UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Posgrados

Design of CMOS physical unclonable functions for hardware security application

Tesis en torno a una hipótesis o problema de investigación y su contrastación

Kevin Andrés Vicuña Barriga

Felice Crupi, Ph.D.
Director de Trabajo de Titulación

Trabajo de titulación de posgrado presentado como requisito
para la obtención del título de Master en Nanoelectrónica

Quito, 16 de diciembre del 2022

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

COLEGIO DE POSGRADOS

HOJA DE APROBACIÓN DE TRABAJO DE TITULACIÓN

Design of CMOS physical unclonable functions for hardware security application

Kevin Vicuña

Nombre del Director del Programa: Luis Miguel Prócel
Título académico: Doctor of Philosophy
Director del programa de: Maestría en Nanoelectrónica

Nombre del Decano del colegio Académico: Eduardo Alba
Título académico: Doctor of Philosophy
Decano del Colegio: Colegio de Ciencias e Ingenierías

Nombre del Decano del Colegio de Posgrados: Hugo Burgos
Título académico: Doctor of Philosophy

Quito, Diciembre 2022

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Nombre del estudiante: Kevin Vicuña

Código de estudiante: 00322858

C.I.: 1718186214

Lugar y fecha: Quito, 16 de diciembre de 2022

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en http://bit.ly/COPETheses.

UNPUBLISHED DOCUMENT

Note: The following graduation project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on http://bit.ly/COPETheses.

DEDICATORIA

Dedicato a tutte le persone che mi hanno affiancato e sostenuto in questi due anni di studio.

AGRADECIMIENTOS

Ringrazio i miei docenti, Felice Crupi, Lionel Trojman, Luis Miguel Procel e Ramiro Taco.

Ringrazio Massimo Vatalaro per avermi seguito nella scrittura del mio elaborato finale.

Ringrazio i miei nonni, Laura Olivo, Gerardo Barriga e Glady Palau.

Ringrazio Katherine Rossella Foglia per essermi stata vicina e avermi motivato e aiutato nello svolgimento di questo elaborato.

Ringrazio mio padre, mia madre e mia sorella per essere sempre stati presenti nel momento del bisogno e avermi motivato a raggiungere questo obiettivo.

RESUMEN

Esta tesis presenta el desarrollo de diferentes funciones físicamente no clonables (PUFs) CMOS para aplicaciones de seguridad hardware. Basándose en la forma en que las variaciones del proceso se traducen en una respuesta binaria, las PUFs pueden clasificarse en diferentes clases. En este proyecto nos centraremos en las estructuras estáticas y dinámicas.

El circuito estático analizado se basa en un divisor de tensión metaestable de cuatro transistores que trabajan en el régimen subumbral junto con un inversor en la etapa de salida. El circuito dinámico que se implementó está basado en un oscilador en anillo, el diseño de este circuito utiliza una topología de celda de retardo que incorpora un transistor de paso tipo PMOS entre dos etapas sucesivas con el objetivo de modificar la variabilidad del tiempo de respuesta de la celda. Se realizó una comparación de las distintas soluciones con el virtuoso TCAD y la tecnología TSMC de 180nm. Analizamos los resultados de los PUF en términos de estabilidad, reproducibilidad, unicidad y consumo de energía.

Palabras clave: Diseño CMOS, seguridad de hardware, Internet de las cosas (IoT), función físicamente no clonable (PUF), divisor de tensión, oscilador en anillo, comparador de base de colapso.

ABSTRACT

This thesis presents the development of different CMOS physically unclonable functions (PUFs) for hardware security application. Based on the way in which process variations are translated into a binary response PUFs can be categorized in different class. In this project we will focus on static and dynamic structures.

The static circuit that was analyzed is based on a metastable voltage divider of four transistors working in the subthreshold regime together with an inverter in the output stage. The dynamic circuit that was implemented is based on a ring oscillator, the design of this circuit uses a delay cell topology that incorporates a PMOS type pass transistor between two successive stages with the aim of modifying the variability of the cell response time. A comparison of the different solutions with the virtuoso TCAD and the 180nm TSMC technology was carried out. We analyze the results of the PUFs in terms of stability, reproducibility, uniqueness and power consumption.

Key words: CMOS design, hardware security, Internet of Things (IoT), physically unclonable function (PUF), voltage divider, ring oscillator, collapse base comparator.

TABLA DE CONTENIDO

Dedicatoria **5**

Agradecimientos **6**

Resumen **7**

Abstract **8**

Tabla de contenido **9**

índice de TABLAS **11**

ÍNDICE DE FIGURAS **12**

Introduction **15**
- **Physical Disorder in Integrated Circuits** **16**
  - Devices Geometry 17
  - Devices Material 18
  - Interconnects Geometry 18
  - Interconnects Material 20
- **Design of a Physically Unclonable Function** **22**
- **Delay Based PUFs** **27**
  - Arbiter 27
  - Ring Oscillator PUF 31
  - Self-timed Rings PUF 33
- **Current-Based PUFs** **35**
  - Current-Based PUFs Using Transistors Arrays 36
  - Current-Based PUFs Using Dynamic Random Access Memories 40
- **Voltage-Based PUFs** **42**
  - SRAM PUFs 43
  - Latch-Based PUFs 46
- **Metrics of PUF Devices** **47**
  - Uniqueness 47
  - Reliability 48
  - Uniformity 49
  - Tamper Resistance 50

Development and simulation methodology **51**
- **Full custom metodology** **51**
  - Technology 51
  - Development software 53

Proposed schemes and results **56**
- **Metastable PUF bitcell** **56**
  - Simulation Results of the PUF Bitcell 69
- **Delay Base PUF** **73**
  - Inverter 75
  - Ring Oscillator 82
  - Calibrated Ring Oscillator 89
  - Calibrated ring oscillator collapse-based comparator 94

.... **96**

... 99
Conclusions.. 100
References.. 102

ÍNDICE DE TABLAS

Table 3.1: Unstable bits with different bias voltages a) 1.8[V], b) [V], c)0.5[V] and d)0.4[V]. **86**

Table 3.2: Unstable bits with bias voltage of 0.5[V] at different temperatures a) 0ºC, b) 25ºC, c) 50ºC and d) 75ºC. **88**

Table 3.3: Unstable bits with different bias voltages a) 1.8[V], b) [V], c)0.5[V] and d)0.4[V]. **91**

Table 3.4: Unstable bits with bias voltage of 0.5[V] at different temperatures a) 0ºC, b) 25ºC, c) 50ºC and d) 75ºC. **92**

Table 3.5: Unstable bits with different bias voltages a) 1.8[V], b) [V], c)0.5[V] and d)0.4[V]. **97**

Table 3.4: Unstable bits with bias voltage of 0.5[V] at different temperatures a) 0ºC, b) 25ºC, c) 50ºC and d) 75ºC. **98**

ÍNDICE DE FIGURAS

Figure 1.1: Devices Geometry Variation. 17

Figure 1.2: Devices Material Variation. 18

Figure 1.3: Interconnects Geometry Variation. 19

Figure 1.4: Interconnects Material Variation. 20

Figure 1.5: The impact of variability on the electrical parameters of VLSI circuits. 21

Figure 1.6: Conceptual Ring Oscillator PUF device. 23

Figure 1.7: Fingerprinting as a physical method of cryptography. 24

Figure 1.8: Architecture for silicon-based PUFs. 25

Figure 1.9: PUF novel solution for cryptography. 26

Figure 1.10: A single challenge arbiter PUF. 27

Figure 1.11: Circuit diagram of an arbiter based on an S-R latch. 28

Figure 1.12: Structure of a multiple bit challenge arbiter PUF 30

Figure 1.14: Muller-C element. 33

Figure 1.15: Three stage self-timed ring (STR). 34

Figure 1.16: Token and bubble movement. 34

Figure 1.17: Smart PUF based on self-timed ring structure. 35

Figure 1.18: A current-based PUF. 36

Figure 1.19: Circuit schematic of the TCO unit array. 38

Figure 1.20: DRAM 41

Figure 1.21: 6T-SRAM Schematic. 43

Figure 1.22: Characteristics for SRAM. 44

Figure 1.23: SRAM Memory Array. 45

Figure 1.24: SR Latch. 46

Figure 1.25: Uniqueness evaluation of a PUF. 48

Figure 1.26: Reliability evaluation of a PUF. 49

Figure 1.27: Uniformity response of a PUF. 49

Figure 2.1: Evolution of the number of transistors every two years according to Moore’s Law. 52

Figure 2.2: Flow of custom design. 54

Figure 3.1: Conceptual diagram of the proposed PUF bitcell. 56

Figure 3.2: Metastability characteristic. 57

Figure 3.3: Transistor-Level scheme of the 4T bitcell. 58

Figure 3.4: PUF bitcell with 2T-core. 62

Figure 3.5: PUF bitcell with 4T-core sizing. 69

Figure 3.6: Statistical distribution of the voltage V_X of the bitcell core. 70

Figure 3.7: Nominal input–output characteristics of the inverter. 71

Figure 3.8: Statistical distribution of the voltage VX of the bitcell core a) FF , ... 72

b) SS, c) FS, d) SF corners. 72

Figure 3.9: PUF bitcell with 4T-core layout. 73

Figure 3.10: Ring Oscillator bitcell. 74

Figure 3.11: Voltage transfer curve of inverter V_{DD} = 2.5[V]. 76

Figure 3.12: Device variations voltage transfer curve of inverter V_{DD} = 2.5[V]. 78

Figure 3.13: Voltage variations voltage transfer curve of inverter V_{DD} = 2.5[V]. 78

Figure 3.14: Inverter gate schematic: a) transistor size $V_M = V_{DD}$, b) minimum 79

transistor size NMOS and PMOS. 79

Figure 3.15: Inverter VTC: a) transistor size VM = VDD , b) minimum transistor .. 80

size NMOS and PMOS. 80

Figure 3.16: Inverter transient response: a) transistor size VM = VDD , b) minimum transistor size NMOS and PMOS. 80

Figure 3.17: Inverter energy per operation: a) transistor size VM = VDD , b) minimum transistor size NMOS and PMOS 81

Figure 3.18: Layout inverter minimum transistor size. 82

Figure 3.19: Proposed RO schematic. 82

Figure 3.20: Simulated distributions of the frequency of proposed RO-base. 83

Figure 3.21: Schematic Test-Bench of RO PUF-base. 84

Figure 3.22: Time response of a) input signal, enables oscillating circuits, b) current consumption of the oscillators, c) response of the first oscillator and 85

d) response of the second oscillator. 85

Figure 3.23: Normalized frequency difference with different bias voltages a) 1.8[V], b) [V], c)0.5[V] and d)0.4[V] 86

Figure 3.24: Normalized frequency difference with bias voltage of 0.5[V] at different temperatures a) 0ºC, b) 25ºC, c) 50ºC and d) 75ºC. 87

Figure 3.25: Layout ring oscillator. 88

Figure 3.26: Stability enhancement of an RO-based PUF. 89

Figure 3.27: Schematic of configurable RO. 90

Figure 3.28: Effect of the PMOS pass transistor between two inverters. 90

Figure 3.29: Normalized frequency difference with respect to the nominal frequency with different bias voltages a) 1.8[V], b) 1[V], c) 0.5[V] and d) 0.4[V]... 91

Figure 3.30: Normalized frequency difference with respect to the nominal frequency with bias voltage of 0.5[V] at different temperatures a) 0ºC, b) 25ºC, c) 50ºC and d)75ºC. 92

Figure 3.31: Comparison of Ring Oscillator and Calibrated Ring Oscillator Distributions. 93

Figure 3.32: Layout calibrated ring oscillator. 94

Figure 3.33: Calibrated ring oscillator collapse-based comparator. 95

Figure 3.34: Time response of a) input signal, enable, b) current consumption of the PUF, c) response in the case that collapses OUT to V_{DD} and d) response in the case that collapses OUT to V_{SS}. 96

Figure 3.35: PUF response with different bias voltages a) 1.8[V], b) 1[V], c) 0.5[V] and d) 0.4[V]. 97

Figure 3.36: PUF response with bias voltage of 0.5[V] at different temperatures . 98

a) 0ºC, b) 25ºC, c) 50ºC and d) 75ºC. 98

Figure 3.37: Layout of the proposed PUF cell. 99

INTRODUCTION

Nowadays, the development of the IoT scenario pushes the demand of preserving information down to the chip level. The security services (i.e., confidentiality, integrity, authentication, nonrepudiation and digital signature) required for preserving data are guaranteed by using a secret key (often called root of trust). However, the problem is the leakage of the key for which malicious users can perform actions that violate some basic requirements of the data transaction (such as data eavesdropping or password breaking) conventionally, the secret key is generated off-chip and stored in a nonvolatile memory (NVM) but this approach requires additional costs and suffers of reverse engineering attacks that may cause the leakage of the key. Ideally, the secret key should be generated in a volatile way when required without storing it in a NVM.

Physically unclonable functions (PUFs) are promising cryptographic primitives which exploits random static phenomena for generating unique, reproducible and random key in a volatile way. A PUF can be seen as a physical device, whose system could be considered simple, taking into account that it has a number of features that are very interesting for security applications. This is because this type of hardware should be able to have a simple evaluation, however, it should be impossible to predict. Another important aspect is that these devices should have relatively easy fabrication, nevertheless, their duplication should be difficult to achieve. The term PUF was first described in (Daihyun Lim and Lee, 2005), in this paper the authors also introduce a new expression “silicon PUF”. Referring to all physically non-countable devices that are built and designed on integrated circuits (ICs). What makes silicon PUFs interesting is that they use the variations in the fabrication process that always exist between ICs to generate completely random responses.

It is important to note that PUFs could be fabricated from any other format than silicon. Some examples of PUF designs using other methods of randomizing results can be found in the literature. The most common would be optical devices, which exploit light scattering (Fournel, 2016), (Dolev, 2015). We can also find radio frequency (RF) designs, which use electromagnetic waves that are emitted from a device while it is operating (Reising, 2015), (Cobb, 2012). Like these examples there are several forms of identification schemes, however, for this project we will focus on silicon PUFs.

Physical Disorder in Integrated Circuits

When talking about integrated circuits we could say that they are synthetic; therefore it should be possible to design all their irregularities both in shape and structure. However, this is not the case in most modern chips, the main reason being variability. When we talk about variability in IC design, we refer to inaccuracies in the manufacturing processes. It also refers to temperature and voltage variations within the chip, causing a change in the performance and power consumption of the circuits (Nassif, 2007). This increases with the scaling of very large scale integrated circuit (VLSI) technologies, despite the improvement this has on performance and power consumption. Two sources of variation can be identified in systems implemented in real life (Narasimhan, 2007), (Nassif S. , 2001). The first factor is environmental, where changes in power, supply voltage, operating temperature and electrical degradation parameters in the devices occur. The second factor is the physical one, where variations in dimensions and structures are observed when manufacturing these devices.

Devices Geometry

The variation of device geometry has some very clear examples, the one we will give the most importance and explain is the MOSFET structure in integrated circuits. This device usually includes the variation of thickness and lateral dimensions.

- Thickness variation occurs in gate oxide width (T(ox)), is a fundamental part and a parameter that is relatively easy to control. This type of effect is usually most evident when changing wafers.
- When talking about the lateral dimension, the channel width and length are taken into account, typically their variations are due to the photolithography process. MOSFETs, especially in scaled devices, tend to be particularly sensitive to the effective channel length (Leff), this type of variation has to be considered as it directly impacts the current characteristic observed at the output (Chandrakasan, 2001).

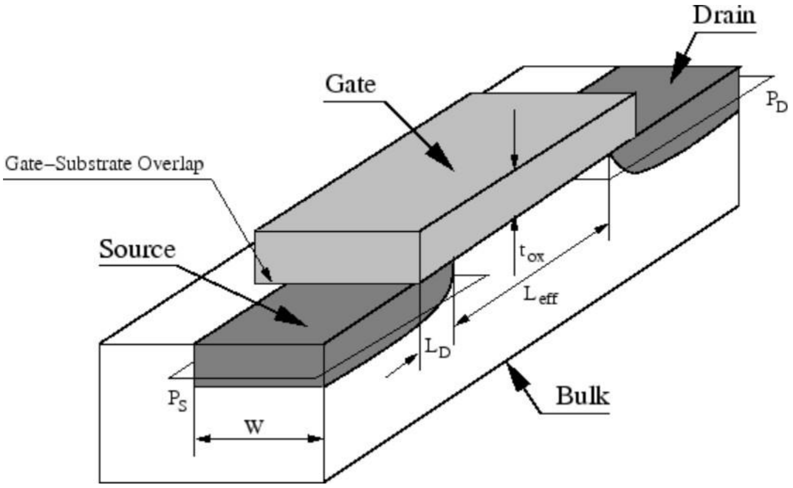


Figure 1.1: Devices Geometry Variation.

Devices Material

Another factor of variation found a lot in MOSFETS is the one that de- pends on the type of material used for the construction of the device including:

- The doping that is introduced in the materials usually has variations due to the dose, angle, energy and other types of variables that are taken into account when making the ion implantation. It is necessary to analyze the technology that is being used, since these deviations can cause losses in the adaptation of NMOS devices against PMOS. This can generate cases in which the variation is found within the wafer and the transistor array that are operating.
- Deposition: These types of changes in material parameters affect the variation in resistance generated at the transistor contacts.

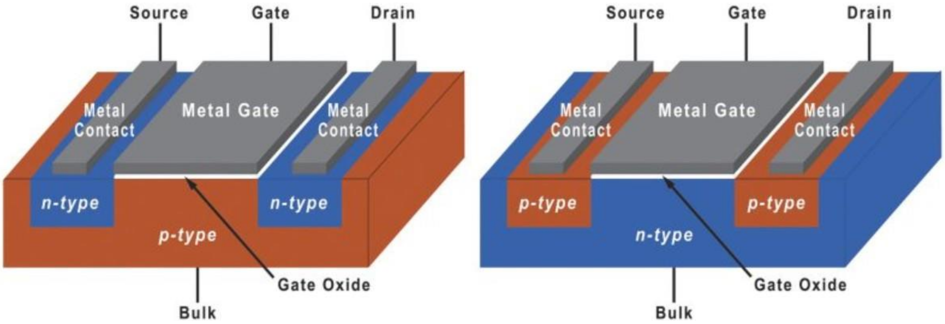


Figure 1.2: Devices Material Variation.

Interconnects Geometry

Following the analysis of the parts of a circuit where variability can be witnessed we find the interconnection geometry where we have the main effects on:

- Line width and line spacing: There are deviations in the width of the lines this is due to the effect of photolithography. The line width directly affects the resistance observed at the input and output of the integrated, the line spacing in turn affects the capacitance. This can result in de- creased performance and increased power consumption due to glitches.

- Metal thickness: The thickness of the interconnect metal is usually de- posited symmetrically and correctly on a wafer, however, a lot of variability is observed when changing the wafer for printing. Although printing on a single silicon wafer is symmetrical, significant changes in metal strength can be observed due to erosion.
- Dielectric height: In order to fabricate the dielectric, the oxide needs to be deposited and polished, however, this process can lead to high device variability. In addition, the chemical-mechanical polarization process (CMP) causes variability in chip performance because the effective density of the topology changes in different parts of the wafer.

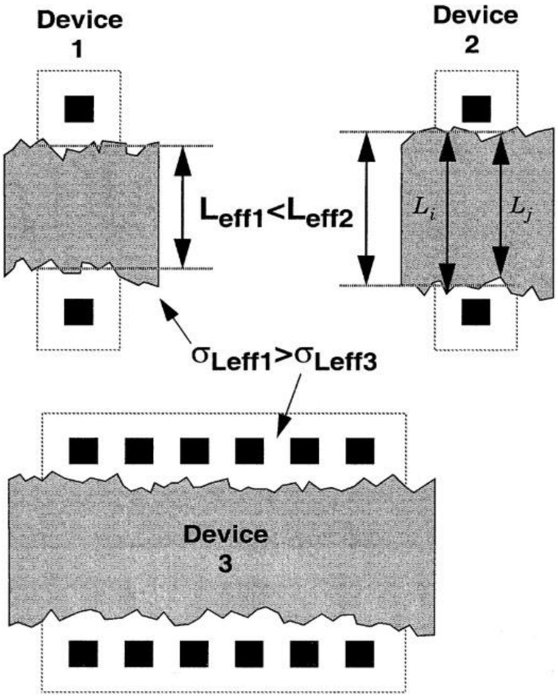


Figure 1.3: Interconnects Geometry Variation.

Interconnects Material

When talking about the interconnection of the devices we must also take into account the material with which they are made, since the following sources of variability can be found:

- Resistivity of the material: Depending on the material used, the resistance of the material changes. The moment where more variation can be observed is when changing wafers.

- Dielectric constant: This type of variation depends on the deposition process, however, this is usually very well controlled.

- Contact and via resistance: This is the largest source of variability as it is very sensitive to the printing process, and a random change is observed if the printing is changed from wafer to wafer.

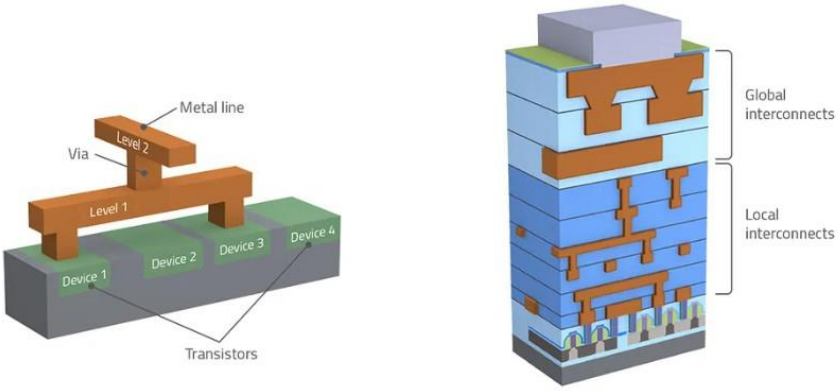


Figure 1.4: Interconnects Material Variation.

The impact of variability is expected to be of great importance for future technologies (Hoefflinger, 2012), making this type of parameter variation an unavoidable feature in VLSI circuits. This means that it will be more difficult to estimate with certainty the exact performance and power consumption of a specific circuit. This is because variations in the physical parameter of the devices will greatly affect the electrical parameters. From Fig.1.5 it can be seen that the variation in the devices increases over the years. The threshold voltage parameter (V(th)) and the IC performance are the most affected, these mea- surements were performed by analyzing the delay generated in the leads and the integrated ones in the work reported in (Hoefflinger, 2012).

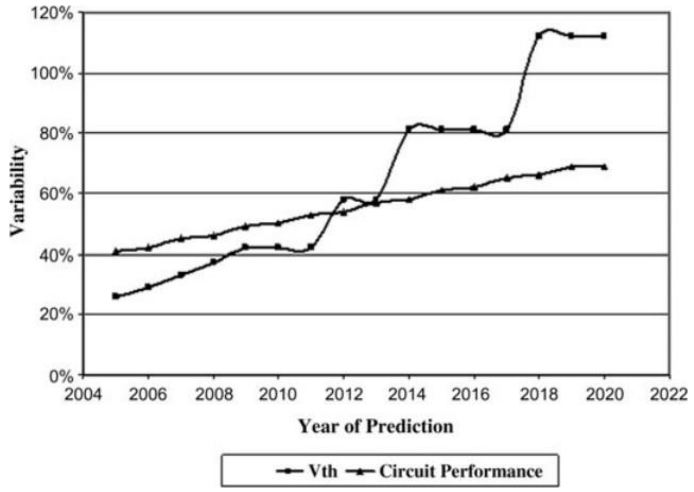


Figure 1.5: The impact of variability on the electrical parameters of VLSI circuits.

This uncertainty when designing integrated circuits means that the worst case scenario is always taken into account, the point at which it would be working below the optimum operating point. To give an example we could consider a thousand processors that were designed to operate at a frequency of 600 [MHz], if the manufacturing process has a variability of 10 % that affects the performance of the circuit. With these conditions if measurements of the devices are made to know the maximum operating frequency, we would find a Gaussian distribution centered at 600 [MHz] and a standard deviation of 20 [MHz]. If we interpret these results, we realize that only half of the processors manufactured operate at the maximum required frequency. Therefore, the true operating frequency at which we could work is 540 [MHz], so that all devices can be used and potential processing time errors can be avoided.

Nevertheless, this is not all bad, as these types of variations are very useful for the design of physical unclonable functions, which promise to be a very interesting solution for cryptography. Being the main answer that could help the security problems faced by embedded systems (Daihyun Lim and Lee, 2005).

Design of a Physically Unclonable Function

It is considered to be physical unclonable function when a cryptographic system is embedded in a physical object such as a silicon chip. This system must be able to receive an input signal which is processed and generates an output signal. The output will be known as response (R) and the input will be called challenge (C). It is necessary to mention that this system must be robust and difficult to clone, even considering the worst case which would be when the organization of the components of this system is known (Halak, 2018).

The basic example with which the design of this type of integrated circuits, the PUF, started is the ring oscillator shown in Fig.1.6. If we analyze this figure, we can see that the oscillator starts to work when it receives a logic 1 in the enable signal. Then, depending on the number of inverters and the physical properties of the technology used for the design of this circuit, there will be an oscillation frequency. This type of oscillator-based PUF only needs a challenge (enable signal) to generate a response (oscillation frequency). What is interesting is that due to the variation of manufacturing processes as we saw before causes a disorder in the semiconductor devices used in the circuit. Causing this PUF in the response to have different oscillation frequencies when implemented in different devices. All these random responses can be used as a unique hardware signature for each physical implementation of this function. Thus achieving non-clonability, since all frequency variations encountered are not due to some software component. In fact, it is all inherent in the manufacturing process due to the lack of control over the geometry and materials used in the devices (Halak, 2018).

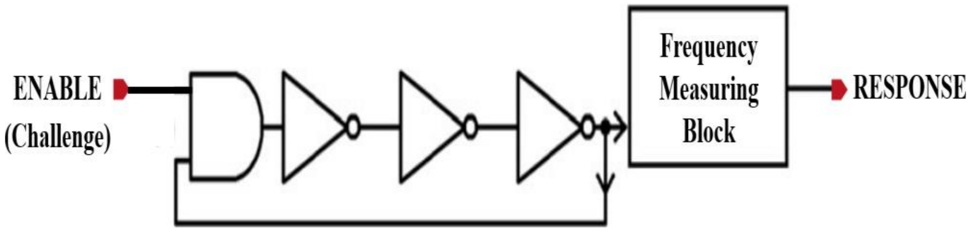


Figure 1.6: Conceptual Ring Oscillator PUF device.

Although the term PUF is relatively new, the thought of identifying objects based on physical properties is not. If we look back through history, some civilizations such as the Babylonian civilization used fingerprints to pre- vent counterfeiting. The way business was conducted at that time was quite interesting, since contracts were signed on clay using the fingerprint to verify that the two parties are the real ones (Xi, 2011). Nowadays, “uniqueness” has been used mainly to identify physical objects, such as nuclear weapons in the cold war. All nuclear warheads are identified by spraying a thin coating layer of a material that endows the object with unique light-reflecting characteristics. These nuclear warheads are then placed under the same light and, due to the imperfection in the coating material layer, generate a unique reflection pattern for each weapon (McFate, 2022).

Figure 1.7: Fingerprinting as a physical method of cryptography.

From these two examples we can say that to design a physical unclonable function it is necessary to have two components. The first is to have a physical disorder that is inherent to each object, the second is to generate a method to evaluate, classify and store this disorder.

Modern integrated circuits have a large number of physical disorders. Since they arise in the process of manufacturing variation, therefore the first ingredient is easily obtained. However, the most difficult thing is to be able to transform this physical disorder into a measurable electrical quantity. The three electrical parameters that can be quantified in an integrated circuit are: current (I), voltage (V) and delay (D). The rest of the measurements that can be calculated are derived from these parameters, for example, the resistance that exists on a conductor is calculated with the voltage drop that exists at its ends divided by the average current that passes through it. In addition, in a PUF circuit it is desired to provide a response that is different for each applied challenge. Taking into account that in this era we live in a digital world, the response provided by the PUF must be digital. For this reason, if we consider the two main requirements of the silicon PUF, we will have the following:

- Considering the semiconductor technology and understanding how the variability of the manufacturing process works, PUFs need to be able to transform these variations into measurable electrical quantities such as current, voltage or delay (Halak, 2018). PUF circuits need to be able to transform these variations into measurable electrical quantities such as current, voltage or delay.

- Every PUF circuit needs to transform its measurable electrical quantity into a digitally quantified response, i.e. a binary response (Fournel, 2016).

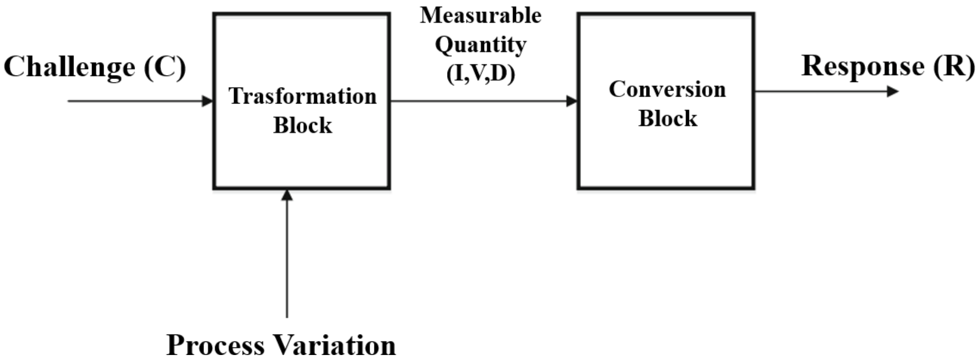


Figure 1.8: Architecture for silicon-based PUFs.

Understanding the above two requirements, an architecture that all silicon PUFs must follow can be proposed as shown in Fig.1.8. By comparing Fig.1.6 and Fig.1.8, we can understand that the oscillator is the transformation block and the frequency measurement block is a conversion block. The first block (transformation) is in charge of converting the challenge and the process variation of technology into a measurable electrical quantity. The second block (conversion) is responsible for converting the measurable electrical quantity into a binary value, i.e., the system response (Halak, 2018).

The above representation of the topology of a PUF allows us to have greater facility at the moment of reusing or designing new circuits. This is mainly because we can perceive the design of a PUF as the construction of two separate blocks. There are currently some examples of designs for transformer and converter blocks in the literature. To give an example of the state of the art for the transformation block we would have: analog digital converter (DC/AC), ring oscillators, current sources, among others. In the conversion block we can find the following circuits: analog to digital converter (AC/DC), time to digital converter, phase decoders, comparators, among others (Halak B. a., 2008), (Nawi, 2016).

PUF devices are classified. Due to the need for measurable electrical quantity, we would have PUFs based on current, voltage and delay. It is necessary to understand that there are PUFs that do not exist within these three categories. To understand this better we would have as an example the devices that are based on the variation of the conduction current or threshold voltage. However, the categorization based on current, voltage and delay provides in a more intuitive way the different design techniques that can be used (Halak B. , 2018).

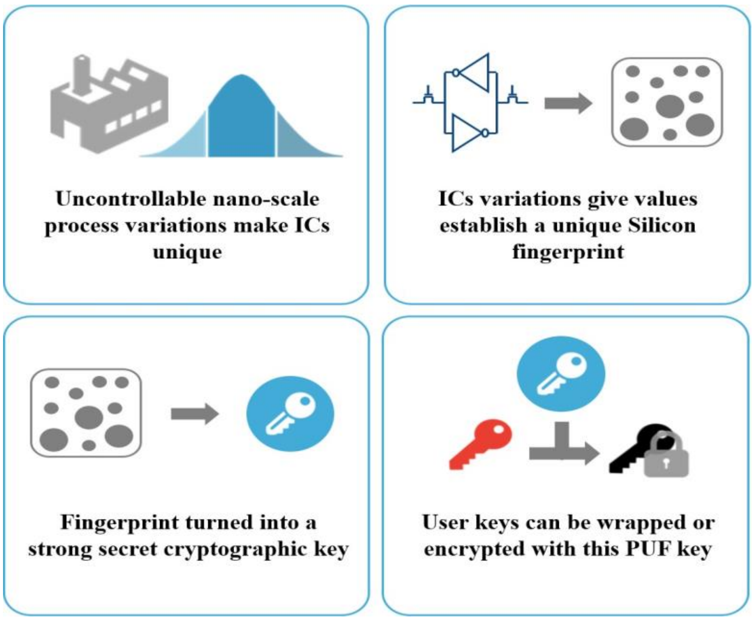

Uncontrollable nano-scale process variations make ICs unique
ICs variations give values establish a unique Silicon fingerprint
Fingerprint turned into a strong secret cryptographic key
User keys can be wrapped or encrypted with this PUF key

Figure 1.9: PUF novel solution for cryptography.

Delay Based PUFs

This type of cell is responsible for transforming the variations of the integrated circuit into delay, and then transforming it into a binary response. The first proposals of PUF structures were based on this type of cells, so there are several solutions using this parameter. If we analyze the literature, we can find architectures based on arbiter (Lin, 2012), ring oscillators (Yin, 2010) and those based on an asynchronous structure (Suh, 2007). In the following, we describe the operation of these types of cells working with the delay response.

Arbiter

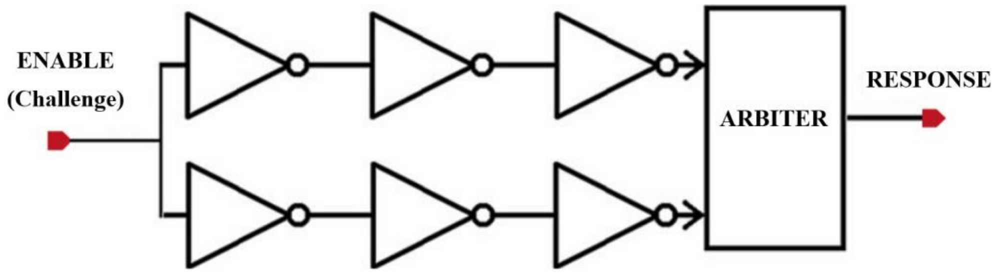


Figure 1.10: A single challenge arbiter PUF.

The structure shown in Fig.1.10 is known as an arbiter-based cell. It is called this way because if we look at the composition of the cell it consists of two parts. The two digital paths that must have the same nominal delay, and the arbiter at the output converts the delay into a digitized response. This cell works by applying an enable signal at the input, which propagates through the two digital paths. As the delayed signals from each of the paths arrive at the inputs of the arbiter, due to intra-die variability either of these signals can be the one that determines the final value of the circuit. The arbiter discretizes the output signal by giving a logic “1” or “0” depending on which path was the fastest. When this circuit is replicated on multiple chips, the response changes due to inter-die variations. In other words, this response is unique for each implementation, so it can be considered a hardware signature (Halak B. , 2018).

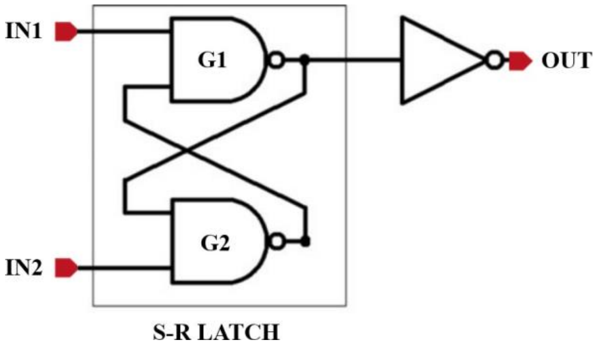


Figure 1.11: Circuit diagram of an arbiter based on an S-R latch.

Then, we will talk about the arbiter, normally for this structure is usually used a Set-Reset latch. This latch is normally built as shown in Fig.1.11 with two cross-coupled gates. The operating principle of this circuit is as follows:

- If both inputs have a low logic level, the output has a low logic level as well.

- If only one of the signals changes state to a logic high level, the response will change its state accordingly and will be blocked. To better explain this process, we look at Fig.1.11, and assume that inputs IN1 and IN2 are in a logic low state. In the first case IN1 changes to a logic high state, causing the output (OUT) to also change its state from low to high and remain at this value even though IN2 changes. Now if we analyze the second case where instead of IN1 it is IN2 that changes its state from low to high. The output will remain in a low logic state because this is how the combinational circuit is described, in the same way as for the previous case the signal is blocked, but in the low logic state.

- If both signals arrive at the circuit in a high state, but there is a small but significant difference in their arrival. The output will assume the value of the signal that arrived first, i.e., logic “1” if IN1 arrived first or logic “0” if IN2 arrived first.

- Special case: if the two signals arrive in a high logic state, but with almost zero delay difference, the output enters a meta-stable state.

$$V(t) = V_O e^{t/\tau} \tag{1.1}$$

The meta-stability is stopped approximately when gates G1 or G2 reach the threshold voltage, therefore in equation 1.1 we could substitute the values by having:

$$t = \tau\, Ln(V_{th}/V_O) \tag{1.2}$$

Taking into account the probability analysis, we can observe that if we have a meta-stability event at $t = 0$, the probability of having a meta-stability event in a longer time is:

$$P_m(t) = e^{-t/\tau} \tag{1.3}$$

Theoretically, one could calculate precisely the moment at which the meta-stability stops. However, to do this requires precise knowledge of the geometry of the circuit, as well as the voltage levels being delivered by gates G1 and G2. All the information needed to do this analysis is very complicated to obtain in practice. As an example, we would have the scenario where it becomes complicated to find out the state of the arbiter; since it is not possible to know if the output is a logic zero because the IN2 signal arrived first or if we are in the case of meta-stability (Halak B. , 2018).

One of the first example that can be found in the literature related to the arbiter-PUF is the one proposed in (Daihyun Lim and Lee, 2005). In this paper, it was suggested to use a design that implements multiplexers connected in series for the logic paths. Each multiplexer has a selection signal, which allows to choose the digital path to be followed. Thus, achieving with this structure that depending on the challenges that are placed in the cell, the path or the delay that can be observed at the output is configured. The number of bits that the challenge will have depends on the number of multiplexers connected in series that the circuit has. Each bit pattern given to the challenge gives a unique configuration to the digital path to be used. Therefore, the response observed at the output is unique for each challenge given to the PUF. The architecture implemented in this paper is shown in Fig.1.12 below.

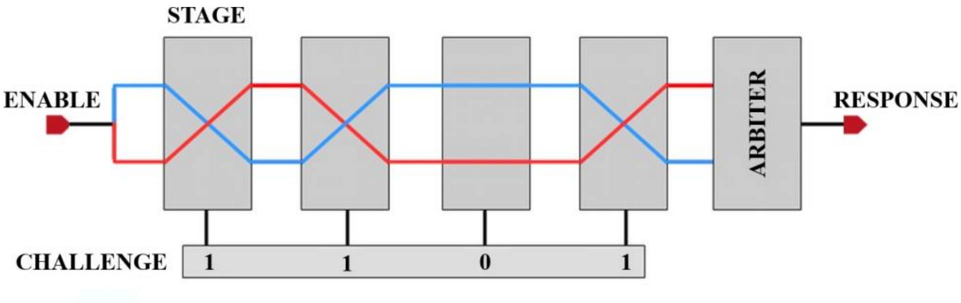


Figure 1.12: Structure of a multiple bit challenge arbiter PUF.

The maximum number of challenge/response pairs that the arbiter-PUF proposed in [1] is obtained by analyzing the number of gates or stages (k) connected to a challenge. The equation that determines the number of responses that can be obtained based on the size of the challenge is the following:

$$CRP = 2^k \qquad \textbf{1.4}$$

The structure that was proposed for the arbiter in the PUF cell of (Daihyun Lim and Lee, 2005) is based on a transparent latch. Nevertheless, the authors reported that due to the asymmetry of the latch, the correct predictability of the response was greatly affected. This is mainly due to the fact that the latch has to favor one path, reason why most of the responses are “0” instead of “1”. To be precise with the article, 90% of the answers were a low logic level, which is not what is desired for this type of circuits. However, the authors made proposals to systematically change the delay of the logic paths. This in order to reduce this problem and to have a 50% probability of obtaining a logic “0” or a logic “1”. In addition, these methods make the PUF intrinsically more predictable, since they are not based on a physical disorder but rather on the designed variations.

The main difficulty in achieving a fully symmetrical chip design is that the designs are done with automated software. These tools just place the components (standard cells) and interconnect them depending on what is favorable for area, power consumption or performance. These tools do not take into account the symmetry required for this type of cells, which depend on their homogeneity when designing them. This limits the degrees of freedom that the designer has to control the behavior of the circuit. Therefore, it is necessary to use tools that allow the manual design of the PUF’s, in order to ensure the maximum possible symmetry in the trajectories of the delays (Halak B. , 2018).

Ring Oscillator PUF

The first appearance of this ring oscillator-based structure was composed of: two multiplexers, two counters, a comparator and a number (k) of ring oscillators (Suh, 2007). The basic structure of this circuit is presented in Fig.1.13, the operation of this PUF is quite simple and consists of the following stages:

- Ring oscillators have a unique frequency that depends on the characteristics of the inverters.

- The two multiplexers are intended to select the two ring oscillators to be compared.
- The counters are used to count the number of oscillations that RO generates in an interval of time.
- The comparator determines which ring oscillator had the highest frequency and thus determines the output value of either logic “0” or “1”.

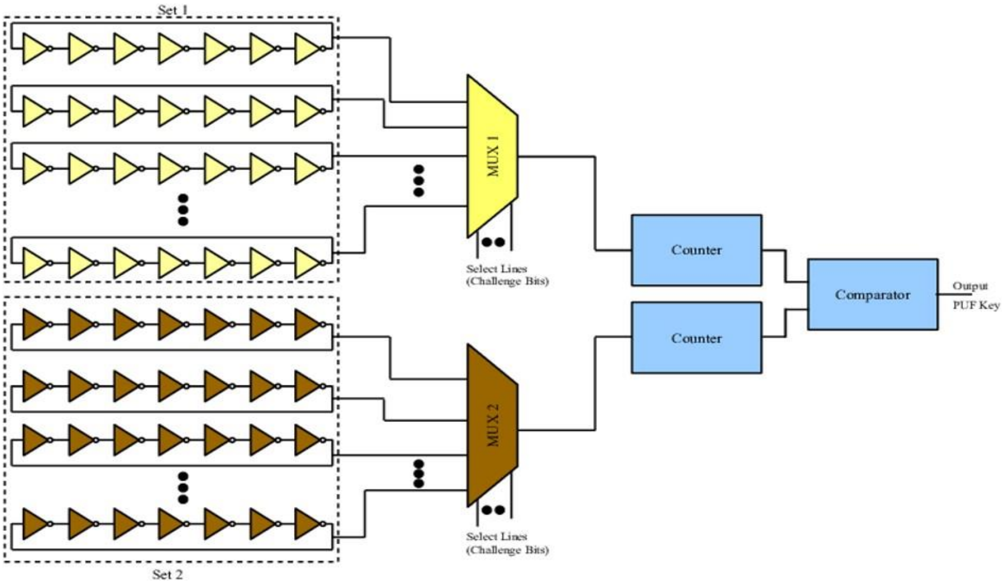


Figure 1.13: Structure of ring oscillator PUF.

Similar to what was observed in the arbiter-based PUF, this cell still has the need for the ring oscillators to have the same nominal delay. However, this architecture allows us to have a response without using an arbiter, thus eliminating the meta-stability problem. Thus giving a higher reliability to the response obtained from this PUF (Suh, 2007).

The maximum number of challenge/response pairs (CRP) is strongly linked to the architecture. In this specific case of the RO-based PUF we can determine that the CRP depends on the number of ring oscillators implemented. So, the equation that determines the CRP would be given by the following expression, where k is the number of RO in the architecture (Suh, 2007).

$$CRP = \frac{k\,(k-1)}{2} \qquad \textbf{1.5}$$

Self-timed Rings PUF

This is another type of architecture proposed for PUFs that base their response on delay, the idea is described in better detail in (Yakovlev, 2012). Its structure is mainly based on the classical ring-oscillator PUF architecture shown in Fig.1.13; however, it uses self-timed cells instead of classical inverters. This type of self-timed cells implements the Muller's C element, which is described in the article as a fundamental building block for asynchronous circuits. A circuit using the C-element and based on the CMOS logic family can be seen in Fig.1.14 below. The principle of operation of this cell is simple and is as follows.

- The output acquires the logic value "1" or "0" when the two inputs have the logic value of "1" or "0" respectively.
- If the respective values are not present, the output remains in the previ- ous state, i.e., it does not change.

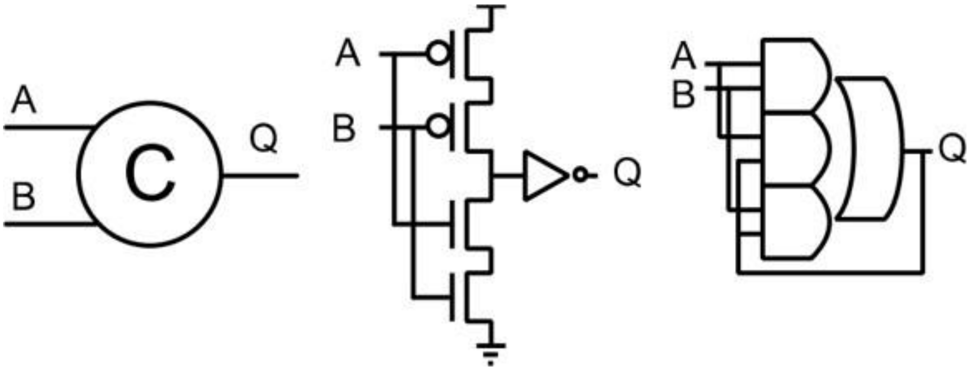


Figure 1.14: Muller-C element.

Fig.1.17 shows a clear example of a self-timed ring (STR), it can be seen that this structure consists of three stages. All the stages are constituted by a C element and an inverter. The connection of these elements is interesting, since the inverter is connected in reverse (R) with the C element and the second input of this cell is connected with the forward stage (F) (Yakovlev, 2012).

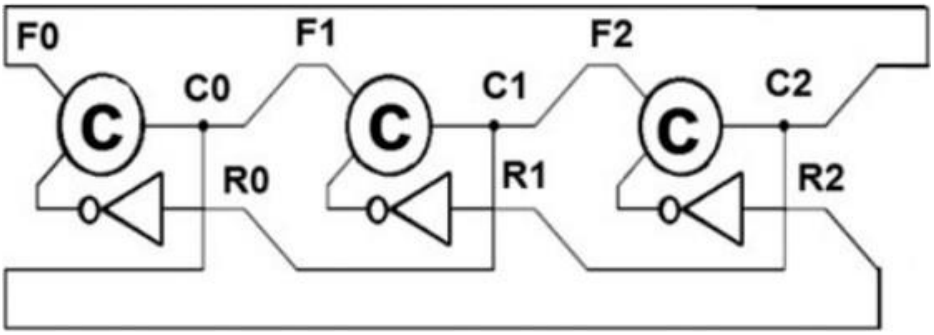


Figure 1.15: Three stage self-timed ring (STR).

In order to better understand how this structure works, the concept of bubbles and tokens must be introduced. We can say that an STR stage has a bubble if and only if the output is equal to the output of the previous stage. The opposite case is when it is considered that there is a token, since the output of the STR is different from the signal of the previous stage (Yakovlev, 2012).

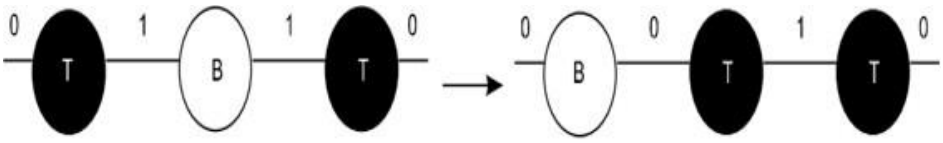


Figure 1.16: Token and bubble movement.

Since self-timed rings have a particular operation, the conditions for this circuit to oscillate are a bit more complicated than the typical RO. The first requirement is that the number of stages the STR has must be greater than or equal to three. The second condition is that the sum of tokens and bubbles must be equal to the number of stages the

STR possesses. To better understand this, the equations underlying these oscillation conditions are described (Yakovlev, 2012).

$$STR \geq 3 \tag{1.6}$$

$$STR = N_b + N_t \tag{1.7}$$

What makes this self-timed ring structure interesting is the increased robustness of the PUF response to environmental variations. Obviously, this benefit comes at a high cost in terms of the silicon area used, it should also be added that this type of structure can suffer from blocking states (Yakovlev, 2012).

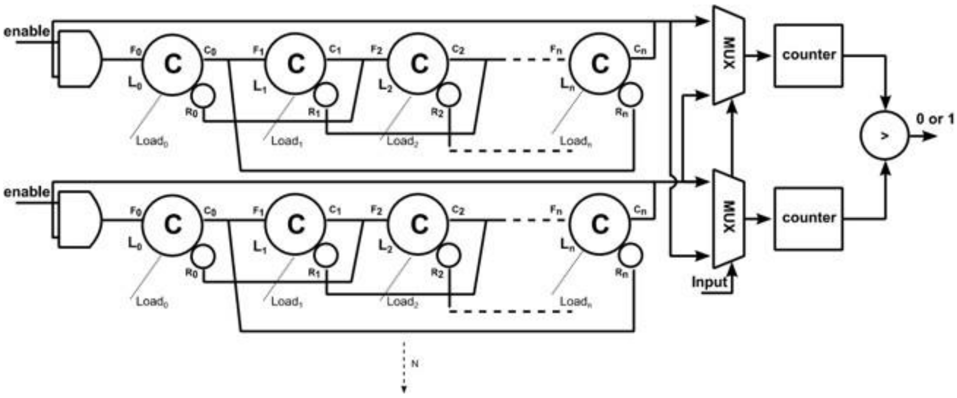


Figure 1.17: Smart PUF based on self-timed ring structure.

Current-Based PUFs

This type of architectures have the ability to convert the variation of manufacturing processes, in a current quantity that can be measured by different circuits. Like all PUFs, this response must be completely digital in a binary system. Two structures based on this type of solution are presented below. The first solution focuses on using the current obtained from transistors that are working in the subthreshold region. The second proposal attempts to capture the leakage currents that exist in a dynamic random access memory (DRAM).

Current-Based PUFs Using Transistors Arrays

In the literature it is possible to find some solutions that are based on the current to generate PUFs, however, one of the first proposals that can be found is the following (Kalyanaraman, 2013). This design is quite interesting since it employs MOSFETS, which is the reason why it is possible to take advantage of the exponential current that is present when the device works in the subthreshold region. In order to operate in this regime it is necessary to use the threshold voltage ($4V_{th}$) and the gate-source voltage (V_{gs}). This in attempt to increase the unpredictability of the PUF operation, the design of this architecture is presented in Fig.1.18 below.

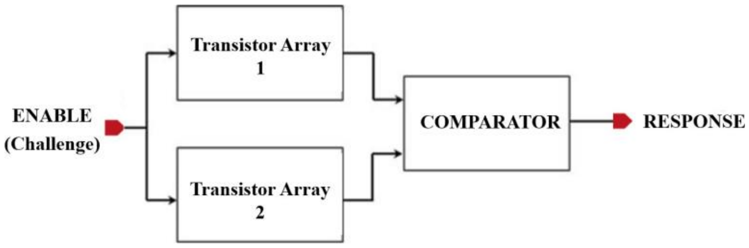


Figure 1.18: A current-based PUF.

The principle of operation of this architecture is simple and is described as follows:

- The challenge signal is applied to two identical arrays of transistors.
- The challenge signals select a number of transistors that send their response to the output of the matrices.
- The outputs of the matrices present a current value, which is compared in order to have a binary response.

The number of challenge/response pairs is of high importance for the PUFs, as well as those based on delay, these also depend on the architecture being used. Since this architecture depends on matrices, the variables k and n are defined as the number of rows and columns that exist in the structure, respectively. Therefore, the expression that defines the CRP is the following (Halak B. , 2018).

$$CRP = 2^{kn} \tag{1.6}$$

Several ways to realize transistor arrays such as the one proposed in the paper (Kalyanaraman, 2013) can be found in the literature. However, this design has some shortcomings, such as the low voltage at the output of the array. Which makes it difficult to develop a comparator that is robust enough to determine a good comparison. Animproved version for this transistor array was proposed in (Mispan, 2015). A simplified block diagram is presented in Fig.1.19 below.

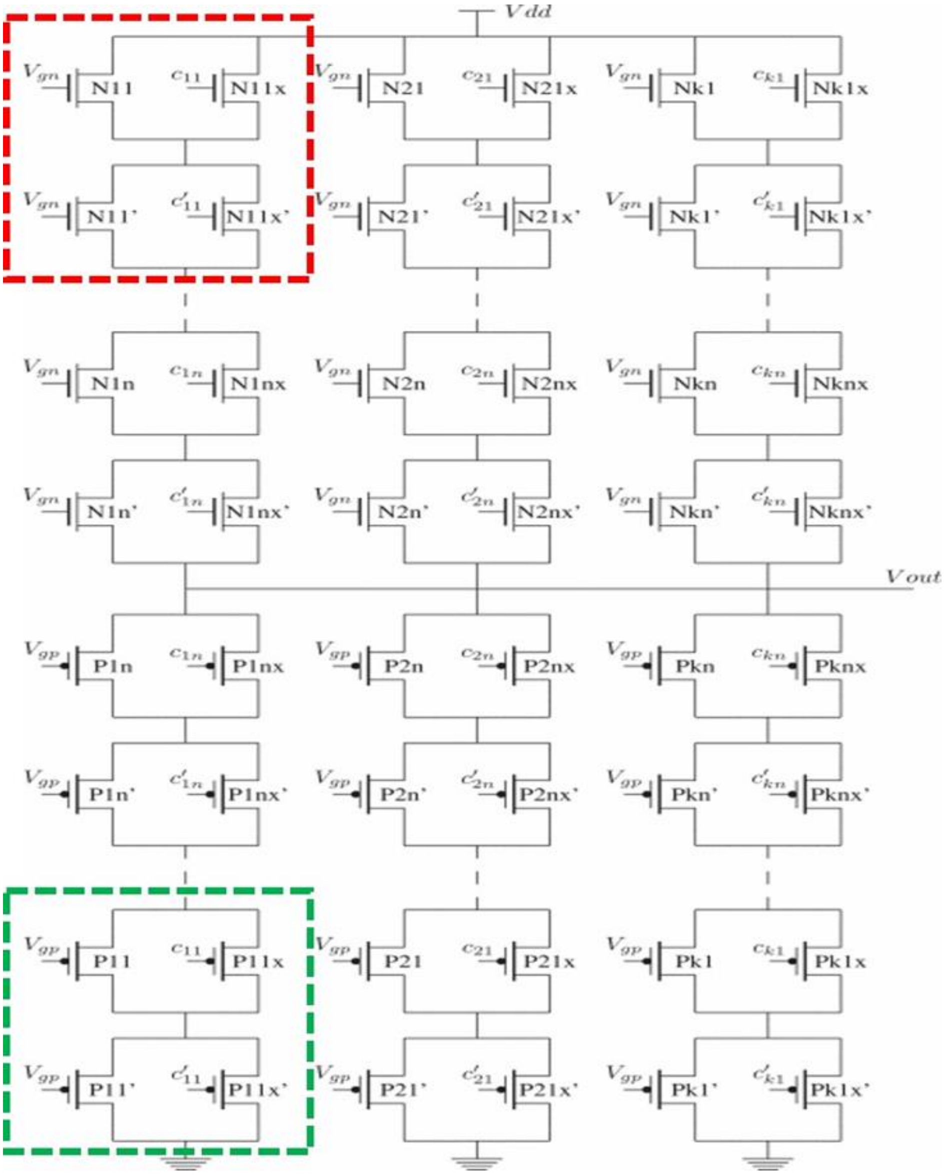


Figure 1.19: Circuit schematic of the TCO unit array.

The matrix shown in Fig.1.19 is made up of *k* columns and *n* rows, everything is built based on the unit cells that are highlighted in the image. By analyzing the unit cells we can realize that it is built with two transistors in parallel. One of these transistors has been designed with the minimum size allowed by the technology, in order to maximize the variability observed in its threshold voltage. This transistor is of critical importance to the unit cell and is referred to as a “stochastic transistor”, e.g., in Fig.1.19 it would be the $N\,11x$ transistor. The function of the second transistor is to operate as a switch, working as a block for the stochastic transistor when it is on, or to allow to deliver the response when the stochastic transistor is off. We can find as an example of the second transistor in Fig.1.19 the $N\,11$ device, we will know it as “switch transistor” (Mispan, 2015).

For each unit cell a single challenge bit is given for both the NMOS transistors and their symmetrical complement PMOS. For example, if we look at Fig.1.19 we can notice that the challenge bit C11 applies to the components marked in green and red.

In the following, we will analyze the fundamental behavior of the matrix and how it detects the values. If the challenge signal is a logic “1” the switch transistors will behave as follows, NMOS and PMOS would be on and off respectively. However, the stochastic transistor will always have a contribution to the response regardless of whether the challenge is “0” or “1”. This is because the inverted and non-inverted versions of each bit are connected to the stochastic transistor. This means that every second stochastic transistor is part of the network, regardless of the binary data that is carrying the challenge signal. This architecture is defined in the literature as “Two Chooses One” (TCO). Finally, the output provided by each matrix depends on the current accumulated by the transistors (Mispan, 2015).

Due to the inherent intra-die variations, it is possible to ensure that the output voltage of one of the arrays is slightly higher than the other. Consequently, dynamic comparators can be used to find this difference, e.g. Op- Amps (Mispan, 2015).

In order to build a current-based PUF, several parameters such as those shown in the previous example must be taken into consideration. First, the correct voltage must be applied to the gates of the stochastic transistors to allow them to operate in the subthreshold region. Moreover, the switch transistors must deliver a negligible sub-threshold current and at the same time they should provide small ON-state resistance. Furthermore, to achieve good performance in unit cells, the dimensions of the switch transistors should be at least ten times those of the stochastic transistors. It must also be understood that the difference in the voltages observed at the output of the arrays has to be sufficiently large to be well detected by the comparator thus avoiding unreliable responses (Halak B. , 2018).

Current-Based PUFs Using Dynamic Random Access Memories

The basic structure of a DRAM cell is constituted by a capacitor along with an access transistor as shown in Fig.1.20. Each basic cell has the capacity to store a single bit of information. Without periodic updates the leakage current from the access transistor may cause the value to be lost. The velocity at which the capacitor is discharged is directly proportional to the leakage current, which depends on the technology in which the design was manufactured. The latter is strongly affected by process variations indeed, cells belonging to the same wafer exhibit different leakage currents. In the literature we can find innovative ideas on how to use memories as PUFs, however, we will use as a basis the work proposed by (Xiong Wenjie, 2016) in which the unique behavior of value decay in DRAM cells induced by process variability was exploited for PUF applications:

- First we need to select the region of the memory that we want to behave as a PUF, this is achieved by defining the initial address and the size we want to analyze.
- The update function is then disabled for this specific part of the memory.
- We proceed to write an initial value to be stored in this region.

- Then access to all cells in the previously selected region is disabled for a defined time (t). During this period of time the load of each unit cell decays proportionally to the leakage current that each cell has.
- After the time that the memory section was disabled, a reading is made to verify which values remained stored. This response is the value given to the PUF response.
- To conclude the process, the previously selected region is returned to normal operation and becomes available for the operating system again.

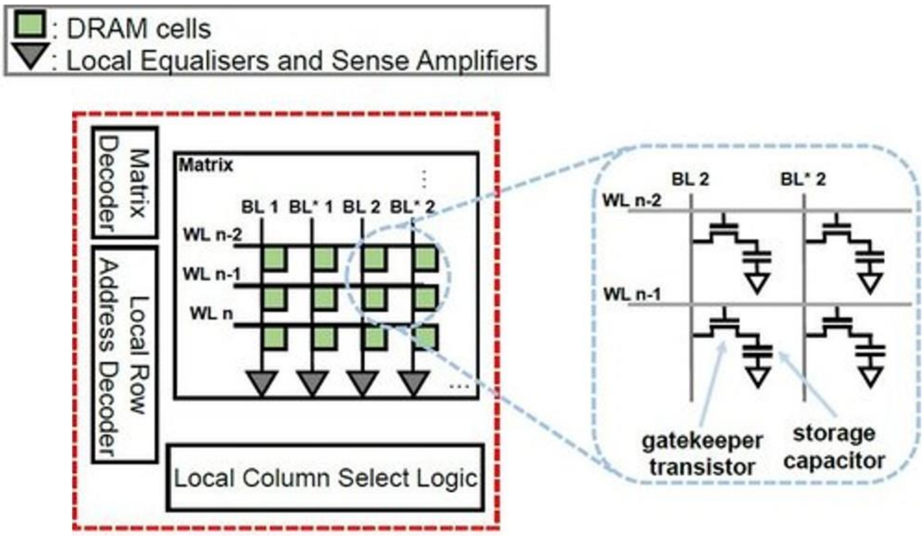


Figure 1.20: DRAM.

Similar to the rest of PUFs, its value of challenge/response pairs depends on its architecture. For this specific type of array, it depends on the size of the memory section selected, and the decay behavior of the DRAM must also be considered. To understand this better, we could imagine the case in which in a DRAM memory there are R spaces reserved to structure the PUFs. Where it is possible to have different N decay periods for each selected space, which would give a unique response for each chip. The expression used to calculate the CRP for this architecture is shown below (Xiong Wenjie, 2016).

$$CRP = RN \tag{1.9}$$

As with all integrated circuits, if you want to implement this DRAM- based architecture, there are several points to consider. The first point to take into account is the careful study of the DRAM memory behavior, in order to understand which are the decay times of the unit cells. To understand this more objectively, let's look at the following case: if the leakage current of the DRAM cells is very high and the discharge time is very long, there is a risk of losing all the data. It is also possible to consider a totally opposite case, where the load decay is too slow and when analyzing these data there is no variation. For both cases the behavior of the PUF would be too predictable, so this design would be a bad implementation for some types of memory (Halak B. , 2018).

Voltage-Based PUFs

This type of PUFs is responsible for transforming manufacturing process variations into voltage values that can be measured and quantified to have a digital response. Two architectures that are quite strong in the current literature of voltage-based PUFs are briefly presented below. The first solution uses as base cell a static random access memory (SRAM), the second circuit on the other hand uses as base a S-R Latch (Halak B. , 2018).

SRAM PUFs

This type of solution, which is based on the use of static random access memories, is found in the literature as the first approaches to a PUF that bases its response on voltage measurement. Initially, this proposal was used in FPGAs in order to generate encryption keys, this type of security is specific to the device and is used to encode bit streams before storing them in an external memory. This specific proposal focuses on data protection when the adversary has the ability to decrypt the storage bitstream. These PUFs fulfill the objective of preventing the attacker from reusing a bitstream to program other FPGAs (Guajardo Jorge, 2007).

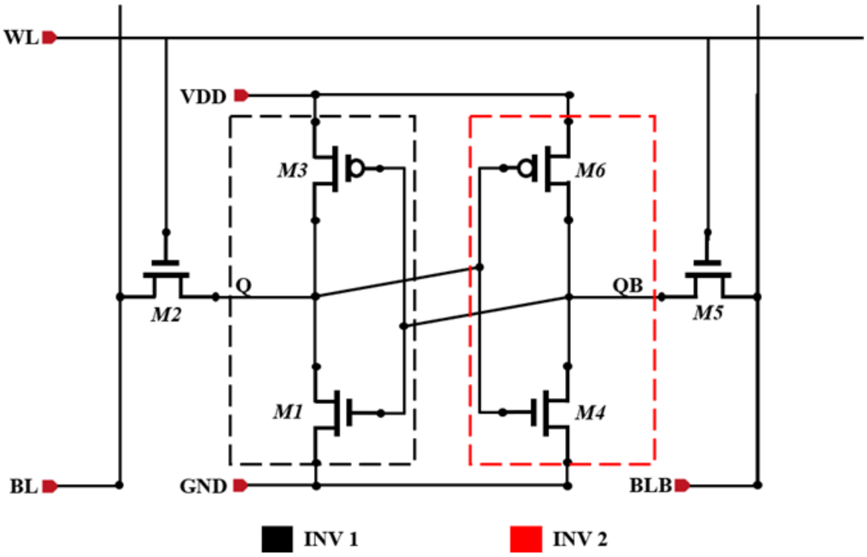


Figure 1.21: 6T-SRAM Schematic.

We will now discuss the behavior of an SRAM-based PUF, for which it is necessary to understand the basic circuitry of this type of memory. Fig.1.21 presents the basic architecture of the six-transistor SRAM, which consists of two cross-coupled inverters and two access transistors. The inverters can be in two logic states “1” or “0”, in order to reach these values it is necessary to use the M2 and M5 transistors. The access transistors M2 and M5 receive the signals from the voltage lines BL and BLB, taking into account that the voltage levels received by these lines must be complementary. This means that if we want a correct storage operation, BL and BLB must receive logic “1” and “0” respectively or vice versa in order to access the cell correctly. The inverters are going to drive inverse states, in other words if we look at Fig.1.21 we have that INV 1 and INV 2 drive Q and QB respectively. The voltage line that we know as WL is the one in charge of selecting the operation that we are going to perform with the memory such as: storage, read and write (Guajardo Jorge, 2007).

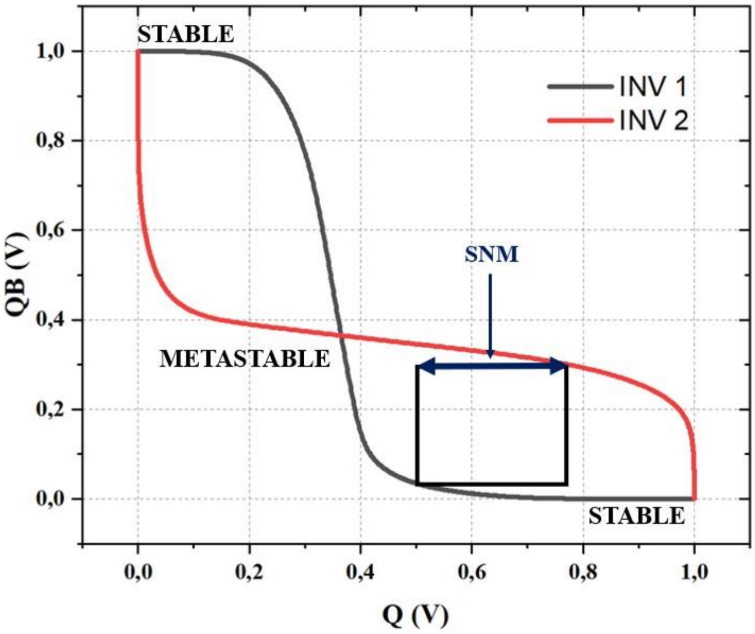


Figure 1.22: Characteristics for SRAM.

When the SRAM cell is biased to the correct voltage values by the access transistors, the cross-coupled inverters start a “power struggle”. The winner of this contest is decided by the difference in the MOSFETS used by the inverters. As can be seen in Fig.1.22 this ends up with three possible operating points for the memory, where two of these states are stable and the third is considered metastable. In the hypothetical case where all the transistors that make up the SRAM cell are perfectly coupled, in theory the state of this memory would be metastable for all time. However, in real life due to manufacturing variation processes, even though at the circuit level the whole structure is designed perfectly coupled. Always one of the two inverters in the SRAM will have a higher conduction current, which is why it is possible to define the initial starting value of the cell (Guajardo Jorge, 2007).

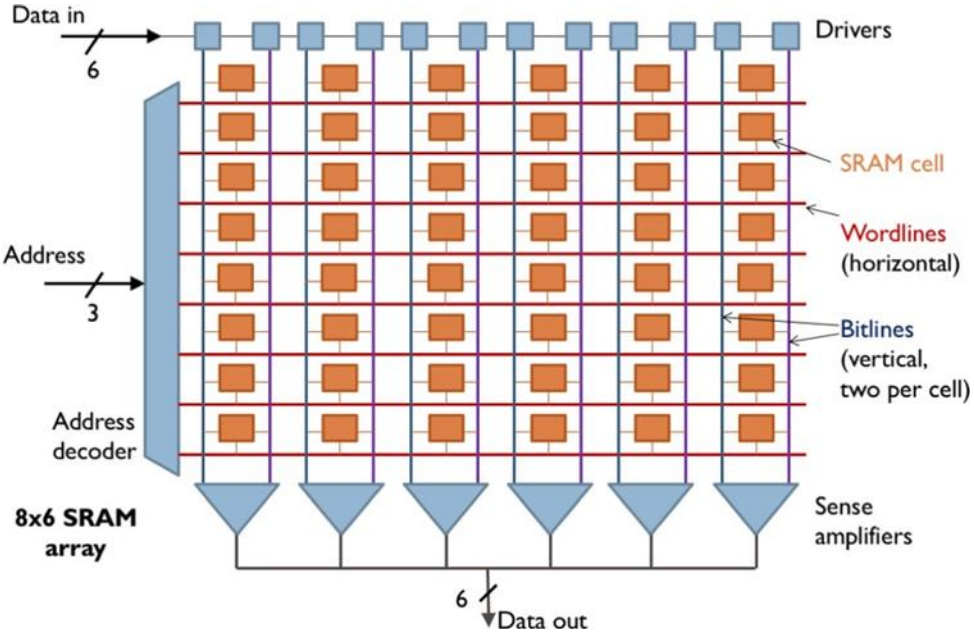


Figure 1.23: SRAM Memory Array.

The basic cells of SRAMs usually have an initial state, which is obtained at the specific moment of turning on these memories. This specific feature is what allows us to use this type of devices for the creation of PUFs. Since the reading of the resulting PUF value depends on the size of the memory, we can say that the CRP is proportional to the size of the device array. In other words, the challenge is the address of the memory to be read and the response is the read values of the beginning of the addressed cells. To understand this more objectively let’s consider the following example, for a 64 megabits byte- addressable memory we have 8 megabits of CRP.

Latch-Based PUFs

This is another type of design that uses voltage variation as a basis for the generation of PUFs based on relatively simple circuits. In this design, it attempts to take advantage of the small variability of threshold voltages that the NOR gates coupled together have (Stanciu, 2016). Fig.1.24 presents a basic schematic of the design of this type of circuit. The principle of operation of this cell is identical to that described for the arbiters in lanches using delay-based PUFs. What is important to mention is when the cell enters a metastable state, since after a certain time a logic “1” or ”0” can be obtained. These logic values will be strongly related to the variations in the driving capability presented by the gates.

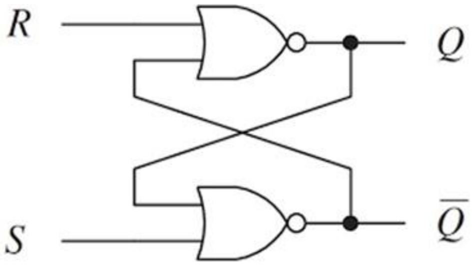


Figure 1.24: SR Latch.

As simple as this circuit may seem, it is not so reliable for the development of a PUF. In order to operate as a cryptographic device, it needs to enter a metastable state. Therefore, it is quite difficult to determine the minimum time it takes for the circuit to overcome this state. In addition, these types of cells are often very susceptible to response variations due to changes in the conditions of their environment (Stanciu, 2016).

Metrics of PUF Devices

In the previous sections of this chapter we discussed the design of PUFs and how to use their physical variability to generate safety systems. However, we have not yet discussed the metrics that these cells must satisfy to guarantee the quality of a PUF design. In addition to the qualities it must have to be suitable for a specific application. For this reason, in this section we will introduce the four metrics with which the physical unclonable functions are evaluated. Where we find the parameters of: uniqueness, reliability, uniformity, and tamper resistant. It is also necessary to mention that the Hamming distance and Hamming weight are used for the calculation of these metrics. For this reason, below you will find the definition of these concepts (Maiti Abhranil, 2013).

Hamming Distance: The Hamming distance d(a, b) between two words a = (a_i) and b = (b_i) of length n is defined to be the number of positions where they differ, that is, the number of (i)s such that $a_i \neq b_i$.

Hamming Weight: Let 0 denotes the zero vectors: 00...0, The Hamming Weight HW (a) of a word a = a_1 is defined to be d(a, 0), the number of symbols $a_i \neq 0$ in a.

Uniqueness

This is a metric to determine the ability of a device to have a unique response. It is a measure in which the ability of a PUF to behave in a unique and distinguishable way from other PUFs that have the same structure on other chips is determined (Maiti Abhranil, 2013).

$$HD_{INTER} = \frac{2}{k\,(k-1)} \sum_{i\,=\,1}^{k-1} \sum_{j\,=\,i+1}^{k} \frac{HD(R_i(n), R_j(n))}{n} * 100\% \qquad \textbf{1.10}$$

To better understand what equation 1.10 is trying to explain, let us consider the example shown in Fig.1.25. In this example we can observe two PUFs receiving the same challenge (011101), to which each device gives a different response. If we analyze the responses given we have the following PUF 1 = 0111000 and PUF 2 = 1111001, where we can see that the two responses differ by 2 bits. With this we can say that out of 7 bits of response 2 bits change therefore the PUFs are unique. However, the inter-chip Hamming distance is small as it is 28% and to ensure that these devices are safe we are looking for a 50% variation between the response of the chips.

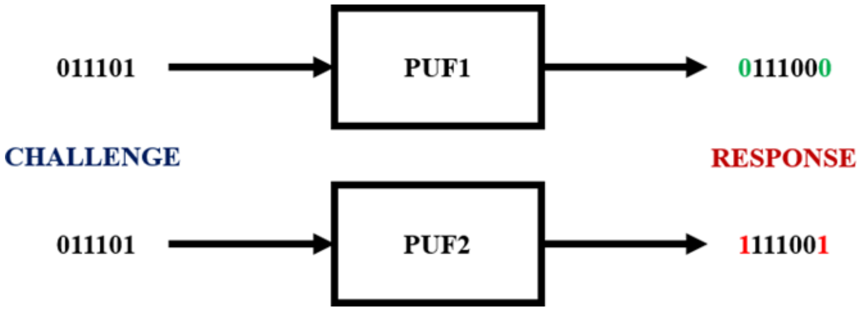


Figure 1.25: Uniqueness evaluation of a PUF.

Reliability

This metric determines the PUF's ability to deliver the same response for a fixed challenge even under different environmental conditions (i.e., temperature and voltage variations). The concept by which this metric is evaluated is the intra-chip Hamming distance (Maiti Abhranil, 2013).

$$HD_{INTRA} = \frac{1}{k} \sum_{i=1}^{k} \frac{HD(R_i(n), R'_i(n))}{n} * 100\% \qquad \textbf{1.11}$$

$$Reliability = 100\% - HD_{INTRA} \qquad \textbf{1.12}$$

In order to have a clearer understanding of this metric, we will consider the example shown in Fig.1.26. As we can see, a PUF is given the same challenge, but its response is evaluated at different temperatures. Ideally, the Hamming distance inside the chip should be 0%. This would guarantee 100% device reliability. However, in real life achieving this goal is very complicated, that is why in Fig.1.26 we have the case where changing the temperature from 273K to 320K the response is modified. This causes the on-chip Haming distance to be 14% and the final PUF reliability to be 86%.

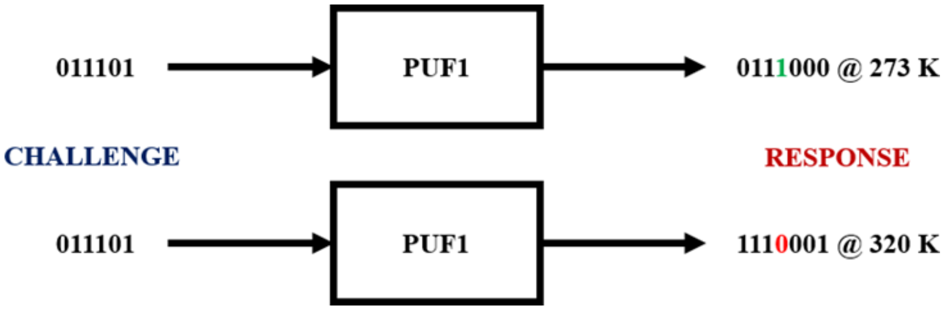


Figure 1.26: Reliability evaluation of a PUF.

Uniformity

This is a metric to determine the unpredictability of the responses that a PUF can give. To calculate uniformity it is necessary to use the concept of Hamming's average weight, where in order to have a completely random response, this value needs to be 50%. In other words, the objective is that the proportion of logic "1" and "0" should be the same in the PUF output (Maiti Abhranil, 2013).

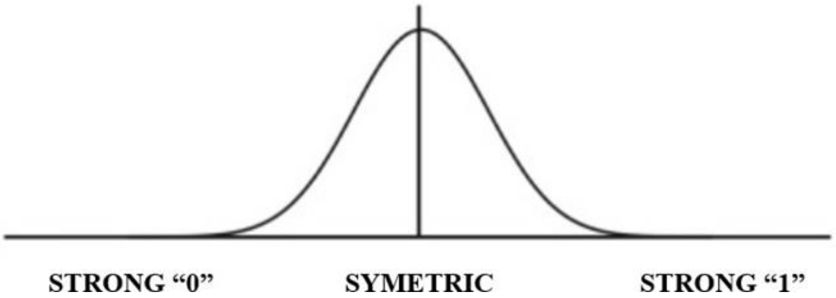


Figure 1.27: Uniformity response of a PUF.

$$Uniformity = \frac{1}{k}\sum_{i=1}^{k} r_i * 100\% \qquad \textbf{1.13}$$

Tamper Resistance

This is a metric that attempts to determine the resistance of a design to manipulation by adversaries. Ideally a PUF changes its response completely if the design or structure is modified by any attack on the device. This can be determined using the Hamming distance between the original chip and the tampered chip. The mathematical expression with which this value can be calculated is given below (Maiti Abhranil, 2013).

$$HD_{AVE} = \frac{1}{CRP}\sum_{L=1}^{CRP} \frac{HD\ (R_i(l), R_i'(l))}{n} * 100\% \qquad \textbf{1.14}$$

To determine that a PUF is tamper resistant the response it should exhibit after the Hamming distance calculation is 50%. Since the response of the PUF is completely different from the one it should have before the manipulation (Maiti Abhranil, 2013).

DEVELOPMENT AND SIMULATION METHODOLOGY

This chapter will describe the design and simulation methodology used for the development of the PUFs. This type of integrated circuits that seek to exploit physical imperfections due to manufacturing variations to generate a hardware signature, strongly depend on the symmetry of the design. For this reason, it was decided to use the full custom methodology for VLSI circuits. The following is a brief description of this design method and the implications necessary for its correct execution.

Full custom metodology

The fully custom design of integrated circuits is of great importance, as it allows the designer to place each transistor and its connections according to the design specifications. The real benefit of using this technique is that maximum performance can be achieved, silicon area can be minimized and power consumption can be improved. However, its application is too laborious and the development time is very long. For that reason, this design technique is limited to integrated circuits that need very efficient performance and mass production. In the following, we will describe the necessary requirements for a fully customized design such as: the transistor technology, the development tool and the complete flow of this methodology

Technology

The transistor technology node refers to the specific semiconductor manufacturing process including design rules. Throughout history the process node name refers to the physical characteristics of a transistor, such as channel length. Lately, due to various marketing issues, this number has lost its meaning. Because the name given to newer technology nodes does not refer to either the gate length or the average pitch. The main reason for maintaining node scaling is Moore's law. Moore's law states that in order to achieve a doubling of transistor density in a specific area, the poly-pin contact (CPP) and minimum metal pitch (MMP) have to increase by 0.7 times between technology nodes. Generally when reference is made to a smaller technology node, smaller device size is usually meant. This means that more transistors can be fabricated in a smaller area, with the benefits of improved performance and reduced power consumption.

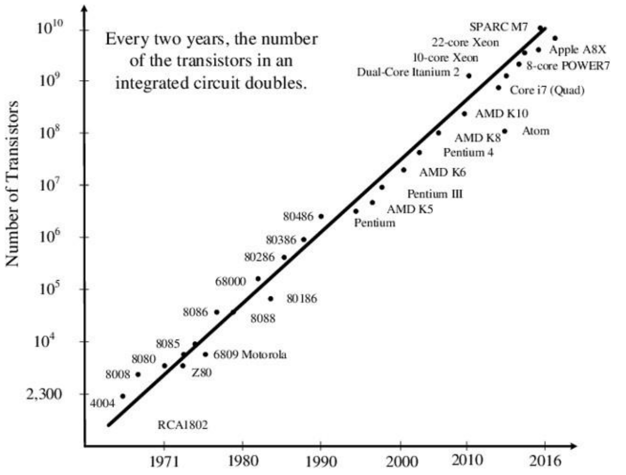


Figure 2.1: Evolution of the number of transistors every two years according to Moore's Law.

Taiwan Semiconductor Manufacturing Company (TSMC) is the world's largest chip manufacturer. Founded in 1987, TSMC was the first company to focus exclusively on semiconductor devices. Although TSMC offers several silicon wafer product lines, it is best known for its logic chip production area. For this thesis and the design of the PUFs, TSMC's 180nm technology node was used. In particular, for the PUF topologies considered in this the- sis, NMOS with regular threshold voltage (RVT) and PMOS with medium threshold voltage (MVT) were used.

Development software

Technological Computer Aided Design (TCAD) is one of the automation branches of electronic design with which the fabrication of semiconductor de- vices and their behavior is modeled. Technology files and design rules are the most important elements for the design of integrated circuits. These software present high accuracy in terms of the manufacturing process technology, variability and operating conditions of the integrated circuit. They are extremely important to accurately determine the performance, yield and reliability of the chip. For this reason, modeling and simulation are a very important aspect in the evaluation of the integrated circuit.

The TCAD takes into account the physical description of the devices considering their configuration, material-related properties and the connections that exist between the physical and electrical model. Since physics-based device modeling is the fundamental part of IC development, TCAD aims to quantify the understanding of the technology. The objective of TCAD is to quantify the understanding of the technology and to abstract that knowledge at the design level, including the extraction of parameters that support the development of the circuit at the electronic level. For this thesis we used the Cadence-Virtuoso TCAD focused on analog circuit design. The following is a brief description of how this software works and what can be done with it.

1. Virtuoso Analog Design Environment

This integrated circuit development environment provides the necessary capacity for the analysis, exploration and verification of a design based on the user’s needs. This allows the designer to achieve all the requirements required for the development of the project, taking into account all the flow involved. This tool is used for the control, management and simulation of analog circuits. It allows the designer to flexibly select the level of customization desired in the integrated circuit.

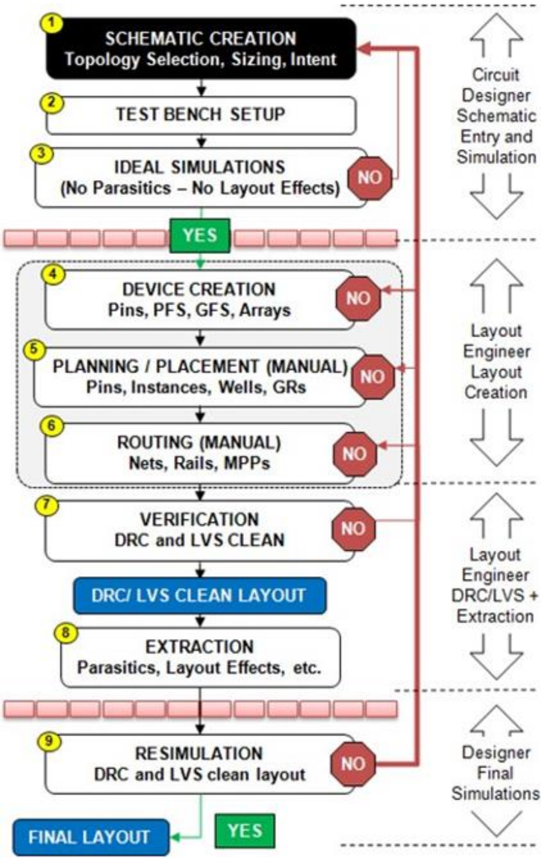

1
SCHEMATIC CREATION
Topology Selection, Sizing, Intent
2
TEST BENCH SETUP
3
IDEAL SIMULATIONS
(No Parasitics – No Layout Effects)
NO
YES
4
DEVICE CREATION
Pins, PFS, GFS, Arrays
NO
5
PLANNING / PLACEMENT (MANUAL)
Pins, Instances, Wells, GRs
NO
6
ROUTING (MANUAL)
Nets, Rails, MPPs
NO
7
VERIFICATION
DRC and LVS CLEAN
NO
DRC/ LVS CLEAN LAYOUT
8
EXTRACTION
Parasitics, Layout Effects, etc.
9
RESIMULATION
DRC and LVS clean layout
NO
FINAL LAYOUT
YES
Circuit Designer Schematic Entry and Simulation
Layout Engineer Layout Creation
Layout Engineer DRC/LVS + Extraction
Designer Final Simulations

Figure 2.2: Flow of custom design.

- First, it is necessary to select the technology to work with.
- Then, design circuit at schematic level using Cadence's schematic editor.
- The circuit is then simulated using the Cadence analog simulation environment. This is done in order to verify that the design is working as desired.
- Once the device specifications have been reached, the circuit is created at the physical level using the Virtuoso layout tool.
- The layout resulting from the implementation must be verified by some geometrical rules that depend on the selected technology. For this purpose, design rule checks (DRC) are used.

- Continuing with the flow, the electrical verification has to be performed, checking for short-circuit faults. This is possible by checking the electrical rules (ERC).
- Afterwards, it is necessary to verify the circuit's schematic and layout agreement. Checking that all devices and connections are correctly made, using the layout versus schematic (LVS) analysis.
- Consequently, it is necessary to extract a network that includes all the parasitic elements of the physical implementation. Using the Parasitic Layout Elements Extraction (PLE) tool.
- Finally, a simulation is performed considering these parasitic elements, in order to have a response that is very similar to real life and that we would find after printing. This last analysis is known as post layout simulation.

2. Analog Design Environment XL

The Analog Design Environment XL is an advanced simulation tool that incorporates Virtuoso. This platform supports comprehensive analysis of multiple designs and is therefore established as the standard in IC verification.

- Supports extensive verification of environmental conditions and operat- ing conditions.
- Analysis of multiple simulations, through testing under the operating conditions of the integrated circuit. Compiling all data in an easy-to-use database.
- It is capable of performing process variation simulations, corners, para- metric sweeps, Monte Carlo and reliability analysis.
- It allows quick debugging of the design by delivering a response that depends on the design stored in the system.

PROPOSED SCHEMES AND RESULTS

In the following chapter we will present the development of two PUF cells. The main point of interest is to understand the benefits of using a dynamic structure and a static structure. The static cell will have a voltage divider as its basis, the dynamic cell will have a ring oscillator as its main component.

Metastable PUF bitcell

The bitcell described here belongs to the voltage based class. In particular, the circuit is based on a sub-threshold voltage divider between two nominally identical series connected sub-circuits (i.e., top circuit TC and bot- tom circuit BC with nominally TC ≡ BC) along with an output inverter for generating the binary response as illustrated in Fig.3.1. The circuital approach intrinsically guarantees randomness from transistor mismatch as well as robustness against inter-die process variations, environmental variations (i.e., voltage and temperature variations) and aging.

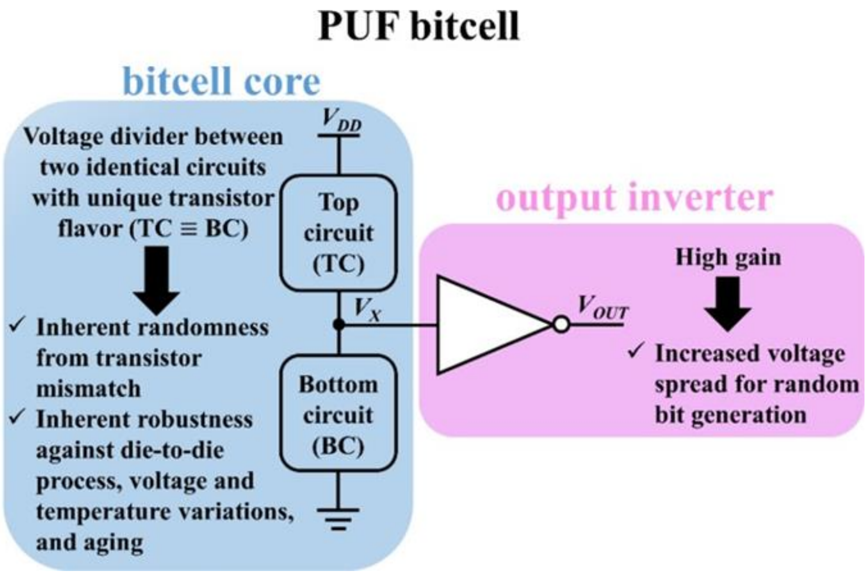
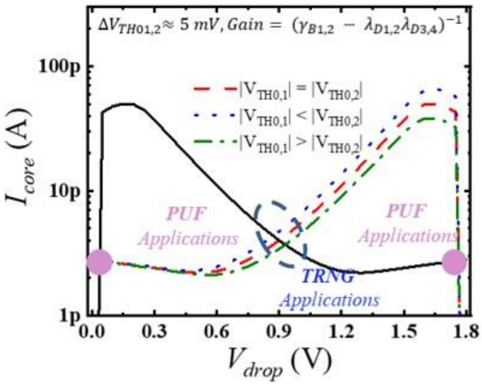

PUF bitcell
bitcell core
Voltage divider between two identical circuits with unique transistor flavor (TC ≡ BC)
✓ Inherent randomness from transistor mismatch
✓ Inherent robustness against die-to-die process, voltage and temperature variations, and aging
V_{DD}
Top circuit (TC)
V_X
Bottom circuit (BC)
output inverter
V_{OUT}
High gain
✓ Increased voltage spread for random bit generation

Figure 3.1: Conceptual diagram of the proposed PUF bitcell.

Fig.3.4 illustrates the transistor-level scheme of the analyzed bitcell which employs a four transistor (4T) sub-threshold voltage divider between two transistors (2T) sub-circuit. Each 2T block is composed of the series of negative-VGS NMOS (i.e., M_3 in the TC and M_4 in the BC with nominally $M_3 \equiv M_4$) with a reverse body-biased PMOS (i.e., M_1 in the TC and M_2 in the BC with nominally $M_1 \equiv M_2$). From the same figure it is also necessary to point out that the body terminal of M_3 and M_4 is connected to the relative source terminal which required the use of deep-n-well regular-voltage transistor (RVT) while for zero-V_{GS} biased PMOS (i.e., M_1 and M_2) medium-voltage transistors (MVT) were used for achieving more variability. Both transistors in each 2T sub-block were upsized (i.e., $W_{1,2} = W_{3,4} = 5$ µm) while keeping low their channel length (i.e., $L_{1,2} = L_{3,4} = 250$ nm which corresponds to a minimum channel length for MVT transistors, M_1 and M_2, and a channel length close to the minimum one for the RVT transistors, M_3 and M_4) to counteract the effect of the parasitic diodes associated to the use of deep-n-well transistors. Conversely, the output inverter was minimum sized because its task is only to digitize the voltage coming from the 4T voltage divider as well as to guarantee an high input impedance for electrically isolating each bitcell since the current flowing in the 4T-core is very small.



☐ Low gain and low stability during TRNG operation

☐ High spread and high stability during PUF operation

Figure 3.2: Metastability characteristic.

According to what said above in the nominal case (i.e., $M_1 = M_2$ and $M_3 = M_4$) the V_X voltage is equal to $V_{DD}/2$ but when mismatch occursM_1 will differ from M_2 as well as M3 from M4 thus leading a straight difference between the two sub-circuits thanks to the positive loops of each of them. This straight difference will result in a deviation of V_X from its nominal value (i.e., $V_{DD}/2$). This concept is well illustrated in Fig.3.2 which shows the current flowing each sub-circuit versus their voltage drop. These characteristics highlight that three operative points exist at the same time for a given mismatch. In particular, the point A and B in the figure represent a logic "1" and a logic "0" value, respectively at the output of the inverter while the point C corresponds to a metastable point. In this way it is possible to get a logic "1" or "0" as output by forcing the circuit in its metastable point and letting the metastability resolve itself according to the mismatch between the two sub-circuits.

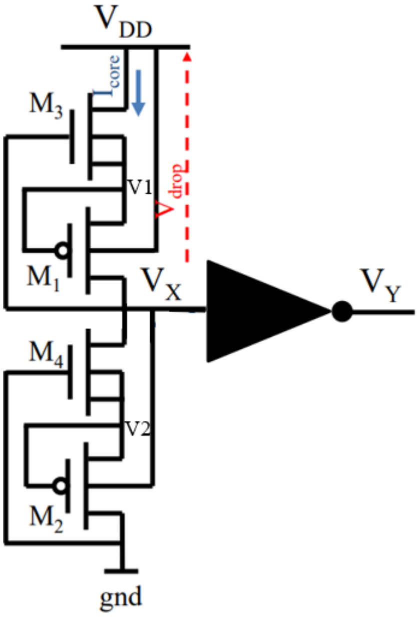


Figure 3.3: Transistor-Level scheme of the 4T bitcell.

To gain deep insight to the 4T-core it is possible to analytically deviate the V_X value in the metastable point by following. Fig.3.3 shows the circuit diagram of the solution proposed in this thesis, using a voltage divider of four transistors (4T). From these few devices connected in series we can notice that two identical subcircuits of two transistors (2T) are formed. These subcircuits are composed of a NMOS and PMOS transistor connected in series. Where the PMOS transistors M_1 and M_2 have a voltage V_{SG} zero and nominally these devices are identical. On the other hand, the NMOS transistors M_3 and M_4 have a negative V_{GS} voltage and likewise nominally identical. Devices with zero V_{GS} serve as major sources of mismatch in terms of threshold V_{th}. To understand this better we can say that the relative voltage difference between transistors M_1 and M_2, depends mainly on the voltage difference between subcircuits TC and BC. This can be best observed by transistors M_3 and M_4 operating with reverse bias, so the mismatch between transistors M_1 and M_2 can be increased. This can be easily verified in the schematic, since the source voltage drain V_{SD} of M_1 and M_2 correspond to $-V_{GS}$ of M_3 and M_4.

With which we can say the following:

- For M_1 to be stronger than M_2, the voltage V_{SD} of M_2 is greater than that of M_1, causing M_4 to be weaker than M_3. Consequently the voltage of subcircuit *BC* decreases, pushing the voltage at node V_x toward V_{DD}.

- For M_2 to be stronger than M_1, the voltage V_{SD} of M_1 is greater than that of M_2, causing M_3 to be weaker than M_4. Consequently the voltage of subcircuit *BC* decreases, pushing the voltage at node V_x toward ground.

Considering that the circuit works in the subthreshold regime, the current of the transistors is determined by this regime. Also because of the connection that the devices have, the effect of drain induced barrier lowering (DIBL) is observed, this effect can be expressed with the following equation.

$$I_{SUB} = I_0 \frac{W}{L} exp\left(\frac{V_{GS} + V_{TH}}{n\, V_T}\right)\left[1 - exp\left(-\frac{V_{SD}}{V_T}\right)\right] \quad \textbf{3.1}$$

From this expression we have the following terms: I_0 is the intrinsic subthreshold current, W and L refer to the transistor width and length respectively, n is the slope factor, V_{th0} is the bias threshold voltage Vth at room temperature (T_{room} = 25ºC) and the DIBL coefficient is given by λ_D. The following approximation can be made in the case where the voltage V_{SD} is greater by at least three times the thermal voltage V_T.

$$I_{SUB} \approx I_0 \frac{W}{L} exp\left(\frac{V_{GS} + V_{TH0} + \lambda_D V_{SD}}{n\, V_T}\right) \quad \textbf{3.2}$$

$$V_T = \frac{kT}{q} \quad \textbf{3.3}$$

This thermal voltage term is determined by the equation 3.3. Where the value of k is Boltzmann's constant, T refers to the absolute temperature in degrees Kelvin and q is the electron charge. The equations for the $M_1 - M_4$ transistor currents are presented below.

$$I_{M3} = I_{03} \frac{W}{L} exp\left(\frac{V_{OUT} - V_1 - V_{VT03} + \lambda_{D3,4}(V_{DD} - V_1)}{n_{3,4}\, V_T}\right) \quad \textbf{3.4}$$

$$I_{M1} = I_{01} \frac{W}{L} exp\left(\frac{V_{TH01} + \lambda_{D1}(V_1 - V_{OUT}) + \gamma_B(V_1 - V_{DD})}{n_{1,2}\, V_T}\right) \quad \textbf{3.5}$$

$$I_{M4} = I_{04} \frac{W}{L} exp\left(\frac{-V_2 + \lambda_{D3,4}(V_{DD} - V_1) - V_{TH04}}{n_{3,4}\, V_T}\right) \quad \textbf{3.6}$$

$$I_{M2} = I_{02} \frac{W}{L} exp\left(\frac{V_{TH02} + \lambda_{D1,2}(V_2) + \gamma_B(V_2 - V_{OUT})}{n_{1,2}\, V_T}\right) \quad \textbf{3.7}$$

For the development of these equations, the slope factor and DIBL of transistors M1 and M2 were named n1,2 and λD1,2 likewise for M3 and M4 have n3,4 and λD3,4. This can be done so since these transistors are nominally identical. It is necessary to mention that between devices M1 and M3 there is voltage V1, also the potential drop V2 is between M2 and M4. To obtain the values of these voltages V1 and V2 we have to equal the current equations of the transistors M1 = M3 and M2 = M4 respectively. Solving these equalities we have the expressions of V1 and V2 in equations 3.8 and 3.9.

$$V_1 = \left[\frac{1}{n_{1,2}+n_{1,2}\lambda_{D3,4}+n_{3,4}\lambda_{D1,2}+n_{3,4}\gamma_B}\right] * \left[-n_{1,2}V_{TH03} - n_{3,4}V_{TH01} + \left(n_{1,2} + n_{3,4}\lambda_{D1,2}\right)V_{OUT} + \left(n_{1,2}\lambda_{D3,4} + n_{3,4}\gamma_B\right)V_{DD} - n_{1,2}n_{3,4}V_T ln\frac{I_{01}}{I_{03}}\right] \qquad \textbf{3.8}$$

$$V_2 = \left[\frac{1}{n_{1,2}+n_{1,2}\lambda_{D3,4}+n_{3,4}\lambda_{D1,2}+n_{3,4}\gamma_B}\right] * \left[-n_{1,2}V_{TH04} - n_{3,4}V_{TH03} + \left(n_{1,2}\lambda_{D3,4} + n_{3,4}\gamma_B\right)V_{OUT} - n_{1,2}n_{3,4}V_T ln\frac{I_{01}}{I_{03}}\right] \qquad \textbf{3.9}$$

To derive the value of VX delivered by the 4T base cell it is necessary to substitute equations 3.8 and 3.9 into 3.4 and 3.5 respectively, obtaining the expression we have in 3.11. To have a point of comparison with other works we can observe the expression 3.10, which is the resultant value of a voltage divider of two transistors that we can observe in Fig.3.4. The equations and results of this PUF cell were presented in (De Rose, 2017).

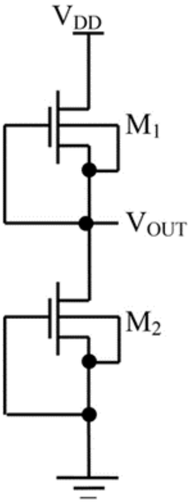


Figure 3.4: PUF bitcell with 2T-core.

$$V_X = \frac{V_{DD}}{2} + \frac{1+\lambda_{D3,4}}{2(\gamma_B - \lambda_{D1,2}\lambda_{D3,4})}\left[V_{TH02} - V_{TH01} + n_{1,2}V_T ln\frac{I_{02}}{I_{01}}\right] * \qquad \mathbf{3.10}$$

$$\frac{\lambda_{D1,2}+\gamma_B}{2(\gamma_B - \lambda_{D1,2}\lambda_{D3,4})}\left[V_{TH03} - V_{TH04} + n_{3,4}V_T ln\frac{I_{04}}{I_{03}}\right]$$

$$V_{OUT} = \frac{V_{DD}}{2} + \frac{V_{TH02} - V_{TH01}}{2\lambda_{D1,2}} + \frac{n_{1,2}V_T ln\frac{I_{01}}{I_{02}}}{2\lambda_{D1,2}} \qquad \mathbf{3.11}$$

These two expressions can be reduced if we take into account the logarithmic term; since the transistors we are working with are nominally identical. Therefore, these terms within the logarithm would have a ratio of one, so all the logarithmic terms would have a value of zero. Looking at expressions 3.12 and 3.13, because the mismatch transistors are not present it is observed that the dominant value of VX and V_{OUT} is $V_{DD}/2$.

$$V_X \approx \frac{V_{DD}}{2} + \frac{1+\lambda_{D3,4}}{2(\gamma_B - \lambda_{D1,2}\lambda_{D3,4})}[V_{TH02} - V_{TH01}] *$$

3.12

$$\frac{\lambda_{D1,2}+\gamma_B}{2(\gamma_B - \lambda_{D1,2}\lambda_{D3,4})}[V_{TH03} - V_{TH04}]$$

3.13

$$V_{OUT} \approx \frac{V_{DD}}{2} + \frac{V_{TH02} - V_{TH01}}{2\lambda_{D1,2}}$$

Considering that M_1 and M_2 are the major sources of mismatch of the deviation of V_X, we can say that expression 3.12 has a dependence on the V_{TH01} and V_{TH02} variation. By performing the respective comparison between 3.13 and 3.12, we can observe that the 4T base cell is more affected in the V_{TH0} difference of M_3 and M_4 devices by the DIBL effect. Mainly by the M_3 and M_4 transistor termination with the $\lambda_{D3,4}$ termination. For devices that do not have a short channel, it can be assumed that $\lambda_D = 1$, since the effect of the short channel is not so pronounced. For this reason the dispersion in V_X increases considerably in the voltage divider working with 4T comparing the results with its 2T counterpart. This can be observed better by deriving a simplified expressions of the standard deviation of V_X and V_{OUT} from the difference of V_{TH0} found in equations 3.12 and 3.13. To derive these expressions in a simpler way we will assume that the variation of the difference of V_{TH0} of M_1 and M_2 are statistically independent, furthermore we will take into account that $1 + \lambda_{D3,4} \approx 1$ in equation 3.12. With these considerations we can say that the standard deviation of voltage at nodes VX and VOUT i.e. σ_{VX} and σ_{VOUT} respectively, are given by the following expressions.

$$\sigma_{VX} \approx \frac{\sigma_{VTH01,2}}{\sqrt{2}(\gamma_B + \lambda_{D1,2}\lambda_{D3,4})} \quad \textbf{3.14}$$

$$\sigma_{VOUT} \approx \frac{\sigma_{TH01,2}}{\sqrt{2}\lambda_{D1,2}} \quad \textbf{3.15}$$

Where the term $V_{TH01,2}$ represents the standard deviation of the difference of V_{TH0} between M_1 and M_2 devices. Taking equation 3.15 into account, it is easy to notice that in the 4T base circuit, we have the additional contribution of $\lambda_{D3,4}$. Resulting in a higher random mismatch amplification between M_1 and M_2, when making the respective comparison with its 2T counterpart.

The effect of the increased mismatch between transistors M_1 and M_2 due to M_3 and M_4, is a very important advantage in the 4T voltage divider. Mainly due to the increase in voltage deviation at node V_X, enhancing the stability of the PUF cell. Taking into consideration that the voltages at V_X that are close to $V_{DD}/2$, are the tensions that potentially generate unstable bits at the inverter output. Mainly when the circuit is affected by various sources either environmental or electrical, such as could be voltage variations in bias, temperature, noise, etc. In fact, a wider voltage spread in V_X is of great benefit in reducing power consumption. This is because voltage values at the output of the base cell that are close to $V_{DD}/2$, produce an increase in the power consumed in the inverter. It is worth mentioning that this voltage divider based circuit with 4T is more robust to voltage variation when compared to its 2T counterpart. This can be explained by taking into account that there is the voltage drop across the reverse bias transistors M_3 and M_4, thus protecting the voltage difference between M_1 and M_2 from the bias variation VDDD. The same reason why this cell also has a higher robustness to noise.

However, to take full advantage of the increased mismatch in transistors M_1 and M_2 due to M_3 and M_4, ensuring that the response generated by the PUF cell is stable. Several design parameters need to be considered with respect to the 4T voltage divider. First, we must take into account the selection of the transistors, considering that these devices must have a low threshold voltage V_{TH}. In order to ensure that the subthreshold currents have a larger contribution in the response compared to the leakage currents in any low temperature mode of operation. Considering the proposal in (De Rose, 2017) since the main sources of mismatch are the M_1 and M_2 transistors, these should be sized with the purpose of having the best trade-off in the relation of equations 3.14, 3.15. In other words, what is sought is to have a high mismatch and a reduced DIBL effect in order to maximize σ_{VX}. To satisfy these conditions, relatively large channel lengths $L_{1,2}$ are often used, resulting in a very low channel modulation $L_{1,2}$. Also, to maximize the mismatch effect caused by the manufacturing processes, the minimum transistor width $W_{1,2}$ is used.

Knowing that M_3 and M_4 transistors help to increase the mismatch of M_1 and M_2, they must be sized taking into account the following requirements. The first requirement is to have a low sensitivity to mismatch, so they need to be sized large. The second requirement is to have adequate conductivity, in order for the relative strength between M_1 and M_2 to be independent of the drain source voltage drop $V_{SD1,2}$. Explaining this in more detail we can say that the conductivity of M_3 and M_4, has to be large enough to guarantee that $V_{SD1,2} > 3V_T$. However the conductivity cannot be too high, preventing the $V_{SD1,2}$ from counteracting the mismatch effect of the V_T H01,2 difference due to DIBL effect. To achieve this it is necessary to use a correct ratio between the width and length of the M_3 and M_4 transistors. This means that the channel length $L_{3,4}$ must be chosen favoring $\lambda_{D3,4}$, therefore the width $W_{3,4}$ must be selected in such a way that it helps to meet the above requirements. As a result of sizing transistors M_3 and M_4, the existing difference in the mis- match of M_1 and M_2 increases. As a result of equating 3.4 and 3.6 neglecting the logarithmic term we have the following equation.

$$V_{SD1} - V_{SD2} = \frac{V_{TH02} - V_{TH01}}{\lambda_{D1,2}} + \frac{\gamma_B (V_{BS2} - V_{BS1})}{\lambda_{D1,2}} \quad \textbf{3.16}$$

From the equation 3.16 we can verify that the difference supports the amplification effect by M_3 and M_4, since the voltage drops in these transistors is determined by $V_{SD1} - V_{SD2} = V_{SG4} - V_{SG3}$. Using this equality we can derive the following equation.

$$V_{SD3} - V_{SD4} = \frac{V_{TH03} - V_{TH04}}{\lambda_{D3,4}} + \frac{V_{TH02} - V_{TH01}}{\lambda_{D1,2}\lambda_{D3,4}} + \frac{\gamma_B (V_{BS2} - V_{BS1})}{\lambda_{D1,2}\lambda_{D3,4}} \quad \textbf{3.17}$$

Taking into account the equation 3.17, it is observed that a high $V_{SD1} - V_{SD2}$ helps to decrease the mismatch effect in transistors M_3 and M_4. This allows to not use such large dimensions for M_3 and M_4 decreasing the voltage difference $V_{SD3} - V_{SD4}$, this means that the first term of the equation 3.17 is negligible compared to the following ones. This is of great importance since it gives stability to the voltage divider cell with 4T, since it is not affected by the variation of voltages or temperatures favoring the mismatch of M_1 and M_2 that determine the cell response. The 4T base has a robust response to voltage variation even in the case of having different bias for M_1 and M_2, this is due to the effect of transistors M_3 and M_4 functioning as shielding. Furthermore, due to the mismatch of the V_{TH} difference between M_1 and M_2 there is the possibility of reversing with the change in temperature, this is independent of the boost provided by M_3 and M_4. In fact, if M_1 and M_2 are considered to have different temperature coefficients. We can say that $k_{T1} \neq k_{T2}$ therefore the usual linear relationship would be given by the following term $V_{TH} = V_{TH0} + k_T (T - T_{room})$. Reason why the difference of $V_{SD1} - V_{SD2}$ when considering the temperature coefficient, presents a strong dependence with these terms as observed in the equation 3.19. This can be understood as the possibility of reversing the difference in the relative strength of M_1 and M_2 upon temperature change.

$$V_{SD1} - V_{SD2} = \frac{V_{TH02} - V_{TH01} + (k_{T2} - k_{T1})(T - T_{room})}{\lambda_{D1,2}} + \frac{\gamma_B(V_{BS2} - V_{BS1})}{\lambda_{D1,2}} + \frac{n_{1,2}V_T}{\lambda_{D1,2}} ln \frac{I_{02}W_2L_1}{I_{01}W_1L_2} \quad \textbf{3.18}$$

From the design considerations for the PUF cell, we can understand that there is a strong trade-off between stability and circuit sizing. It is also necessary to mention that this depends on the technology selected for cell development. In fact, when the technology is reduced, the mismatch in terms of k_T increases. Thus the possibility of having a correct response of the PUF is affected in a greater way in the face of temperature variation. Also by reducing the technology node the DIBL effect is affected, this is because it increases the short channel effect λ_D. This is an undesirable effect as it prevents a large deviation of voltage values at node V_X. Therefore we can say that to keep the instability low it is necessary to work more on the CMOS process scaling, which could lead to an unfavorable trade-off in stability and area.

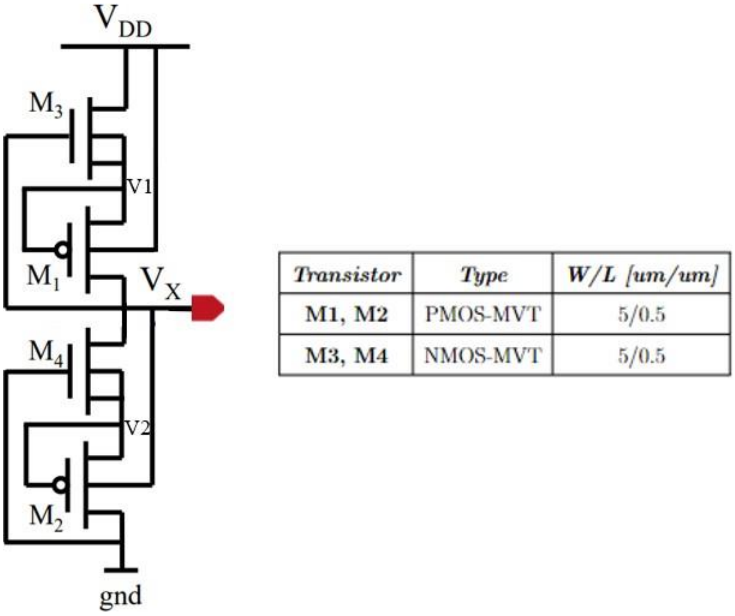
Thus exploiting the random mismatch between the complementary transistors. In addition, to transform this voltage value into a digital value, a high transconductance inverter will be used to minimize the range of unsteady values. It is also necessary to emphasize that because the cell is working in subthreshold regime, the circuit can work at a supply voltage causing the reduction in power consumption.

Fig.3.1 presents the PUF concept that was implemented in this thesis, where the base cell works with a voltage divider working in the subthreshold region. The block consists of two nominally identical subcircuits that are connected in series, for better understanding we will call the top part of the divider TC and the bottom part BC. The random variation that exists in the transistors due to the manufacturing process causes the voltage at node V_X to be random as well. It is also necessary to mention that working with nominally identical subcircuits in the base cell provides robustness against PVT variations (De Rose, 2017). After the base cell, an inverter is connected, in order to provide a high impedance at node Vx and a high gain to digitize the voltage Vx (Alvarez, 2016). Because the PUF requires a binary value, the value observed at the output of the inverter (V_{OUT}) represents the bit generated by the PUF base cell.

The metastable behavior was also confirmed by DC simulations. Since the output voltage V_X is at a metastable value the voltage values it can acquire are discrete. Therefore the working range at the output of this node is small. Like other metastable solutions once the circuit is assessed in a stable point no flips occurs even under noisy conditions but during the metastability resolution noise and/or different environmental conditions may affect the result thus leading to bit flips during different evaluations. Indeed when the mismatch in terms of transistor V_{TH} is high the output voltage will settle at the same value even under noisy or different environmental conditions but in the cases of small mismatch the output may change its value in different evaluations (i.e., external noise could be higher then the threshold voltage mismatch) and under voltage or temperature variations (i.e., the threshold voltage mismatch may flip varying the environmental conditions). A more in-depth study, assisted by both simulations and measurements, will be done in future.

Simulation Results of the PUF Bitcell

For the development of the above analysis, Cadence-Virtuso simulations of the 4T voltage divider based PUF bit cell have been performed. TSMC 180nm technology was used, with the transistor size and flavors as shown in Fig.3.5. The mid-threshold voltage devices known as (MVT) were used for this circuit, because these transistors allow good compensation in terms of the $\sigma V_{TH01,2}/\lambda_{D1,2}$ ratio.



Transistor	*Type*	*W/L [um/um]*
M1, M2	PMOS-MVT	5/0.5
M3, M4	NMOS-MVT	5/0.5

Figure 3.5: PUF bitcell with 4T-core sizing.

Fig.3.6 shows the distribution of VX in the 4T cell implementation, this result was obtained from 5000 Monte Carlo simulations with nominal bias V_{DD} = 1.8V and at 25ºC temperature. From this image we can clearly see the probability of having “0”, logic “1” and unstable bits. The unstable bits refer to the range of voltage values that fall within the unstable region of the inverter (Alvarez, 2016). This can best be seen in Fig.3.7, where the range of unstable values fall in the width of the noise margin.

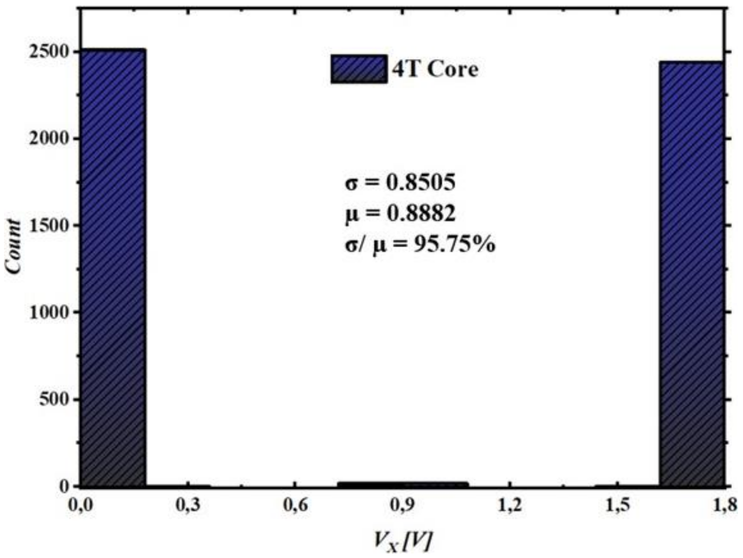


Figure 3.6: Statistical distribution of the voltage V_X of the bitcell core.

The results presented in Fig.3.6 show a good performance of the 4T base cell. However, this simulation was performed with transistors under typical conditions. To understand if the voltage divider is robust it is necessary to perform a PVT analysis, the acronym refers to process, voltage and temperature. The process analysis has to do with the variations of the manufacturing process, because not all transistors behave the same. This is mainly due to the lithography process and implantation of dopants in the devices, reason why there are typical, fast and slow transistors. This is the nomenclature with which the corners are going to be known, referring to the fact that the fast and slow corners have higher or lower carrier mobilities than the normal ones respectively. Since we are using CMOS technology it is necessary to generate a nomenclature that depends on two letters to understand which transistor is operating in a specific corner. Therefore the first letter refers to NMOS devices, the second letter specifies PMOS devices. To better understand this we propose the following example, Fig.3.6 shows the results of a TT simulation, referring that the NMOS transistor operates in a typical way and the PMOS as typical.

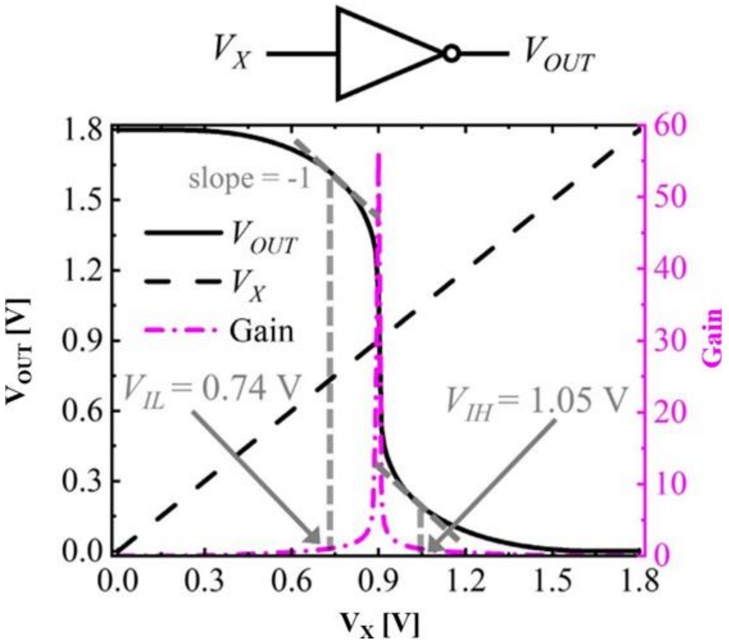


Figure 3.7: Nominal input–output characteristics of the inverter.

For this reason below in Fig.3.8. the results of 5000 Monte Carlo simulations at 25ºC are shown. Showing the effect of varying the different corners on the transistors, in order to verify how robust the cell is when the devices are at FF , SS, FS and SF.

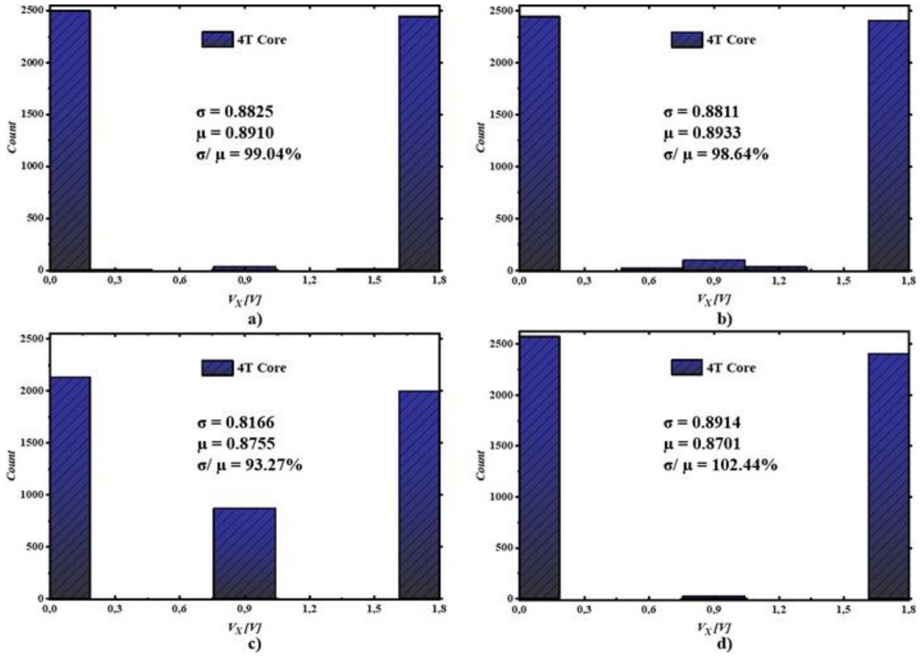


Figure 3.8: Statistical distribution of the voltage VX of the bitcell core a) FF , b) SS, c) FS, d) SF corners.

From these results we can say that the PUF base cell is robust to manufacturing process variation. Since it presents good results for most of the corners. However, we notice that when the NMOS transistors work fast (F) and the PMOS transistors work slow (S), the voltage divider increases the amount of values entering the instability range. On the other hand, the best operating corner is with the transistors in FF.

Finally, one of the most important parameters is the physical implementation of the cell. Therefore, the physical design was carried out using the layout creation tool in Cadence, taking into account all the design considerations for 180nm technology.

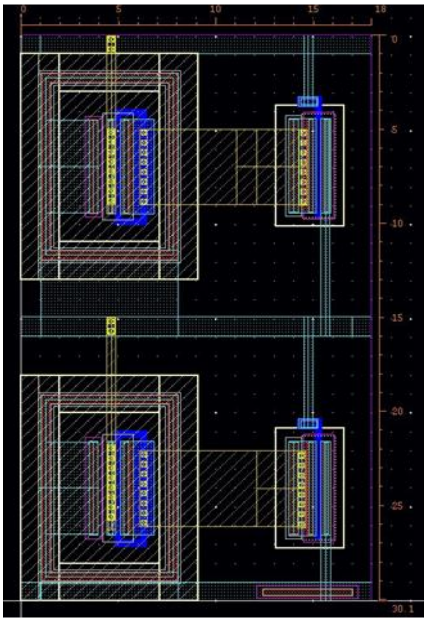


Figure 3.9: PUF bitcell with 4T-core layout.

Fig.3.9 presents the development of the 4T voltage divider cell at the physical level. It is necessary to mention that all transistors have their respective connections, what makes the development of the cell difficult is the bulk connection of the M 2, M 3 and M 4 transistors. Of these transistors the NMOS

Delay Base PUF

Ring oscillators (ROs) are among the first cells used for PUF development, especially for hardware signature generation. RO PUFs are one of the most popular architectures for security circuit development in the literature. However, due to the nature of the circuit the RO is usually electronically noisy. The generation of robust results is one of the main problems of this type of architectures. The objective of modifying this PUF cell is to maximize robustness, since no analytical solutions have been developed to solve this problem. What is proposed is to increase the spread of frequency values that can be generated in an oscillator due to the variation of manufacturing processes, thus achieving a simpler comparison between the frequency values. The goal of this thesis is to design an oscillator circuit which works in subthreshold regime to achieve a higher variability ensuring at the same time a good robustness under environmental variations (i.e., voltage and temperature variations). To do this it is necessary to understand the operation of the oscillator and the sizing that was given to the cells that make up this structure.

In the operation of a ring oscillator the most important parameter is the gate delay. In any physical device, it is impossible to achieve instantaneous switching as long as there is charge and discharge time. For inverters designed with CMOS, the gate capacitance must be charged before current can flow from source to drain. For this reason in a ring oscillator the inverters are in a constant state change depending on the charge and discharge time of the gates. Therefore it is easy to understand that by increasing the number of inverters in the chain the ring oscillator increases the total delay causing the oscillation frequency to decrease.

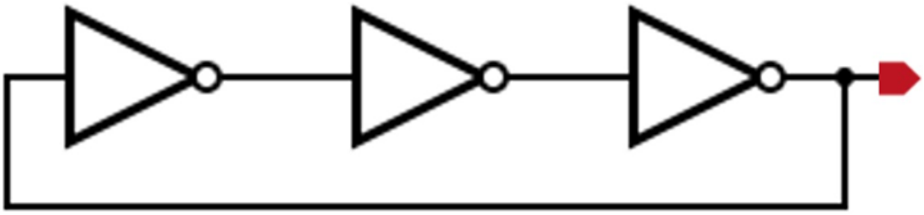


Figure 3.10: Ring Oscillator bitcell.

This ring oscillator belongs to the family of oscillators that depends on the delay of the base cell. This oscillator is composed of inverters that work as an amplifier, in addition this gate is the one that adds the delay to the input signal. The ring oscillator uses an odd number of inverters as in Fig.3.10 in order to give the effect of an inverting amplifier with gain greater than one. Instead of having a delay element, all the inverters are placed in a ring that contributes to the propagation of the signal, for this reason this structure has the name of ring oscillator. It is necessary to consider that adding pairs of inverters increases the total delay, causing the oscillation frequency to decrease. Also the change in the bias voltage in the inverters modifies the delay of the inverters, thus also modifying the frequency of the oscillator. Below is the equation for the oscillation frequency of a ring oscillator, where n is the number of inverters and t is the propagation delay of an inverter.

$$f = \frac{1}{2tn} \tag{3.19}$$

Inverter

The inverter is the base cell for the design of digital circuits and has a great importance for the development of integrated circuits. For the development of this thesis it also has a very high importance, since it is the main gate of the ring oscillator. Next, a brief explanation of the main parameters for the design of this base gate will be given. Specifically for the development of the PUF considering that we want a good relation between variability and robustness.

The voltage transfer characteristic (VTC) of an inverter provides the necessary information about the most important parameters of the gate. Mainly we could highlight the efficiency, noise margin and threshold voltage. What stands out about the VTC in an inverter is the narrow transition zone observed in the voltage at the output of the gate. Due to the high gain at the switching transition, mainly when the NMOS and PMOS transistor are in saturation. When the gate is in this region of operation, any small variation at the input can be a large change at the output.

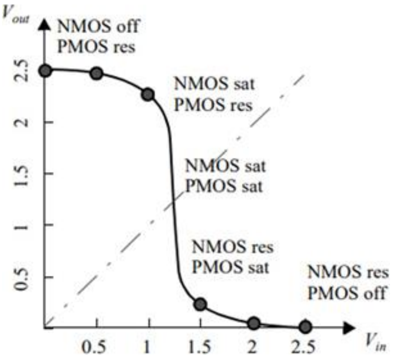


Figure 3.11: Voltage transfer curve of inverter $V_{DD} = 2.5$[V].

The switching threshold (V_M) can be defined as the value where $V_{IN} = V_{OUT}$. This data can be obtained graphically as presented in Fig.3.11 extrapolating the data from the intersection of $V_{IN} = V_{OUT}$ with the V_{TC} characteristic. However, it can also be obtained analytically, by analyzing the operating point where both the NMOS and PMOS transistor are in saturation. This can be better understood in the expression shown below, it is necessary to mention that the channel modulation effect is not being considered.

$$V_M = \frac{\left(V_{THn} + \frac{V_{DSn}}{2}\right) + r\left(V_{DD} + V_{THp} + \frac{V_{DSp}}{2}\right)}{1 + r} \tag{3.20}$$

$$r = \frac{k_p V_{DSp}}{k_n V_{DSn}} \tag{3.21}$$

In order to move the threshold voltage V_M you need to have more control over the equation 3.21. That is, if you need to move the value of V_M up, you must design the PMOS transistor wider. On the other hand, if we want V_M to be close to the value of gnd, the NMOS transistor must have a larger dimension. If we take this consideration into account, it is possible to derive an expression that allows us to determine the sizes to be used in the inverter transistors. Below is an expression showing the ratio of the transistors to place the V_M value where desired.

$$\frac{\left(\frac{W}{L}\right)_p}{\left(\frac{W}{L}\right)_n} = \frac{k_n' V_{DSn}\left(V_M - V_{THn} - \frac{V_{DSn}}{2}\right)}{k_p' V_{DSp}\left(V_{DD} - V_M - V_{THp} - \frac{V_{DSp}}{2}\right)} \tag{3.22}$$

When designing these gates, it is necessary to take into consideration the wide range of temperatures in which the inverter must work without affecting its response. However, due to the characteristics of the inverter current, they do not result in a significant variation with temperature change. This is demonstrated with the equations 3.20 to determine the threshold voltage V_M for switching, showing that the sizing of the transistors does not significantly affect this value. Fig.3.12 shows what happens to the VTC characteristics when the transistors do not have a nominal behavior. In this image we can observe the effects of having the case where the NMOS device is better than the PMOS, also the scenario where the NMOS transistor is worse than the PMOS. What can be observed from this comparison is that there is a shift in the switching threshold, however, the gate operation is not affected. Therefore, it can be stated that the performance of this gate is robust despite the high variability in the manufacturing process and over a wide temperature range.

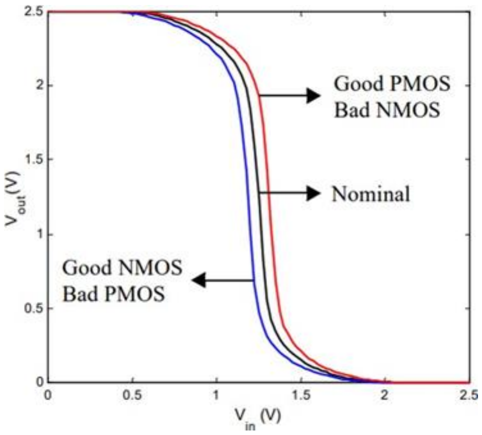


Figure 3.12: Device variations voltage transfer curve of inverter V_{DD} = 2.5[V].

Another parameter that we must take into account in the development of inverters is what happens with their behavior in the face of voltage scaling. The equation 3.23 allows us to say that the inverter gain in the transition region increases with voltage reduction. It is easy to understand that the relationship that exists in transistors makes the switching threshold voltage proportional to the bias voltage of the circuit. Fig.3.13, shows the behavior of the VTC characteristic to the supply voltage variation. Indicating that the inverter behavior is correct even with bias values with voltages below the threshold voltage.

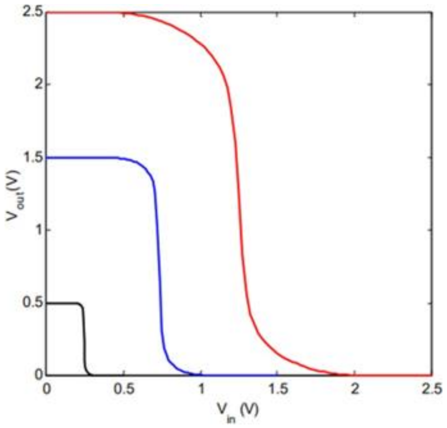
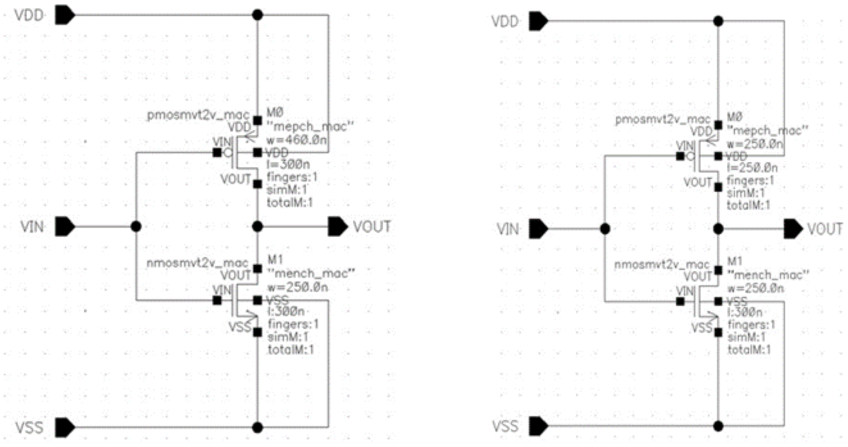


Figure 3.13: Voltage variations voltage transfer curve of inverter V_{DD} = 2.5[V].

$$\frac{\delta V_{OUT}}{\delta V_{IN}} = -\frac{1}{I_D(V_M)}\frac{k_n V_{DSn} + k_p V_{DSp}}{\lambda_n - \lambda_p} \tag{3.23}$$

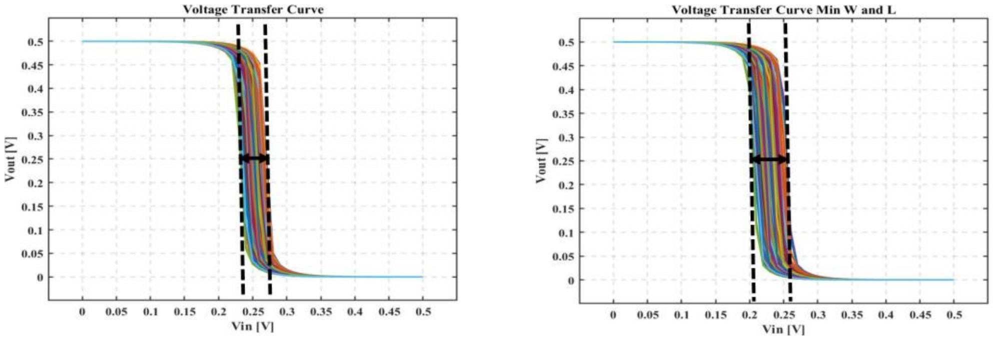
In order to develop the ring oscillator, two inverters with two different sizes were characterized. This with the purpose of verifying with which sizing the inverter acquires greater variability in its response. Fig.3.14 shows the schematic of the inverters in the virtuoso TCAD, together with their sizing. Two sizing were used here, the first one leads to a V_M equal to $V_{DD}/2$ and the second one is a minimum sizing for the NMOS and PMOS transistors.



a) **b)**

Figure 3.14: Inverter gate schematic: a) transistor size $V_M = V_{DD}$, b) minimum transistor size NMOS and PMOS.

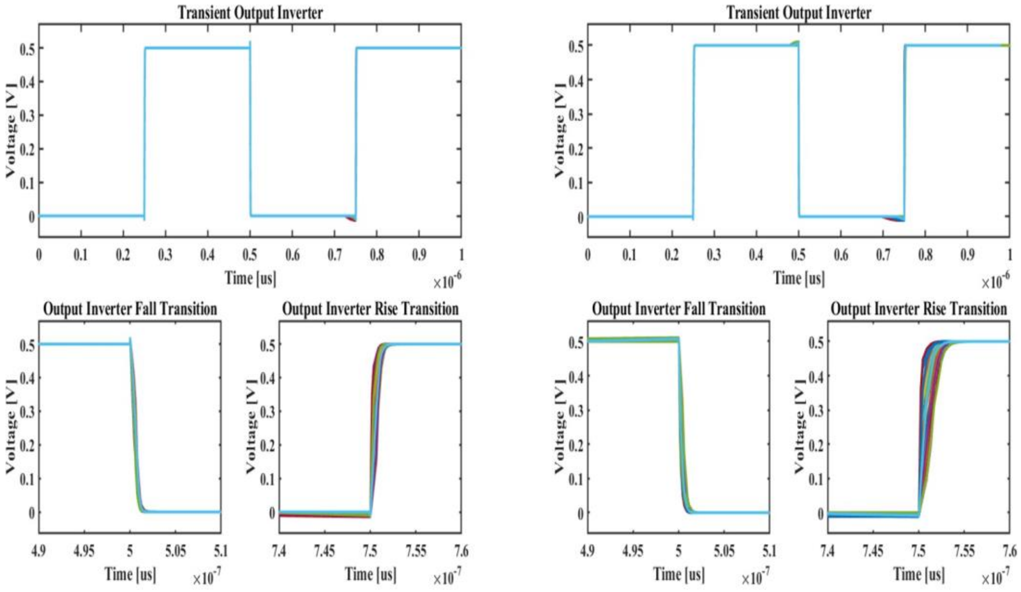
To continuing with the analysis, 1000 points of Monte Carlo simulations were performed for each inverter, configuring the simulation with the mismatches of the manufacturing process. Also the polarization used in the devices is 0.5[V], in order to operate in the subthreshold region increasing the variability of the response and decreasing the energy consumption of the cell.



a) b)

Figure 3.15: Inverter VTC: a) transistor size VM = VDD , b) minimum transistor size NMOS and PMOS.

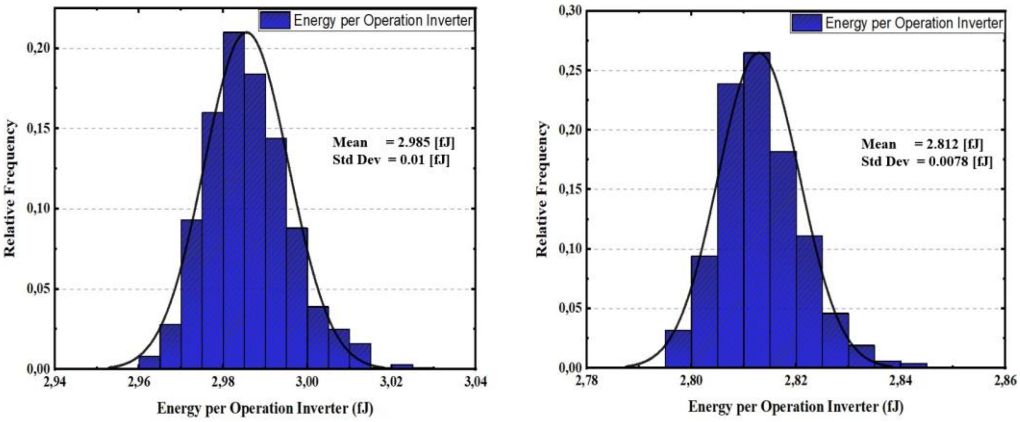
Fig.3.15. shows the result of the DC analysis of the inverter, the main graph that can be analyzed is the VTC characteristic. What is interesting to mention from these results, is that when the inverter works with the minimum sizing of the transistors the variability caused by the manufacturing processes increases, this effect is of great benefit for the development of the PUF cell.



a) b)

Figure 3.16: Inverter transient response: a) transistor size VM = VDD , b) minimum transistor size NMOS and PMOS.

In Fig.3.16 we have instead the time response, which is important to analyze since it allows us to understand the behavior in response to the change of the input signal. The first thing we can observe is that, in both inverters, the response obtained is robust and reaches the desired logic values. However, the response based on the minimum sized transistors shows a higher variability in the signal rise delay. This condition is beneficial for PUF cells, since a high variability in the rise signal delay will correspond to a higher variability in the oscillation frequency and hence in a simpler comparison between two different ring oscillators.



a) b)

Figure 3.17: Inverter energy per operation: a) transistor size VM = VDD , b) minimum transistor size NMOS and PMOS

In addition, another point that was considered in the analysis is the energy consumption of the inverters. This analysis was carried out as shown in Fig.3.17. The expected result is that the inverter with minimum sizing has a lower energy consumption with the cost of higher delay.

As this design must have a physical implementation. Fig.3.18 shows the layout of the inverter with minimum transistor dimensions. Minimum sizing were used for optimizing the variability of the manufacturing process. It is necessary to mention that the design presented considers minimum distances and sizes in the connections and spaces between materials. The entire development of the cell takes into account the design rules of TSMC 180nm technology.

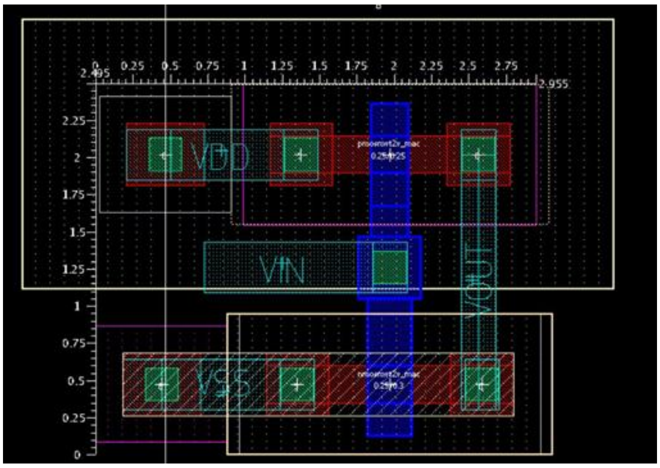


Figure 3.18: Layout inverter minimum transistor size.

Ring Oscillator

The ring oscillator is the structure in responsible of transforming the variability of the manufacturing process into a measurable electrical quantity. In this particular case this electrical quantity is the delay, however, for ease and robustness it is preferred to use the frequency response of this circuit. In order to measure the frequency, what is done is to count how many oscillations there are in a predetermined period. It must be taken into account that the classical oscillator is composed of an odd number of inverters that have a negative feedback as shown in Fig.3.10.

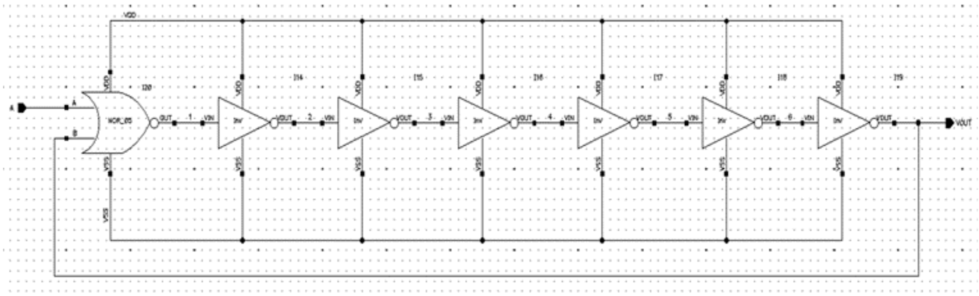


Figure 3.19: Proposed RO schematic.

Fig.3.19 presents the ring oscillator scheme that was implemented for the development of the PUF base cell. From this scheme we can highlight that it is a chain of gates composed by a NOR at the beginning followed by six inverters. The NOR combinational circuit is used in order to receive an input signal, which allows the circuit to start oscillating for a given period. This architecture can oscillate because the NOR behaves like an inverter when it has the correct logic signals at the input. Due to the behavior of the NOR it is possible to say that when we have the logic signal that allows the circuit to oscillate, we have seven inverters in chain therefore this requirement is fulfilled. In addition of having the feedback of the last inverter to one of the inputs of the NOR. Seven stages were chosen because of the trade off between silicon area, operating frequency and power consumption.

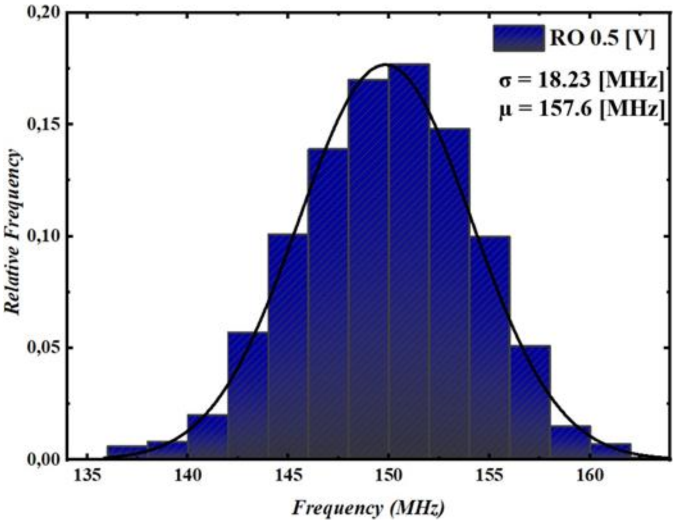


Figure 3.20: Simulated distributions of the frequency of proposed RO-base.

In Fig.3.20 we can observe the result of 1000 Monte Carlo simulations. The conditions used were a bias voltage of 0.5[V] at a temperature of 25ºC. With this plot it is possible to understand the wide variability of frequencies that can be obtained from an oscillator due to the manufacturing process. Showing that the ring oscillator with inverters having transistors with minimum dimensions and operating a subthreshold has a good performance for a PUF cell base.

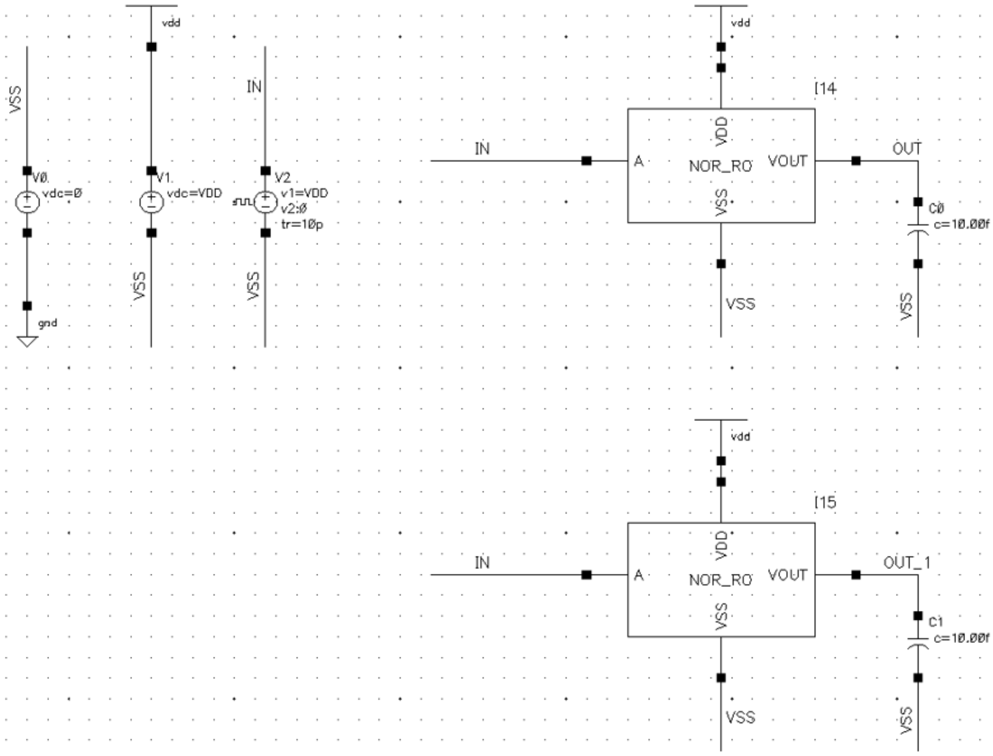


Figure 3.21: Schematic Test-Bench of RO PUF-base.

Once we have observed that the designed oscillator has a good performance and generates a wide range of frequencies, we proceed to develop the PUF. It is necessary to understand that this type of cells needs two identical oscillators, which are used to generate the frequencies that will be compared in the future. Using this concept, a simulation test was designed to measure the operating frequency of two identical ring oscillators under the same conditions. This can be clearly seen in Fig.3.21, where we can notice that the two devices have the same polarization, temperature, oscillation signal and load.

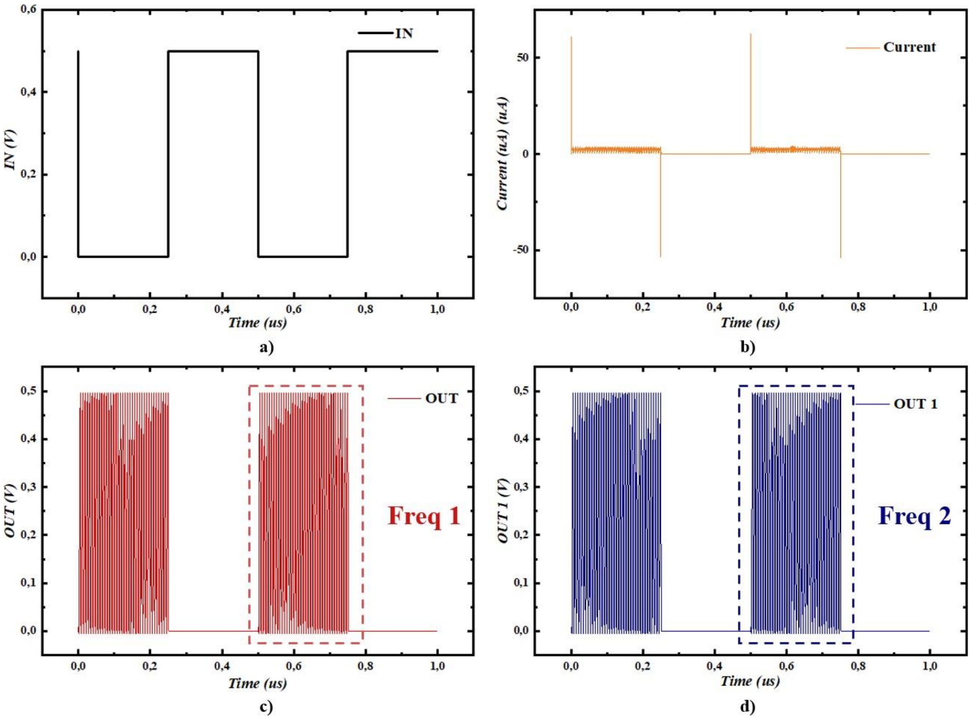
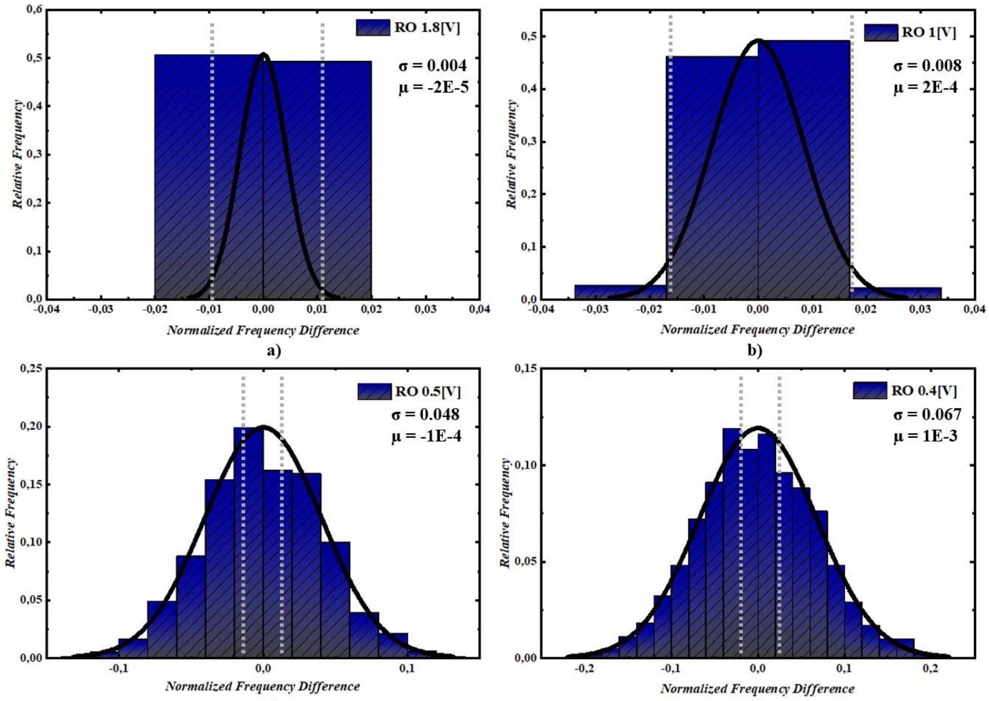


Figure 3.22: Time response of a) input signal, enables oscillating circuits, b) current consumption of the oscillators, c) response of the first oscillator and d) response of the second oscillator.

Then, to explain the operation of the designed simulation, Fig.3.22 is presented. Where we can observe the behavior of the ring oscillators before an input signal, understanding that when this has the value of “0” logic the rings starts to oscillate. Then, using a period and counting the number of oscillations, the frequencies of the ring oscillators are extracted. It is also necessary to mention that at the moment of oscillation the highest current consumption of the cell occurs. That is why the implementation of the NOR gate at the beginning helps to improve the consumption, since the circuit starts working only when it is needed.



a) b)

c) d)

Figure 3.23: Normalized frequency difference with different bias voltages a) 1.8[V], b) [V], c)0.5[V] and d)0.4[V].

Voltage [V]	**1,8**	**1**	**0,5**	**0,4**
Flipping Bits [%]	0	0,1	0	1,2
Unstable Noisy Bits [%]	93,1	79,8	30,4	18,6
Unstable Bits [%]	93,1	79,9	30,4	19,8

Table 3.1: Unstable bits with different bias voltages a) 1.8[V], b) [V], c)0.5[V] and d)0.4[V].

Once we have an analysis scheme for the base of the PUF cell, we are interested in studying the variability of the difference between the two generated frequencies. Therefore, 1000 Monte Carlo simulations were performed under different conditions to analyze the behavior of the cell at different voltages and temperatures. Fig.3.23 shows the results when varying the polarization voltage, noting that when the cell operates in the subthreshold region the variability increases. If we compare the standard deviation of the response 1.8V and 0.5V we can observe a change from σ = 0.0043 to σ = 0.048 respectively. Therefore we can conclude that by working in subthreshold region the standard deviation increases of 91.6% when comparing the responses at 1.8V and 0.5V. This is mainly due to the fact that in subthreshold region the relationship between the current and the threshold voltage (i.e., one of the most affected parameters by process variations) is exponential.

Robustness under temperature variations was subsequently analyzed considering a range from 0−75ºC which represents a typical operating range found in literature. Fig.3.24 shows the frequency difference of the two identical oscillators working at a bias voltage of 0.5[V].

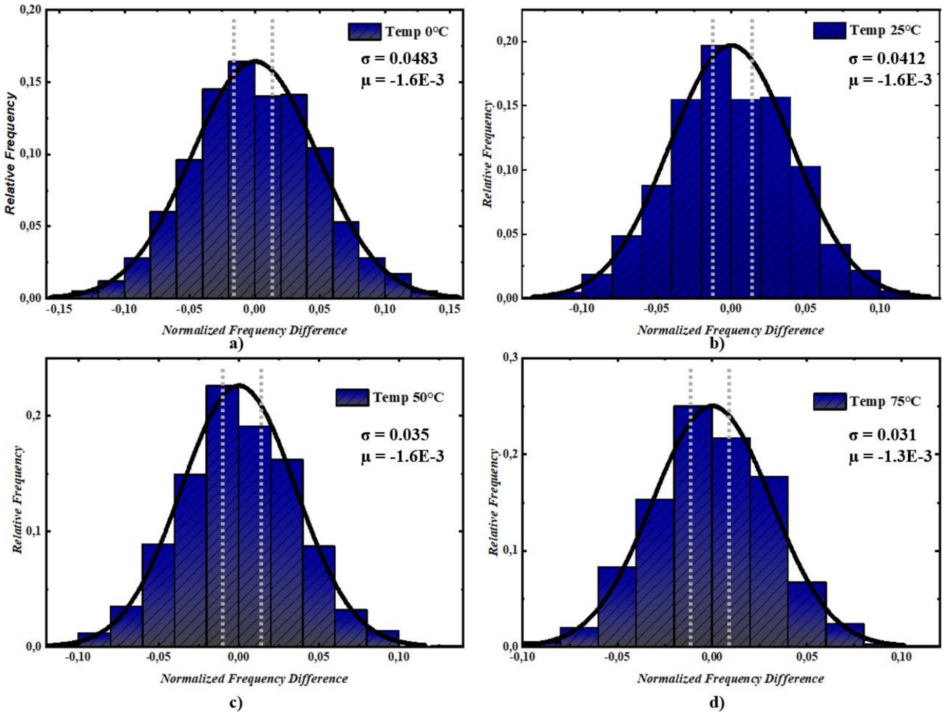


Figure 3.24: Normalized frequency difference with bias voltage of 0.5[V] at different temperatures a) 0ºC, b) 25ºC, c) 50ºC and d) 75ºC.

Temperature [°C]	0	25	50	75
Flipping Bits [%]	0,7	0	0,9	1,9
Unstable Noisy Bits [%]	32,5	37,2	40,3	46,7
Unstable Bits [%]	33,2	37,2	41,2	48,6

Table 3.2: Unstable bits with bias voltage of 0.5[V] at different temperatures a) 0°C, b) 25°C, c) 50°C and d) 75°C.

The physical implementation of the ring oscillator was also developed. Since we can have complete control of the design as we are using the full custom methodology. A perfectly symmetrical cell was developed, where the gate spacing is the same and the minimum dimensions are based on the design rules of the 180 nm technology. Fig.3.25 shows the cell as seen in a physical implementation, it is necessary to note that the silicon area is 75.5um^2.

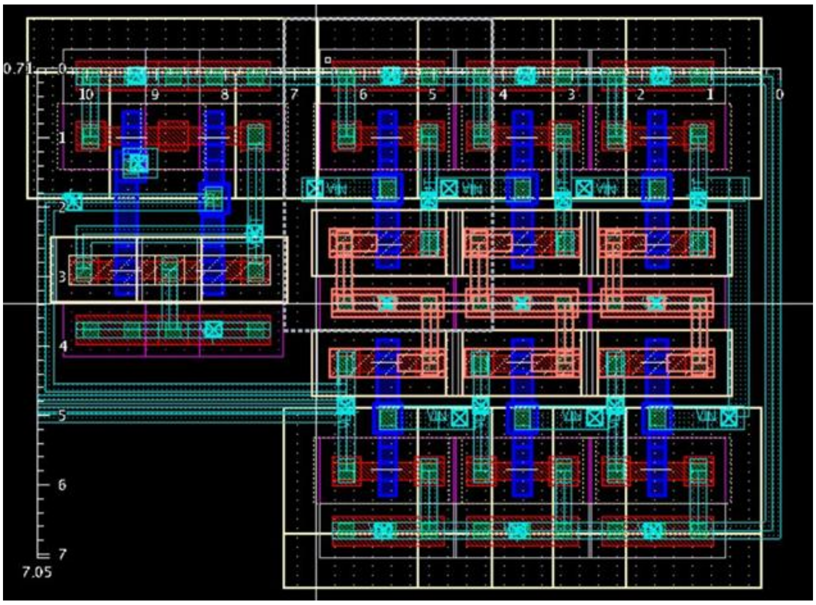


Figure 3.25: Layout ring oscillator.

Calibrated Ring Oscillator

The conventional ring oscillator based PUF structure is capable of generating "1" bit output responses. This is achieved by comparing the frequency difference in a pair of RO. However, if the frequency difference in the pair of RO is small, the bit resulting from the comparator can be inverted due to noise conditions or environmental variations as well illustrated in Fig.3.26. From this picture we can understand that variations in the environmental conditions may cause a failure in the cell response by inverting the mismatch (i.e., the VT H difference, for example, could be inverted under different environmental con- ditions). This is the reason why in the literature the idea of using configurable ROs is proposed, with the objective of obtaining a pair of ROs that have a larger difference in their frequencies, improving the reliability of the PUF.

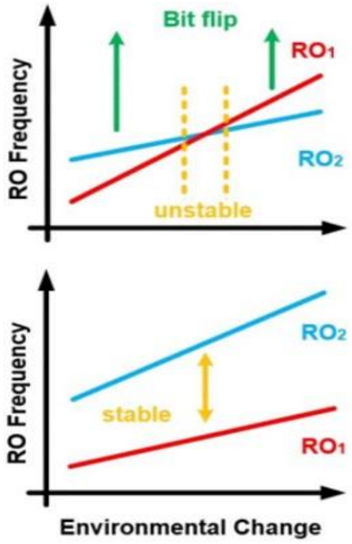


Figure 3.26: Stability enhancement of an RO-based PUF.

According to what said above a delay cell topology using a mismatch enhancement technique is proposed to amplify the variability and improve the stability of the PUF. Various design techniques can be used to adjust the signal delay, such as sizing the transistors or using different logic families. Fig.3.27 presents the proposed ring oscillator composed of a NOR gate, six inverters and six PMOS pass transistors placed between the inverters.

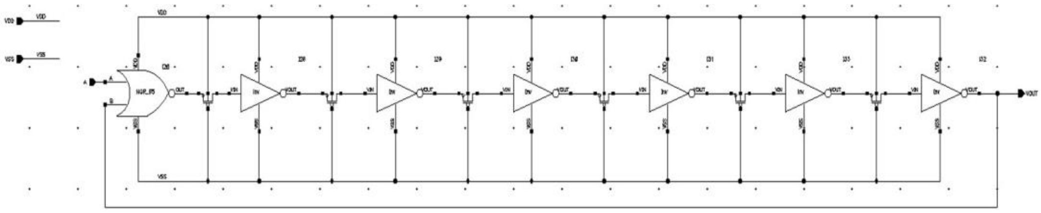


Figure 3.27: Schematic of configurable RO.

Such basic idea was previously proposed in 3.27 in which a NMOS pass transistor were used. To understand the operation of the PMOS pass transistor, we have Fig.3.28 showing the effect of this device on the signal. Because the PMOS pass transistor is unable to deliver a low voltage signal X, its output Y exhibits a slow transition when Y node must be discharged from V_{TH} to 0 as well illustrated in Fig.3.28. The uncertainty of the transition time due to variability in the devices causes the Z signal that is the response of an inverter connected to the pass transistor to be delayed. The sizing chosen for the pass transistor is the minimum size allowed by the technology, mainly to continue with the reasoning that with these devices there is more variability.

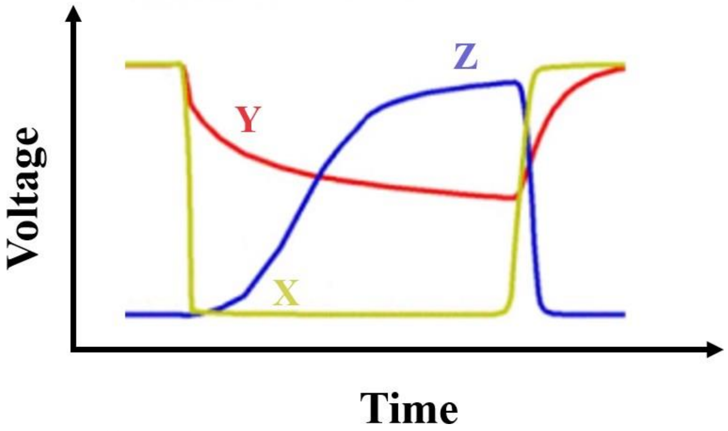


Figure 3.28: Effect of the PMOS pass transistor between two inverters.

With the proposed ring oscillator we proceed to perform the same analysis of the previous section. Doing 1000 Monte Carlo simulations analyzing the behavior of the PUF base cell under voltage and temperature variation. Fig.3.29 shows the frequency difference between two identical ring oscillators, varying the bias voltage. We can observe a similar effect to the conventional ring oscillator, showing that when operating in the subthreshold region the frequency deviation increases.

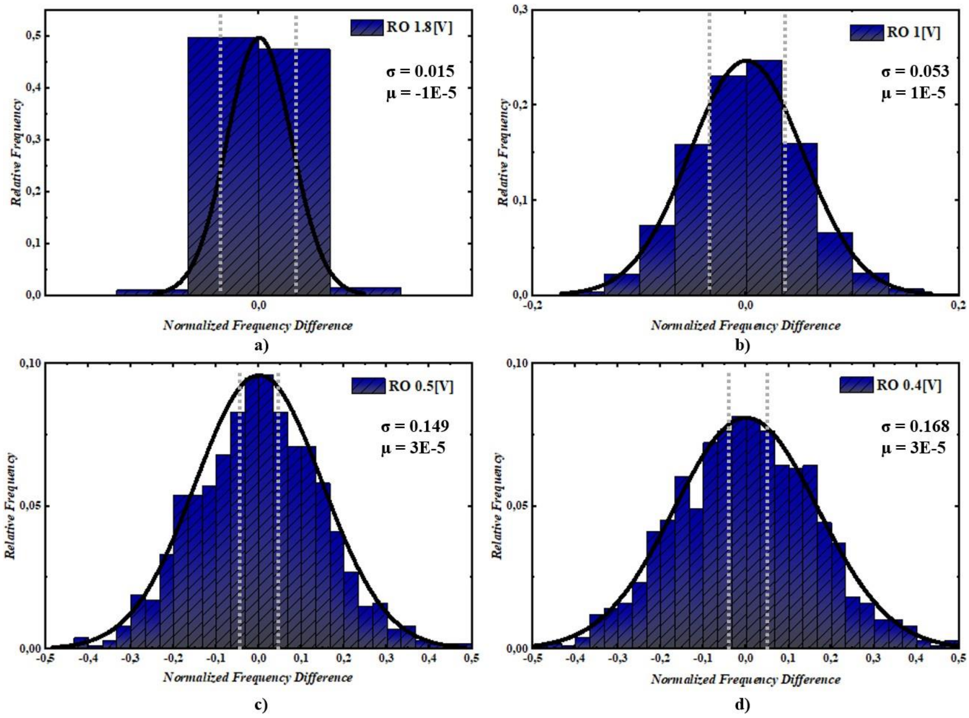


Figure 3.29: Normalized frequency difference with respect to the nominal frequency with different bias voltages a) 1.8[V], b) 1[V], c) 0.5[V] and d) 0.4[V].

Voltage [V]	**1,8**	**1**	**0,5**	**0,4**
Flipping Bits [%]	0,5	0,6	0	0,9
Unstable Noisy Bits [%]	85,6	53,9	14,8	9,3
Unstable Bits [%]	86,1	54,5	14,8	10,2

Table 3.3: Unstable bits with different bias voltages a) 1.8[V], b) [V], c)0.5[V] and d)0.4[V].

To analyze the behavior of the base cell using the proposed oscillator, the temperature was varied from 0 – 75ºC using a fixed bias voltage of 0.5[V]. In order to verify how changing this parameter affects the variability of the frequency difference in the pair of oscillators. Fig.3.30 allows us to understand that the cell behaves without modifying the values as much when the temperature is changed. However, we notice an effect in the reduction of the standard deviation if we compare the results at 0ºC and 75ºC.

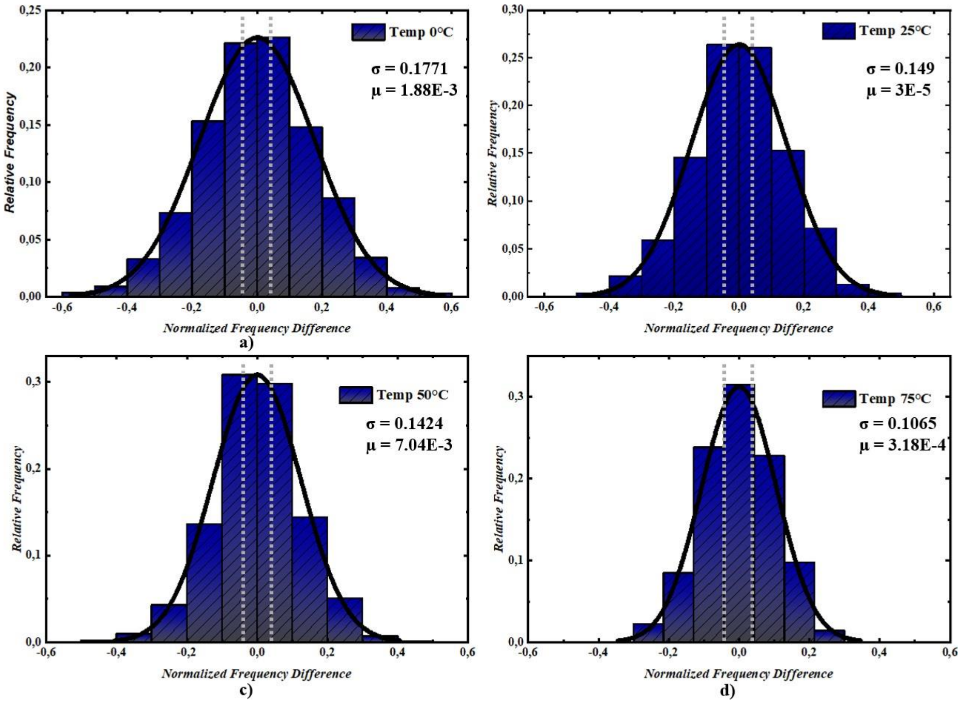


Figure 3.30: Normalized frequency difference with respect to the nominal frequency with bias voltage of 0.5[V] at different temperatures a) 0ºC, b) 25ºC, c) 50ºC and d)75ºC.

Temperature [°C]	**0**	**25**	**50**	**75**
Flipping Bits [%]	1,4	0	1,3	2,1
Unstable Noisy Bits [%]	28,2	29,5	31,4	31,8
Unstable Bits [%]	29,6	29,5	32,7	33,9

Table 3.4: Unstable bits with bias voltage of 0.5[V] at different temperatures a) 0ºC, b) 25ºC, c) 50ºC and d) 75ºC.

In order to understand the benefit of using pass transistors between the inverters, the two oscillators were compared. To make this comparison we used data extracted from the distributions showing the difference in oscillator frequencies. Fig.3.31 presents two trend curves, in black is the distribution of the classical ring oscillator and in red is the proposed configurable oscillator with pass transistors. It is easy to observe that the objective of increasing the variability of the frequency difference is achieved by the proposed architecture.

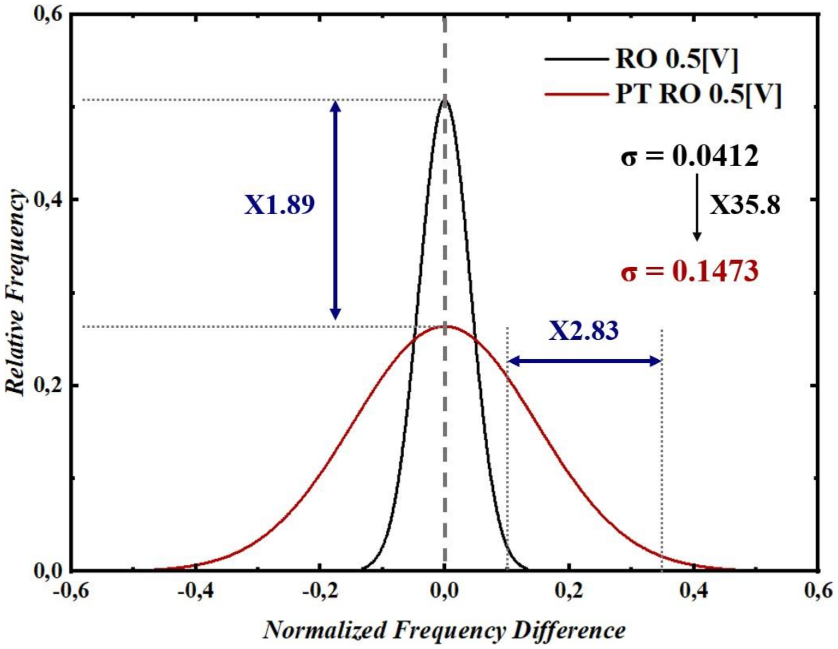


Figure 3.31: Comparison of Ring Oscillator and Calibrated Ring Oscillator Distributions.

Taking into account that the standard deviation goes from σ = 0.0412 to σ = 0.1473, i.e. it has an increase of X35.8 thus benefiting the PUF cell. It is necessary to mention that this simulation is under the following conditions 0.5[V] bias voltage and 25ºC temperature.

The physical implementation is very important, for this reason it was also designed. We must consider that this oscillator has more complications because we have the pass transistor. However, being a PMOS type, the bulk polarization is not complicated to manage at the physical level. The physical scheme can be seen in Fig.3.32, it must be taken into account that it was designed with the minimum dimensions accepted by the design rules of the technology. Obtaining an oscillator cell with an area of 113.50um^2, which if we compare it with the normal oscillator we have an increase in area of X1.5.

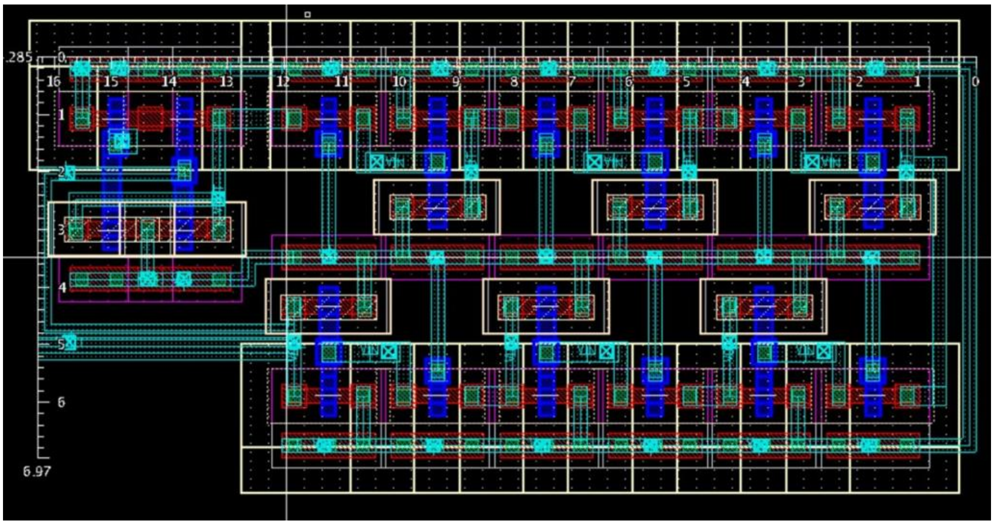


Figure 3.32: Layout calibrated ring oscillator.

Calibrated ring oscillator collapse-based comparator

A PUF is composed of two parts, the element in charge of extracting a measurable electrical quantity that depends on the variation of the manufacturing process and the element in charge of converting this measurement into a binary response. At the moment we have the proposal of a calibratable ring oscillator, however, it is necessary to have the frequency comparison stage of the oscillator pair. Since we are looking to realize PUFs that have a low power consumption, it is necessary to design a low power comparator. This component must be able to automatically scale its conversion power based on the difference of the input signal. It is necessary to take into account that the higher the resolution the energy efficiency of comparators tends to be poor. Mainly due to the noise introduced by the comparator, wasting energy in processes that have a simple response.

For this work we used a comparator based on the collapse of the ring oscillator, which in the literature is known as edge pursuit comparator (EPC). It is capable of automatically scaling the comparator energy, depending on the difference of the input signals without the need for external control. It is also efficient in the aspect of being able to adapt its energy consumption depending on each conversion.

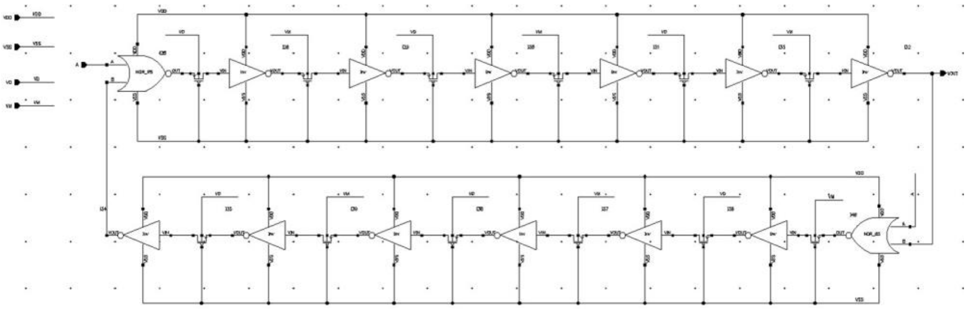


Figure 3.33: Calibrated ring oscillator collapse-based comparator.

In Fig.3.33 we have the schematic of the complete PUF structure, taking into account the comparator using the EPC. This schematic consists of two NOR gates, twelve inverters and twelve pass transistors. The comparator initiates a comparison when the IN signal takes the value of logic “0” in the two NOR gates. This generates two propagating edges in the oscillator, traveling around the comparator until one overtakes the other collapsing the oscillation. We have two input signals V_M and V_D, which are applied consecutively in the bulk of the pass transistors. Causing a change in the delay of the pass transistors, thereby modulating the propagation edges. The propagation delay is controlled by the voltage applied, i.e. V_M causes the high to low transition to be fast, the low to high transition to be slow, and the voltage on V_D does the opposite.

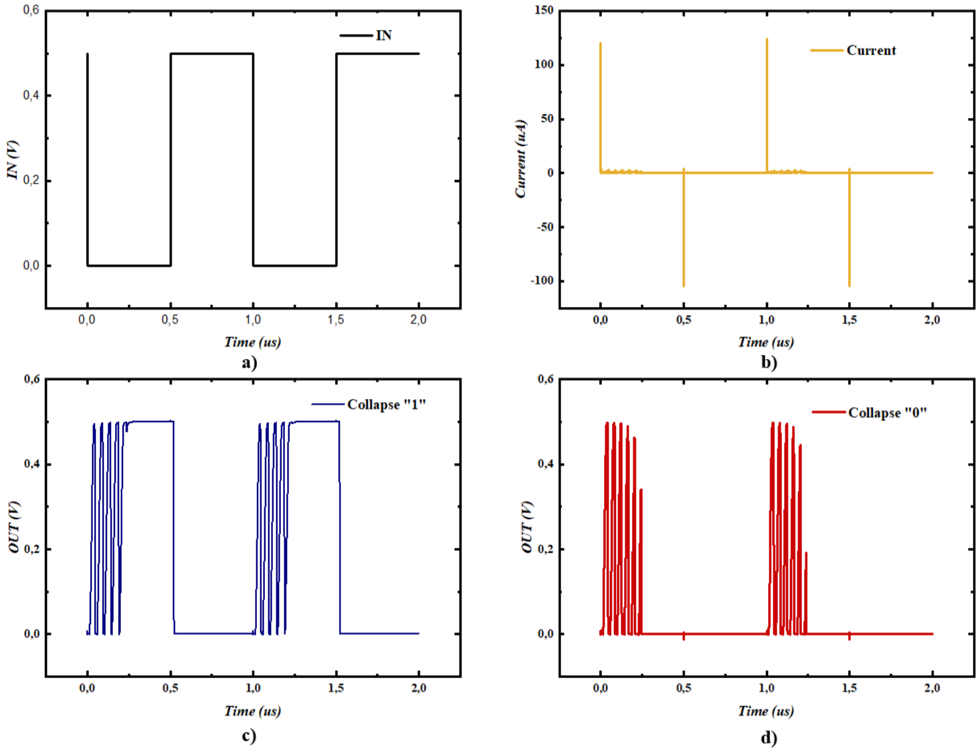


Figure 3.34: Time response of a) input signal, enable, b) current consumption of the PUF, c) response in the case that collapses OUT to V_{DD} and d) response in the case that collapses OUT to V_{SS}.

When one propagating edge exceeds the other, the comparator causes the oscillation to collapse and the output OUT is placed at the voltage of V_{DD} or V_{SS}. This depends on which bode has been slower, Fig.3.34 shows the two collapse cases for both logic “0” and “1”. What was sought in this work is that the voltage difference between V_M and V_D is zero, causing the propagation delay to be similar at the edges and the difference depends on the fabrication processes.

To understand if the proposed PUF idea is a solution that can be implemented in future works, 1000 Monte Carlo simulations were performed. Obtaining the distributions where we can clearly observe the behavior of the cell to the variations of the environment. Here we are already going to see a binary response because it is the result of the whole PUF process, the generation of frequencies by the oscillators and the comparison of these using the EPC. To avoid failures in the signal collapse time at the logic value, it was decided to use a period X1.2 as long as the minimum collapse period of the circuit for the configured V M and V D voltages. Fig.3.35 presents the different distributions obtained by varying the bias voltage from 0.4[V] to 1.8[V].

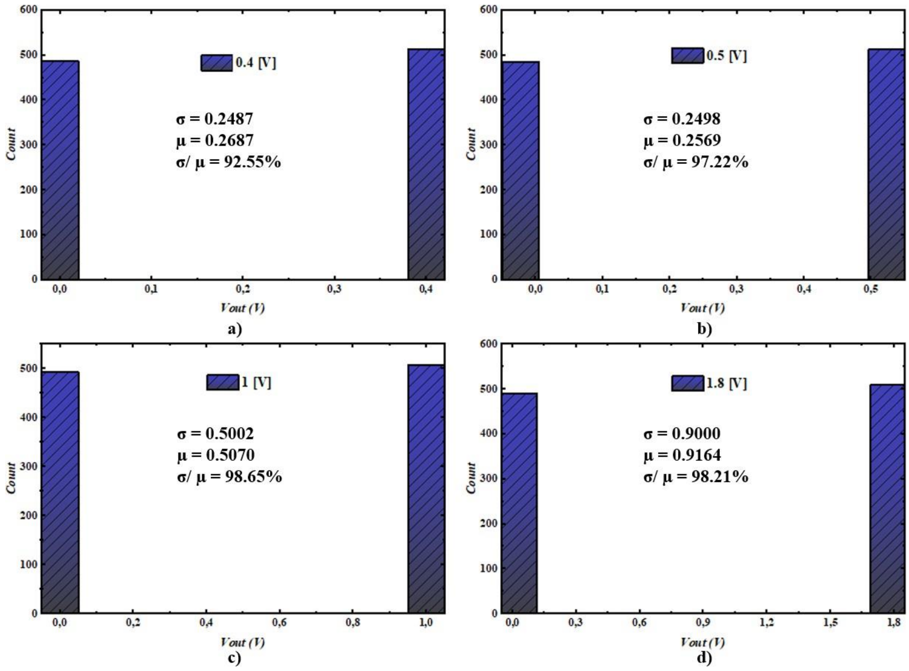


Figure 3.35: PUF response with different bias voltages a) 1.8[V], b) 1[V], c) 0.5[V] and d) 0.4[V].

Voltage [V]	**1,8**	**1**	**0,5**	**0,4**
Flipping Bits [%]	25,3	9,7	0	1,6
Unstable Noisy Bits [%]	1,6	1,3	1,1	0,7
Unstable Bits [%]	26,9	11	1,1	2,3

Table 3.5: Unstable bits with different bias voltages a) 1.8[V], b) [V], c)0.5[V] and d)0.4[V].

A temperature variation was also performed to see the trend of the distri- butions. Fig.3.36 presents the effects of varying the temperature from 0ºC to 75ºC, showing that the distributions are affected by the temperature change. This is mainly due to the voltages V_M and V_P since being fixed these modify the time response of the pass transistors. To improve this effect in the future, it is proposed to use biasing based on the temperature change to avoid such a detrimental effect.

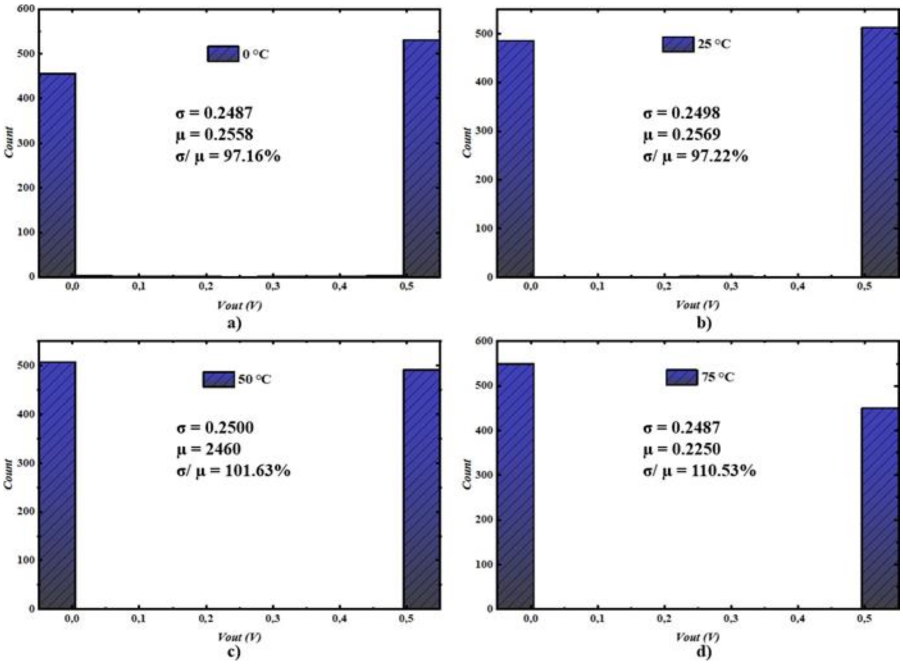


Figure 3.36: PUF response with bias voltage of 0.5[V] at different temperatures a) 0ºC, b) 25ºC, c) 50ºC and d) 75ºC.

Temperature [°C]	**0**	**25**	**50**	**75**
Flipping Bits [%]	5,2	0	2,2	6,4
Unstable Noisy Bits [%]	3,4	1,1	0,3	2,2
Unstable Bits [%]	8,6	1,1	2,5	8,6

Table 3.4: Unstable bits with bias voltage of 0.5[V] at different temperatures a) 0ºC, b) 25ºC, c) 50ºC and d) 75ºC.

In addition to the proposed cell, the physical implementation was carried out, it is necessary to mention that for this particular structure the greatest possible symmetry was sought. In addition to using the minimum dimensions allowed by the design rules allowed by the technology. The total area of the complete cell is 547.53um^2.

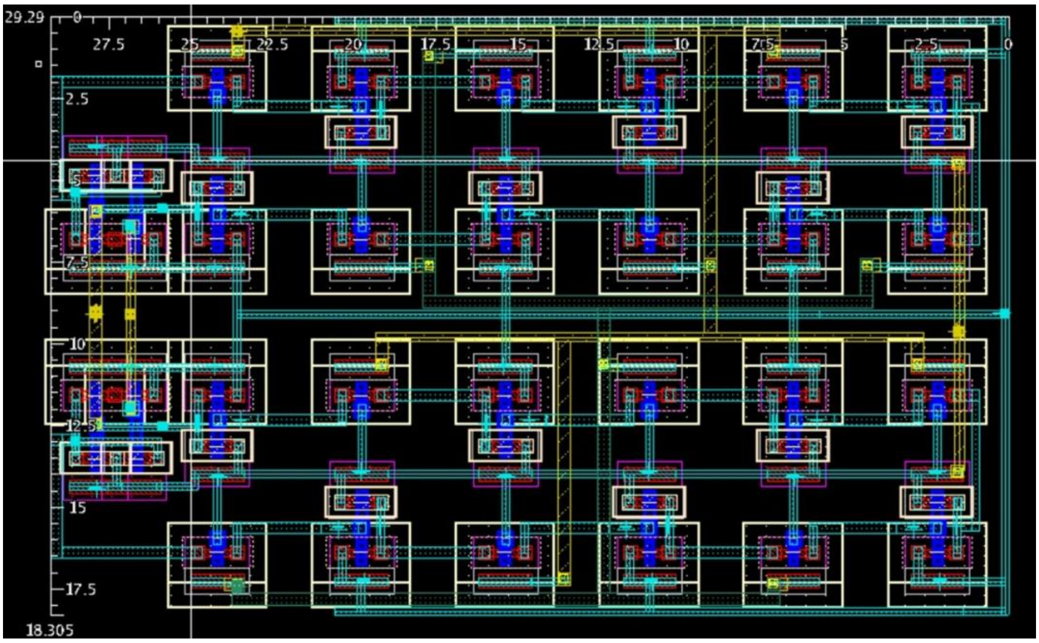


Figure 3.37: Layout of the proposed PUF cell.

CONCLUSIONS

Physical disorder is a phenomenon present in all materials, this can be observed even at nanometer scales in the form of irregular structures of physical objects. The continuous scaling of semiconductors has made it increasingly difficult to precisely control the dimensions of manufactured devices. This is known as manufacturing process variations, these can be considered as manifestations of physical disorder giving unique characteristics to all circuits. PUFs are integrated circuits that take advantage of this phenomenon, in order to generate a unique digital identifier for each circuit. There are several techniques to design a PUF, however, they all share the same design principles. These circuits must consist of two stages, the first stage converts the manufacturing process variation to a measurable electrical quantity (e.g., voltage, current and delay). The second stage aims to transform this measurable electrical quantity into a binary response. Using this as a starting point, two PUF cells were developed, one based on voltage variability and the other using delay.

The first cell uses as a basis the metastability that can be obtained from a voltage divider consisting of two identical circuits. From this solution, an analysis was developed to determine the main physical parameters that affect the output voltage, allowing the cell to work in a metastable region. The benefit of working with this cell is that it provides us with the metastable stage, so we would have a binary response at the output. The circuit was designed both schematically and physically. In addition, a DC response with nominal parameters was obtained, showing that the response of the PUF gives us a good relationship of the distribution of logical “1” and “0”.

This cell has a more experimental parameter, focusing on simulation to determine if the implemented circuit improvements affected the robustness of the PUF. This cell uses as a basis the delay variation in a ring oscillator caused by the manufacturing processes. In this cell, we sought to increase the delay variability by sizing the circuit and working at a subthreshold voltage. Since the main difficulty faced by ring oscillator based cells is the comparison of operating frequencies. A calibratable oscillator was used in order to increase the standard deviation of the frequency difference. What makes this structure interesting is that it uses PMOS pass transistors between the inverters and depending on the voltage applied in the bulk of these devices the frequency of the circuit can be regulated. This architecture gives us the facility to use a comparator that uses the same oscillators as a base, using the edge pursuit comparator EPC technique. Thus showing an improvement in robustness when compared to a classical ring oscillator using a comparator. It is necessary to mention that this cell can be improved in the future if the transistor bulk is calibrated taking into account the temperature variation.

REFERENCES

Alvarez, A. B. (2016). Static physically unclonable functions for secure chip identification with 1.9--5.8\% native bit instability at 0.6--1 V and 15 fJ/bit in 65 nm. *IEEE Journal of Solid-State Circuits*, 763-775.

Chandrakasan, A. a. (2001). Impact of Physical Technology on Architecture. En *Design of High-Performance Microprocessor Circuits* (págs. 2-24). doi:10.1109/9780470544365.ch1

Cobb, W. E. (2012). Intrinsic Physical-Layer Authentication of Integrated Circuits. *IEEE Transactions on Information Forensics and Security*, 14-24. doi:10.1109/TIFS.2011.2160170

Daihyun Lim and Lee, J. a. (2005). Extracting secret keys from integrated circuits. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 1200-1205. doi:10.1109/TVLSI.2005.859470

De Rose, R. a. (2017). A physical unclonable function based on a 2-transistor subthreshold voltage divider. *International Journal of Circuit Theory and Applications*, 260-273.

Dolev, S. a. (2015). Optical PUF for Non Forwardable Vehicle Authentication. En *2015 IEEE 14th International Symposium on Network Computing and Applications* (págs. 204-207). doi:10.1109/NCA.2015.25

Fournel, T. a. (2016). Towards weak optical PUFs by random spectral mixing. En *2016 15th Workshop on Information Optics (WIO)* (págs. 1-3). doi:10.1109/WIO.2016.7745572

Guajardo Jorge, K. S.-J. (2007). FPGA Intrinsic PUFs and Their Use for IP Protection. En *Cryptographic Hardware and Embedded Systems - CHES 2007* (págs. 63-80). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:978-3-540-74735-2

Halak, B. (2018). Physically Unclonable Functions: Design Principles and Evaluation Metrics. En *Physically Unclonable Functions : From Basic Design Principles to Advanced Hardware Security Applications* (págs. 17-52). Cham: Springer International Publishing. doi:10.1007/978-3-319-76804-5_2

Halak, B. a. (2008). Fault-Tolerant Techniques to Minimize the Impact of Crosstalk on Phase Encoded Communication Channels. *IEEE Transactions on Computers*, 505-519. doi:10.1109/TC.2007.70825

Hoefflinger, B. (2012). ITRS: The International Technology Roadmap for Semiconductors. En B. Hoefflinger, *Chips 2020: A Guide to the Future of Nanoelectronics* (págs. 161-174). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-23096-7_7

Kalyanaraman, M. a. (2013). Novel strong PUF based on nonlinearity of MOSFET subthreshold operation. En *2013 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)* (págs. 13-18). doi:10.1109/HST.2013.6581558

Lin, L. a. (2012). Design and Validation of Arbiter-Based PUFs for Sub-45-nm Low-Power Security Applications. *IEEE Transactions on Information Forensics and Security*, 1394-1403. doi:10.1109/TIFS.2012.2195174

Maiti Abhranil, G. V. (2013). A Systematic Method to Evaluate and Compare the Performance of Physical Unclonable Functions. En *Embedded Systems Design with FPGAs* (págs. 245-267). New York, NY: Springer New York. doi:10.1007/978-1-4614-1362-2_11

McFate, S. N. (2022). Getting Out of the Starting Block. *Scientific American*, 61-67. Obtenido de http://www.jstor.org/stable/24987511

Mispan, M. S. (2015). TCO-PUF: A subthreshold physical unclonable function. En *2015 11th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME)* (págs. 105-108). doi:10.1109/PRIME.2015.7251345

Narasimhan, A. a. (2007). Impact of Variability on Clock Skew in H-tree Clock Networks. *8th International Symposium on Quality Electronic Design (ISQED'07)*, (págs. 458-466). doi:10.1109/ISQED.2007.88

Nassif, S. (2001). Modeling and analysis of manufacturing variations. *Proceedings of the IEEE 2001 Custom Integrated Circuits Conference (Cat. No.01CH37169)*, (págs. 223-228). doi:10.1109/CICC.2001.929760

Nassif, S. a. (2007). High Performance CMOS Variability in the 65nm Regime and Beyond. *2007 IEEE International Electron Devices Meeting*, (págs. 569-571). doi:10.1109/IEDM.2007.4419002

Nawi, I. M. (2016). The influence of hysteresis voltage on single event transients in a 65nm CMOS high speed comparator. En *2016 21th IEEE European Test Symposium (ETS)* (págs. 1-2). doi:10.1109/ETS.2016.7519300

Reising, D. R. (2015). Authorized and Rogue Device Discrimination Using Dimensionally Reduced RF-DNA Fingerprints. *IEEE Transactions on Information Forensics and Security*, 1180-1192. doi:10.1109/TIFS.2015.2400426

Stanciu, A. a. (2016). Analysis and Evaluation of PUF-Based SoC Designs for Security Applications. *IEEE Transactions on Industrial Electronics*, 5699-5708. doi:10.1109/TIE.2016.2570720

Suh, G. E. (2007). Physical Unclonable Functions for Device Authentication and Secret Key Generation. En *2007 44th ACM/IEEE Design Automation Conference* (págs. 9-14).

Xi, K. a. (2011). A fingerprint based bio-cryptographic security protocol designed for client/server authentication in mobile computing environment. *Security and communication networks*, 487-499.

Xiong Wenjie, S. A. (2016). Run-Time Accessible DRAM PUFs in Commodity Devices. En *Cryptographic Hardware and Embedded Systems -- CHES 2016* (págs. 432-453). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:978-3-662-53140-2

Yakovlev, J. M. (2012). Self-timed Physically Unclonable Functions. *IFIP International Conference on New Technologies, Mobility and Security (NTMS 2012)*, (págs. 1-5). doi:10.1109/NTMS.2012.6208707

Yin, C.-E. D. (2010). LISA: Maximizing RO PUF's secret extraction. En *2010 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)* (págs. 100-105). doi:10.1109/HST.2010.5513105