.g., TT, FS, FF, etc.) separately. In other words, the mismatch experiments are executed over each corner individually. This approach ensures that cells have the same characteristics of the chip region they are located but with local variations.

Table 4.3 presents MEC estimates normalized to those of the MPC for TT, SS, FF, FS and SF corners operating at 250 mV at 27°C and using 1,000 Monte Carlo samples. Note that as the MPC generates two distinct SNM values, this work chose the lowest because this is the standard safe course. This notation evidences that MEC criteria always produces higher SNM values than MPC, e.g., 4.6% for INV-INV at TT corner. Accordingly, this table demonstrates that the *mean value* of MEC is, on average, 4.3% higher than those of MPC for TT, FF and SS corners. On FS and SF corners, this difference

increases to 18%.

If a designer, therefore, proposes optimizations for a cell library based on MPC results, he would adopt a strategy 18% more conservative than if he had used MEC. On the other hand, if the Monte Carlo data for MPC technique were based on its highest estimate the designer would equally impose too relaxed constraints. For those reasons, this work suggests that the MEC, also known as maximum-square, is a suitable technique for evaluating combinational cells at subthreshold. Furthermore, MEC does not force the designer to choose between two discrepant values, and it is a method already widely tested in SRAM memory cells.

Table 4.3: Ratio between MEC and MPC SNM results (SNM_{MEC}/SNM_{MPC}). Monte Carlo simulations at 250 mV and 27 °C.

Cell Pair	TT	FF	SS	FS	SF
INV-INV	4.6%	4.4%	5.0%	18.3%	16.1%
INV-NAND2	4.5%	4.3%	5.0%	18.4%	16.5%
INV-NOR2	4.8%	4.4%	5.0%	19.1%	15.9%
NAND2-NOR2	4.6%	4.3%	5.0%	19.3%	16.5%
NAND2-NAND2	4.6%	4.4%	5.0%	18.6%	17.2%
NOR2-NOR2	4.1%	4.3%	5.0%	19.4%	15.8%
INV-NAND3	4.6%	4.4%	5.0%	18.7%	16.9%
INV-NOR3	3.6%	3.9%	5.0%	17.6%	18.0%
NAND3-NOR3	2.8%	4.5%	5.0%	18.2%	18.8%
NAND3-NAND3	3.2%	4.4%	5.0%	19.2%	17.8%
NOR3-NOR3	3.5%	4.9%	5.0%	16.5%	19.5%
NAND2-NAND3	4.4%	4.1%	5.0%	18.9%	17.6%
NOR2-NAND3	4.4%	4.1%	5.0%	19.4%	16.7%
NAND2-NOR3	4.0%	4.6%	5.0%	17.8%	18.5%
NOR2-NOR3	3.8%	4.2%	5.0%	18.4%	17.9%
Average	4.1%	4.3%	4.6%	18.5%	17.3%

4.3 SNM PVT Considerations

Hitherto, the experiments focused on defining a consistent simulation setup, capable of diminishing the SNM inherited pessimism. This section, however, provides insights into the noise margin behavior with temperature and process variations. Moreover, the results here further support the decision on using MEC for measuring the SNM of circuits operating at low voltages.

Considering Monte Carlo MEC mean estimates at TT corner as the baseline, moving to corner: (i) SS increases 6%, (ii) FF decreases 8.7%, (iii) F(p)S(n) decreases 14.8%,

and (iv) S(p)F(n) decreases 12.7% the available noise margin. These numbers evidence that the most stringent corners that a designer has to verify are when nMOS and pMOS transistors suffer opposite effects, i.e., FS and SF. The VTC shift in those cases negatively sums up, and the overall wings become distorted. This behavior, which is depicted in Figure 4.7, explains why MPC yields much lower values than MEC at these corners. To maximize the rectangle area, MPC prioritizes one dimension, e.g., y-axis, and thus produces very distinct high and low noise margins. In consequence, the designer has to choose between a high and a low noise margin estimate, opposed to MEC that offers a single median value.

Figure 4.7: Corner comparison for an inverter pair at 250 mV.
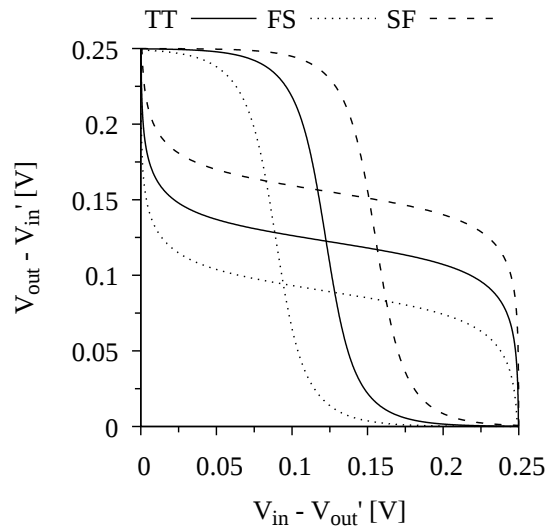
Figure 4.8: Monte Carlo SNM mean value versus temperature, normalized to SNM at 27°C.
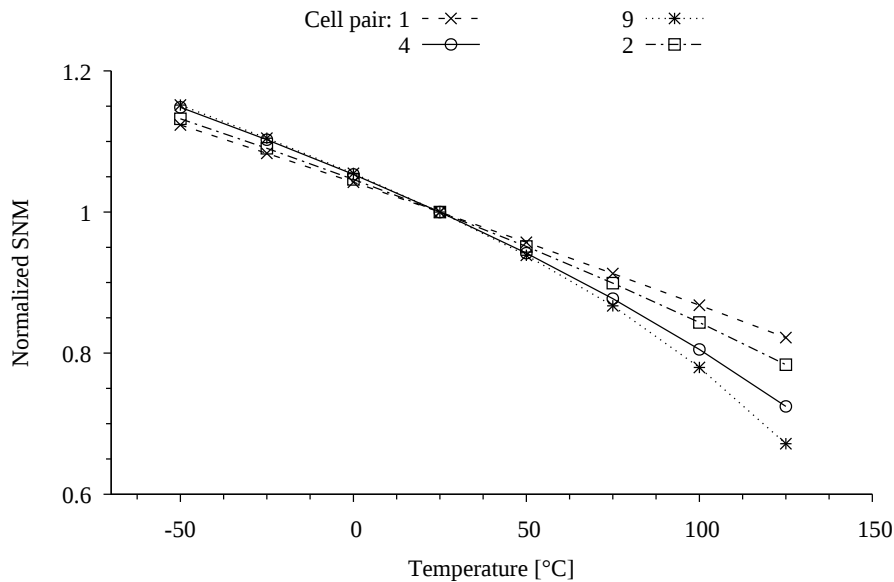
Temperature also impacts on SNM estimates. As it increases, the Monte Carlo SNM mean decreases and vice versa, as depicted in Figure 4.8. For a military standard, for example, it is necessary to account an extra 30% SNM reduction at its extreme right point. It is relevant to point out here that if MPC were used instead of MEC, this value would be even higher, mainly at FS and SF corners. Accordingly, the technique choice has a significant impact on SNM estimation and thus must be appropriately chosen when using it to evaluate cell robustness.

4.4 Chapter Summary

In summary, this Chapter defined an SNM simulation methodology and criteria for combinational cells operating at extremely low voltages. Experiments showed that the improper simulation setup can lead to 70% more pessimistic results, which is an alarming value given that literature refers to SNM as a pessimist metric. Hence, regarding the SNM of digital CMOS inverter, NAND and NOR gates with multiples inputs, the derivations presented here provide the following insights:

- Simultaneously switching multiple inputs is unrealistic, i.e. too pessimistic.
- When switching only one input, the one closest to the output yields the worst-case.
- Subthreshold circuit analysis should consider the MEC criteria because it weights equally NM_H and NM_L, and behaves better in the presence of PVT variations.

5 SUBTHRESHOLD DESIGN ANALYSIS

The increasing demand for portable applications and autonomous systems shifted system design strategy to prioritize energy consumption instead of performance (ORGUC et al., 2017). At the circuit design level, a widely adopted approach is to reduce the supply voltage to lower values, i.e., sub and near-threshold regions. This technique is a very effective technique because the dynamic power dissipation in an integrated circuit is quadratically proportional to the supply voltage (Figure 5.1). Nevertheless, scaling the voltage of the power supply has a negative impact on available noise margin (Figure 5.2), thus making digital logic circuits more sensitive to noise. Additionally, as the voltage reduces there is an exponential increase in the delay, as depicted in Figure 5.1. Wire interconnect minimization also plays an important role in noise margin. The reason is that since wire delay is now comparable to gate delay, the interconnect noise impact on signal characteristics and system performance became significant (KAESLIN, 2008). Hence both voltage scaling and technology minimization negatively affects the overall noise margin of digital circuits.

This Chapter, therefore, investigates how SNM is affected by existing subthreshold cell design techniques. Moreover, it compares different approaches to identify the critical points an engineer has to consider when designing a noise robust cell. Accordingly, Section 5.1 defines the sub/near/super-threshold classification used throughout the experiments, which determine the transistor operation region as a function of its supply voltage. Section 5.2 presents state-of-art subthreshold design techniques and Section 5.3 evaluates their SNM relationship with the supply voltage.

Figure 5.1: Power and delay relationship with supply voltage.
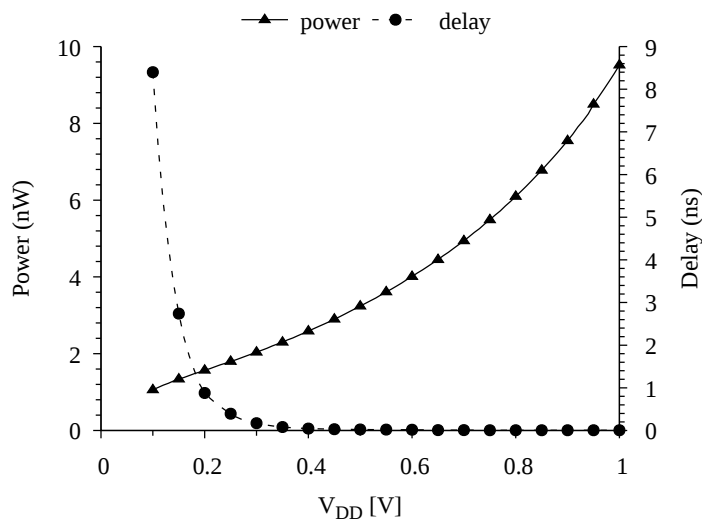
Figure 5.2: Static Noise Margin relationship with supply voltage.
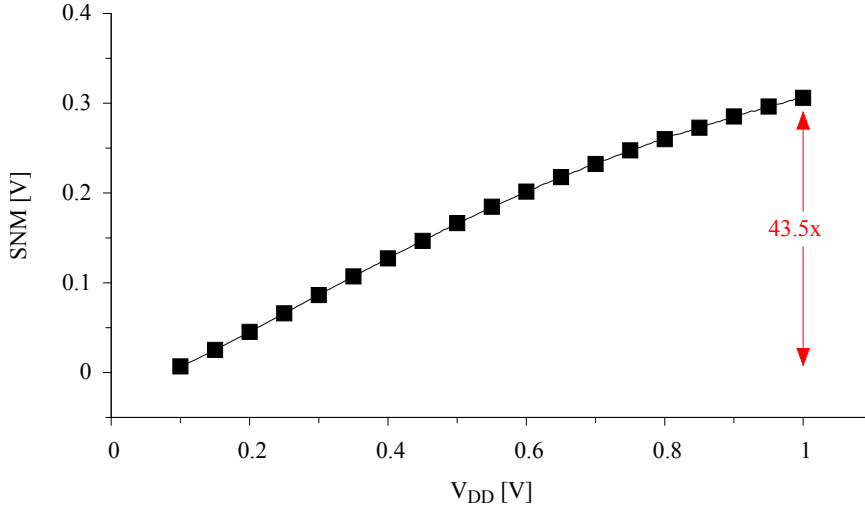
5.1 Regions of Operation

The movement of electrons and holes, i.e., the carriers, within a semiconductor determine the MOS transistor operation (SEGURA; HAWKINS, 2004). These carriers are in constant motion, however, since their movement is random and no preferred direction is collectively chosen, an unbiased semiconductor will not register net current. Therefore, a force is required for net movement of charge to occur. The two primary carrier movement mechanisms in solids are drift and diffusion.

Drift is carrier movement due to an electric or magnetic field. An electric field $\mathcal{E}$ applied to a semiconductor causes electrons in the conduction band to move in the direction opposite to the electric field, whereas holes in the valence band move in the same direction of this field. Electron and hole current densities have the same direction, and both contribute to the current. Diffusion, in contrast, is a thermal mechanism that moves particles from high-density macroscopic regions to low-density ones, so that in the final situation, the particle distribution in space is uniform. Electrons and holes diffusing in a solid are moving charged particles that create a diffusion current.

These two components comprise the transistor drain current as in the following equation

$$I_{DS} = I_{DS1} + I_{DS2} \tag{5.1}$$

where I_{DS1} and I_{DS2} are the drift and diffusion currents respectively. The influence on total I_{DS} of each depends on how the transistor gate V_{GB} is biased with respect to the

bulk. The V_{GB} value, thus, determines how intense (weak, moderate or strong) is the inversion layer the transistor is operating at, as depicted in Figure 5.3. These levels can be modeled in different ways, according to the MOSFET electrical model used. In TSI-VIDIS it is seen that in strong inversion $I_{DS} \approx I_{DS1}$ the current transport in the channel is mainly controlled by the drift component. In weak inversion, the current in the transistor is mainly controlled by the diffusion component, since $I_{DS} \approx I_{DS2}$. However, in moderate inversion both components I_{DS1} and I_{DS2} are important; both drift and diffusion contribute to the total current in the transistor channel. These characteristics lead to simplified transistor models that describe its behavior when operating in specific regions. There is a wealth of transistor electrical models in the literature, and each one treats the definition of moderate inversion differently. In the simulations in this thesis, for 130nm CMOS, the simulator employs the BSIM4v4.

In the strong inversion model of more simplified analytic models, the current follows polynomial behavior (parabolic laws in SPICE level 1, or -alpha laws in Sakurai's model), whereas in the weak inversion model it respects an exponential law. Accordingly, it can be identified the transistor region of operation by approximating the total current to their expected behavior. This work uses this curve fitting concept to define sub, near and superthreshold regions as a transistor is operating in, respectively, weak, moderate and strong inversion.

Figure 5.3: Total current and its components for the nMOS transistor in IBM 130nm.
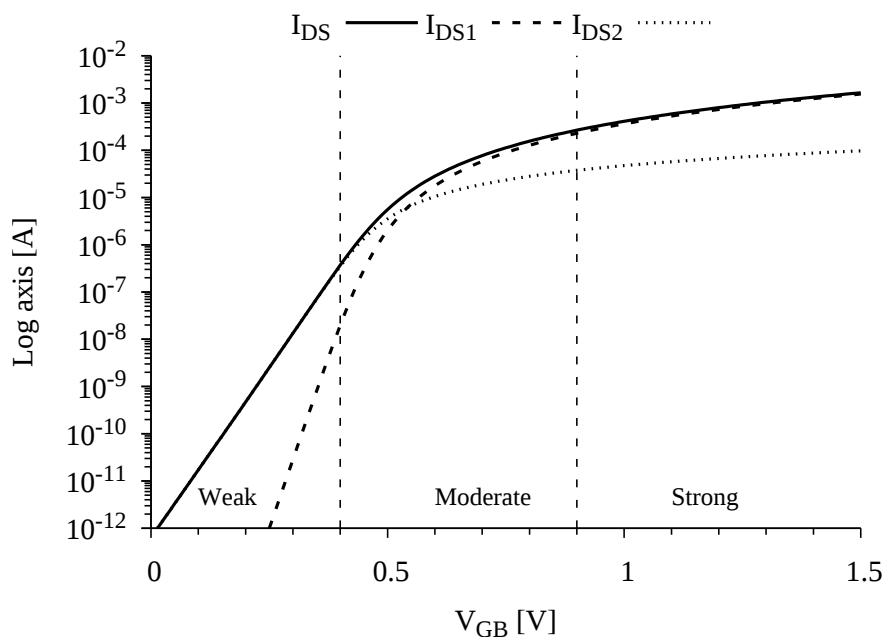
5.2 Design Techniques

The characteristics of MOS transistors in the subthreshold region significantly differ from those in the superthreshold. Transistors are more sensitive to PVT variations, suffer stronger side effects, e.g., reverse short-channel effect, and possess a different current mechanism, as depicted in Figure 5.3. In addition, traditional digital design strategies are optimized for circuits operating in superthreshold and mostly concentrate on enhancing performance. This goal is no longer the dominant paradigm, and in IoT, for example, devices prioritize power consumption instead. Therefore, to design optimal subthreshold circuits, it is crucial to adopt new techniques that consider those aspects.

In this context, Keane et al. (KEANE et al., 2008) proposes an optimal transistor stack factor to increase current drivability for subthreshold circuits. For such, the authors reformulate the traditional Logical Effort (SUTHERLAND; SPROULL; HARRIS, 1999) to account for the differences on this regime, and define a stacking factor of $[1+\alpha(n-1)]$, where n is the number of series transistors and α is given by the following equation:

$$\alpha = e^{\frac{\lambda_d V_{DD}}{(1+\gamma)V_T}} \tag{5.2}$$

where λ_d is the Drain-Induced Barrier Lowering (DIBL) coefficient, γ is the body effect coefficient, and V_T is the thermal voltage. Simulation results demonstrate that the proposed stacking factor provides an improvement of up to 10% on critical path delay.

Extending Keane's work, Kim et al. (KIM et al., 2007) suggest an optimal channel length to achieve high drive current, low device capacitance, less sensitivity to random dopant fluctuations, better subthreshold swing, and improved energy dissipation. His analysis demonstrates that the optimum length is not the technology minimum due to the influence of the channel length on the transistor threshold voltage. In the devices in which the reverse short channel effect (RSCE) is present, the long-channel devices have lower V_{th} than for shorter channel lengths. This lower V_{th} results in higher drain current in the subthreshold region for the longer channel transistor (compared to transistors with the same W/L ratio but shorter L). In particular, for 130nm bulk CMOS technology used, this RSCE is present - and it is attributed to use of the so-called "halo" implant in the channel near the source and drain of the MOSFETs. Our simulation experiments with these devices confirm that increasing the nMOS channel length leads to power and delay improvements. This improvement is due to this second-order effect on V_{th} called RSCE.

The pMOS transistor, however, should use the minimum channel length because the RSCE is significantly weaker than that of nMOS. According to the authors, in more advanced nodes both transistor will benefit from this effect.

Nabavi et al. (NABAVI; RAMEZANKHANI; SHAMS, 2016) adopts a different strategy and proposes an optimum W_p/W_n ratio, i.e., β, that results in the maximum frequency of operation. To determine its value, authors introduce three techniques which include capacitance over current (COC) simulations, deriving an analytic expression, and ring oscillator simulations. Experiments show that these approaches provide similar β values and their results allow delay and energy improvements for a 32-bits carry look-ahead adder. Calhoun et al. (CALHOUN; WANG; CHANDRAKASAN, 2005) also introduce a methodology to define β, but as the ratio that achieves equal pMOS/nMOS current. This method aims to enhance design robustness and to reduce the minimum supported supply voltage. Nonetheless, for TSMC 180nm, the proposed sizing implied on energy consumption overheads.

In summary, these strategies demonstrate the necessity of considering subthreshold effects to design optimal circuits. By exploiting these differences from superthreshold, these works were capable of improving overall device performance and energy consumption.

5.3 Voltage Scaling

The Authors (KEANE et al., 2008; KIM et al., 2007; NABAVI; RAMEZANKHANI; SHAMS, 2016; CALHOUN; WANG; CHANDRAKASAN, 2005) explore distinct transistor properties and evaluate their impact on power consumption, area, and delay. Their analysis, however, does not include noise metrics, which are essential for subthreshold circuits. This Section, therefore, assesses relationship between SNM and supply voltage in the presented techniques. Besides a direct noise comparison, this investigation allows an engineer to understand which design parameters are more critical to enhance overall circuit SNM.

To estimate noise, this Section uses the SET tool, considering process and mismatch variations, and voltage scaling. Moreover, the experiments use the same technologies employed by the authors, which are 65nm, 130nm and 180nm, to evaluate the SNM of classical digital CMOS gates.

5.3.1 Analysis for 65nm

Keane (KS) and Nabavi (NS) provide specific guidelines of their methodologies for classical CMOS gates in 65nm technology. The summary of their results is described in Table 5.1. Following these definitions, Figure 5.4 depicts the SNM Monte Carlo results obtained by executing the SET tool for three gates (INV, NAND, NOR) and each method. In the left row of this figure, the Y-axis represents the relative SNM mean (μ_{SNM}/V_{DD}) of 1,000 Monte Carlo samples. In the graphs located in the right row, the Y-axis represents the coefficient of variation (σ/μ). This parameter defines the extent of variability in relation to the mean of the population, and thus helps to visualize the data spread dependence on the supply voltage. In addition, each graph is marked according its inversion region (Section 5.1) as a function of V_{DD}.

Table 5.1: Subthreshold design summary for 65nm technology.

Parameter	KS	NS
L_{pMOS}	0.060 μm	0.060 μm
L_{nMOS}	0.060 μm	0.060 μm
β	1.5	1.0
Logical Effort	$[1 + \alpha(n-1)]$	n
α_{pMOS}	1.940	-
α_{nMOS}	1.545	-

As Figure 5.4 demonstrates, there is no significant visual distinction between both approaches for this particular 65nm technology . This is due to transistor properties that influence the static noise margin behavior. More specifically, SNM is directly correlated to the difference on pMOS and nMOS threshold voltage and current values. To balance these parameters, one possible approach is to increase the transistor width. Nonetheless, in ST 65nm, this was adjusted during fabrication through advanced techniques, e.g., strain engineering (DIMOULAS et al., 2007), and thus minimum size transistors offer approximate absolute threshold voltages (Figure 5.5) and currents of same magnitude. Due the aforementioned details, therefore, it is not noticeable the impact of these design choices on the SNM. In more mature technologies, such as 130nm and 180nm, where it is not available the same advanced fabrication techniques, there will be significant differences on their threshold and, consequently on noise, as the next two Sections will cover.

Figure 5.4: Comparison of 1k samples Monte Carlo simulation of two subthreshold sizing approaches using ST 65nm.

(a) INV

(b) INV

(c) NAND

(d) NAND

(e) NOR

(f) NOR
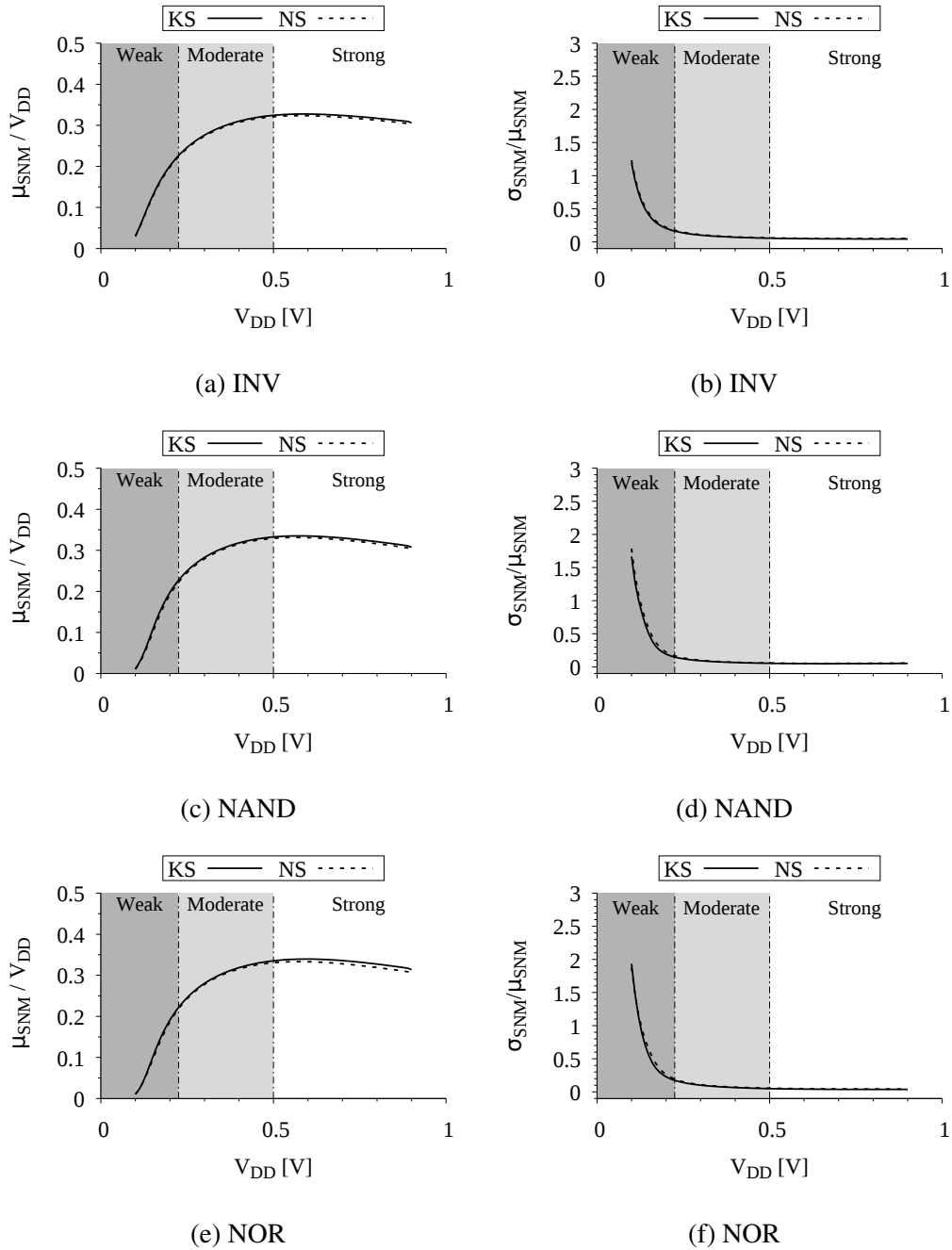
Table 5.2: Subthreshold design summary for 130nm technology.

Parameter	KS	NS	KSE
L_{pMOS}	0.120 µm	0.120 µm	0.120 µm
L_{nMOS}	0.120 µm	0.120 µm	0.340 µm
β	1.5	1.0	1.5
Logical Effort	$[1 + \alpha(n - 1)]$	n	$[1 + \alpha(n - 1)]$
α_{pMOS}	1.567	-	1.567
α_{nMOS}	1.335	-	1.335

Figure 5.5: Threshold voltage versus width for nMOS and pMOS transistor in CMOS technologies ($V_{GS} = V_{DS} = 250mV$).
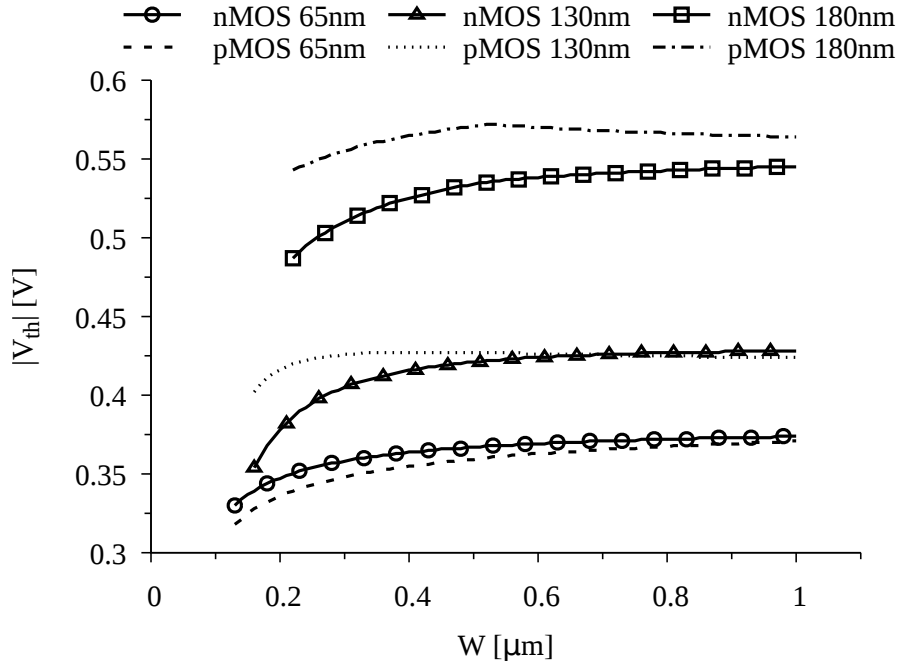
5.3.2 Analysis in 130nm

For 130nm three authors provide the results of their methodologies, which are Keane (KS), Nabavi (NS) and Kim (KSE). Table 5.2 summarizes the specific sizing for each.

This Table indicates three main differences among the design strategies. Kim's asymmetrical channel length in the fourth column, Keane's subthreshold logical effort in the second column, and Nabavi's transistor width ratio β in the third column. Comparing KSE against KS, which only differ on L_{nMOS}, the results in Figure 5.6 imply that the channel length asymmetry has a negative impact on μ and σ/μ. The reason for this is that the SNM depends on the balance between pMOS and nMOS transistor current. This aspect is correlated to the graphical representation that produces the SNM estimations, i.e., the butterfly plot. If these currents are well balanced, the curves will be at $V_{DD}/2$, and the resulting plot will generate maximum SNM. The increment on L_{nMOS}, however, only increases nMOS current and thus aggravates the inherited current imbalance due pMOS lower mobility. Graphically, this increment represents a shift towards the right that indicates that nMOS transistor is stronger than the pMOS (RABAEY; CHANDRAKASAN; NIKOLIć, 2007). This phenomenon is critical at lower supply voltages and can degrade up to 2.3 times the μ value. On average, this strategy leads to 20% and 15% worst μ and

σ/μ respectively.

Figure 5.6: Comparison of 1k samples Monte Carlo simulation of three subthreshold sizing approaches using IBM 130nm

(a) INV

(b) INV

(c) NAND

(d) NAND

(e) NOR

(f) NOR
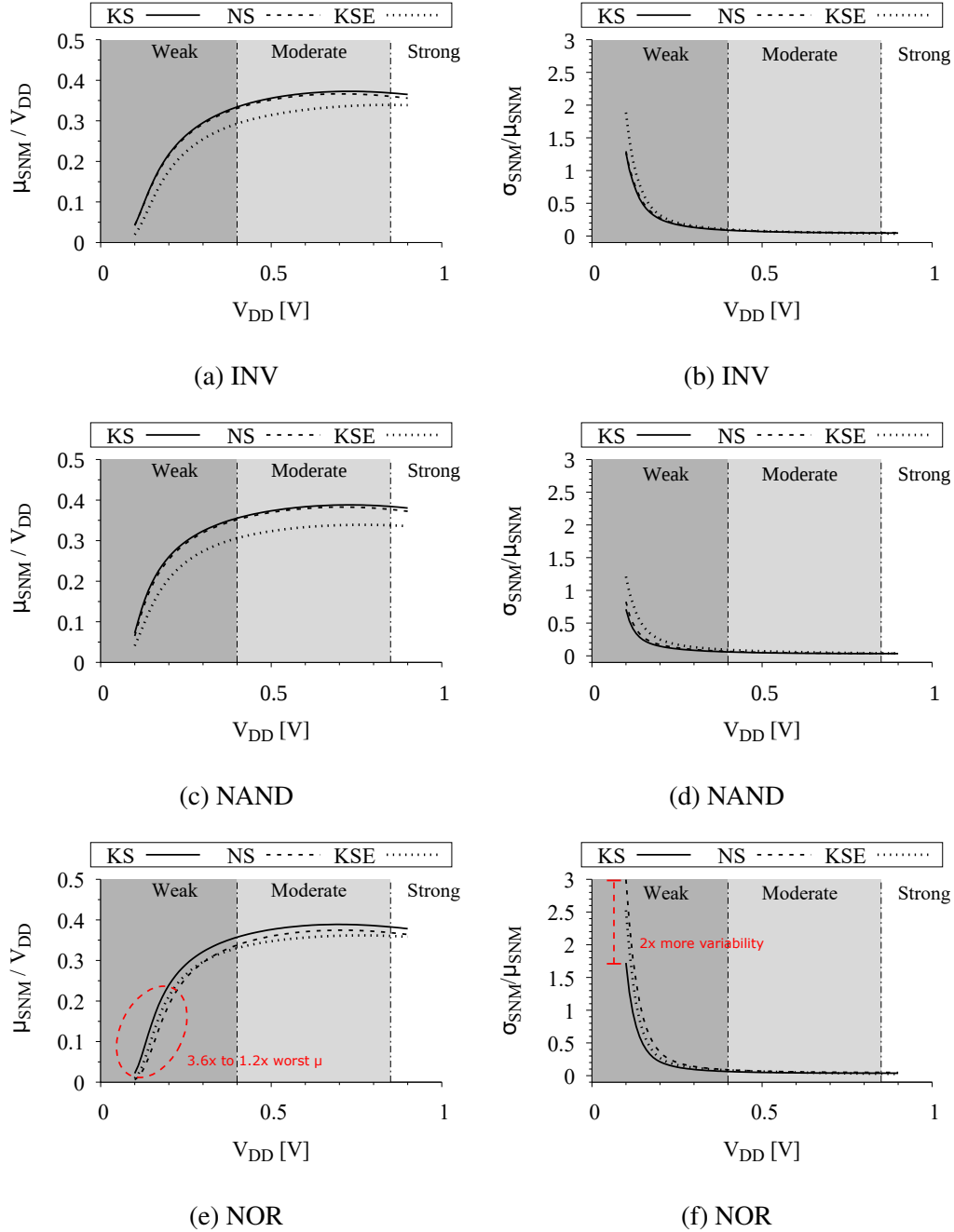
Keane's subthreshold logical effort and Nabavi's β does not seem to have an impact in the SNM of INV and NAND gates. For the NOR gate, however, there is an evident improvement in Keane's approach (KS). This difference indicates that the mentioned current unbalance is more critical in gates with series pMOS transistors. Therefore, Keane's stack factor and higher β compensates the current decrease of series pMOS transistors and provides up to 3 times better μ and 70% less σ/μ than Nabavi's.

5.3.3 Analysis in 180nm

In the 180nm technology, only Nabavi (NS) and Calhoun (CS) provide their sizing values. The values for each approach is summarized in Table 5.3 and the SNM comparison is depicted in Figure 5.7 in the same format as the other figures.

Contrarily to the previous analysis, the β is the only different parameter to consider here. In that regard, Calhoun's strategy provides better μ and σ/μ than the Nabavi's. Below the threshold voltage, CS can achieve up to 6.7 times less spread and 2.5 higher SNM than NS. Since Calhoun's approach for determining β is to balance pMOS and nMOS currents this result was expected. As previously discussed, the SNM relies on the VTC position and such approach leads it to the optimal point. Nonetheless, the chosen β is significantly bigger than NS and would be necessary to evaluate its impact on other parameters.

5.3.4 General Analysis

Decreasing the supply voltage significant reduces the power dissipation of digital circuits at the cost of an exponential increase in the delay. Moreover, Figures 5.4, 5.6 and 5.7 have shown that the relative SNM has close to a logarithmic dependence on V_{DD} for 65nm, 130nm, and 180nm technologies. Accordingly, scaling the voltage has adverse effects on both delay and SNM. Those figures further demonstrate that there is an optimum point for SNM operation in the near-threshold regime, i.e., moderate inversion. This region is a well-known optimum point for energy and also with the best performance, area, and energy trade-offs (BORTOLON et al., 2016). Considering these technologies, the maximum relative SNM, i.e., SNM/V_{DD}, occurs around 0.5V, 0.7V and 0.9V for ST 65nm, IBM 130nm and TSMC 180nm respectively.

Table 5.3: Subthreshold design summary for 180nm technology.

Parameter	NS	CS
L_{pMOS}	0.180 µm	0.180 µm
L_{nMOS}	0.180 µm	0.180 µm
β	2.330	12.0

This Section results indicate that it is possible to explore cell design metrics to enhance SNM. Besides the 65nm that has particular fabrication steps to enhance transistor

Figure 5.7: Comparison of 1k samples Monte Carlo simulation of three subthreshold sizing approaches using TSMC 180nm

(a) INV

(b) INV

(c) NAND

(d) NAND

(e) NOR

(f) NOR
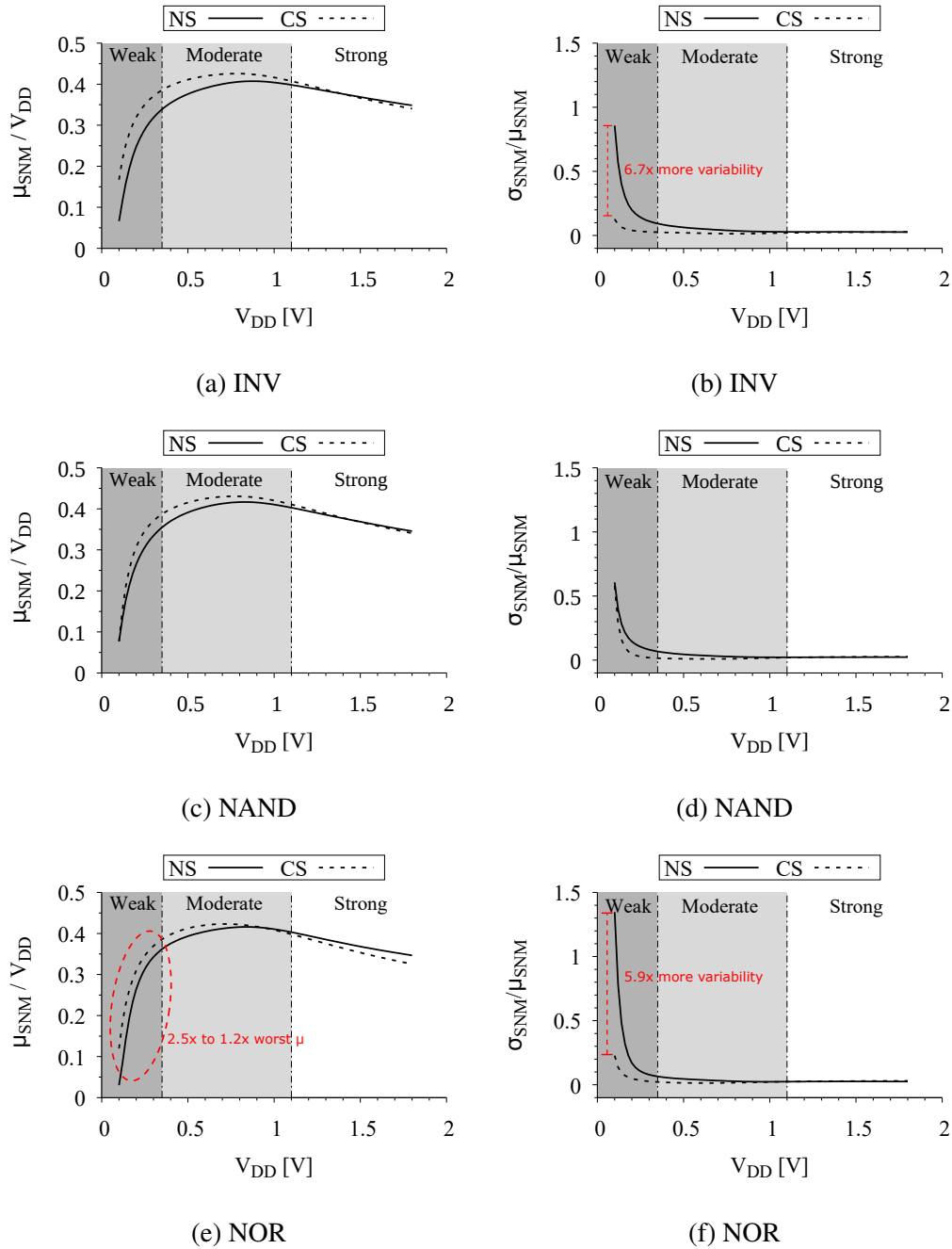
characteristics, a general consideration extracted from 130nm and 180nm technologies is that the impact of subthreshold sizing on SNM is more critical for gates with series pMOS transistors, such as the NOR gate. Usually, pMOS transistors drive less current than nMOS due to their mobility, and thus incrementing their width is necessary to increase the SNM. Consequently, it is necessary to consider this characteristic when designing gates with series pMOS transistors.

5.4 Chapter Summary

This Chapter provides insights on a few relevant aspects to take into account to design SNM-aware cells. Accordingly, the experiments present the following conclusions:

- Operating in the near-threshold region produces better relative SNM values.
- Digital CMOS cell design in near-threshold must examine carefully the current imbalance between pMOS and nMOS transistor stacks, to be able to enhance the SNM.
- SNM is more critical in gates with series pMOS transistors.

6 SNM-AWARE CELL DESIGN

In the semiconductor industry, standard cell design is a method of designing VLSI circuits using pre-defined logic functions called cells. These cells are mapped through Computer-Aided Design (CAD) tools to a circuit definition, usually defined in a hardware description language (HDL). Besides converting a description to a circuit, known as cell mapping, the tool places and routes all cells while ensuring the physical constraints of the technology, e.g., metal spacing. Consequently, this approach significantly relies on the engineer who made the design of these logic functions. In traditional methodologies, the engineer aims to balances performance, power dissipation, and area, according to a certain criteria, to achieve better results after CAD tools process HDL circuit definitions. Nonetheless, for subthreshold devices, these parameters are not enough to ensure that the correct behavior of the final circuit. Therefore, noise tolerance is an essential metric for cell-based approaches.

Static noise margin (SNM) is a common metric to design memory cells, e.g., SRAM, however not until Kwong et al. (KWONG; CHANDRAKASAN, 2006) it was considered for the design of classical CMOS gates. Kwong proposed a cell design approach and demonstrated that upsizing is necessary to mitigate the degraded output voltage levels caused by supply voltage scaling. Beiu et al. (BEIU et al., 2013b; BEIU et al., 2013a) propose a different design strategy to maximize the SNM, and demonstrate the trade-offs between power, delay, and SNM. In (ALIOTO, 2010), Alioto presents a thorough analysis of the DC behavior of subthreshold CMOS logic and proposes an analytic model for an inverter static noise margin. Later, Tajalli et al. (TAJALLI; LEBLEBICI, 2011) and Olivera et al. (OLIVERA; PETRAGLIA, 2017) extend Alioto's expression to consider, respectively, DIBL and body-bias effects. Moreover, Tajalli further explores the design of a ring-oscillator comparing different trade-offs between power, delay, and SNM. Although, there are authors that explore SNM on the design of combinational cells, their analysis is restricted to simple circuits, such as, NAND, NOR, and ring oscillators. Hence, it is not clear the impact of their approach on real circuits.

This Chapter presents the **main contribution** of this work, which is a methodology to design combinational cells considering their SNM. Moreover, it **proposes** a novel strategy to evaluate SNM of complex combinational circuits along with traditional performance, power dissipation and area metrics. The remaining of this Chapter is organized as follows. Section 6.1 begins the discussion defining the set of logic functions, and driving

strengths that are available in the proposed standard cell design. Section 6.2 presents the cell design strategy to maximize the static noise margin of combinational circuits. Afterwards, Section 6.3 explains the library characterization flow, while Section 6.4 presents the results from a CAD tool after the logic synthesis step. Experiments are carried through extensive simulations in TSMC 180nm, IBM 130nm, and ST 65nm technologies.

6.1 Cell Library Selection

The set of standard cells selected for the target library follows the statistical based analysis of cell usage devised by Gibiluka et al. (GIBILUKA, 2016). The goal of this analysis is to restrain the number of logic functions, i.e., types of cells, in a library to the essential to synthesize digital circuits without compromising generality. To determine these cells, the author initially synthesizes 58 benchmarks, using a commercial library that contains many logic functions, and evaluates how many times each cell type is employed in each circuit. Afterwards, the author removes the least used logic functions from the library and repeats the synthesis. This process is repeated until finding the minimum number of logic functions that can be used while completing the synthesis of all benchmarks. The idea behind this strategy is to reduce the variety of cells by exploiting the synthesis tool capability of implementing the functionality of least used cells with the most common.

According to GIBILUKA, it is preferred to use a minimum cell library because it reduces the number of required computations in the characterization process (CADENCE, 2015), which is the focus of his work. Nonetheless, those experiments further provide insights on the number of cells that a library requires to synthesize multiple digital circuits. The proposed library, therefore, contains only the fundamental logic functions that are INV, NAND2, NOR2, AOI12, AOI22, OAI12 and OAI22, according to his analysis. The XOR2 is not included in the library as its function can be implemented with AOI22. Notice that there are other essential logic functions, but as they have three (or more) series transistors, they are not used because transistor stacks are critical when operating at near/subthreshold region. This decision has little impact on synthesis since only three sequential circuits out of all benchmarks did not complete the logical synthesis in (GIBILUKA, 2016) and those are not employed here. Moreover, these three circuits did not synthesize with a two-input library only because they require flip-flops with asynchronous reset, thus requiring three inputs.

Besides the cell variety offered in a standard cell library, the synthesis tool requires multiple driving strengths, i.e., sizing, to explore design trade-offs and meet timing constraints. Hence, this work devises a small set of strengths: X1, X2, X4, and X6, that are designed according to the minimum cell dimensions, i.e., X1. The X2, for example, is two times bigger than X1, while others follow a similar logic. Overall, the target cell library contains 20 cells and is summarized in Table 6.1

Table 6.1: Target standard cell library for this work.

Cell	X1	X2	X4	X6
INV	✓	✓	✓	✓
NAND2	✓	✓	✓	✓
NOR2	✓	✓	✓	✓
AOI12	✓		✓	
AOI22	✓		✓	
OAI12	✓		✓	
OAI22	✓		✓	

6.2 Cell Library Design

This Section presents the dissertation's main contribution. First, Section 6.2.1 investigates which design parameters affect the SNM by correlating its graphical representation to the physical characteristics of transistors. Once these parameters are identified, Section 6.2.2 graphically evaluates the theoretical limits of the butterfly-based criteria. Afterwards, Section 6.2.3 presents the SNM-aware combinational cell design strategy. The experiments use IBM 130nm technology as a reference for the methodology explanation, but the procedure was also executed to ST 65nm and TSMC 180nm. The results for the other technologies are presented in Appendix B.

6.2.1 SNM Design Parameters

MEC (Section 2.4.1.2) defines the maximum tolerable noise as the biggest inscribed square in the smaller wing of a butterfly plot. Therefore, to enhance the noise margin of a cell, it is necessary to increase this wing's size. Increasing the size of one wing, however, reduces the other, as they share the same area, as depicted in Figure 6.1.(a). This characteristic, hence, may cause the other wing to decrease up to the point that it becomes the smaller wing, bringing us to the problem's original state.

Figure 6.1: Relationship between the SNM and the graphical aspects of a butterfly plot.

(a) V_{mid}

(b) Slope
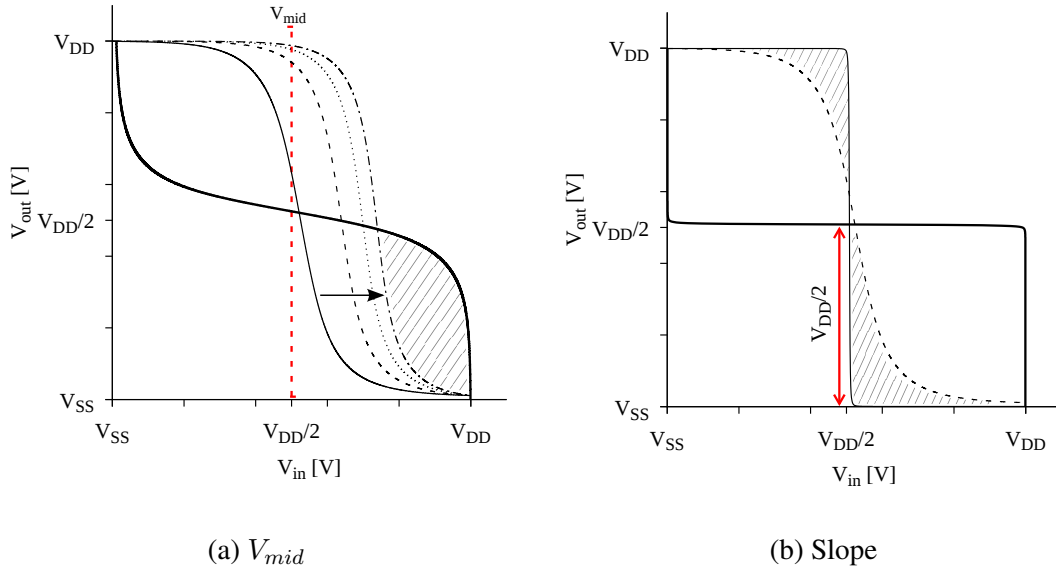
Since this criterion always uses the smaller wing, the only way to achieve maximum SNM is by having wings with the same size. To achieve this condition, the VTC position must be set in the middle of the graph for both mirrored and non-mirrored curves. More specifically, the VTC position (V_{mid}) is defined as the value at the output (V_{out}) when the input (V_{in}) is half the supply voltage. The following equation illustrates this definition.

$$V_{mid} = V_{out}(V_{DD}/2) \tag{6.1}$$

Figure 6.1.a demonstrates that as V_{mid} moves further from $V_{DD}/2$, the right wing area significantly reduces. This shift in the VTC unbalances the sizes between wings, making the right one limit the SNM (shaded area). Chapter 5 demonstrated that the VTC position depends on the ratio between pMOS and nMOS currents. If the current of the nMOS transistor is stronger than that of the pMOS, the VTC will be shifted towards the right of the ideal V_{mid}, e.g., similar to Figure 6.1.a. On the other hand, if pMOS transistor current is stronger than the VTC will shift towards the left. To adjust these currents, a common strategy is to correctly size the transistor's width ratio W_p/W_n, referred as β.

Besides the VTC position, the SNM is also correlated to its slope, as depicted in Figure 6.1.b. Smaller slopes have a negative impact on the available wing area (shaded), and thus on SNM. Therefore, ideal VTC instantly changes from logic 1 to 0, i.e., slope = ∞. Note that the slope represents the gain of the inverter and depends on the conductance and resistance its transistors. As a consequence, the slope depends on the transistor characteristics, i.e., width and channel length. For digital CMOS inverters, however, the re-

sistance has an inverse relationship with the conductance. Increasing the transistor width reduces its resistance and increases its conductance. Therefore, the slope remains approximately constant, since it is equal to the product of these two parameters. Consequently, it is not possible to adjust the slope, but only the position of the VTC.

Overall, the SNM depends on the current of pMOS and nMOS transistors. In order to balance their ratio and enhance SNM, this work explores the β parameter. In addition, the analysis assess the channel length impact on SNM, as current sometimes benefits from small increases due to specific side-effects, as it will be demonstrated later.

6.2.2 Theoretical SNM Limits

Unlike V_{mid}, it is not possible to achieve the ideal infinity slope to maximize the noise margin. Nonetheless, the ideal values for slope and V_{mid} allows drawing the SNM theoretical limits. If both conditions are met, then the VTC is represented by a step function similar to the sharp transition in Figure 6.1.(b). Consequently, the maximum SNM that a cell pair can achieve is $V_{DD}/2$.

The experiments in this Chapter assess noise as a metric relative to the supply voltage, i.e., SNM/V_{DD}. Accordingly, given that the maximum theoretical SNM is half the supply, then the maximum relative SNM is 50%. This concept is crucial to understand that any little improvement in the noise margins, e.g., from 20% to 25%, are significant since their reference is not 100%, but 50%.

6.2.3 Design Methodology

6.2.3.1 Transistor Width Ratio

The discussion in Section 6.2.1 demonstrated that it is necessary to adjust the current ratio between pMOS and nMOS transistors to enhance the SNM of a digital CMOS logic function. Usually, this adjustment is done by choosing the nMOS width (W_n) and evaluating which pMOS width (W_p) yields the target feature, which is the maximum SNM here. Therefore, W_p is defined as the product between W_n and β (Definition 1) that

produces the maximum SNM:

$$W_p = \beta W_n \Big|_{\beta \to max(SNM)} \tag{6.2}$$

Nonetheless, as β varies according to W_n, and digital cell libraries have many cell strengths, there will be many W_ns. Accordingly, before further advancing this discussion, the methodology's first step is to define the nMOS width for all driving strengths.

Since this work intends to design devices for IoT applications, the design choices must have a strong commitment to power consumption, besides the SNM. Usually, the cell area is directly correlated to the power consumption. The bigger the cell is, the more it consumes. For this reason, the minimum cell strength, i.e., X1, is devised to operate with minimum size, as defined by each technology. The other cell strengths are a product of the minimum and their strength factor, as illustrated in the first row of Table 6.2 for IBM 130nm.

Definition 1. β: optimum width ratio between nMOS and pMOS transistor that gives the maximum SNM of a cell.

Definition 2. $\beta_{opt}*$: local optimal β, producing the maximum SNM for a particular nMOS width.

Definition 3. β_{opt}: global optimum transistor ratio, β that produces the maximum sum of SNM on all cell strengths.

Table 6.2: IBM 130nm nMOS transistor width for each strength and their $\beta_{opt}*$

	X1	X2	X4	X6
W_n (μm)	0.160	0.320	0.640	0.960
$\beta_{opt}*$	7	2	1.3	1.1

Returning to the previous discussion, β (Definition 1) depends on W_n and thus it has multiple $\beta_{opt}*$ (Definition 2), because they produce the maximum SNM for a particular nMOS width, and are illustrated in Table 6.2 for IBM 130nm. Even though it is possible to use a different β for each cell strength, this is not a desirable approach for two reasons. First, it increases design complexity, and, second, it disrupts the logic of digital synthesis process. More specifically, CAD tools explore cell strength availability to balance delay, power dissipation, and area trade-offs specified by the engineer. If each cell is designed with their maximum SNM, there will be a significant impact on these other parameters. The β for cell X1 in Table 6.2, for example, generates an area penalty

higher than X2, and thus might not even be used. Moreover, non-critical paths will not benefit from either area and power reduction from using cells with less driving strength. Therefore, it is beneficial to use a single β and leave to the tool the duty to balance these trade-offs.

To determine the β_{opt} (Definition 3), it is desirable to find the β that produces the maximum sum of SNM on all cell strengths. In other words, since it is not possible to maximize the SNM for all strengths (i.e., X1, X2, X4, and X6), β_{opt} must be the value that has the best SNM trade-offs. The following equation defines this criterion.

$$\beta_{opt} = X \mid X \rightarrow max(\sum_{i=1}^{N} SNM_i) \tag{6.3}$$

where N represents the number of cell strengths. Thence, to define the X, Figure 6.2 depicts the normalized sum of SNMs of an inverter using IBM 130nm. In this Figure, the closer the normalized sum is to 1 (dashed lines), the closer each inverter strength is to their local maximum. Consequently, the chosen β_{opt} for this technology is 2, which is the curve's global maximum.

Figure 6.2: SNM Trade-offs for different inverter strengths versus β.
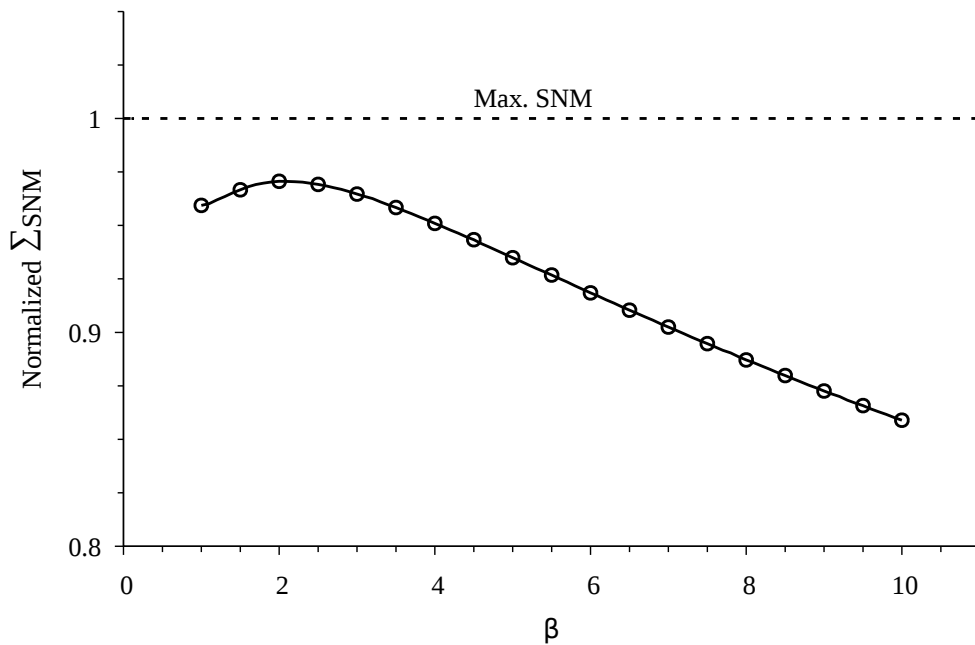
6.2.3.2 Delay and Power Considerations

Analogous to the SNM analysis, it is crucial to also evaluate the relationship between β, delay, and power dissipation. A possible way to define a cell delay is the average

of high-to-low and low-to-high transitions, as in the following equation:

$$T_D = \frac{T_{HL} + T_{LH}}{2} \tag{6.4}$$

Following this definition, it is desirable to select the β that produces the minimum delay for all cell strengths. Similar to the previous analysis, it depends on a particular W_n, and thus it is possible to use the same approach of Figure 6.2. Figure 6.3 presents the normalized sum of delays for all cell strengths and defines the optimum delay-β such that

$$\beta'_{opt} = Y \mid Y \rightarrow max(\sum_{i=1}^{N} T_{D_i}) \tag{6.5}$$

The simulation setup considers a single inverter driving an output capacitance proportional to a fan-out of 4, i.e. $C_{out} = 4C_{in}$, and its input is generated by another inveter. Increasing the output capacitance for each β is necessary because pin capacitance depends on this parameter. Accordingly, this Figure demonstrates that β'_{opt} is around 2 (this point is slightly less than at 1). These results do not necessarily coincide with the SNM β_{opt}, as it is shown for other CMOS bulk technologies (Appendix B).

Figure 6.3: Inverter T_D delay trade-off versus β for IBM 130nm.
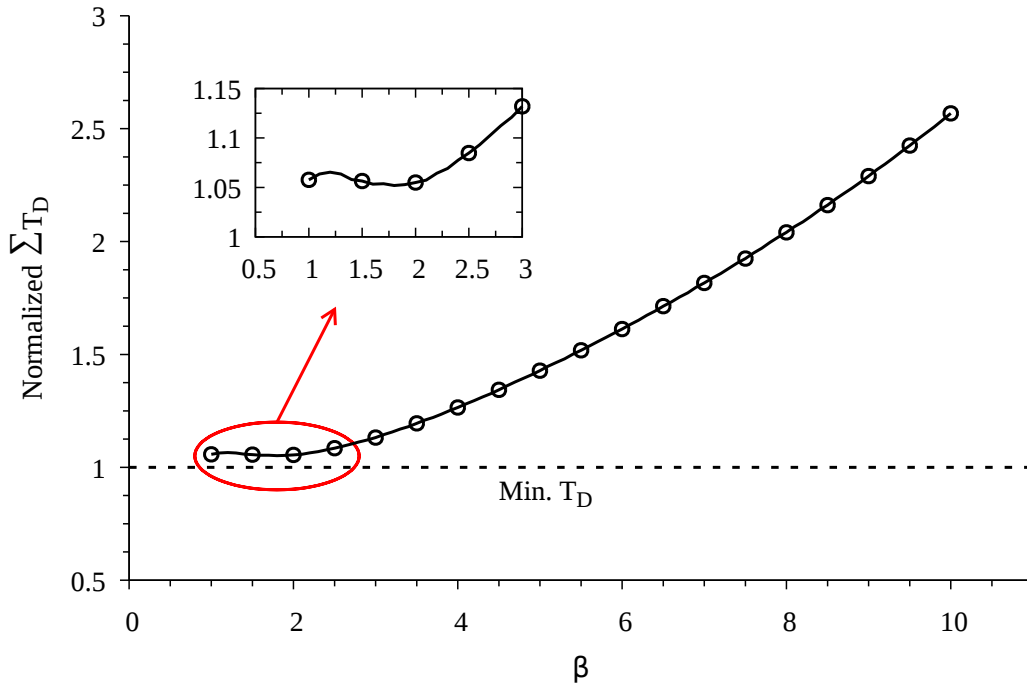
Figure 6.4 assess the parameter β impact on the Power-Delay-Product (PDP) of an inverter for the same simulation setup. Since it is desirable to minimize both product terms, the Figure demonstrates that minimum β produces the best PDP. Nonetheless,

using the optimum value for SNM, and coincidentally delay, does not have a significant impact on the PDP. Overall, the results demonstrate that both delay and power rely on minimizing the capacitance is crucial. Increasing β well-above the chosen β_{opt} of 2 for IBM 130nm produces significant impacts on power and delay.

Figure 6.4: Inverter PDP trade-off versus β for IBM 130nm.
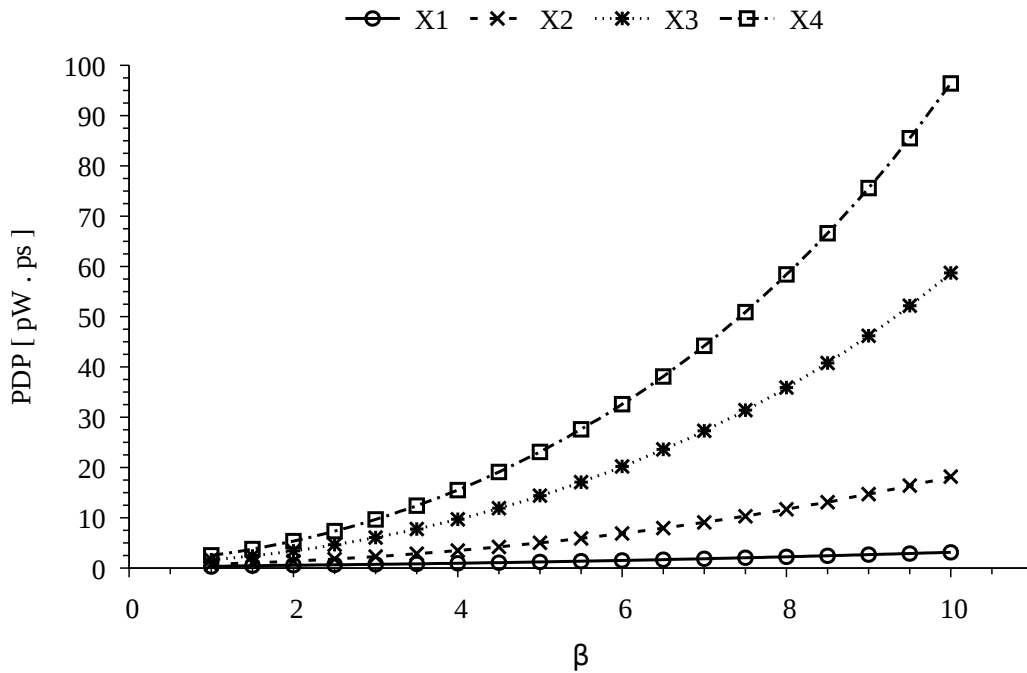
6.2.3.3 Channel Length

Chapter 5 presented Kim's (KIM et al., 2007) work, which demonstrates that increasing the transistor channel length (L) can enhance the drain current. Their result confirms that this increment had a positive effect on both power consumption and delay. Other authors also observe the same benefits by upsizing L. Gupta et al. (GUPTA et al., 2004) shows that there is an $L_{opt} > L_{min}$ which can be used to minimize leakage power, and Venugopal et al. (VENUGOPAL; CHAKRAVARTHI; CHIDAMBARAM, 2006) proposes another L_{opt} that minimizes delay. The same concept of upsizing L was also used in SRAM cells (HANSON et al., 2008), where the suggestion was to use it in combination with optimizing the doping profile. Beiu et al. (BEIU et al., 2013a) demonstrates that tuning V_{th} through increments on L is beneficial to enhance SNM of advanced technologies.

More generally, these authors confirm that increasing the channel length can have a positive impact on other parameters when operating the subthreshold region. Therefore,

this Section determines the L_{opt} that produces the maximum SNM, which is defined by the following equation

$$L'_{opt} = Z \mid Z \to max(\sum_{i=1}^{N} SNM_i) \tag{6.6}$$

This equation states that the optimum channel length is the one that produces the best SNM trade-off between all cell strengths. This analysis is similar to the previous β and T_D. Figure 6.5 shows that $L_{opt} = 0.250\mu m$, and further increasing its value deteriorates the SNM. Since L also affects the current, this analysis evaluates if this increase does not have a severe impact on it. For such, this work adopts a strategy similar to Nabavi et al. (NABAVI; RAMEZANKHANI; SHAMS, 2016), which evaluates the Current-over-Capacitance (COC) dependence on the transistor width. Here, however, this metric is employed to evaluate the channel length. Using the capacitance in the analysis is a way to assess the current drive capability and the cell area. Although this Figure could be extended to a multi-dimensional problem, considering multiple transistor width and β, it was observed, from data inspection, that the optimum channel length was approximately the same for other configurations. Therefore, it is only presented here the graph for the transistor width and βdefined previously.

Figure 6.5: Inverter channel length trade-off versus SNM for IBM 130nm.
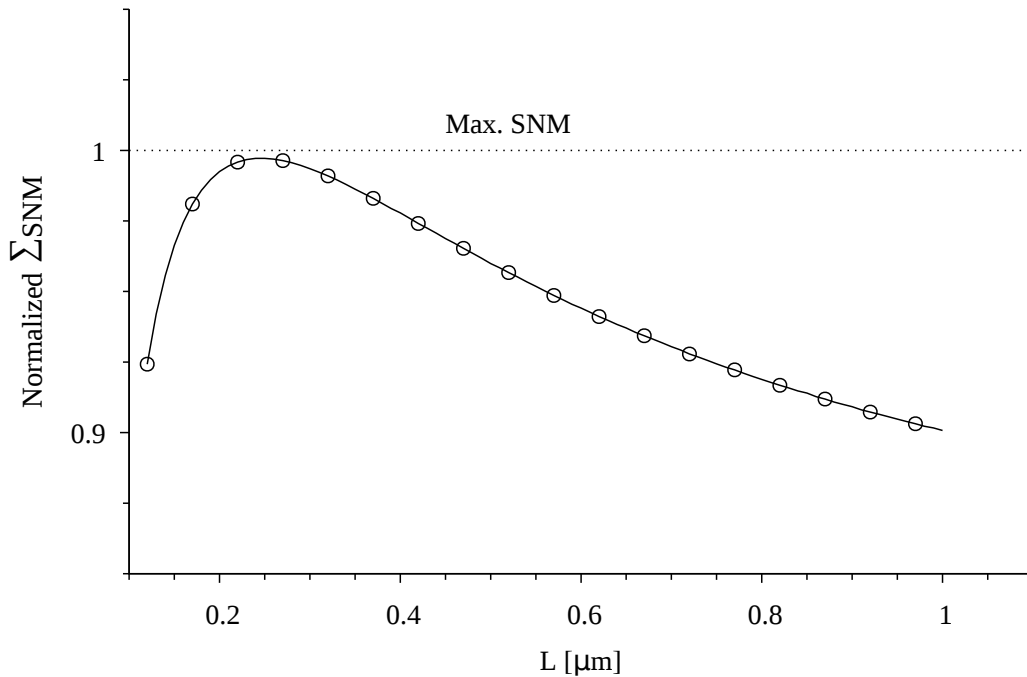
Figure 6.6 demonstrates that increasing L for IBM 130nm also has a positive impact of COC values. The peak value for nMOS transistor is 0.340 μm and, 0.22 μm for

pMOS transistor. Using the optimum nMOS length reduces pMOS current, but the opposite has no adverse effects on the nMOS transistor. It is possible to use either the SNM optimum length, i.e., 0.250 μm, or the COC optimum because none significantly impact the other. In this case, the COC was prioritized since its increments on pMOS are not as sharp as for the SNM.

Figure 6.6: Current-over-Capacitance (COC) versus transistor channel length.
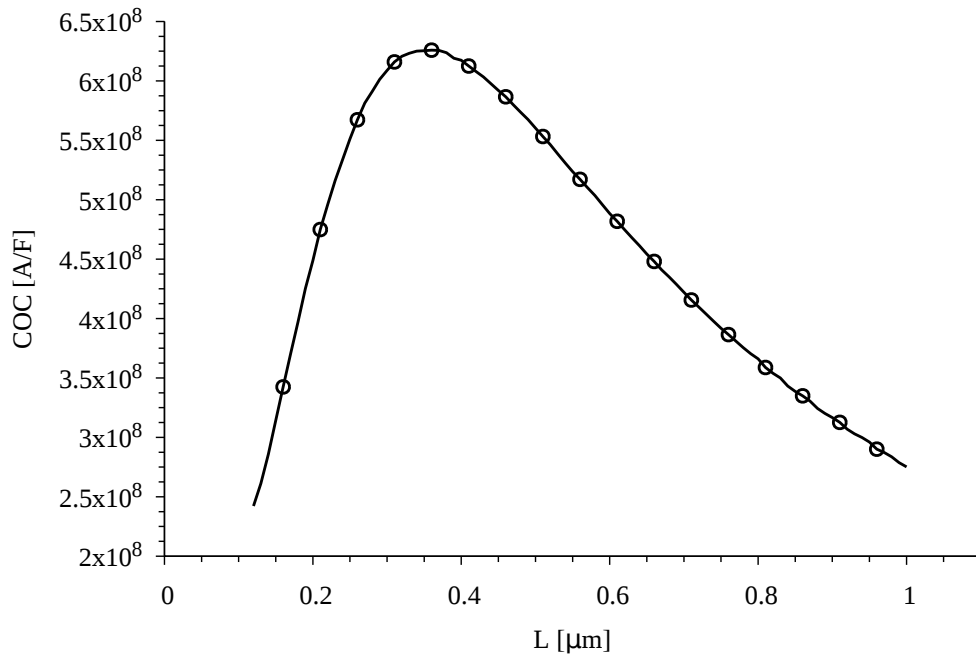
(a) nMOS.
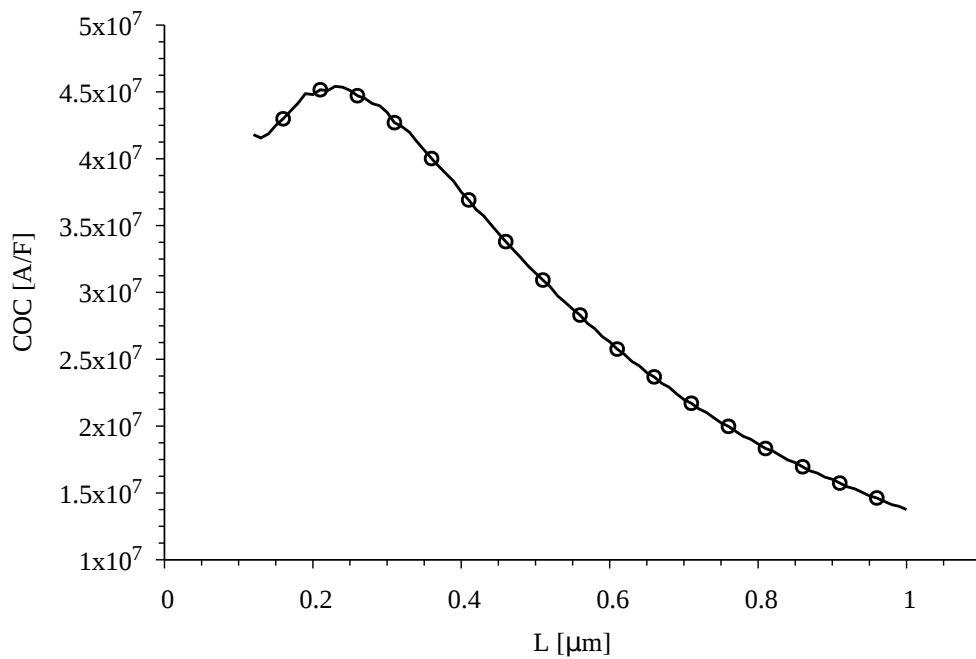
(b) pMOS.

6.2.3.4 Logical Effort

Logical Effort (SUTHERLAND; SPROULL; HARRIS, 1999) is a design strategy to select the size of series transistors. This approach intends to compensate the added parasitic elements of stacked logic cells, to match its delay with an inverter. The original publication is mainly for superthreshold circuits, but Keane et al. (KEANE et al., 2008) has already extended this definition to account the differences from super to subthreshold. Therefore, there is no purpose on re-defining the methodology. Nonetheless, this work uses Keane's strategy because their work is used for comparison in the results section. Series transistors are upsized by traditional rule-of-thumb (RABAEY; CHANDRAKASAN; NIKOLIć, 2007) that multiples transistor size by the number of series components.

6.2.3.5 Design Summary

This Section presents the SNM-aware design strategy used throughout the remaining of this work. For clarity purposes, all considerations were devised for IBM 130nm. However, the same analysis and plots for ST 65nm and TSMC 180nm are available in Appendix B. Results from this Section and the Appendix are summarized in Table 6.3.

Table 6.3: Summary of parameters for all technologies for the SNM-aware CMOS cell design.

Parameter	65nm	130nm	180nm
L_{opt}	0.060 µm	0.220 µm	0.180 µm
β_{opt}	1.2	2.0	3.4
Logical Effort	2	2	2

Overall, the design strategy prioritizes enhancing the SNM of logic cells, but it also evaluates traditional power and delay metrics. Their intent is to avoid that SNM-based decisions harm those metrics, thus analyzing both may produce better results. The strategy presented here can be summarized in the following five-step process:

- **Step 1:** Define nMOS transistor width (W_n) for all strengths in the library.

- **Step 2:** Find the local optimum pMOS/nMOS width ratio $\beta_{opt}*$.

- **Step 3:** Evaluate the global optimum β_{opt} that has the maximum $\sum_{i=1}^{N} SNM$, where $i = 1..N$ represents all cell strengths.

- **Step 4:** Analogous to the SNM-β_{opt}, evaluate the optimum β for cell delay. Adjust β_{opt} according to its impact on this metric.

- **Step 5:** Find SNM local and global optimum channel length, and evaluate its impact

on current using COC plots. Adjust optimum channel length accordingly.

6.3 Library Characterization

One of the main steps of cell-based IC design flow is the cell mapping and inter-connecting to assemble a circuit. This step is called logic synthesis and takes as input a target circuit, which is usually supplied as a register transfer level (RTL) specification written in a hardware description language (HDL). Besides an HDL description, the synthesis uses timing models that guide cell selection and optimization to achieve design-specific timing constraints. These models contain relevant electrical characteristic of each cell (i.e., timing, power, and noise), and are the result of the electrical characterization process. EDA companies provide tools to execute the library characterization (CADENCE, 2015) automatically. These tools usually take as input characterization settings and a SPICE-level netlist containing references to specific transistor models, resistance, and capacitance. Its primary output is the synthesis input models, which is a database that has information about the logic function, area, timing arcs, and power dissipation for all cells. A comprehensive analysis of how the logic synthesis uses this data and the format of the database is outside the scope of this work. In addition, library characterization uses non-linear delay models (NLDM), matching those provided by the foundries of each technology. Therefore, current delay models, such as the Effective Current Source Model (ECSM) and the Composite Current Source Model (CCS), are not discussed in this work (GOYAL; KUMAR, 2005).

According to Gibiluka and Bortolon et al. (BORTOLON et al., 2016) using li-braries characterized at multiple voltage levels allows better trade-offs for voltage scaling (VS) applications. In the latter, authors propose a VS-aware synthesis flow that produces 20% higher clock frequencies in the subthreshold region, with only a 5% penalty in superthreshold, and higher energy savings in both regions. These results show that synthesizing with cells characterized at lower voltages helps the tool to produce better circuits. Correctly characterizing a cell library at multiple voltages, however, is not a straightforward approach. Foundries provide their library only at the nominal voltage and two variations at plus and minus ten percent of its value. Therefore, it is necessary to use the multi-voltage characterization (MV) flow devised in GIBILUKA to characterize the proposed standard cell library to the target subthreshold voltage. The author, nonetheless, developed a framework for the Encounter Library Characterizer (ELC), which is Cadence

deprecated version of Liberate (CADENCE, 2015). Hence, a **minor contribution** of this work is the conversion of the characterization flow devised by Gibiluka to comply with Cadence tool modifications.

In addition to the electrical characterization, the SET tool is used to correlate the SNM with all cells pairs in a library. This information is not available through the standard characterization process, and SET saves it in a standard noise format (SNF) file. The file contains only two columns that specify the cell pair, e.g., INVX1-INVX1, and their SNM value. The next Section further explains the purpose of this file in the logic synthesis.

6.4 Logic Synthesis

One of the main output from logic synthesis is an estimation of the critical path delay, power dissipation, number of cells and area. These numbers are calculated with the information from the database, i.e., the electrical characterization, and the result from cell mapping and interconnecting. Besides giving an idea of the final system characteristics, these parameters are used to guide the synthesis optimization process. For instance, the tool has to calculate the delay of all paths and determine if the critical path delay, i.e., the slower path, respects the timing constraints. If its value is higher than the specified, the tool tries to modify the system until it is either equal or lower. Sometimes these constraints are too strict, and they must be relaxed to allow the synthesis to complete. Since this work compares the synthesis results from different design strategies, it is necessary to adopt an approach to ensure a fair comparison among them. Therefore, the experiments stress the synthesis process until the critical path slack, i.e., the difference from the maximum specified value, is equal to zero.

Another aspect that requires attention for a fair comparison is the area estimation. This metric depends on the cell layout, however, due to the number of libraries tested, only the cell schematic was implemented. Accordingly, the area is based on their design parameters, i.e., transistor width and length, as in the following equation

$$Area = \sum_{i=0}^{N_n} W_n L_n + \sum_{i=1}^{N_p} W_p L_p \tag{6.7}$$

where N_n and N_p are the number of nMOS and pMOS transistors respectively. This simplification is not very accurate, but is useful for comparison purposes.

The logic synthesis, however, does not provide any SNM estimate for the circuit.

Consequently, this work proposes a post-synthesis analysis that extracts the mean static noise margin (μ or μ_{SNM}) and coefficient of variation (σ/μ) of all used cells, as in the following equations:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} SNM \qquad (6.8)$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=0}^{N} |SNM - \mu|^2} \qquad (6.9)$$

where N is the number of cell pairs in the resulting circuit. These equations give an overview of the overall circuit SNM, and how much cells deviate from their mean. Higher μ values imply on more SNM robust circuits, while lower σ/μ mean that used cells have closer SNM values. This last metric is important to ensure that all cells fall into a SNM range specified by the engineer. To calculate their values, after the synthesis optimization, an script iterates over all cell pairs and reads from the SNF file their respectively SNM. More specifically, the scripts identify all interconnected cells and, using the SNF file, determines the SNM for each existing pair.

6.4.1 Results

This Section compares the subthreshold design strategies for IBM 130nm, which were presented in Chapter 5. Table 6.4 summarizes these approaches, their reference and abbreviation. The methodology proposed in this Chapter is referenced as Noise Optimized (NO). The benchmarks considered in these experiments are the combinational circuits from ISCAS 85 suit (HANSEN; YALCIN; HAYES, 1999). These circuits have from four hundred to six thousand cells.

Table 6.4: Summary of the libraries for comparison in IBM 130nm.

Author	Abbreviation	Reference
Bortolon et. al (ours)	NO	This Document
Keane et. al	KS	(KEANE et al., 2008)
Kim et. al	KSE	(KIM et al., 2007)
Nabavi et. al.	NS	(NABAVI et al. 2016)

The comparison is separated into three tables, which show the proposed approach improvements (green cells) and disadvantages (red cells) with regard to others. Tables 6.5, 6.6, and 6.7, respectively, depicted the NO values versus KS, NS, and KSE. Percentage

values are adjusted to represent results according to the parameter, i.e., column. For example, while a +60% delay improvement means faster circuits, i.e. lower delay, the same percentage for SNM means more robust circuits, i.e., higher SNM.

Table 6.5: Normalized SNM-aware design synthesis results to Keane's approach, i.e. NO/KS.

Percentage Improvement/Loss [%]

Design	Leak. P.	Dyn. P.	Total P.	#Cells	Area	Delay	μSNM	σ/μ
c1355	-7.01	-11.86	-11.85	16.01	-88.24	60.63	10.30	42.44
c1908	7.40	2.23	2.24	29.26	-66.92	60.87	10.72	38.28
c2670	-3.48	-1.02	-1.02	16.64	-81.94	62.49	10.93	37.47
c3540	17.71	-1.59	-1.56	35.52	-66.23	59.01	10.96	34.47
c432	2.71	-4.98	-4.97	24.06	-88.46	67.07	10.99	34.32
c499	-4.02	-6.66	-6.66	17.71	-78.68	58.70	10.79	34.96
c5315	6.35	-0.32	-0.31	34.74	-64.01	60.24	10.88	34.52
c6288	18.79	0.86	0.88	32.57	-70.98	60.30	11.35	32.67
c880	10.46	1.59	1.60	28.46	-54.93	61.10	11.34	32.56

In general, the proposed approach exhibits higher $μ_{SNM}$ regardless the other technique, thus indicating that the design choices met their goal. The improvements are around 11% for KS and 12% for NS, respectively, Tables 6.5 and 6.6, but increase to 23% for KSE, Table 6.7, due the asymmetrical channel length impact on current balance. The design SNM variability (σ/μ) is lower, i.e., better, in comparison with KS and KSE (30% to 60%), but a little higher than NS (1% to 7%). The cell pairs in Nabavi's (NS) library have closer SNM estimates to each other, thus implying on less variability. Nonetheless, their absolute value is much lower than our method and thus does not compensate the small variability improvement. The Tables also evidence a tight relationship between the delay and SNM. Balancing pMOS and nMOS current to produce better SNM had a 40-60% improvement. Evidently, increasing pMOS current, which usually has lower mobility than the nMOS, balances the average delay T_D.

On the other hand, these improvements come with a moderate to severe area penalty (50% to 160%), because the proposed strategy increases both transistor width and channel length. Moreover, NO consumes 3% and 6%, on average, more power than KS and NS respectively. In spite of that, leakage power reduced in some designs due the channel length increase as demonstrated in (GUPTA et al., 2004). From all approaches, only KSE has worst area and power consumption as their channel length is much higher than NO.

Table 6.6: Normalized SNM-aware design synthesis results to Nabavi's approach, i.e. NO/NS.

Percentage Improvement/Loss [%]

Design	Leak. P.	Dyn. P.	Total P.	#Cells	Area	Delay	µSNM	σ/µ
c1355	-1.96	-15.84	-15.82	28.29	-150.98	65.01	13.00	-0.77
c1908	1.33	-4.11	-4.11	33.70	-146.59	66.73	12.78	-4.74
c2670	-12.16	-5.07	-5.08	21.91	-160.26	67.48	12.58	-4.26
c3540	17.22	-5.42	-5.38	42.11	-131.10	64.14	12.62	-7.14
c432	13.56	-8.12	-8.07	35.70	-151.28	68.22	12.59	-7.82
c499	-4.39	-11.93	-11.91	30.88	-143.00	65.00	12.46	-6.73
c5315	7.69	-5.24	-5.22	38.36	-132.01	66.14	12.49	-7.17
c6288	13.72	-5.76	-5.73	42.23	-133.54	64.13	12.88	-2.13
c880	8.67	-6.16	-6.14	37.52	-131.58	66.68	12.87	-2.27

Table 6.7: Normalized SNM-aware design synthesis results to Kim's approach, i.e. NO/KSE.

Percentage Improvement/Loss [%]

Design	Leak. P.	Dyn. P.	Total P.	#Cells	Area	Delay	µSNM	σ/µ
c1355	39.02	36.90	36.91	17.68	9.22	45.61	23.75	59.36
c1908	40.72	16.44	16.48	23.30	4.82	45.41	23.88	55.74
c2670	36.63	13.60	13.64	10.55	15.85	46.70	23.85	54.45
c3540	55.23	26.52	26.59	35.92	36.94	44.86	23.68	55.68
c432	45.81	17.25	17.33	23.89	17.65	48.93	23.87	55.58
c499	38.28	32.38	32.39	22.01	8.65	45.63	24.18	55.49
c5315	48.55	15.75	15.81	32.37	23.13	46.55	24.01	54.72
c6288	56.67	44.39	44.41	26.23	50.35	47.38	20.37	62.39
c880	47.95	16.62	16.69	28.92	26.91	43.19	20.39	62.24

6.5 Chapter Summary

This Chapter presented a thorough discussion ranging from the cell library creation to its validation. More specifically, the text provides specific details about the cell selection for the library, the electrical characterization, the design choices and the logic synthesis results. The contributions of this Chapter are two-fold. The major is the SNM-aware design methodology using IBM 130nm CMOS bulk technology to illustrate, while the minor is simple, but novel, approach to calculate the SNM of complex digital circuits. Literature had only evaluated thus far the SNM for simple pair of cells, e.g. NAND-NOR. The proposed methodology *achieved its main goal*, which was increasing the circuit SNM robustness. Additionally, for IBM 130nm, there was a meaningful improvement in delay with a severe and low penalties in area and power, respectively. Appendix B completes this Chapter discussion for ST 65nm and TSMC 180nm CMOS bulk technologies. Moreover, the adaptation of part of Gibiluka's characterization flows can also be considered as a minor contribution.

7 CONCLUSIONS

This Dissertation investigated SNM as a metric of robustness for near and sub-threshold combinational digital circuits. This investigation comprises several steps of digital IC design ranging from the gate design and characterization to the logic synthesis. Additionally, conclusions are validated through extensive simulation in three commercial CMOS bulk technologies: ST 65nm, IBM 130nm, and TSMC 180nm. The original contributions of this work can be summarized in three items: *(i)* the development of a tool to assess the SNM of logic gates; *(ii)* the definition of specific guidelines to estimate the SNM of subthreshold cells; and *(iii)* a methodology to design SNM robust circuits and evaluate the SNM of complex circuits.

The first contribution is the SNM estimation tool (SET) that assesses the SNM of any combinational cells with the same unateness. This tool considers process, voltage and temperature variations which are essential attributes to evaluate circuits operating in the sub and near-threshold. In addition, it sorts and organizes results, providing meaningful data for the user.

The second contribution is the definition of a simulation setup and criteria to consistently estimate SNM considering PVT variations. The simulation setup should switch only one input, whichever is closer to the output rail, to avoid unrealistic SNM values. The suggested criteria is the maximum-equals because it weights equally high and low noise margins, and hence behaves better when considering PVT variations. Other criteria demonstrated that prioritizing one noise component produces more pessimistic results in such conditions.

Finally, the main contribution is the SNM-aware cell design and the evaluation of SNM of complex circuits after the logic synthesis. This study indicates the primary design parameters, and their physical meaning, that influence the SNM, and proposes a systematic approach to enhance its value. Interestingly, the resulting standard cell library had meaningful improvements in delay besides the expected SNM. Nevertheless, there was a considerable area penalty given that the sizing strategy creates larger cells. Despite that, power consumption did not substantially increase since this parameter was considered at design time. Moreover, the voltage scaling analysis suggests an optimum SNM operating point in the near-threshold region. The study shows that, for simple gates, the SNM relative to the supply voltage increases in this region regardless of the implemented logic function.

7.1 Future Work

This work opens many topics for future research and it points to ways to tackle them with the SET tool. The most straightforward topics are extending SET to consider sequential cells, and using more complex benchmarks in logic synthesis. Developing these suggestions should prove to be simple, as their framework is almost complete. Another valuable investigation is the impact of body-bias techniques in the SNM, and how to employ them in voltage scaling circuits. Furthermore, two other studies were briefly explored throughout the Dissertation development but were not included in this document. The first is to extend the maximum-equals criteria to allow the evaluation of cells that have different unateness. Partial results confirm that this is possible through the overlap of VTC curves that compose positive unate cells, but it requires further modeling and simulation. The name given to this approach is dragonfly method. The second is to analyze the SNM of cells paths during the logic synthesis to detect and change cell pairs with low SNM. Its strategy is to embed the SNM inside the timing tables and let the static timing analysis tool to implicitly remove cell pairs with SNM lower than specified by the designer.

REFERENCES

AL-FUQAHA, A. et al. Internet of things: A survey on enabling technologies, protocols, and applications. **IEEE Communications Surveys & Tutorials**, v. 17, n. 4, p. 2347–2376, 2015.

ALIOTO, M. Understanding dc behavior of subthreshold cmos logic through closed-form analysis. **IEEE Transactions on Circuits and Systems I: Regular Papers**, v. 57, n. 7, p. 1597–1607, Jul 2010.

BECER, M. R. et al. Post-route gate sizing for crosstalk noise reduction. In: **International Symposium on Quality Electronic Design**. [S.l.: s.n.], 2003. p. 171–176.

BEIU, V. et al. Enabling sizing for enhancing the static noise margins. In: **International Symposium on Quality Electronic Design**. [S.l.: s.n.], 2013. p. 278–285.

BEIU, V. et al. On upsizing length and noise margins. In: **International Semiconductor Conference**. [S.l.: s.n.], 2013. v. 2, p. 219–222.

BENNETT, W. R. **Electrical Noise**. New York, 1960. 1-30 p. Accessed February 22, 2018. Available from Internet: <http://hdl.handle.net/2027/wu.89042783910>.

BORTOLON, F. T. et al. Design and analysis of the hf-risc processor targeting voltage scaling applications. In: **Symposium on Integrated Circuits and Systems**. [S.l.: s.n.], 2016.

CADENCE. **Noise-Aware Timing Analysis**. [S.l.], 2001. 9 p.

CADENCE. **Virtuoso Liberate Reference Manual**. [S.l.], 2015.

CALHOUN, B.; WANG, A.; CHANDRAKASAN, A. Modeling and sizing for minimum energy operation in subthreshold circuits. **Journal of Solid-State Circuits**, v. 40, n. 9, p. 1778–1786, Sep 2005.

DE, V.; VANGAL, S.; KRISHNAMURTHY, R. Near threshold voltage (ntv) computing: Computing in the dark silicon era. **IEEE Design & Test**, v. 34, n. 2, p. 24–30, April 2017.

DENNARD, R. et al. Design of ion-implanted mosfet's with very small physical dimensions. **Journal of Solid-State Circuits**, v. 9, n. 5, p. 256–268, Oct 1974.

DIMOULAS, A. et al. (Ed.). **Advanced Gate Stacks for High-Mobility Semiconductors**. [S.l.]: Springer, 2007. 1–19 p. (Advanced Microelectronics, 27).

DING, L.; MAZUMDER, P. Dynamic noise margin: Definitions and model. In: **International Conference on VLSI Design**. [S.l.: s.n.], 2004. p. 1001–1006.

DOWNING, R.; GEBLER, P.; KATOPIS, G. Decoupling capacitor effects on switching noise. **IEEE Transactions on Components, Hybrids, and Manufacturing Technology**, v. 16, n. 5, p. 484–489, Aug 1993.

DRESLINSKI, R. G. et al. Near-threshold computing: Reclaiming moore's law through energy efficient integrated circuits. **Proceedings of the IEEE**, v. 98, n. 2, p. 253–266, Feb 2010.

EDENFELD, D. et al. 2003 technology roadmap for semiconductors. In: **IEEE Computer Society**. [S.l.: s.n.], 2004.

ESMAEILZADEH, H. et al. Dark silicon and the end of multicore scaling. In: . [S.l.]: ACM Press, 2011. p. 365.

GAO, T.; LIU, C. L. Minimum crosstalk channel routing. In: **International Conference on Computer Aided Design**. [S.l.: s.n.], 1993. p. 692–696.

GIBILUKA, M. **Analysis of Voltage Scaling Effects in the Design of Resilient Circuits**. Thesis (Dissertation) — PUCRS, Porto Alegre, BR, Mar 2016.

GOYAL, R.; KUMAR, N. **Current Based Delay Models: A Must For Nanometer Timing**. [S.l.], 2005. 8 p.

GRAY, P. R. **Analysis and Design of Analog Integrated Circuits**. 5. ed. [S.l.]: Wiley, 2010. 881 p.

GUPTA, P. et al. Selective gate-length biasing for cost-effective runtime leakage control. p. 327–330, Jun 2004.

HANSEN, M. C.; YALCIN, H.; HAYES, J. P. Unveiling the iscas-85 benchmarks: a case study in reverse engineering. **IEEE Design Test of Computers**, v. 16, n. 3, p. 72–80, 1999.

HANSON, S. et al. A low-voltage processor for sensing applications with picowatt standby mode. In: **IEEE Journal of Solid-State Circuits**. [S.l.: s.n.], 2009. v. 44, n. 7, p. 1145–1155.

HANSON, S. et al. A low-voltage processor for sensing applications with picowatt standby mode. **Journal of Solid-State Circuits**, v. 44, n. 4, p. 1145–1155, Apr 2009.

HANSON, S. et al. Nanometer device scaling in subthreshold logic and sram. **IEEE Transacions on Electron Devices**, v. 55, p. 175–185, Jan 2008.

HAUSER, J. R. Noise margin criteria for digital logic circuits. **IEEE Transactions on Education**, v. 36, n. 4, p. 363–368, nov. 1993.

HEYDARI, P.; PEDRAM, M. Ground bounce in digital vlsi circuits. **IEEE Transactions on Very Large Scale Integration Systems**, v. 11, n. 2, p. 180–193, Apr 2003.

HILL, C. F. Noise margin and noise immunity in logic circuits. In: **Microelectronics**. [S.l.: s.n.], 1968. v. 1, p. 16–21.

HIMANSHU, K. et al. Near-threshold voltage (ntv) design — opportunities and challenges. In: **Design Automation Conference**. San Francisco: IEEE, 2012. p. 1149–1154.

JIANG, I. H. R.; CHANG, Y.-W.; JOU, J.-Y. Crosstalk-driven interconnect optimization by simultaneous gate and wire sizing. **IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems**, v. 19, n. 9, p. 999–1010, Sep 2000.

KAESLIN, H. **Digital Integrated Circuit Design: From VLSI Architectures to CMOS Fabrication**. [S.l.]: Cambridge Univ. Press, 2008. 886 p.

KATOPIS, G. A. Delta-i noise specification for a high-performance computing machine. In: **Proceedings of the IEEE**. [S.l.: s.n.], 1985. v. 73, p. 1405–1415.

KAUL, H.; SYLVESTER, D.; BLAAUW, D. Active shields: A new approach to shielding global wires. ACM Press, p. 112, 2002.

KEANE, J. et al. Stack sizing for optimal current drivability in subthreshold circuits. **IEEE Transactions on Very Large Scale Integration (VLSI) Systems**, v. 16, n. 5, p. 598–602, May 2008.

KIM, T. H. et al. Utilizing reverse short-channel effect for optimal subthreshold circuit design. **IEEE Transactions on Very Large Scale Integration Systems**, v. 15, n. 7, Jul 2007.

KWONG, J.; CHANDRAKASAN, A. P. Variation-driven device sizing for minimum energy sub-threshold circuits. In: **Symposium on Low Power Electronics and Design**. [S.l.]: IEEE, 2006. p. 8–13.

LOHSTROH, J. Static and dynamic noise margins of logic circuits. **Journal of Solid-State Circuits**, v. 14, n. 3, p. 591–598, Jun 1979.

MOORE, G. E. No exponential is forever: but "forever" can be delayed! In: **International Solid-State Circuits Conference**. [S.l.: s.n.], 2003.

NABAVI, M.; RAMEZANKHANI, F.; SHAMS, M. Optimum pmos-to-nmos width ratio for efficient subthreshold cmos circuits. **IEEE Transactions on Electron Devices**, v. 63, n. 3, p. 916–924, Mar 2016.

OLIVERA, F.; PETRAGLIA, A. Analytic modeling of static noise margin considering DIBL and body bias effects. In: **International Symposium on Circuits and Systems**. [S.l.: s.n.], 2017. p. 1–4.

ORGUC, S. et al. 0.3 v ultra-low power sensor interface for emg. In: **European Solid State Circuits Conference**. [S.l.: s.n.], 2017. p. 219–222.

OTT, H. W.; OTT, H. W. **Electromagnetic compatibility engineering**. [S.l.]: John Wiley & Sons, 2009. 475-476 p.

RABAEY, J. M.; CHANDRAKASAN, A. P.; NIKOLIć, B. **Digital Integrated Circuits: A Design Perspective**. [S.l.]: Prentice-Hall, 2007. 761 p.

SALMAN, E. **Switching Noise and Timing Characteristics in Nanoscale Integrated Circuits**. Thesis (Thesis) — University of Rochester, Rochester, New York, 2009.

SEEVINCK, E.; LIST, F.; LOHSTROH, J. Static-noise margin analysis of mos sram cells. **IEEE Journal of Solid-State Circuits**, v. 22, p. 748–754, 1987.

SEGURA, J.; HAWKINS, C. F. **CMOS electronics: How it Works, How it Fails**. New York: IEEE Press ; Wiley-Interscience, 2004. 356 p. ISBN 978-0-471-47669-6.

SHAFIQUE, M. et al. The eda challenges in the dark silicon era: Temperature, reliability, and variability perspectives. In: **Design Automation Conference**. [S.l.]: ACM Press, 2014. p. 1–6.

SHEPARD, K. L.; CHOU, K. Cell characterization for noise stability. In: **Custom Integrated Circuits Conference**. [S.l.: s.n.], 2000. p. 91–94.

SUTHERLAND, I. E.; SPROULL, R. F.; HARRIS, D. **Logical Effort: Designing Fast CMOS Circuits**. San Francisco, Calif: Morgan Kaufmann Publishers, 1999. 1-80 p.

TAJALLI, A.; LEBLEBICI, Y. Design trade-offs in ultra-low-power digital nanoscale cmos. **IEEE Transactions on Circuits and Systems I: Regular Papers**, v. 58, n. 9, p. 2189–2200, Sep 2011.

TSIVIDIS, Y. **Operation and Modeling of the MOS Transistor**. New York, NY, USA: McGraw-Hill, Inc., 1987. 1-200 p.

VENUGOPAL, R.; CHAKRAVARTHI, S.; CHIDAMBARAM, P. Design of cmos transistors to maximize circuit fom using a coupled process and mixed-mode simulation methodology. **Electron Device Letters**, v. 27, p. 863–865, Oct 2006.

ZHANG, J.; FRIEDMAN, E. Effect of shield insertion on reducing crosstalk noise between coupled interconnects. IEEE, p. 529–532, May 2004.

ZURADA, J. M.; JOO, Y. S.; BELL, S. V. Dynamic noise margins of mos logic gates. In: **International Symposium on Circuits and Systems**. [S.l.: s.n.], 1989. p. 1153–1156 vol.2.

APPENDIX A — SET DEVELOPER GUIDE

Even though the SET provides a wide range of functions, there is space left for enhancements and other features. Perhaps some new estimation metrics, a graphical interface, different types of analyses, and so on. For this reason, this Section details the internal behavior and structure of SET execution flow depicted in Figure 3.1.

A.1 The Software

The software, written in C-language, is responsible for estimating the SNM of a cell pair using the technique selected through the `-t` flag. As illustrated in Figure 3.1, this is an internal block of the flow and, hence, the *user* should only access it through the wrapping scripts. Using the standalone version can lead to incorrect behavior because the script is responsible for several consistency checks. The *developer*, on the other hand, should be capable of using this block independently to insert in his projects. For such, its command line syntax is:

```
snm [bf|mpc] [size] [vdd] [file_name]
```

where, `bf|mpc` references the previously discussed techniques, `size` is the number of points in the DC analysis, `vdd` is the supply voltage, and `file_name` is the DC simulation output.

Considering that this command is executed without the scripts, the aforementioned consistency checks must be done otherwise. First, the software expects a DC output that complies with the three column format outlined in Figure A.1. The file header (lines 1-5) content is not relevant; however, the software must know the number of lines it comprises to separate it from the VTC data. Second, the number of entries in this file must match the `size` argument. For example, in Figure A.1 there must be `size` entries between the 6th and the last lines of the file. Finally, the supply voltage, i.e. `vdd`, must match that used in the DC simulation.

Given that all these simple constraints are met, the software is ready execute the flow illustrated in Figure A.2. After loading the Voltage Transfer Characteristic (VTC) curves, it verifies the spacing between each point, i.e. precision, and interpolates those that do not met a certain threshold. More specifically, to estimate the SNM it is necessary to find the exact point where the VTCs under analysis cross. This usually lays in the

Figure A.1: DC analysis sample output containing the VTCs.

```
1  * * STATIC NOISE MARGIN TOOL
2  ******  DC Analysis ( dcrun )  tnom=27.0 temp=27.0
3  ******
4  x
5     dc            v(_gate_1)        v(_gate_2)
6      0          249.616e-03       249.616e-03
7     500e-06    249.606e-03       249.606e-03
8     750e-06    249.600e-03       249.600e-03
9     1e-03      249.595e-03       249.595e-03
10
11        ...
```

transition region between logic level zero and one, where even small changes in the input lead to significant changes in the output. This lead to less points in this region and hence it becomes harder to detect the cross section between curves, as depicted in Figure A.3. Increasing the number of points in the DC simulation is an alternative not as efficient as interpolating because SPICE simulators, such as Cadence® Spectre®, evenly distribute their points and thus most of the new ones will be located in the extremes. Once the VTCs are properly tuned, the software estimates the SNM and outputs the result.

A.2 The Bash Scripts

While the software is the SET core, the scripts are the wrappers that pre- and post-process the software input and output respectively. Accordingly, this Subsection explains the subtleties that compose this outer shell as devised in Figure 3.1.

A.2.1 Pre-Processing

The pre-processing block is responsible for configuring the environment according to the user input. Initially, it loads the SPICE simulator, which is Cadence Spectre® in this dissertation, and processes the user parameters. This defines the operating mode, i.e., MAM or SAM, and thus the bash routines that all blocks should execute. For example, if the user specified a parametric analysis for temperature, the DC Analysis knows that there will be several SPICE simulations to execute and the Post-processing must adapt the output to a format that is easy to read.

Figure A.2: Software detailed flowchart.
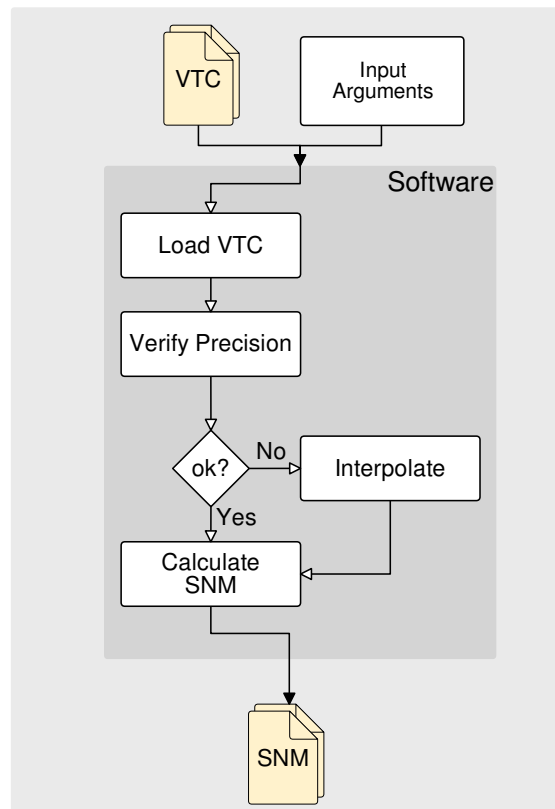
This block also reads the configuration file and parses all the information for the electrical simulation. From it, it first verifies that the specified transistor model and cell library exists, and informs the user if there are any problems to reach those files. Given the path is correct, it creates the working directory and the SPICE input with the information inside the .cfg. If the tool is running on MAM with multiple cell pairs, different directories are created to separate the results. On the other hand, for parametric analysis the same input files is reused but altering the selected parameters, i.e. either voltage or temperature. Moreover, for distributed systems where storage is synchronized through the network, this block further ensures that all processing is local to avoid flooding the network. After configuring the environment, the Pre-processing block invokes the DC Analysis block passing the necessary information, e.g. directory location and technique type.

A.2.2 DC Analysis

The DC Analysis executes the SPICE simulations according the specified analysis and formats the output file. If the parametric simulation is active, this block runs multiple DC simulation to extract the VTC for different voltage/temperature values. Afterwards, it

Figure A.3: Example of an abrupt response from a cell, leading to lesser points in the transition region.
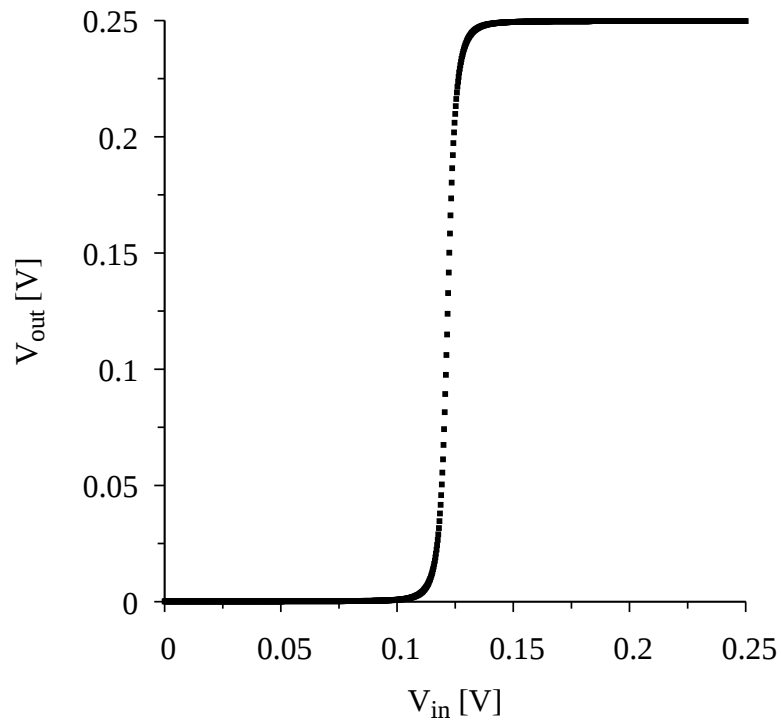
adapts the outfile file format to comply with that expected by the software. For Spectre®, the unique change is to transform numbers from S.I. unit system to scientific E-notation. This modifications depend on the simulator and thus should be adapted for other vendors.

A.2.3 Post-processing

Once all outputs from the DC Analysis have been processed by the software, the Post-processing block is invoked. This block handles the output from the SNM estimation software before forwarding to the user. This is specially important for the Monte Carlo analysis where the software produces multiple SNM estimations. Therefore, this block gathers all results and calculates the mean, the standard deviation and the variance. Additionally, it generates ready-to-plot logs for a visual inspection of the samples spread.

APPENDIX B — SNM-AWARE DESIGN RESULTS

This Appendix presents the results for the SNM-Aware design for TSMC 180nm and ST 65nm CMOS bulk technologies. The discussion follows the same outline of Chapter 6, however it is synthesized to the critical points. The reader is encouraged to refer to this Chapter for a thorough analysis of the proposed design methodology.

B.1 TSMC 180nm

The first step of the SNM-aware design methodology is to define the nMOS transistor width (W_N) for all strengths in the library. Afterward, the second step is to find the local optimum width ratio (β^*) for pMOS/nMOS transistors. The local optimum represents the point that, for a giving W_N, maximizes the static noise margin. Table B.1 shows the results of these two steps for TSMC 180nm.

Table B.1: TSMC 180nm nMOS transistor width for each strength and their local optimum β.

	X1	X2	X4	X6
W_n (μm)	0.220	0.440	0.880	1.320
$\beta_{opt}*$	10	7.6	3.9	3.2

Figure B.1: SNM Trade-offs for different inverter strengths versus β.
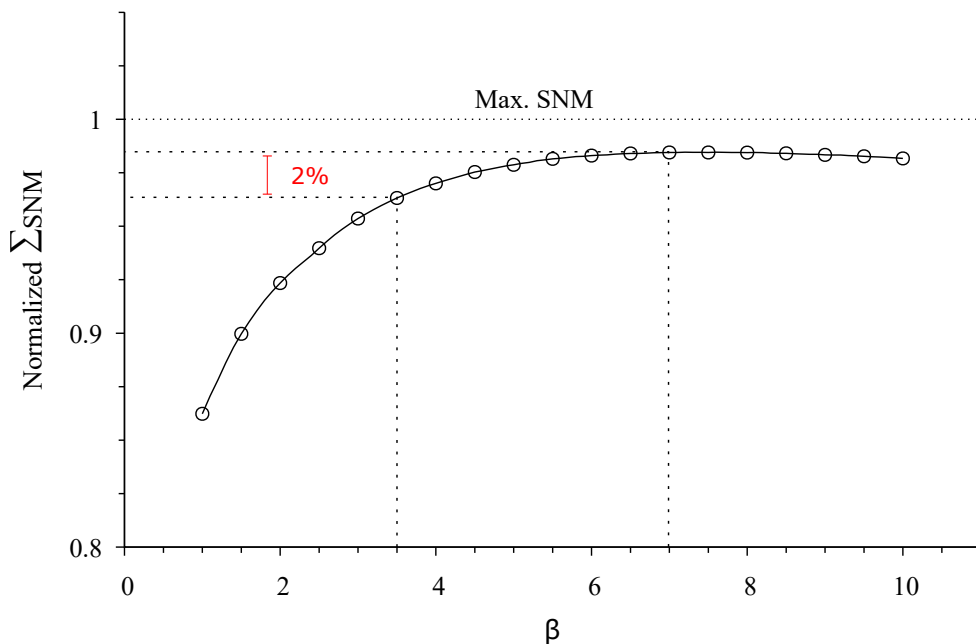
Once the local optimum is defined for every nMOS width, the next step finds the global optimum width ratio (β_{opt}) that produces the maximum noise margin for all driving strengths. This analysis depicts what β value offers the best SNM trade-offs for all W_N. Accordingly, Figure B.1 demonstrates that for TSMC 180nm the β_{opt} value is very high, i.e., around 7 to 8. This β, hence, would imply on a significant area penalty even for the minimum nMOS width, i.e., W_P would be $7 * W_N$.

Figure B.2: Inverter T_D delay trade-off versus β for TSMC 180nm.
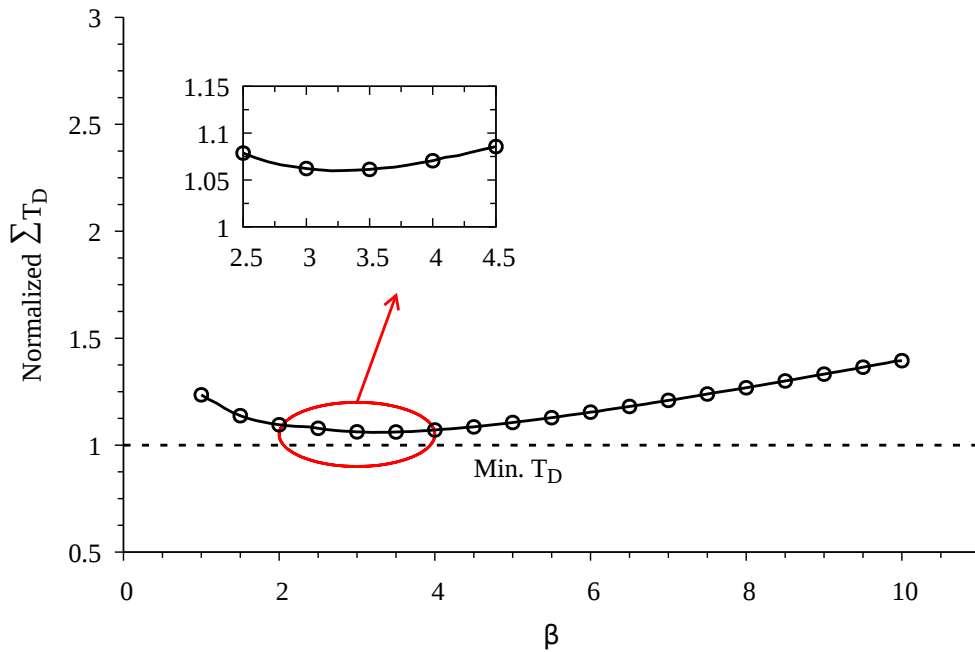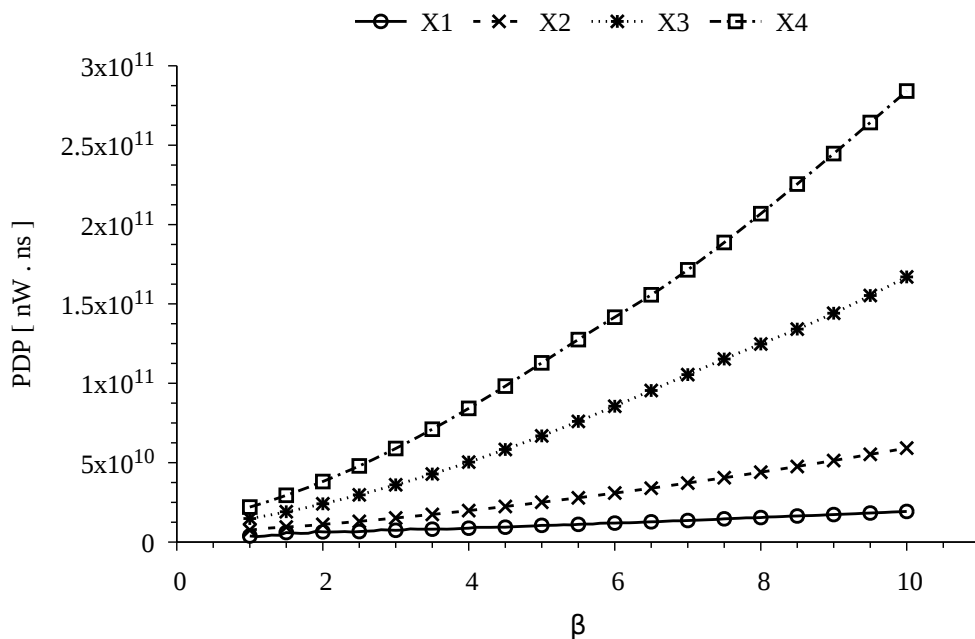

Figure B.3: Inverter PDP trade-off versus β for TSMC 180nm.

The SNM-aware design methodology, nonetheless, further evaluates the β implication on timing characteristics. Figure B.2 shows that the optimum width ratio for timing lies between 3 and 3.5. Beyond this range the delay trade-offs, i.e., the sum of delays for each W_N, reduces. Therefore, instead of using the optimal β for SNM which degrades the timing characteristics, the β_{opt} is chosen for delay. The penalty from the highest $\sum SNM$ is only 0.020 and, besides optimizing delay, the area is reduced, thus implying on less power dissipation, as depicted in Figure B.3.

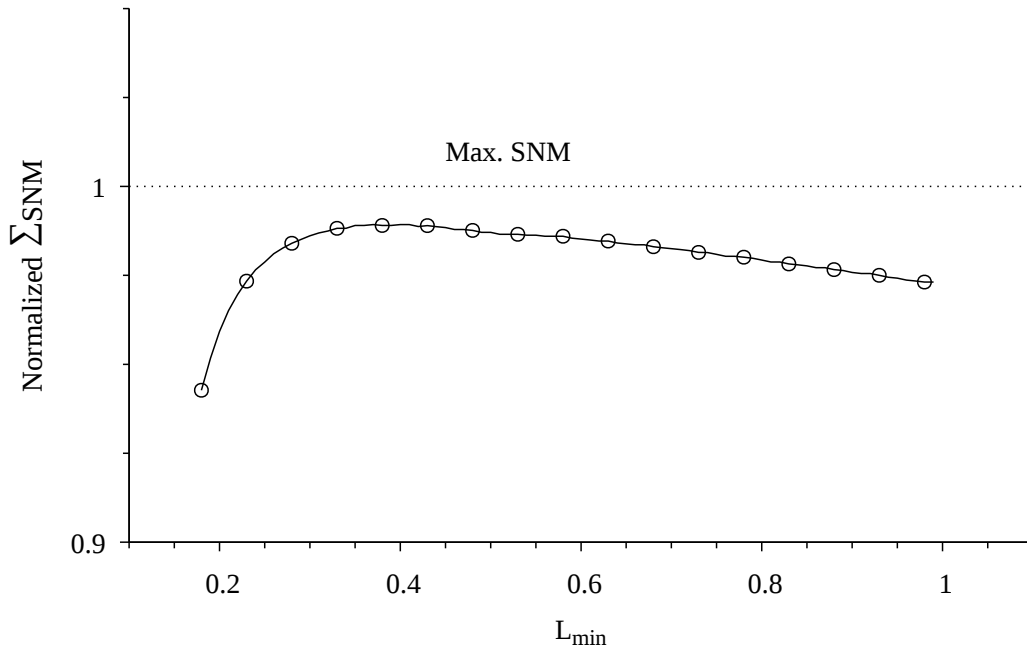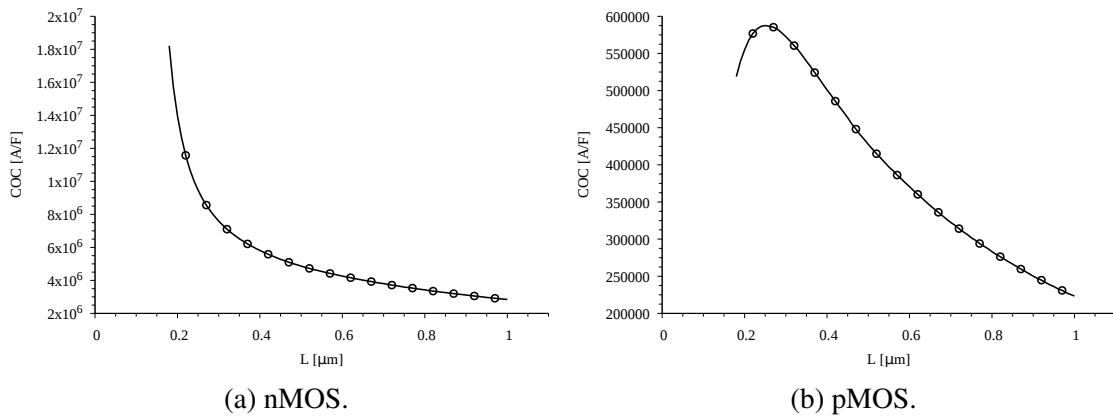Figure B.4: Inverter channel length trade-off versus SNM for TSMC 180nm.

Figure B.5: Current-over-Capacitance (COC) versus transistor channel length.

(a) nMOS.

(b) pMOS.

Finally, the last step from the proposed methodology is to evaluate the channel length impact on the SNM and current over capacitance (COC). Figure B.4 shows that the optimum channel length (L_{opt}) is close to 0.35 µm. On the other hand, the COC for

nMOS transistor has an exponential decrease with small increments in L, as depicted in Figure B.5.(a). The pMOS COC current increment is too small compared to the nMOS decrease to justify any increase in the channel length. Therefore, to avoid decreasing nMOS current, L_{opt} is set to the minimum channel length.

Table B.2: Normalized SNM-aware design synthesis results to Nabavi's approach, i.e., NO/NS, for TSMC 180nm.

	Percentage Improvement/Loss [%]							
Design	Leak. P.	Dyn. P.	Total P.	#Cells	Area	Delay	µSNM	σ/µ
c1355	25.39	12.65	12.65	33.05	34.07	-13.20	2.62	33.42
c1908	10.96	-0.01	-0.01	15.17	5.15	1.54	2.97	32.89
c2670	-7.17	-14.60	-14.60	3.35	-43.24	-4.44	3.16	31.80
c3540	-35.25	-28.71	-28.71	-24.37	-69.54	10.57	3.43	33.09
c432	-19.40	-20.37	-20.37	-10.13	-42.86	6.97	3.42	33.12
c499	15.53	17.34	17.34	16.59	16.74	3.06	3.42	34.71
c5315	-11.83	-10.96	-10.96	-4.35	-40.43	6.95	3.55	33.01
c6288	-1.90	-9.84	-9.83	3.27	-29.60	2.75	3.65	29.28
c880	-11.15	-16.16	-16.16	-0.77	-34.41	0.81	3.66	29.27

Table B.3: Normalized SNM-aware design synthesis results to Calhoun's approach, i.e., NO/CS, for TSMC 180nm.

	Percentage Improvement/Loss [%]							
Design	Leak. P.	Dyn. P.	Total P.	#Cells	Area	Delay	µSNM	σ/µ
c1355	40.44	53.38	53.38	13.77	74.03	5.65	-4.28	6.62
c1908	31.90	46.17	46.17	1.03	69.93	17.99	-4.01	5.42
c2670	33.31	46.82	46.82	3.11	65.81	16.64	-4.09	2.89
c3540	27.05	48.61	48.60	-11.37	63.26	27.37	-3.84	6.06
c432	29.74	45.50	45.50	-1.81	68.97	21.43	-3.93	5.57
c499	29.44	51.04	51.04	-5.14	61.59	19.36	-3.87	7.73
c5315	37.09	53.14	53.13	0.33	66.26	26.28	-4.18	5.12
c6288	38.52	55.69	55.69	3.54	64.35	27.08	-3.06	10.68
c880	29.62	46.87	46.87	-7.20	63.56	26.84	-3.02	7.09

The resulting library is compared against Nabavi et. al and Calhoun et. al, respectively, in Tables B.2 and B.3. In general, the proposed approaches exhibits higher μ_{SNM} for Nabavi, but smaller than Calhoun. The reason is that Calhoun defined β to equalize pMOS and nMOS currents, thus yielding better SNM. Nonetheless, Calhoun's strategy consumes significant more area, dissipates more power and has slower circuits. Moreover, since the proposed strategy balances the SNM among all transistor width, the SNM spread σ/µfrom Calhoun is worst. Compared against Nabavi, the proposed design methodology consumes more power (8% on average) and more area (21% on average), but achieves better delay on the majority benchmarks (2% on average).

B.2 ST 65nm

The ST 65nm bulk CMOS technology has well balanced pMOS and nMOS threshold voltage and currents at the minimum width, as indicated in Chapter 5 discussion. Therefore, the local optimum width ratio for available W_N is concentrated at one, as depicted in Table B.4. Figure B.6 further shows that the best SNM trade-off, i.e., $\sum SNM$, is near one.

Table B.4: ST 65nm nMOS transistor width for each strength and their local optimum β.

	X1	X2	X4	X6
W_n (µm)	0.140	0.280	0.560	0.840
$\beta_{opt}*$	3.3	1.3	1	1

Figure B.6: SNM Trade-offs for different inverter strengths versus β.
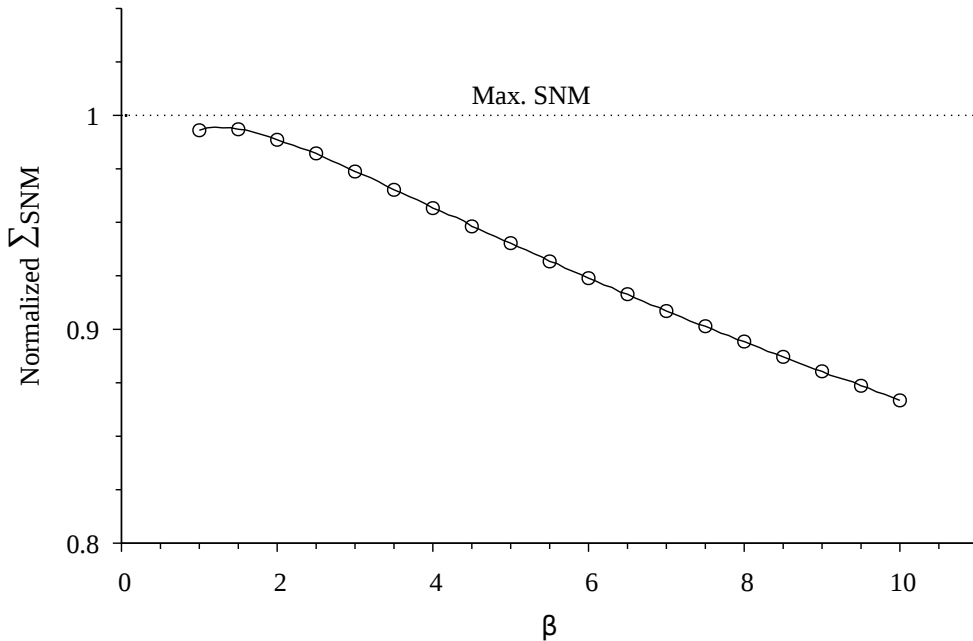
In fact, the influence of the advanced manufacturing process is noticeable even in the delay (Figure B.7) and power consumption (Figure B.8). This process leaves little space for improving cell characteristics based on β. Accordingly, the proposed library uses β_{opt} equals to 1.2, which has slightly higher $\sum SNM$ than the minimum 1.

Similarly to TSMC 180nm, it is not beneficial to increase the channel length to get higher SNM due the exponential decrease in current in ST 65nm. Consequently, the minimum channel length is also used for this technology. The resulting library has little

Figure B.7: Inverter T_D delay trade-off versus β for ST 65nm.
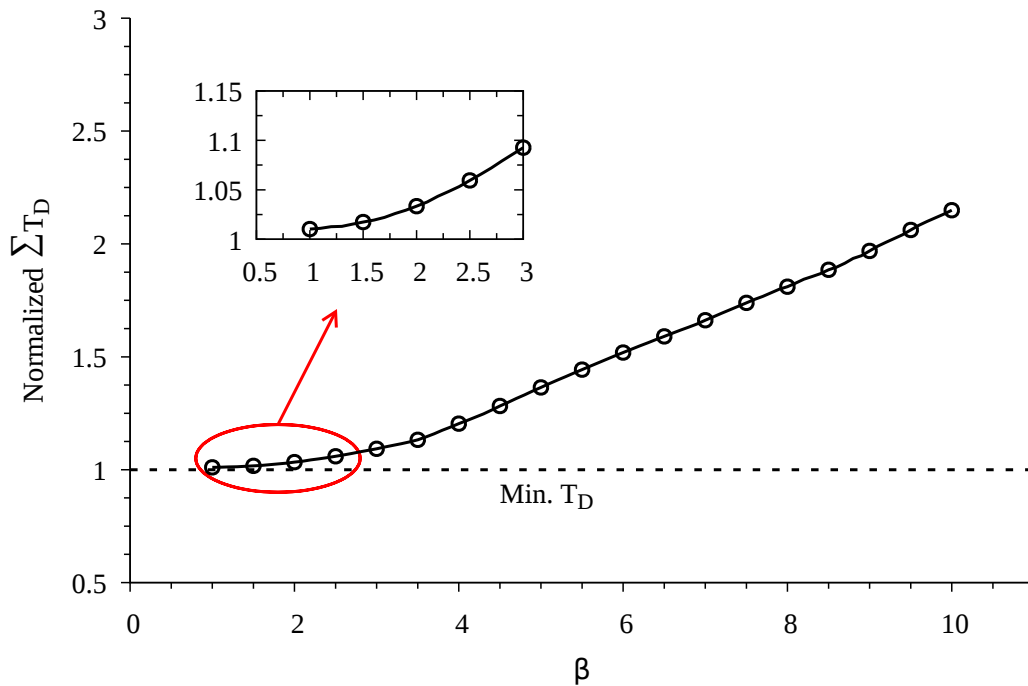
Figure B.8: Inverter PDP trade-off versus β for ST 65nm.
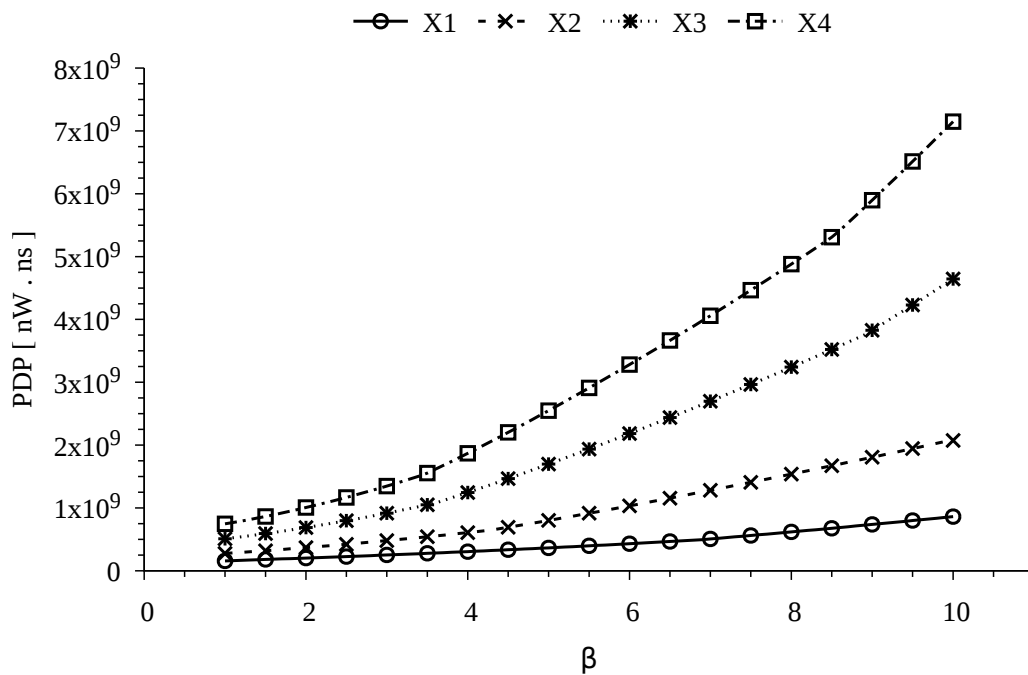
difference with Nabavi and Keane's strategies. In spite of that, the logic synthesis outcome is very different.

The comparison against Nabavi et. al and Keane et. al are shown, respectively,

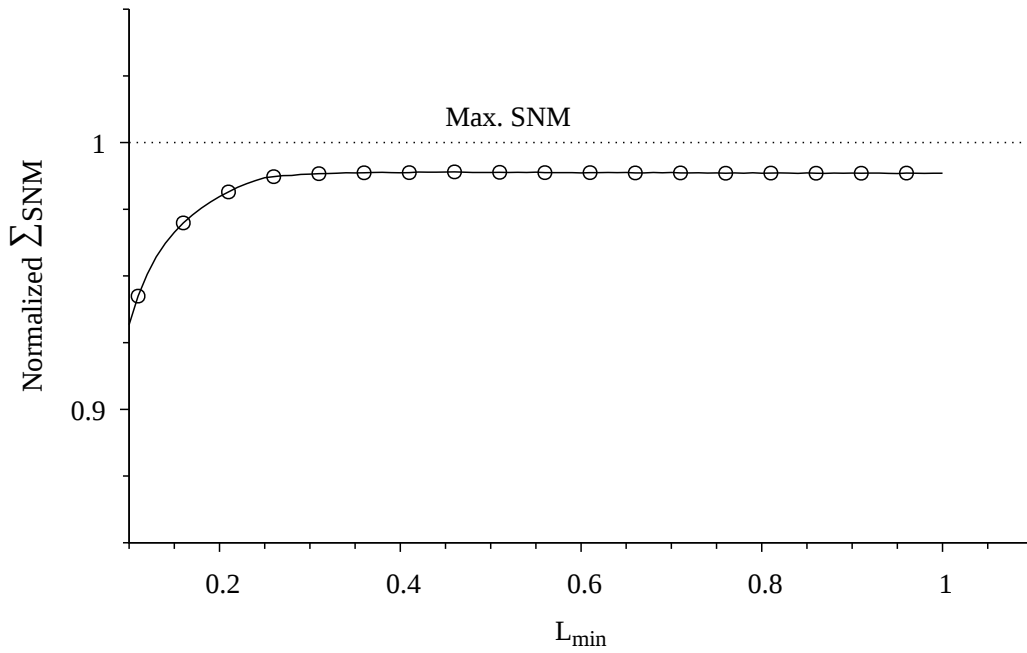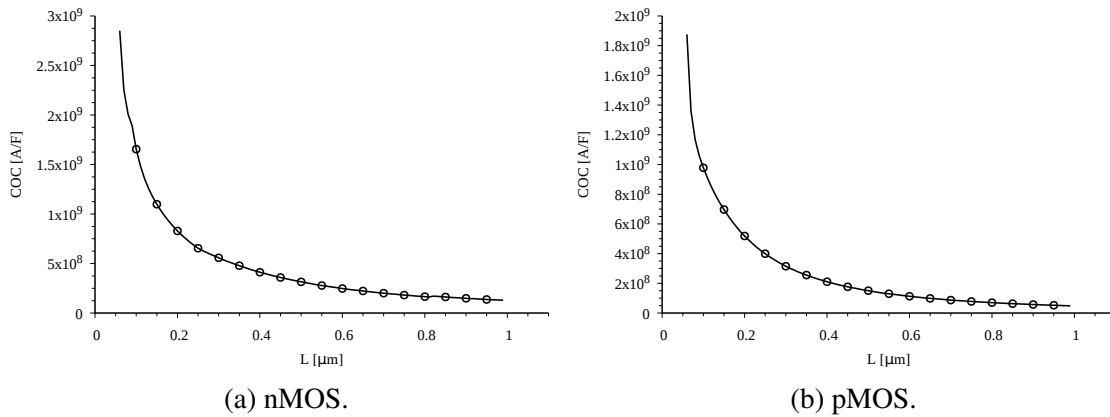Figure B.9: Inverter channel length trade-off versus SNM for ST 65nm.

Figure B.10: Current-over-Capacitance (COC) versus transistor channel length.

(a) nMOS.

(b) pMOS.

in Table B.5 B.6. In general, the proposed approaches exhibits higher μ SNM and better σ/μ. Nonetheless, the slight difference from proposed library βof 1.2 to Nabavi's of 1 had a meaningful impact on other parameters. Table B.6 shows that Nabavi's library has, on average, 8.03%, 19.18%, and 2.72% better power consumption, area, and delay respectively. On the other hand, the proposed library had overall better results compared against Keane's approach.

Table B.5: Normalized SNM-aware design synthesis results to Nabavi's approach, i.e., NO/NS, for ST 65nm.

Percentage Improvement/Loss [%]

Design	Leak. P.	Dyn. P.	Total P.	#Cells	Area	Delay	μSNM	σ/μ
c1355	-26.93	-13.79	-14.87	-21.85	-33.33	0.46	0.13	13.95
c1908	27.11	10.54	11.93	32.69	33.33	-23.09	0.39	16.36
c2670	-7.94	0.01	-0.69	-2.03	-24.00	1.05	0.25	14.72
c3540	-15.18	-8.79	-9.40	-10.37	-32.35	1.39	0.20	11.36
c432	-14.43	-2.39	-3.80	-19.41	-12.50	2.62	0.22	11.67
c499	-4.29	-36.30	-33.87	16.84	-6.25	-14.29	0.08	9.58
c5315	-24.82	-10.45	-11.46	-21.96	-60.61	9.38	0.11	9.81
c6288	-22.49	-11.12	-11.82	-20.61	-37.36	0.00	0.19	10.88
c880	-13.30	-8.91	-9.35	-8.90	-18.75	-2.13	0.19	10.87

Table B.6: Normalized SNM-aware design synthesis results to Keane's approach, i.e., NO/KS, for ST 65nm.

Percentage Improvement/Loss [%]

Design	Leak. P.	Dyn. P.	Total P.	#Cells	Area	Delay	μSNM	σ/μ
c1355	15.97	14.96	15.05	4.39	16.98	3.04	2.22	20.51
c1908	31.94	21.38	22.22	30.00	33.33	-14.29	1.78	15.86
c2670	13.46	14.34	14.26	-1.55	20.51	5.00	1.94	18.72
c3540	15.36	12.75	13.02	1.88	16.67	7.79	2.19	21.62
c432	20.66	19.13	19.33	12.65	35.71	-1.32	2.17	22.14
c499	40.38	3.20	6.64	42.03	56.41	-20.05	1.94	21.49
c5315	14.73	13.08	13.21	4.23	11.67	9.38	1.95	22.05
c6288	-1.63	13.07	12.21	-30.78	-9.65	16.67	2.86	36.24
c880	6.09	9.24	8.92	-8.44	5.00	6.53	2.85	35.87