Current Mode, Array-Based Physical Unclonable Function Circuit

A Thesis

submitted to the designated by the Assembly
of the Departure of Computer Science and Engineering
Examination Committee

by

Dimosthenis Georgoulas

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN DATA AND COMPUTER SYSTEMS ENGINEERING

WITH SPECIALIZATION
IN ADVANCED COMPUTER SYSTEMS

University of Ioannina
School of Engineering
Ioannina 2024

Examining Committee:

- **Georgios Tsiatouchas**, Professor, Department of Computer Science and Engineering, University of Ioannina (Advisor)
- **Vassilios Tenentes**, Assistant Professor, Department of Computer Science and Engineering, University of Ioannina
- **Georgia Tsirimokou**, Assistant Professor, Department of Computer Science and Engineering, University of Ioannina

DEDICATION

I dedicate this thesis to my family, who was always there for me, even on the tough days.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

CMSA: Current Mode Sense Amplifier

CRP: Challenge-Response Pair

HVT: High Voltage Threshold

IoT: Internet of Things

LVT: Low Voltage Threshold

ML: Machine Learning

POK: Physically Obfuscated Key

PUF: Physical Unclonable Function

SVM: Support Vector Machine

ABSTRACT

Dimosthenis Georgoulas, M.Sc. in Data and Computer Systems Engineering, Department of Computer Science and Engineering, School of Engineering, University of Ioannina, Greece, June 2024
Current Mode Array Physical Unclonable Functions.
Advisor: Georgios Tsiatouchas, Professor.

Nowadays, ensuring the security of electronic devices is imperative, particularly within the technology and digital device domain where prioritizing the protection of data integrity and confidentiality is crucial. Toward this direction, various techniques have been proposed, among which is the Physical Unclonable Function (PUF). PUF, is a hardware-based technique which takes advantage of process variations and generates a unique response to a given input. PUFs play a crucial role in securing applications related to authentication, transactions, and IoT devices by generating unique cryptographic keys ensuring device identification. Unfortunately, voltage and temperature variations can adversely affect most of the proposed PUF circuits, leading to unreliable decisions and making the device vulnerable.

In this thesis, a current mode PUF circuit is proposed, utilizing reduced power supply dependency current generation cells. A matrix of n rows and m columns of cells generates process variation dependent currents which are used to create unique keys. Furthermore, a current mode sense amplifier (CMSA) is exploited which takes as input the currents from two columns of the array and determines the winning column based on the greater current. Every column drives a pair of CMSAs. The ability to activate multiple rows along with the reduced supply dependency current generation cells, make the PUF strong and reliable to voltage variations respectively.

In the evaluation of the proposed PUF circuit, the 90nm technology of UMC was exploited. The PUF design and simulation were carried out using the Virtuoso and Spectre platforms of Cadence. To validate the robustness and reliability of the PUF, an extensive analysis was conducted through multiple Monte Carlo simulation sessions, of 10,000 runs each, considering temperature and voltage variations to assess the PUF's performance across diverse operating conditions. Results show a reliability of 96.65% with respect to voltage fluctuations and 97.45% with respect to temperature variations, along with a uniqueness of 49.99% and uniformity of 49.85% on average.

Εκτεταμενη Περιληψη

Δημοσθένης Γεωργούλας, Δ.Μ.Σ. στη Μηχανική Δεδομένων και Υπολογιστικών Συστημάτων, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πολυτεχνική Σχολή, Πανεπιστήμιο Ιωαννίνων, Ιούνιος 2024
Μη Κλωνοποιήσιμες Φυσικές Συναρτήσεις σε Λειτουργία Ρεύματος
Επιβλέπων: Γεώργιος Τσιατούχας, Καθηγητής

Σήμερα, η επίτευξη της ασφάλειας των ηλεκτρονικών συσκευών είναι επιτακτική, ιδίως στον τομέα της τεχνολογίας και των ψηφιακών συσκευών, όπου προτεραιότητα είναι η προστασία της ακεραιότητας και της εμπιστευτικότητας των δεδομένων. Προς αυτή την κατεύθυνση, διάφορες τεχνικές έχουν προταθεί, μεταξύ των οποίων βρίσκεται και η μη κλωνοποιήσιμη φυσική συνάρτηση (physical unclonable function - PUF). Το κύκλωμα PUF αξιοποιεί μια τεχνική βασισμένη στο υλικό η οποία εκμεταλλεύεται τις διακυμάνσεις των παραμέτρων των κυκλωματικών στοιχείων ενός κυκλώματος, οι οποίες προκύπτουν κατά την κατασκευαστική διαδικασία και δημιουργεί μια μοναδική απόκριση σε μια δεδομένη είσοδο. Τα PUF παίζουν πολύ σημαντικό ρόλο στην ασφάλεια σε εφαρμογές σχετικές με αυθεντικοποίηση, συναλλαγές και συσκευές IoT, δημιουργώντας μοναδικά κρυπτογραφικά κλειδιά που εξασφαλίζουν τον προσδιορισμό της συσκευής. Δυστυχώς, οι μεταβολές της τάσης τροφοδοσίας και της θερμοκρασίας επηρεάζουν αρνητικά πολλά από τα προτεινόμενα κυκλώματα PUF, οδηγώντας έτσι σε αναξιόπιστες αποκρίσεις του κυκλώματος καθιστώντας την συσκευή ευάλωτη.

Στην παρούσα θέση, προτείνεται ένα κύκλωμα PUF που λειτουργεί με βάση το ρεύμα, χρησιμοποιώντας κελιά για τη δημιουργία ενός ρεύματος σχεδόν ανεξάρτητου από μεταβολές της τάσης τροφοδοσίας. Η δομή ενός πίνακα από n γραμμές και m στήλες από κελιά δημιουργούν ρεύματα επηρεαζόμενα από

κατασκευαστικές διακυμάνσεις τα οποία μπορούν να αξιοποιηθούν για την παραγωγή μοναδικών κλειδιών. Επιπλέον, αξιοποιείται ένας ενισχυτής αίσθησης ρεύματος (CMSA), ο οποίος λαμβάνει ως είσοδο τα ρεύματα από δύο στήλες αποτελούμενες από κελιά και αποφασίζει την νικητήρια στήλη σύμφωνα με το μεγαλύτερο ρεύμα. Κάθε στήλη οδηγεί δύο CMSAs ταυτόχρονα. Η δυνατότητα ενεργοποίησης πολλαπλών γραμμών από κοινού με την μειωμένη εξάρτηση των δημιουργούμενων ρευμάτων από την τάση τροφοδοσίας προσφέρουν ένα PUF ισχυρό και ανθεκτικό στις διακυμάνσεις της τάσεως τροφοδοσίας.

Η αποτίμηση του προτεινόμενου PUF κυκλώματος έγινε με χρήση της τεχνολογίας των 90nm της UMC. Η σχεδίαση και προσομοίωση του κυκλώματος έγινε με χρήση των εργαλείου Virtuoso και Spectre της Cadence. Για να ελεγχθεί η ανθεκτικότητα και η αξιοπιστία του προτεινόμενου κυκλώματος PUF, πραγματοποιήθηκε εκτενής ανάλυση Monte Carlo προσομοιώσεων σε πολλαπλές ενότητες των 10,000 επαναλήψεων λαμβάνοντας υπόψιν θερμοκρασιακές μεταβολές αλλά και μεταβολές στην τάση της τροφοδοσίας. Τα αποτελέσματα δείχνουν πως το κύκλωμα προσφέρει αξιοπιστία 96.65% ως προς τις μεταβολές της τάσης τροφοδοσίας και 97.45% ως προς τις μεταβολές της θερμοκρασίας, με μοναδικότητα 49.99% και ομοιομορφία 49.85%, κατά μέσο όρο.

CHAPTER 1

INTRODUCTION

1.1 Objectives

As technology evolves, new demands for data confidentiality are emerging. Data encryption has revolutionized how we process data by using cryptographic keys for encryption. Unfortunately, these cryptographic keys are typically stored in memory, making the system vulnerable to attacks that can leak the stored keys. To address this vulnerability, a different approach to data security is gaining prominence generating keys on demand using a specific hardware design instead of storing them in memory. This approach is achieved by utilizing the Physical Unclonable Function (PUF), a circuit capable of generating a unique response to a given challenge for a specific chip. However, PUF circuits are sensitive to voltage and temperature variations, which makes it challenging to create a robust and reliable PUF with a low latency, low silicon area, and low power consumption.

PUF designs are not completely immune to attacks. Depending on the number of possible Challenge-Response Pairs (CRPs), a PUF is categorized as weak or strong, with the latter being more desirable. With the rise of machine learning, new threat models are emerging that target both weak and strong PUFs.

The main objective of this thesis is to highlight the importance of PUF designs in security and data confidentiality while constructing a robust and reliable PUF that is less sensitive to voltage variations, an aspect which is crucial in PUF designs. To achieve this, we propose a PUF that shifts from the conventional voltage mode to a current mode approach. Our approach utilizes an array configuration of reduced supply dependency current generation cells, known for their stability against voltage variations. By using an array of such cells, we aim to create a strong and scalable PUF capable of generating a large number of possible CRPs. Furthermore, operating in current mode allows us to leverage inherent advantages in robustness and reliability, with a dual focus on reduced power consumption and response latency. The proposed PUF can generate a response on demand, making it an ideal candidate for applications that frequently use cryptographic keys.

To validate the performance of our proposed PUF design, extensive simulations were performed across a range of temperatures (0°C, 27°C, and 80°C) and voltages (0.9V, 1V, and 1.1V) considering different numbers of activated rows (8, 128 and 248) in the array, according to the applied challenge. These evaluations were carried out through 15 simulation sessions, each comprising 10,000 Monte Carlo runs. To provide a clearer view of the proposed circuit's performance, we performed comparative analyses with other state-of-the-art PUFs in the field.

1.2 Thesis Outline

The structure of the thesis is organized as follows:

In Chapter 2, an introduction to Physical Unclonable Functions (PUFs) is provided, exploring their usage in various applications and offering insight into state-of-the-art PUF designs. Furthermore, we explain the metrics Uniqueness, Reliability, and Uniformity used for evaluation. A categorization of PUFs into weak and strong types is presented, along with an examination of various threat models.

In Chapter 3, we propose a current mode array PUF circuit that utilizes reduced power supply dependency current generation cells. This chapter details the key components of the circuit (cells, buffers, and comparators), including a deep dive into transistor-level design. We also describe the operation phases of the circuit, which include discharge-equalization, activation and sensing phase.

In Chapter 4, a presentation and discussion of the simulation results is provided. We performed multiple Monte Carlo simulations under various environmental conditions to evaluate our PUF in terms of silicon area, power consumption, challenge-response pairs, latency, reliability, uniqueness, and uniformity. Finally, we compare our PUF with other state-of-the-art designs.

The thesis concludes in Chapter 5, where an overview of the proposed PUF is provided, and future work is suggested.

CHAPTER 2

PHYSICAL UNCLONABLE FUNCTIONS

The primary objective of this chapter is to provide a comprehensive understanding of Physical Unclonable Functions (PUFs), fundamental components that ensure the security and integrity of electronic devices. PUFs find extensive applications, primarily in cryptography. Such circuits generate unique keys, leveraging the inherent physical variations in electronic components, and can be used to enhance data security and confidentiality. The strength or weakness of a PUF is often defined based on the number of challenge response pairs (CRPs). While numerous PUF designs exist, the majority operate in voltage mode, often utilizing pre-existing SRAM cells. In contrast, there is a noticeable absence of PUFs operating in current mode. Most of the implementations follow an array-like structure, a characteristic shared by our proposed implementation. The evaluation of PUFs involves the utilization of multiple metrics.

2.1 Introduction to Physical Unclonable Functions

A physical unclonable function is an entity that uses production variability to generate a device-specific output which usually is a binary number. This output can be seen as the fingerprint of a device. A PUF is made of several components defined by local parameter variations. The differences between the components are called local mismatches. Depending on the PUF approach these local parameters are combined, compared or directly read out to generate the binary output. Since the variation of the components cannot be controlled from the outside, a PUF cannot be replicated. This fact makes it unclonable. Depending on the application, the PUF output depends on an input signal. Hence, a PUF is a function. To what extend the input signal influences the output differs between the various PUF approaches. The input (challenge) may alter the internal combination of the mismatching components which changes the output (response). The input may also define which of the components should be used to generate the output [1].

2.2 Weak and Strong PUFs

The number of challenge response pairs (CRPs) defines the strength of a PUF. Strong PUFs are typically used for authentication, while weak PUFs are used for key storage. A weak PUF can only support a small number of challenge-response pairs, and in some cases, it may only support a single challenge. This type of PUF is also known as Physically Obfuscated Key (POK) [2] and it is vulnerable to tampering and brute force attacks. A strong PUF can support a large enough number of challenges such that complete determination/measurement of all challenge–response pairs (CRPs) within a limited timeframe is not feasible [3]. Furthermore, an attacker can gain physical access to the weak PUF for any given time because of the low scaling and the limited pairs. The strength of a PUF is generally determined by how the number of potential CRP scales with the increasing PUF size. In general, if the number of CRPs supported by the PUF scales exponentially with its size, it is considered strong, while linear or polynomial increases typically correspond to weak PUFs [4].

2.3 Security Threats

Security vulnerabilities are inherent in nearly every circuit, including Physical Unclonable Function (PUF) circuits. Due to their critical role, defending PUF circuits against various types of attacks is imperative. In the literature, numerous attacks targeting such circuits have been documented. In [5], some of the most common PUF attacks are mentioned including Machine Learning (ML) Attacks and Side-Channel Attacks.

2.3.1 Machine Learning (ML) Attacks

With the enhancement of computer capabilities and the advancement of Machine Learning (ML) algorithms, a new avenue for PUF attacks has emerged. ML algorithms can solve complex problems that may be challenging for humans and detect unusual patterns while predicting values based on existing data. Certain robust PUFs, notably Arbiter PUFs, appear vulnerable to ML-based attacks. As we will explore later, Arbiter PUFs exploit path delays to generate responses. However, a notable constraint of such PUFs lies in their sequential structure, where each cell's response depends on the preceding component. By modeling the delay of each subsequent cell starting from the arbiter, ML techniques can predict delays along the path created by previous cells, facilitating a reverse engineering process. With knowledge of some challenge-response pairs, the PUF can be effectively modeled. ML modeling attacks employed in this context include Support Vector Machines (SVMs), Logistic Regression (LR), and Evolution Strategies (ES).

2.3.2 Side-Channel Attacks

A Side-Channel attack is a type of attack that targets the physical circuit of a PUF (Physical Unclonable Function). With this attack, hackers can extract basic information like timing, power usage, electromagnetic signals, and even sound. By using techniques like Simple Power Analysis (SPA) or Differential Power Analysis (DPA), anomalies in power usage during cryptographic processes can be detected. By collecting and analyzing these power traces, hackers can eventually uncover the inner digital keys.

The authors classify side-channel attacks into four categories based on how they interact with the target circuit: passive, active, semi-invasive, and hybrid attacks.

Passive attacks, a key component of side-channel attacks, involve observing and gathering information from a target without altering it. Power side-channel attacks, demonstrated by Mahmoud et al.'s work, focus on extracting sub-response information from XOR Arbiter PUFs and Lightweight PUFs by analyzing power consumption patterns. These PUFs encode responses with XOR functions to protect secret messages, but power consumption increases with the generation of more '1's due to the latch-based arbiter's power consumption characteristics. Timing side-channel attacks are not yet demonstrated on PUF circuits but could potentially provide additional information about response bits by analyzing timing variations in cryptographic operations.

Active attacks, unlike passive ones, involve actively tampering with system resources or disrupting their normal functioning, often seen in network attacks. These tactics aim to change the target or its surroundings beyond their usual behavior, causing noticeable shifts in system performance that attackers can observe. Common active attack methods include fault injection, which tries to reveal cryptographic keys or mess with program flow to bypass integrity checks. Delvaux and Verbauwhede introduced an active attack on PUFs [6]. They focused on creating a model for PUF repeatability, which relies on inter noise sources. It is important to note the difference between noise and variability here. While neither is great in regular electronic circuits, PUFs take advantage of process variability for security. So, noise becomes the perfect candidate for fault injection. Temperature changes, voltage fluctuations, and other sources contribute to noise in PUF circuits, ultimately reducing the repeatability of Challenge-Response Pairs. Delvaux and Verbauwhede's study delve into using noise to understand variability for generating response bits. They analyze the fraction of responses that end up as '1' for a given CRP, establishing a probability distribution function (PDF) for repeatability through Repeatability Measurements. They suggest using statistical methods like the Least Mean Square (LMS) Method and Differential Measurements Method to study the model. Their findings show that response repeatability can be exploited as a side channel for modeling strong PUFs.

2.4 Evaluation Metrics

To evaluate the performance of a PUF and compare it with other implementations, a set of key evaluation metrics must be employed. Three widely used metrics are Uniqueness, Reliability and Uniformity. Before providing an explanation of these evaluation metrics, it is essential to define certain structural components that play a crucial role in the equations of these metrics. These components include Hamming Distance and Hamming Weight [7].

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.

2.4.1 Uniqueness

Starting with Uniqueness, we can define it as the measure of the ability of one PUF instance to have a uniquely distinguishable behavior compared with other PUFs with the same structure implemented on different chips. The uniqueness metric is evaluated using the "Inter-chip Hamming Distance". If two chips, i and j ($i \neq j$), have n-bit responses, $R_i(n)$ and $R_j(n)$, respectively, for the challenge C, the average inter-chip HD among k chips is defined as in [7].

$$HD_{\text{INTER}} = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \frac{HD\left(R_i(n), R_j(n)\right)}{n} \times 100\%$$

Ideally, Uniqueness should be close to 50%.

2.4.2 Reliability

The next metric is called Reliability and is a measure of the ability of the PUF to generate a consistent response R for a challenge C, regardless of any changes in the conditions of the environment such as the ambient temperatures and voltage supply. The reliability metric is evaluated using the "Intra-chip Hamming Distance". If a single chip, represented as i, has the n-bit reference response $R_i(n)$ at normal operating conditions and the n-bit response $R'_i(n)$ at different conditions for the same challenge C, the average intra-chip HD for k samples/chips is defined as in [7]:

$$HD_{\text{INTRA}} = \frac{1}{k}\sum_{i=1}^{k}\frac{HD\left(R_i(n), R_i'(n)\right)}{n} \times 100\%$$

and so, the Reliability of the PUF can be defined as:

$$Reliability = 100\% - HD_{\text{INTRA}}$$

Ideally, the value of Reliability should approach 100% meaning that the Intra-chip Hamming Distance should be close to 0%. In the literature, HD_{INTRA} can be mentioned as Bit Error Rate (BER) [8].

2.4.3 Uniformity

Also, Uniformity is a crucial metric and estimates how uniform the PUF's responses are and is defined as the proportion of 0's and 1's in the response bits of a PUF. This percentage in a truly random response is 50% and it can be calculated using the average Hamming Weight of the responses as follows [7]:

$$Uniformity = \frac{1}{k}\sum_{i=1}^{k} r_i \times 100\%$$

where k is the total number of responses and r_i is the Hamming Weight of the *i*th response.

2.5 Proposed PUF Implementations

A variety of PUF architectures exists in the literature, each with its own unique properties. In [9], a taxonomy of PUF designs is introduced, providing a thorough overview of existing Physical Unclonable Function (PUF) implementations. This paper systematically categorizes PUFs based on their implementation concepts and highlights several prominent architectures that have received considerable attention in the field. Some of the proposed PUFs are Ring Arbiter PUF, Ring Oscillator PUF, SRAM PUF and SCA-PUF. Besides the ones mentioned in the taxonomy, there are a few more SRAM PUF architectures worth mentioning for study purposes. In the following sections, we will take a closer look at them to understand their pros and cons.

2.5.1 Arbiter PUF

A novel Arbiter PUF (APUF) is proposed in [10], and it is placed among other famous PUF models. Authors describe this PUF as delay-based and weak because it generates only one bit response. The key components of the APUF include eight switching elements (SEs) and eight selecting modules (SMs), both illustrated in Figure 2.1, along with an arbitration unit.

The SEs create two distinct paths, and based on the voltage Ci, the path randomly switches between the X-state and T-state, as depicted in Figure 2.2. The Ci voltage is determined by the Selecting Modules, which consist of three inverters. The input to the SMs is the challenge provided to the PUF, set at half the Vdd. Process variations lead to the creation of different paths and delays in the circuit's transistors. An arbiter placed at the end of the chain outputs either '0' or '1' depending on which path was faster.

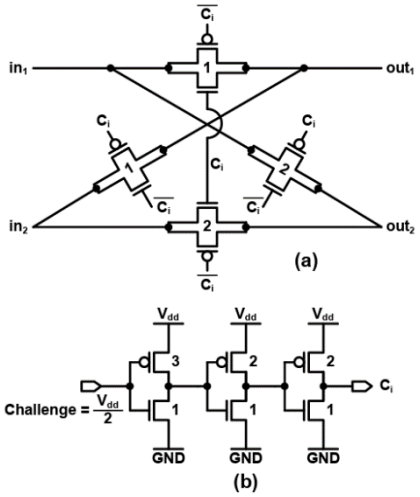


Figure 2.1: Components of the single-bit PUF cell, (a) Switching element, (b) Selecting module. [10]

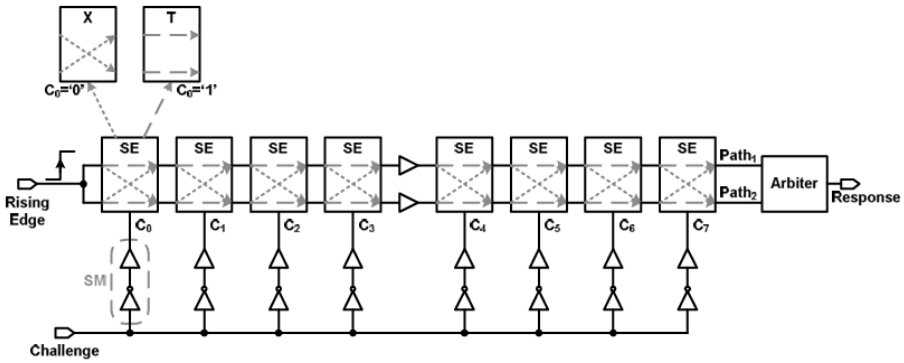


Figure 2.2: The single-bit PUF cell consisting of eight switching elements, eight selecting modules, and an arbiter. [10]

2.5.2 Ring Oscillator PUF

The Ring Oscillator PUF (RO PUF) [11] is a PUF implementation based on delay loops, specifically ring oscillators. Its simplicity is what makes it popular, although it is considered slower, larger, and more energy-consuming compared to the Arbiter PUF as discussed earlier. However, RO PUF offers higher reliability and is easier to implement for both Application-Specific Integrated Circuits (ASICs) and Field Programmable Gate Arrays (FPGAs), particularly in secure processor designs.

The RO PUF consists of multiple identical delay loop circuits that oscillate at a particular frequency. Each delay loop block contains an AND gate and n number of inverters. Due to process variations, each delay loop operates with a slightly different frequency, leading to the uniqueness required for a PUF. Counter circuits are used to count the oscillator cycles, and a comparison is made to generate a response.

However, the circuit's simplicity makes RO PUF vulnerable to model-building attacks, where attackers can learn timing patterns from multiple input-output responses. To mitigate such attacks, XOR or MUX components can be incorporated to increase the complexity of the internal paths. An illustration of the RO PUF architecture is presented in Figure 2.3.

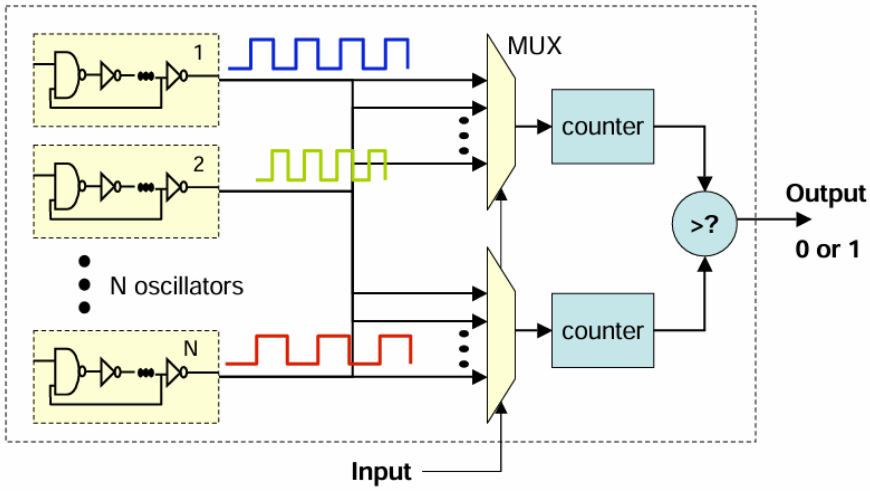


Figure 2.3: Illustration of the Ring Oscillator PUF architecture. [11]

2.5.3 SCA-PUF

Previously, we discussed about PUF implementations based on time delays (Arbiter PUF, RO PUF). Another type is the current-mode PUFs, which leverage current differences and translate them into corresponding bits. In [12], a robust subthreshold

current array (SCA) PUF is proposed, designed to withstand machine learning attacks. The overall architecture of the proposed PUF is depicted in Figure 2.4.

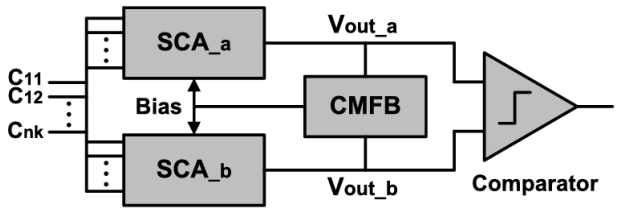


Figure 2.4: The architecture of SCA-PUF consisting of a pair of arrays, a comparator, and a common-mode feedback (CMFB) circuit. [12]

To begin with, the architecture comprises two n by k subthreshold current arrays (SCA_a, SCA_b) with inputs Cnk, followed by common-mode feedback (CMFB) and a comparator. As illustrated in Figure 2.5, each column features a primary PMOS transistor at the top and multiple unit cells.

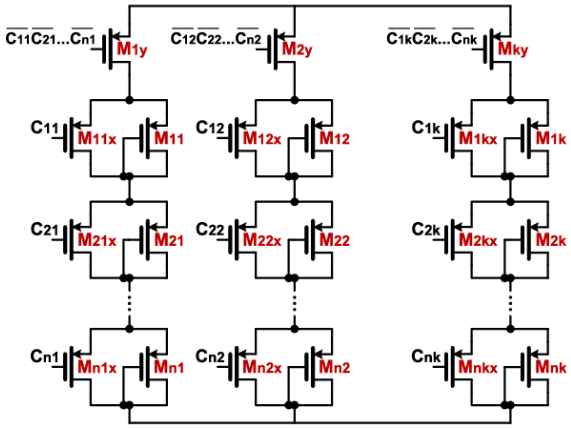


Figure 2.5: The subthreshold current array consists of n rows and k columns of controllable unit cells. [12]

Each unit cell comprises two PMOS transistors. The Mijx transistor switches are labeled as non-stochastic, while the diode connected Mij transistors are deemed stochastic. By 'stochastic', the authors refer to transistors with the maximum amount of Vth variability, achieved by using minimum-sized transistors. Conversely, Mijx switches are sized to minimize their variability. The roles of transistors Mijx and Cij signals are significant. When Cij = 1, transistors Mij contribute to determining the output voltage, but the same is not true when Cij = 0. Both arrays are driven by the same Cij signals, so we anticipate the two arrays to yield equal output voltages. However, due to process variations, Vth values are not identical, resulting in different output voltages and yielding a random binary response after comparing the two output voltages.

2.5.4 PTAT PUF

In [2], a different type of current mode PUF is proposed. This implementation is very different from the previous SCA-PUF. To begin with, this PUF is considered as weak PUF due to the low number of CRPs. On the other hand, is considered to have a very high reliability of 99.55% at various temperatures between 0°C - 80°C (with constant Vdd) and the PUF is also operates very well at the wide range of voltages 0.6V - 1.2V. The structure of the PUF is also an array-like with the unit cells following the Proportional To Absolute Temperature (PTAT) principle. The structure of the described PUF is shown in Figure 2.6.

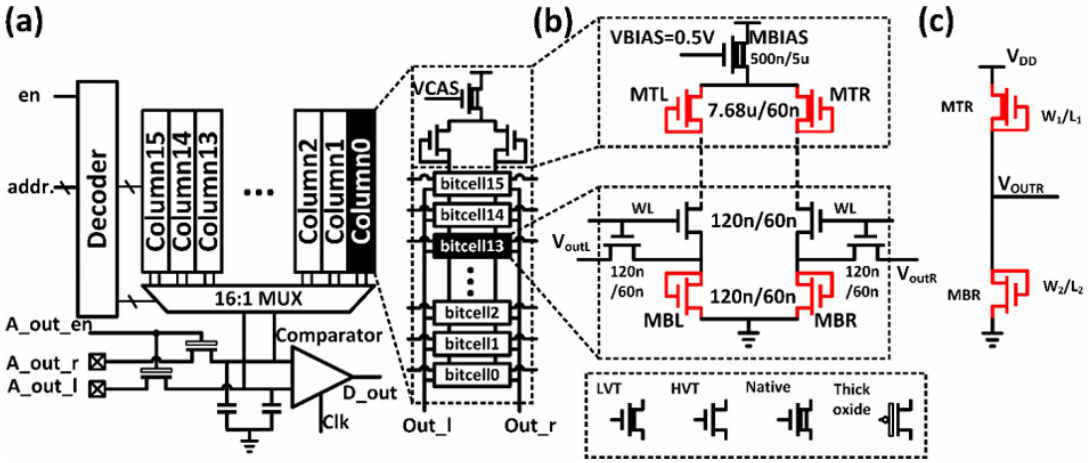


Figure 2.6: (a) The 256-bit PUF. It is composed of a bitcell array, an address decoder, an analog multiplexer, and a 1-bit comparator. (b) A bitcell and a shared header in a column. (c) A PTAT generator, two of which form a single PUF bitcell. [2]

Each bitcell comprises four transistors and all bitcells of the same column share the same MTL and MTR transistors as shown in Figure 2.6 (a) and (b). A pair of two MTR and MBR, construct a PTAT generator with the MTL, MTR, MBL and MBR operating in subthreshold region. The current that passes through them can be defined as:

$$I_{\text{sub}} = \mu C_{\text{ox}} \frac{W}{L} (m-1) V_t^2 \exp\left(\frac{V_{\text{gs}} - V_{\text{th}}}{m V_t}\right) \times \left(1 - \exp\left(\frac{-V_{\text{ds}}}{V_t}\right)\right)$$

Equation 1: Subthreshold current

where μ is carrier mobility, Cox is sheet oxide-capacitance density, W, L are the width and length, Vth is threshold voltage, m is subthreshold slope, Vgs/Vds is gate–source/drain–source voltage, and Vt is the thermal voltage.

Due to the equal size transistors on the opposite sides of the bitcell, and the column respectively, the subthreshold current on both sides must be identical. The bitcells are connected in stack and hence, we can solve for the V_{OUTR}.

$$V_{\mathrm{OUTR}} = \underbrace{V_{\mathrm{th2}} - \frac{m_2}{m_1} V_{\mathrm{th1}} - K_{V_{\mathrm{th}}}(T_0)}_{V_{\mathrm{th\ determined}}} + \underbrace{K_{V_{\mathrm{th}}} \cdot T + m_2 \frac{kT}{q} \ln\left(\frac{\mu_1}{\mu_2} \cdot \frac{C_{\mathrm{ox1}}}{C_{\mathrm{ox2}} \cdot \frac{W_1 L_2}{W_2 L_1} \cdot \frac{m_1 - 1}{m_2 - 1}}\right)}_{\text{temperature dependent}}$$

Equation 2: Calculation of V_{OUTR}

where the subscripts 1 and 2 represent MTR and MBR, respectively, $K_{V_{\mathrm{th}}}$ is a combined constant of the temperature dependencies of the Vths of MTR and MBR, T_0 is the reference temperature, k is the Boltzmann constant, T is temperature, and q is an electron charge. In this derivation, we assume Vds is sufficiently larger than Vt, which allows us to eliminate the second exponential term of Equation 1.

As we can observe in Equation 2, there are two parts, the temperature-dependent and the temperature-independent. The second part is proportional to temperature and the slope is defined by the sizes of MTR and MBR transistors along with the $K_{V_{\mathrm{th}}}$. So, the difference of the Vout (ΔVout) can be defined as:

$$\Delta V_{\mathrm{out}} = V_{\mathrm{outL}} - V_{\mathrm{outR}} = (V_{\mathrm{th2L}} - V_{\mathrm{th2R}}) - \left(\frac{m_{2L}}{m_{1L}} V_{\mathrm{th1L}} - \frac{m_{2R}}{m_{1R}} V_{\mathrm{th1R}}\right) + K_\Delta T$$

Equation 3: Calculation of ΔV_{out}

where the subscript L and R represent the left and right output of a PUF bitcell, respectively, and K_Δ is the difference in the temperature slopes between the two PTAT generators. K_Δ is supposedly very small as the generators are identically sized and symmetrically layout-ed.

The authors, confirms via simulations, that the second and third terms of the Equation 3 are negligibly small and that ΔV_{out} is mostly determined by the V_{th} difference of the bottom devices. The variations in V_{th} of MBL and MBR transistors, follows the same normal distribution with zero mean and same standard deviation, and so the ΔV_{out} is consider having also the same normal distribution improving the uniqueness of the PUF. Finally, the output is determined by a classic sense amplifier comparator with a PMOS differential pair and an NMOS cross-coupled latch.

2.5.5 SRAM PUF

Another intriguing type of PUF is the Static Random-Access Memory PUF (SRAM PUF) [13]. The SRAM PUF is particularly noteworthy as it utilizes the already present SRAM cells in modern integrated circuits and Field Programmable Gate Arrays (FPGAs). SRAM comprises an array of cells constructed with six transistors, known as the 6T SRAM cell. This storage unit is composed of two cross-coupled inverters and two access transistors, as illustrated in Figure 2.7.

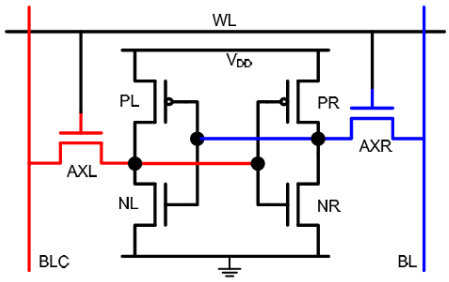


Figure 2.7: Six transistor SRAM cell. [13]

To store a bit in SRAM, the word line (WL) must be enabled, and either a voltage must be injected from bit line complement (BLC) or bit line (BL) to set the two inverters to a stable state, representing '1' and '0' respectively. Conversely, during read operations, the WL is enabled without applying any voltage to the cell, as done during the write phase. To leverage SRAM as a PUF, cell readings are taken only during the startup phase, where no prior writes have occurred. Due to manufacturing variabilities, the transistors within SRAM cells are not identical, resulting in imbalance and leading to unpredictable meta-stable states in the cross-coupled inverters, storing either '1' or '0' bits.

In [13] only one row is activated at a time, meaning that only a single cell is activated in each column. Therefore, activating multiple rows (Wordlines) simultaneously is not feasible in this implementation which makes the PUF weak.

2.5.6 SiCBit-PUF

The authors in [14] propose a novel Strong in-Cache Bitflip PUF (SiCBit-PUF) computation method for extracting static entropy from SRAM arrays, leveraging bitflips. This PUF architecture allows for the simultaneous activation of multiple Wordlines, with multiple cells driving the Bitlines. By activating many cells at the same time, they can perform in-memory computing allowing two bitwise operations, logic-AND and logic-NOR. After conducting research, they concluded that the bit-

flips that occur during those in-memory computations are not stochastic but systematic. Taking advantage of this, a strong PUF can be created where the initial state of the cells can produce a random response after bitwise operations. For their PUF implementation, they suggest the bitline architecture, as shown in Figure 2.8, instead of the wordline architecture because it is possible to increase the Challenge-Response Pairs and thus create a strong PUF.

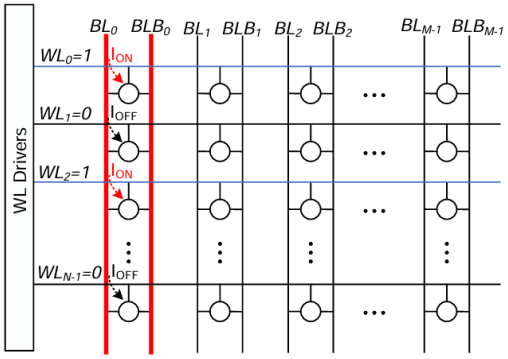


Figure 2.8: The Bitline Computing Architecture with two Wordlines enabled and the collision of the two rows on the bitlines. [14]

The operation of the PUF proceeds as follows when two wordlines are enabled. Initially, the two cells have pre-stored bits, assuming '1' and '0' are stored respectively in this example. Then, a precharge phase occurs by charging the Bitlines to Vdd. Subsequently, the wordlines are activated, and the active cells of the same column begin to interact on the bitlines. During this phase, current begins to flow from one cell to another through the bitlines. The currents are influenced by the resistance of the cross-coupled inverters (PMOS and NMOS) and the two access transistors. Afterward, the two cells reach a meta-stable condition, and bit-flips may occur. The paths of the currents are depicted in red color in Figure 2.9. Ultimately, the output of the PUF will be the result of a logic-AND operation between the stored values of the two cells. The authors also suggest an error correction method to increase the reliability.

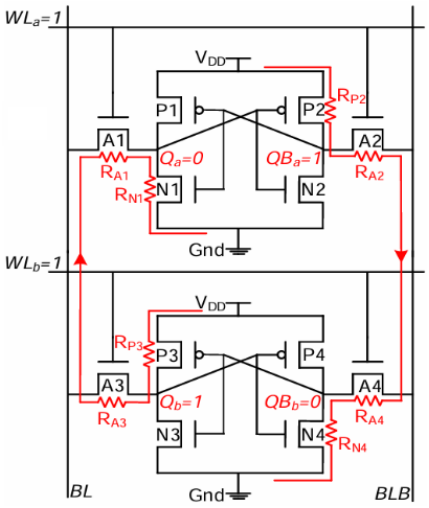


Figure 2.9: Colliding currents in a 2-cell SRAM memory circuit during PUF mode. [14]

2.5.7 SRAM Array Current Based PUF

The paper [15] introduces a novel Physical Unclonable Function leveraging random variations in SRAM cells' read access current, arising from manufacturing process variations. The authors propose a method to translate the analog read current of an SRAM array into robust binary signatures, employing a standard 6-Transistor (6T) SRAM cell as depicted in Figure 2.10 (a).

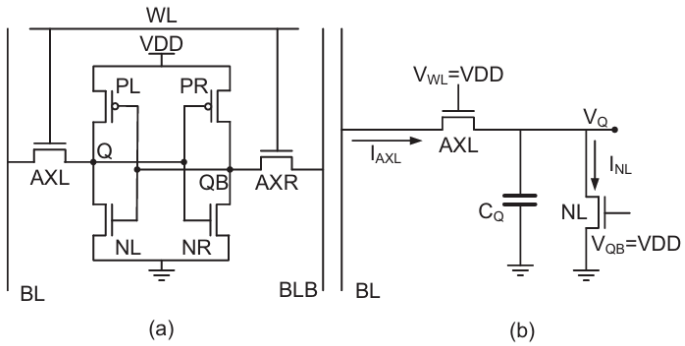


Figure 2.10: (a) Structure of a typical 6T SRAM cell. (b) Relevant parts of the SRAM cell circuit during a read operation when the cell content is 0. [15]

At the beginning, the SRAM cells have random stored values due to process variations. If a cell is initially storing bit 0, resulting in $V_Q = 0$ (and $V_{QB} = Vdd$), during the read phase, a current flow from BL through AXL as shown in Figure 2.10 (b), causing a slight increase in V_Q while BL decreases by the same amount. The read current is measured in a corresponding time interval Δt, within an optimal time, resulting in an average current ($\overline{IDDT}$). That current can be calculated as:

$$I_{NL} = (\mu C_{ox})\left(\frac{W}{L}\right)_{NL}\left[(V_{DD} - V_{tn})V_Q - \frac{1}{2}V_Q^2\right]$$

It has been observed that the read current varies depending on the stored logic value. By calculating the difference between $\overline{IDDT}$ and $\overline{I'DDT}$ (the complementary average read current from the BLB), we determine that a $\Delta\overline{IDDT} \geq 0$ indicates a result of 1, while a $\Delta\overline{IDDT} < 0$ indicates a result of 0. Following multiple Hspice simulations, the time window Δt is adjusted to optimize uniqueness and robustness. For the comparisons a sense circuit can be used.

2.5.8 PUF-CIM

Nowadays, compute-in-memory (CIM) is slowly gaining prominence within the realms of edge computing and deep neural networks (DNN). Pre-existing integrated circuits such as SRAM can be used in the domain of neural networks. In [16], an interesting idea of a PUF-CIM is proposed, utilizing the SRAM providing CIM capabilities and lightweight DNN model protection. Currently, SRAM is used only as storage and to perform a calculation, data must be transfer back and forth to the Arithmetic Logic Unit (ALU), consequently a bottleneck in data transfer rates and an increase in power consumption appears. The maths behind neural network calculations are simple because they consists of a dot product between input vector and weight vector. After that, the dot product is summed, and the output is generated through the activation function. These calculations are also happening in a XOR operation.

This simple perspective arrives new confidentiality issues. The weights of a neural network model are stored in the SRAM and so they must be kept private and secure otherwise X-ray photoelectron emission can be performed and cause model leakage. To address these issues, authors proposed a new 10 transistors (10T) SRAM cell which is shown in Figure 2.11.

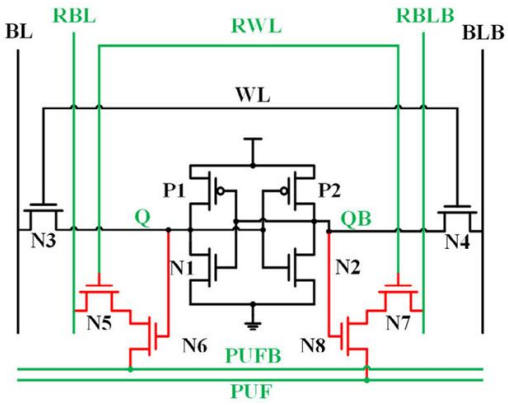


Figure 2.11: The 10T SRAM cell. [16]

This cell can encrypt the stored weights and perform XOR operation which is also binary multiplication. Initially, the SRAM acts as PUF and keys are being generated. To do this, a circuit called time-to-digital converter (TDC) is used. The SRAM cells are written into high state (Q is 1, QB is 0) and bit cells in the same WL are written with the same weight. Also, BLB0 and BLB1, as shown in Figure 2.12, are precharge to Vdd.

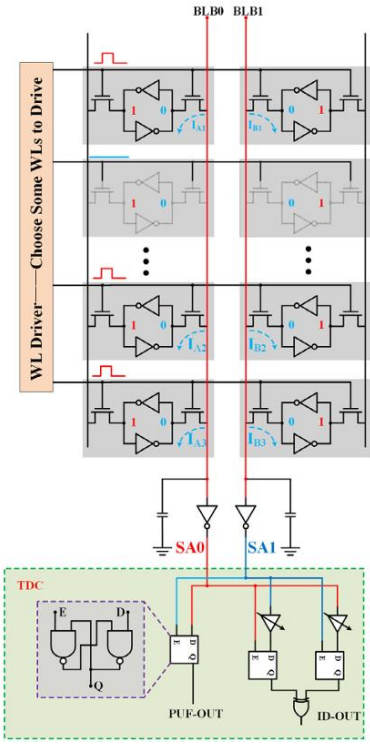


Figure 2.12: PUF mode operation and TDC block. [16]

After that, a discharging phase takes place where WL is selected and so QB turns on. The voltage of BLB0 and BLB1 will decrease with different rates due to transistor mismatches and TDC computes the difference in discharge time between BLB0

and BLB1. If $(T1 - T0) < 0$, where T1 - T2 is the discharge time difference between two columns, then the response is 0, elsewhere the response is 1. Authors mention that there is a tradeoff among area penalty, CRP number and BER of the proposed PUF. Furthermore, the TDC also generates an ID value which tells if the PUF value is unstable or not. To find this value, they perform the same method as before finding the difference in discharge time between BLB0 and BLB1 but considering a threshold value (Td) which is calculated based on multiple Monte Carlo simulations. If $|T1 - T0| \leq Td$ then the ID result is 0, indicating an unstable PUF value, otherwise the PUF value is stable.

The encryption can be performed using the XOR operation and can be expressed as follows:

$$\overline{W_e} = \overline{W \oplus PUF} = \bar{W} \cdot PUF$$

Where W_e is the encrypted binary weight, W is the original binary weight and PUF is the PUF response. Now we can do the binary multiplication with the encrypted weight and the given input as follows:

$$M = IN \cdot (\overline{W_e} \oplus PUF) = IN \cdot \bar{W}$$

where M is the result of multiplication calculations and IN the input value.

2.5.9 SPUF

The paper [17], introduces an SRAM-based PUF leveraging pre-existing SRAM. Although is considered a weak PUF, it features a higher number of Challenge-Response Pairs (CRPs) compared to other SRAM PUFs thanks to its array scalability. However, the PUF's array size is constrained, limiting the CRPs. To address this, the authors propose a modified SRAM cell design, maintaining the standard 6 Transistors (6T) but with different arrangement as illustrated in Figure 2.13.

In a typical SRAM cell, only one Worline (WL) exists, controlling both gate transistors, but this configuration differs here. The WL is split into two separate lines: WLL controls the left gate NMOS transistor and WLR controls the right gate NMOS transistor.

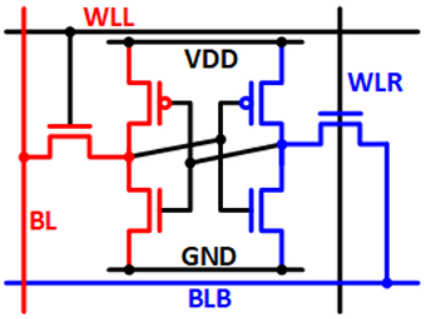


Figure 2.13: The SPUF unit schematic. [17]

This SRAM cell supports PUF, read and write modes. During read mode, the stored value of a bit-cell is accessed by activating the WLL and WLR signals for the specific cell. Similarly, in write mode, WLL and WLR are enabled, with BL or BLB set to '1' or '0' respectively, or vice versa.

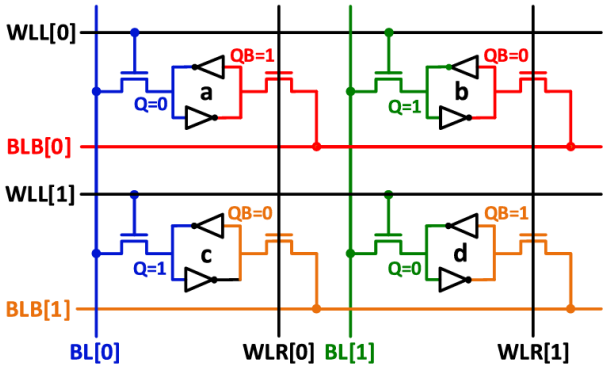


Figure 2.14: The SPUF unit 2 × 2 array for one step PUF operation. [17]

To operate the circuit in PUF mode, two vertical right-wordlines (WLR[0] and WLR[1]) and two horizontal left-wordlines (WLL[0] and WLL[1]) must be activated simultaneously, involving four bit-cells as depicted in Figure 2.14. These bit-cells are paired through four bit-lines (BL[0], BL[1], BLB[0], and BLB[1]), with their voltages determined by conflicting cells and stored values {Q(a), Q(c)}, {Q(b), Q(d)}, {QB(a), QB(b)}, and {QB(c), QB(d)}. For instance, BL[0] will be '0' or '1' if both QB(a) and QB(c) are '1' or '0', respectively. However, dissimilarities in transistors may lead to bit flips, overriding weaker bit-cells with stronger ones.

Depending on the number of rows and columns, we can expand this PUF procedure to include neighboring groups of 4 bit-cells. Initially, as described in the previous paragraph, bit flips may occur, potentially affecting the adjacent group of 4 bit-cells, and so forth. Subsequently, a read operation occurs to determine the PUF response based on the stored value of one of the participating cells.

CHAPTER 3

THE PROPOSED CURRENT MODE ARRAY PUF

In this thesis, we propose a novel implementation of an array type Physical Unclonable Function (PUF) circuit operating in current mode. Our primary objective is to develop a robust PUF resilient by design to voltage variations. To achieve this goal, we adopted a cell topology that is based on a current generation circuit with reduced dependency on power supply variations. The mentioned circuit is known for its capability to maintain currents almost independent of power supply variations. Utilizing multiple cells, we construct an expandable array, which forms the foundational architecture of our implementation. By increasing the number of rows and columns, we enhance the number of Challenge-Response Pairs (CRPs), thereby strengthening the PUF's security. The PUF operation requires the use of current comparators to compare pairs of column currents within the array, generating a response.

3.1 PUF Cells

The basic component of our PUF design is a reduced supply dependency current generation cell topology. This PUF cell is capable of providing a current that its value is almost independent of power supply voltage and is determined by the transistor sizes. Supply independent current generation circuits are frequently employed for generating bias currents in A/D or D/A converters. Within the framework of [18], an in-depth analysis of such circuitry is conducted.

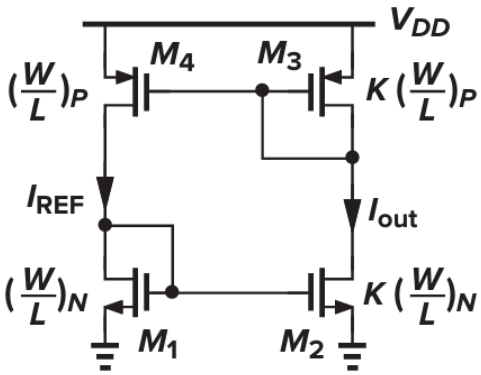


Figure 3.1: Reduced supply dependency current generation circuit. [18]

In Figure 3.1, transistors M1, M2, M3 and M4 construct the reduced supply dependency current generator. The circuit operation has a small dependence on power supply and the current magnitude is determined by the M1-M4 transistors. Typically, this circuit consists of a PMOS-type current mirror on top (M3 and M4) and an NMOS-type current mirror on the bottom (M1 and M2), connected in a complementary manner. The idea behind this scheme is the ability of the circuit to bias itself. This is the case, since both diode connected transistors (M1 and M3) are fed by coupled current source, the corresponding currents are relatively independent from the power supply. In Figure 3.1, neglecting the channel length modulation phenomenon, it stands that $I_{out} = K \cdot I_{REF}$. Under typical scenarios, with all transistors being identical in size, we anticipate an equal current on both sides of the circuit given that K=1. However, variations in the manufacturing process will naturally lead to an imbalanced topology. In such case, $K \neq 1$ thereby creating a current that deviates from the expected value, which aligns with our PUF concept.

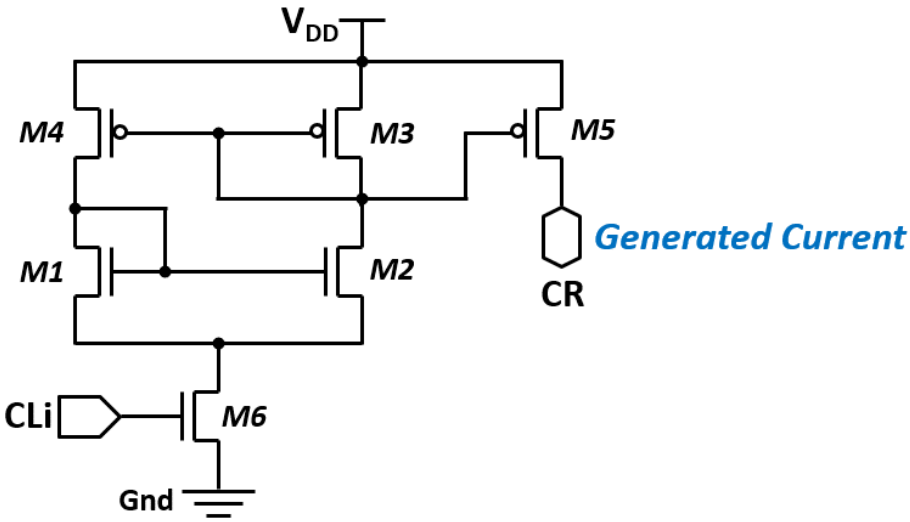


Figure 3.2: The proposed current generation cell, of reduced supply dependency, which is based on the current mirror pair with an extra transistor (M6) driven by the 'CLi' signal for cell activation according to the challenge input.

Our PUF cell requires controllability, necessitating the inclusion of a switch transistor. Hence, we utilize an NMOS switch transistor (M6) positioned at the bottom of current generation unit, as shown in Figure 3.2. This device also mitigates the current leakage in the standby mode of operation. When transistor M6 is OFF, no current passes through the cell and so the internal nodes (drains of M3 and M4) are set to Vdd. By turning M6 ON, a current begins to flow from both branches defined by the transistor sizes. Finally, the current that is generated from the circuit cell is mirrored to the bitline through a third current mirror consisting of M3 and M5.

Transistor M5 plays a crucial role in our design. As multiple cells are simultaneously activated, driving a common bitline, it is responsible to replicate the cells' operating current and sum it with all other unit cells. This approach effectively isolates the cells from the rest of the cells of the array and the comparators. All six transistors in the cell have the minimum size and high threshold voltage to keep silicon area and leakage current as low as possible.

3.2 PUF Array

In paragraph 2.3, we discuss some PUF implementations, and we concluded that array-type PUFs have a potential to be strong due to high number of rows and columns and therefore CRPs. Our PUF adopts the concept of an array PUF, which aligns perfectly with the principles of a current-mode PUF. As a current-mode PUF, we can define a PUF circuit which responses arise from the current differences, generated by the structural cells due to process variations on them. Adopting the cell topology of Figure 3.2, the cells or the array generate currents almost independent of the power supply. We hypothesize that these power supply fluctuations uniformly impact all transistors across the PUF array. Keeping that in mind, we can compare pairs of columns at any power supply voltage as their current difference remains theoretically the same. In Figure 3.3, we can see how the proposed PUF cells are arranged in an array structure.

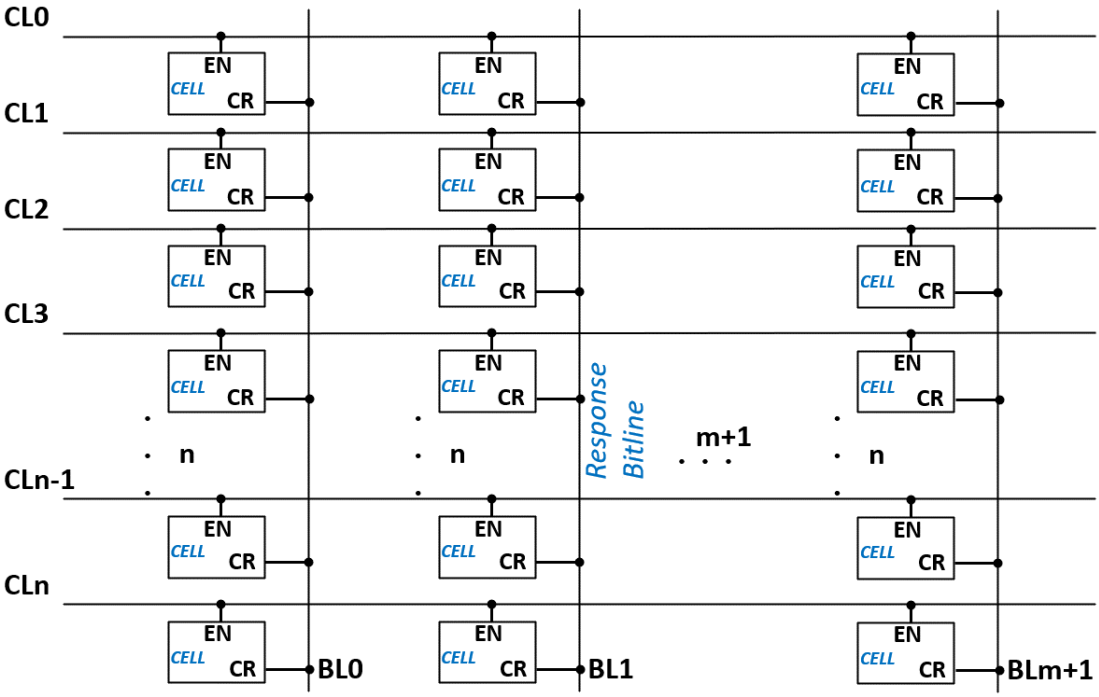


Figure 3.3: The proposed PUF array consisting of n challenge lines (CLn) and m+1 bitline columns (BLm+1).

Our PUF array consists of n rows (CL) and m+1 columns (BL) of cells. As a challenge to our PUF we define the activated rows-lines (CL) of the array. Our

implementation supports multiple rows activation at the same time, retrieving an m+1 bits response. A detailed analysis of the CRPs will be provided later in this thesis. To retrieve the responses, current comparators have been used and the overall PUF architecture is shown in Figure 3.4.

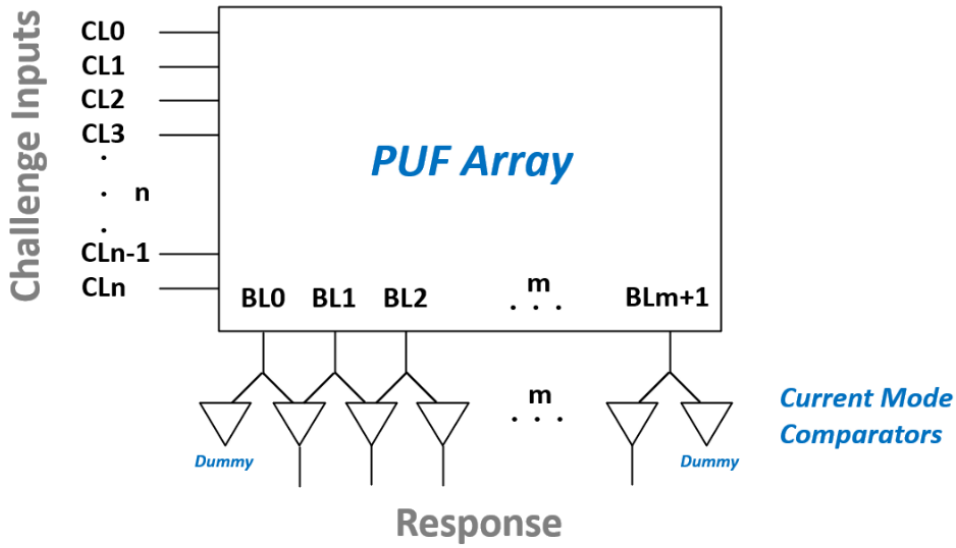


Figure 3.4: Architecture of the proposed Current Mode Array PUF with the n challenge inputs, m+1 bitline responses and m+2 comparators.

In the next paragraph, we will discuss the comparators that we use in our implementation. However, it is important to mention that each bitline drives two comparators and the bitline current gets splitted and passes through both comparators. In other words, we need m+2 comparators to generate a m-bit response, using m effective comparators and 2 extra 'dummy' comparators for the edge columns.

3.3 PUF Array Peripherals

When applying a challenge, a signal is required to drive m+1 cells within the same row. However, the high capacitance necessitates a circuit to mitigate delays. To address this challenge, we employ buffers, as drivers, to drive all signals across the circuit, given the large number of transistors activated by each signal. Another key component for the PUF operation is the comparator. As we mentioned earlier, comparators are used to get responses and evaluate a PUF. These two components are crucial for the overall PUF performance since buffers contribute negatively in terms of time, area and energy consumption while comparators can influence the PUFs responses and hence reliability.

3.3.1 Buffers

While specific buffers are designed for distinct purposes, they all adhere to the same fundamental principle. Figure 3.5 depicts the buffer that receives the challenge input 'CLi' for cell activation and forwards it to the corresponding PUF row. It features two cascading connected inverters, with the second inverter driving the m+1 minimum-sized M6 NMOS transistors of the cells in a row.

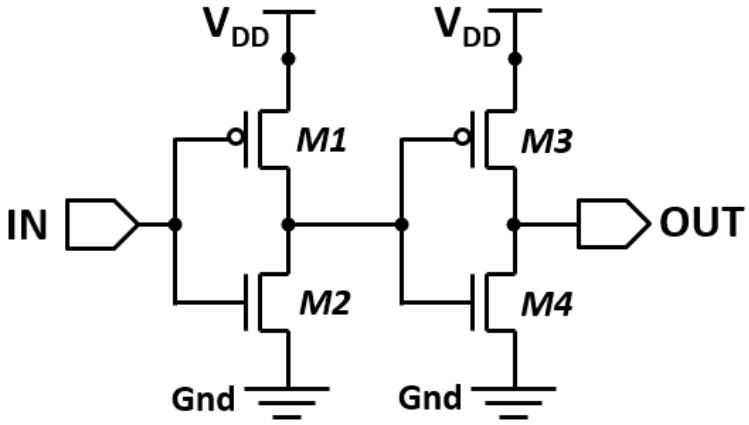


Figure 3.5: Buffer topology.

3.3.2 The Current Comparators

Since the PUF operates in current mode, which means that it performs the comparison of current differences, we choose a comparator known as the Current Mode Sense Amplifier (CMSA) [19]. This circuit compares currents between two-bit lines (columns - BL). It is important to mention that every column of the array drives two comparators simultaneously. As can be observed in Figure 3.4, a column drives the right input of its left placed comparator and also drives the left input of its right placed comparator. Even a small current difference applied to the comparator can contribute to a rapid response. We can identify the column with the higher current value, referred to as the 'winner column'. Figure 3.6, depicts the Current Mode Sense Amplifier employed in our PUF, which exhibits partial differences from Blalock's circuit [19] regarding transistor sizes and signal timings.

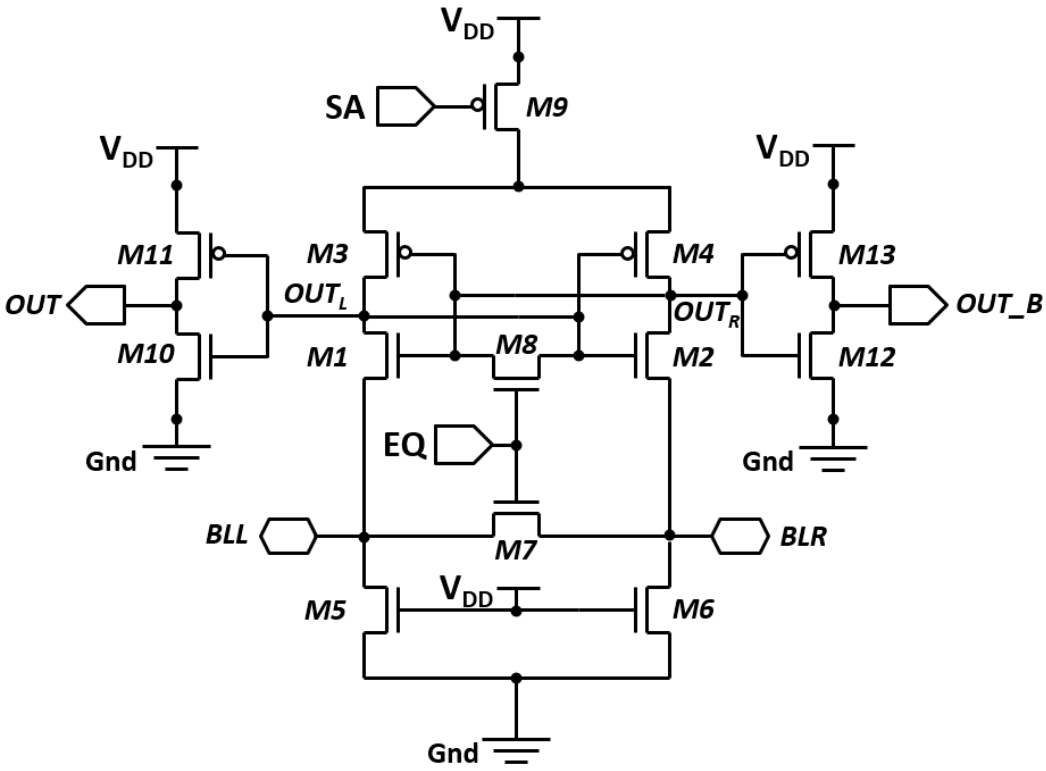


Figure 3.6: Current Mode Sense Amplifier.

A typical structure of such a CMSA consists of two inverters (M10 - M11, M12 - M13), a cross-couple latch (M1, M2, M3 and M4), equalization transistors (M7 and M8) and a pair of NMOS transistors (M5 and M6) which serve as active resistors.

3.3.3 Operation Phases

The operation of the proposed PUF is divided into three distinct phases: discharge-equalization, activation, and sense phase. An overall plot of all used signals is depicted in Figure 3.7. Initially, the PUF remains inactive until a challenge is applied. This idle state, known as the discharge-equalization phase, lasts for at least 500ps, during which the 'EQ' signal is set high. In this phase, the circuit discharges any remaining voltage on the bitlines through the bottom transistors (M5 and M6) of the comparators in Figure 3.6, while ensuring node equalization to prevent biasing.

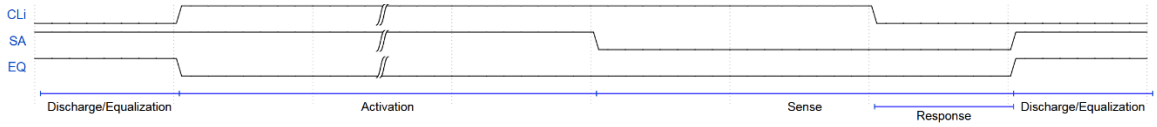


Figure 3.7: Indicative control signal waveforms.

Following the discharge-equalization phase, a challenge is applied to the PUF, marking the beginning of the activation phase. Activation occurs when one or more 'CLi' signal(s) turns to high activating every PUF cell in the selected row(s) by setting on corresponding transistors M6 in Figure 3.2. The current mirror, as discussed in paragraph 3.1, feeds current, through M5, to the pertinent bitline. Once a stable current flows on the bitlines, the circuit transitions to the next phase.

The final phase involves the operation of the comparator. Setting the 'SA' signal low, the CMSA is enabled and rapidly generates a response by comparing the currents of the associated bitlines. Depending on this comparison, the comparator's output (OUT) provides the final digital response.

Afterward, the circuit returns to the idle state (discharge-equalization phase), preparing the PUF for the next session by setting both 'EQ' and 'SA' to logical '1'.

CHAPTER 4

SIMULATION RESULTS

This chapter provides an in-depth exploration of the performance of our proposed PUF. We evaluate the circuit performance based on the common PUF metrics and compare its characteristics with other state-of-the-art PUFs.

4.1 Simulations and Results

To evaluate the performance and robustness of our PUF, we conducted exhaustive Monte Carlo simulations to account for process variations. Our primary objectives were to assess the PUF's reliability under different environmental (thermal) conditions and voltage levels, as well as to analyze its sensitivity to variations in the number of activated rows by a challenge. We executed 15 simulations sessions, each comprising 10,000 runs, to ensure an extended exploration of the PUF's behavior. In our evaluation, we considered a range of temperatures, including 0°C, 27°C, and 80°C, to simulate a variety of operating conditions. Furthermore, we varied the supply voltage levels ±10% of the nominal value 1V (from 0.9V to 1.1V), to assess the PUF's performance under different power supply conditions. To provide a comprehensive analysis, we selected three distinct numbers of rows activated by a challenge,

8, 128 and 248, as they represent varying levels of complexity for the PUF. The proposed PUF array consists of 256 rows and 65 columns.

A small example of 10 Monte Carlo simulations is depicted in Figure 4.1 in which we observe that half of the responses rapidly goes to '0' and the other half goes to '1', as we expected.

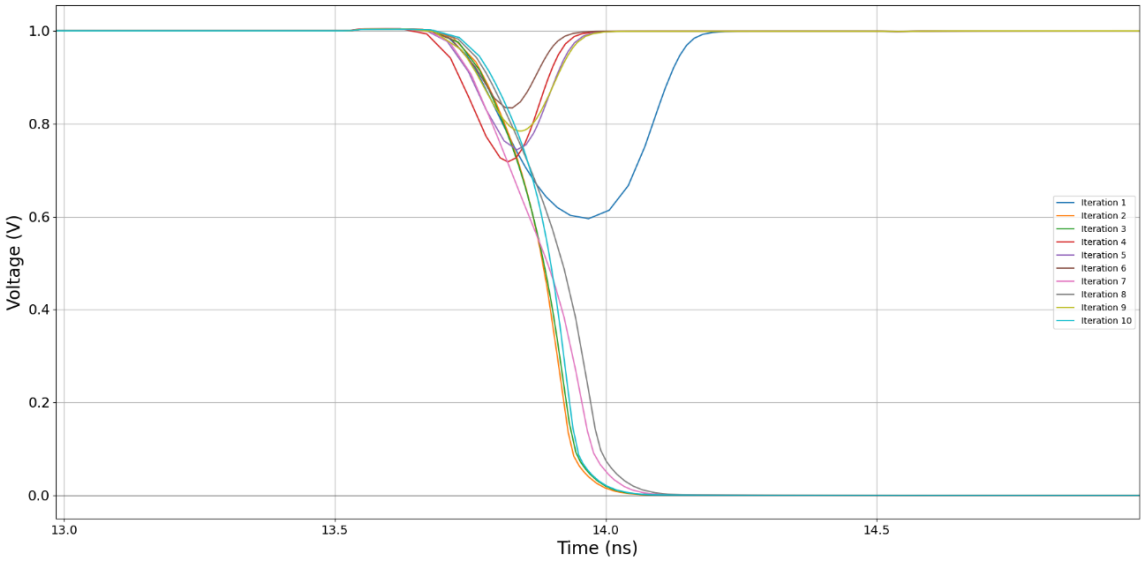


Figure 4.1: 10 indicative Monte Carlo simulations at nominal voltage and temperature levels of 1V and 27°C respectively, with a challenge activating 128 rows.

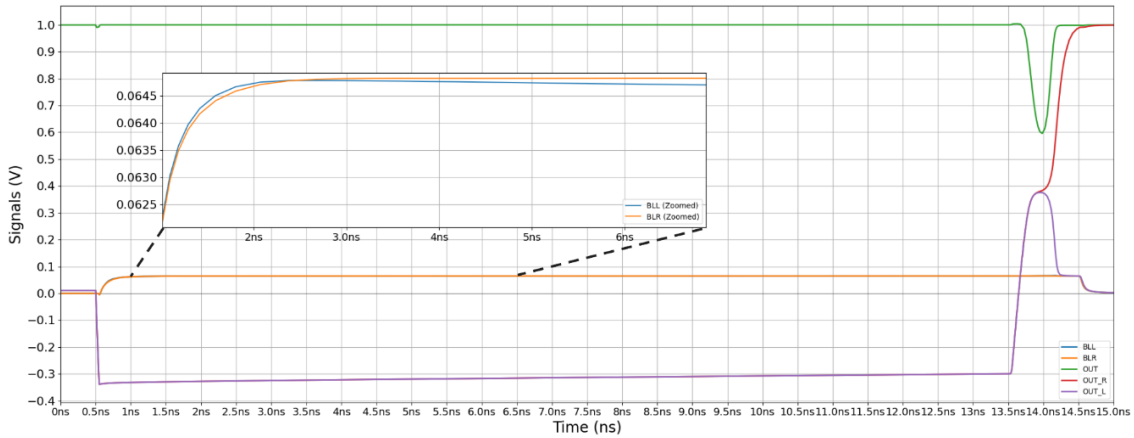


Figure 4.2: Simulation example at 1V and 27°C with 128 activated rows.

Figure 4.2 focuses on the first iteration of the Monte Carlo simulation of the Figure 4.1 and depicts a simulation example using a challenge of 128 activated rows at the nominal voltage (1V) and temperature (27°C). In the zoomed subplot, we notice that initially, the 'BLL' charges faster than 'BLR', but after a few nanoseconds, 'BLR' takes the lead. This example demonstrates that our PUF requires some time to stabilize during the activation phase, typically taking between 1ns to 13ns to settle.

Additionally, time is needed for the bitlines to create voltage differences for the comparator to operate correctly. In this example, stabilization begins in less than roughly 3ns. However, for challenges with fewer activated rows, where the current is lower and fewer PUF cells contribute, it takes more time to establish a stable voltage. Conversely, when more rows are activated, stabilization occurs more quickly. If 'BLL' is greater than 'BLR', 'OUT_L' goes to '1' and 'OUT_R' goes to '0', and vice versa. Consequently, 'OUT' switches to '0' or '1' respectively. In this example, 'BLR' is greater than 'BLL', causing 'OUT' to switch to '1'. However, since the challenge signal(s) 'CLi' remains active, the PUF cells are generating current continuously, resulting in a continuous feed to the comparator and preventing the outputs (OUT and OUT_B) from stabilizing to a clear '0' or '1'. To address this issue, we allow 1ns for the comparator to reach a stable state at the sense phase and then deactivate the input challenge signal by setting 'CLi' signal(s) to low at 14.5ns, making the comparator finally generate a clear response. The worst-case delay for the PUF response is estimated to be 15ns.

The types and sizes of the transistors in the PUF's structural components, including PUF cells, comparators, and buffers, influence the PUF's overall performance, necessitating careful selection of these transistors. For our PUF implementation, we use the commercial UMC90nm technology. Starting with the PUF cell, we select minimum size transistors with a width (W) of 120nm and a length (L) of 80nm to minimize the required silicon area. Additionally, we choose high threshold voltage (HTV) transistors to reduce static power consumption.

In contrast to PUF cells, the choice of the appropriate transistors of the comparator is more complicated. We aim to use a comparator that has low power consumption, fast response and does not influence the PUF response. In terms of timing, we choose low voltage threshold transistors (LVT) on the cross-couple inverters and the two inverters to minimize the sense phase and get a fast response. For the equalization transistors M7 and M8 we choose HVT transistors to minimize the leakage current between the bitlines. The size of M5 and M6 transistors of the comparator significantly influences the PUF's performance. As a compromise, we select transistors of wide size, with a width of 2.4um. The defined size of both transistors satisfies to a great extent most of the PUF challenge alternatives. The area occupied by our

current mode sense amplifier is 8.5 times that of a PUF cell and a detailed table outlining all transistor specifications is provided in Table 4.1.

Table 4.1: Sizes and types of comparator transistors.

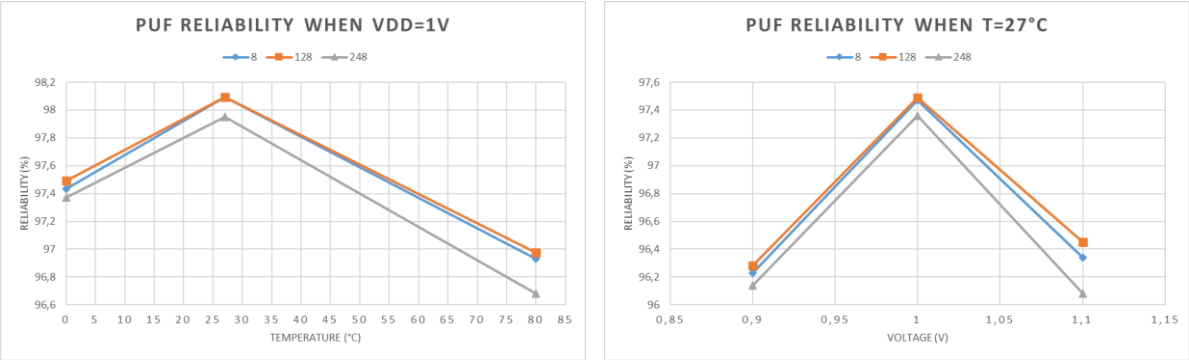
Id	Width	Length	Type
M1, M2, M10, M12	120nm	80nm	NMOS LVT
M3, M4, M9, M11, M13	120nm	80nm	PMOS LVT
M5, M6	2.4um	80nm	NMOS LVT
M7, M8	120nm	80nm	NMOS HVT

Less impactful but equally important is the buffer. Its delay is measured at roughly 50ps, which is deemed sufficiently small to warrant using only two inverters instead of a larger cascading sequence. Additionally, all transistors in our buffer have LVT transistors, thereby enhancing speed. The specification of transistors is detailed in the Table 4.2. For the first inverter, we utilized the minimum size of NMOS transistor paired with a PMOS transistor four times larger than the NMOS. This adjustment accounts for the fact that the mobility of holes is about four times less than that of electrons in the LVT UMC90nm technology that we use. Finally, the second inverter is approximately 3.6 times larger than the first inverter.

Table 4.2: Sizes and types of buffer transistors.

Id	Width	Length	Type
M1	480nm	80nm	PMOS LVT
M2	120nm	80nm	NMOS LVT
M3	1.73um	80nm	PMOS LVT
M4	430nm	80nm	NMOS LVT

Throughout the simulations, we monitored key performance metrics such as reliability, uniqueness, uniformity, response time, power consumption, CRPs number and occupied silicon area.



(a) (b)

Figure 4.3: PUF reliability results under different voltage (a) and temperature (b) conditions for 8, 128 and 248 activated rows per challenge input.

Table 4.3: Tables of PUF reliability on various temperatures, voltages and number of activated rows. Table (a) shows the reliability at nominal voltage V=1V, (b) shows the reliability at nominal temperature T=27°C.

V=1V	Reliabilities(%)		
Rows / T(°C)	8	128	248
0	97,435	97,49	97,37
27	98,095	98,095	97,95
80	96,93	96,975	96,68

(a)

T=27°C	Reliabilities(%)		
Rows / V(V)	8	128	248
0,9	96,23	96,28	96,14
1	97,47	97,49	97,36
1,1	96,34	96,45	96,08

(b)

The reliability results of our PUF are presented in both the Figure 4.3 and the Table 4.3. These results were derived by considering various temperatures and voltage values, followed by calculating the corresponding averages by the Monte-Carlo runs. As the results show, the highest reliability is always achieved when a challenge of 128 activated rows is applied in both temperature and voltage variations. On the other hand, this PUF presents worst-case reliability of 96,68% and 96,08% during temperature variations and voltage fluctuations respectively. Overall, our PUF reliability ranges between 96,68% and 98,085% under temperature variations with an average value of 97,45% and between 96,08% and 97,49% under voltage variations

with the average value of 96,65%. Furthermore, the average measured uniqueness and uniformity of the PUF were 49,99% and 49,85% respectively. It is noteworthy to mention that, for each challenge, the variation in reliability during voltage variations, Figure 4.3 (b), is less than 0.8%. This behavior was anticipated, as the PUF utilizes current generation cells with reduced supply dependency.

To evaluate how strong is the PUF, we must calculate the number of CRPs. In our implementation we can activate multiple rows simultaneously using any combinations of the 256 rows and a 64-bit response will be generated. So, in general terms we can define the CRPs number as:

$$CRPs = \sum_{k=1}^{n} \binom{n}{k} = 2^n - 1$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ and describes the number of combinations of n available rows and k chosen rows. Consequently, using n=256 we get a number of CRPs equal to $2^{256} - 1$. From the Table 4.3, we observe a trade-off between the number of activated rows and reliability, and it seems that reliability tends to decrease for more or less than 128 activated rows. Considering the large number of CRPs, it is feasible to exclude challenges with fewer or more activated rows, particularly focusing on challenges with exactly 128 activated rows (n/2), which yields the maximum number of CRPs. This number is equals to $\binom{256}{128} \approx 5.768658823 \times 10^{75}$.

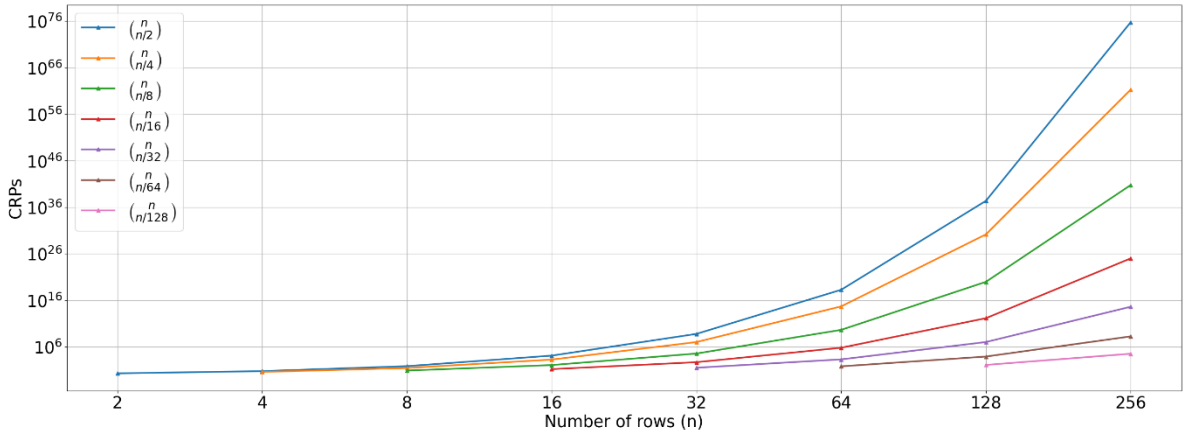


Figure 4.4: Plot of number of PUF array rows versus number of CRPs, choosing challenges with a number of activated rows ranging from n/2 to n/128.

The Figure 4.4, depicts that the number of possible CRPs increases exponentially as the number of available rows (n) in the array doubles, while the number of activated rows by a challenge range between n/2 and n/128. Initially, we observe that choosing challenges with n/2 activated rows yields the maximum number of possible CRPs $\binom{n}{n/2}$. Furthermore, by decreasing the array rows and consequently decreasing the silicon area of the PUF array, we can use challenges with more activated rows, increasing power consumption while achieving greater number of CRPs. For example, a PUF with an array of 64 rows and a challenge of 32 (n/2) activated rows achieves a significantly greater number of CRPs compared to a PUF with an array of 256 rows and a challenge of 8 (n/32) activated rows.

Furthermore, we also evaluated the performance of the comparator in terms of decision correctness. It was observed that for challenges with fewer activated rows (8), our comparator makes incorrect decisions in nearly 1% of the responses. However, when using challenges with more activated rows (128 and 248), the comparator always provides correct decisions under temperature and voltage variations. As a result, the adopted comparator we employ does not impact the overall reliability of the PUF when a challenge with a high number of inputs is applied.

When it comes to power consumption, we conducted multiple measurements depending on the number of activated inputs (challenge input). Our measurements covered the buffers, the PUF array, and the comparators. As indicated in previous sections, employing a challenge of 128 activated rows produced the most advantageous results. Thus, for our energy consumption assessments, we consider a challenge involving 128 activated rows. Thus, at 1V and 27°C, the resulting energy consumption was measured at 20.1pJ/bit (1.26pJ/bit for 8 activated rows, 38.93pJ/bit for 248 activated rows).

For the required silicon area estimation of the proposed PUF, we consider the area of the transistor gates ($W \times L$). Our PUF array comprises 256 rows and 65 columns of PUF, with each cell consisting of 6 transistors of minimum size (W=120nm and L=80nm). Thus, our PUF array occupies an estimated silicon area equal to 958.464μm². Utilizing 256 buffers to match the number of rows in the PUF array, we calculate an estimated total buffer area equal to 205.184μm². The current comparators that we use, occupies an estimated silicon area equal of 32.3136μm². Therefore, the total estimated silicon area occupied by our PUF is 1195.9616μm².

4.2 Comparisons

From the results in the previous chapter, we conclude that our PUF is characterized as strong, of low power consumption, low occupied silicon area and of high reliability especially with respect to voltage variations as it was one of the main goals of this thesis. It is important to note that our measurements on silicon area and the power consumption include both the buffers and the comparators, aspects not typically mentioned in other works. However, note that we are not considering aging or error correction methods. Most of the state-of-the-art PUF implementations that were described in Section 2, are gathered in the Table 4.4.

In many designs, details such as total area per bit and power consumption per bit are often omitted, making it challenging to compare our implementation with proposed PUFs in terms of silicon area and power consumption. In comparison to other PUFs, our proposed design demonstrates significant advancements in these aspects, as it combines a robust PUF with high reliability. Unlike many other PUFs that rely on SRAM topologies, which typically operate in PUF mode only during startup, our circuit can be exploited at any time. Additionally, while half of the designs are characterized as weak PUFs, which are susceptible to attacks as discussed in paragraph 2.3, most array type PUFs are considered strong PUFs due to their array structure, that makes them scalable. We follow the same aspect of array-like PUFs, which, in our case, provides a significant number of Challenge-Response Pairs (CRPs), enhancing the security and reliability of our PUF.

Table 4.4: Comparison table between state-of-the-art PUFs and the proposed one.

Design	APUF [10]	RO PUF [11]	SCA-PUF [12]	PTAT [2]	SRAM PUF [13]	SiCBit-PUF [14]	SRAM Array Current Based PUF [15]	PUF-CIM [16]	SPUF [17]	**This work**
Technology	45nm	FPGA 90nm	130nm	65nm	FPGA	32nm	45nm	55nm	65nm	**90nm**
Operation mode	Voltage	Voltage	Current	Current	Voltage	Voltage	Current	Voltage	Voltage	**Current**
Strongness	Weak	Weak	Strong	Weak	Weak	Strong	Strong	Weak	Weak	**Strong**
Total area (μm^2)	2168	NA	44700	7.42/bit	NA	NA	NA	395×10^3	12580	**1195.9616**
Temperature range (°C)	0~100	-20~120	-20~80	0~80	-20~80	0~100	10~85	-40~125	-10~80	**0~80**
Vdd range (V)	0.9~1.1	1.08~1.2	1.08~1.32	0.6~1.2	NA	0.9~1.1	NA	1.1~1.3	0.5~1	**0.9~1.1**
Reliability (T=c) (%)	97.01	NA	NA	99	NA	95.2†	NA	NA	NA	**96.65**
Reliability (V=c) (%)	94.49	NA	NA	96.5	96.43	97.4†	94.93†	NA	NA	**97.45**
Avg. Reliability (%)	NA	99.52	91†§ 99†•	NA	NA	98.2•	NA	98.9‡	97§	**NA**
Uniqueness (%)	49.99	46.15	49.9	50.01	49.97	49.99	49.97	49.97	49.47	**49.99**
Uniformity (%)	50.094	NA	52.8	49.3	NA	49.74	NA	49.76	50.11	**49.85**
Consumption (pJ/b)	~0.1036	NA	11	0.548	NA	NA	NA	0.248	0.081‡	**20.1**
CRPs	NA	Few	$\sim 2^{65}$ ‡	NA	110	16.76×10^6	NA	601080390	8.37×10^{17} ‡	$\mathbf{2^{256} - 1}$

†: **Worst case**, ‡: **Best case**, §: **Nominal Conditions**, •: **Discarded CRPs**

CHAPTER 5

CONCLUSIONS

In this thesis we propose a new design for Physical Unclonable Function (PUF) circuit, with interesting characteristics among state-of-the-art PUF schemes. We present a strong and reliable, with respect to temperature and voltage variations, current mode array PUF, with a special care on the power consumption and the occupied silicon area. The maximum number of possible CRPs is equal to $2^n - 1$ for an array of n rows. According to the simulation results on the design of the PUF in a 90nm technology, the estimated reliability under temperature variations is 97.45%, while the estimated reliability under voltage variations is 96.65%. The uniformity and the uniqueness of the PUF have been measured to be 49.85% and 49.99% respectively. A trade-off appears between CRPs and reliability, as the reliability slightly decreases when a challenge with more or fewer than n/2 activated rows is chosen. Maximum reliability is achieved when exactly n/2 rows are activated. The PUF is tuned to perform reliably for challenges with n/2 activated rows, which also provides a large number of CRPs. Since the proposed PUF is based on a power supply independent topology, its reliability is less sensitive on voltage variations. It is noteworthy to mention that our PUF is an array type PUF, meaning that is scalable.

As a future work, we may consider alternative current mode comparator topologies for the evaluation of the proposed PUF since this circuit may influence to some degree PUF's performance. Furthermore, different array architectures may be explored, exploiting the same supply independent current generation concept.

BIBLIOGRAPHY

[1] C. Böhm and M. Hofer, Physical unclonable functions in theory and practice, Springer Science \& Business Media, 2012.

[2] J. Li and M. Seok, "Ultra-compact and robust physically unclonable function based on voltage-compensated proportional-to-absolute-temperature voltage generators," vol. 9, pp. 2192--2202, 2016.

[3] C. Herder, M.-D. Yu, F. Koushanfar and S. Devadas, "Physical unclonable functions and applications: A tutorial," vol. 8, pp. 1126--1141, 2014.

[4] T. McGrath, I. Bagci, Z. Wang, U. Roedig and R. Young, "A puf taxonomy," vol. 1, 2019.

[5] X. Xu and W. Burleson, "Hybrid side-channel/machine-learning attacks on PUFs: A new threat?," in *IEEE*, 2014.

[6] J. Delvaux and I. Verbauwhede, "Fault injection modeling attacks on 65 nm arbiter and RO sum PUFs via environmental changes," vol. 6, pp. 1701--1713, 2014.

[7] H. Basel, "Physically Unclonable Functions—From Basic Design Principles to Advanced Hardware Security Applications. DOI: 10.1007".

[8] J. Yao, L. Pang, Z. Zhang, W. Yang, A. Fu and Y. Gao, "Design and Evaluate Recomposited OR-AND-XOR-PUF," 2021.

[9] A. Al-Meer and S. Al-Kuwari, "Physical unclonable functions (puf) for iot devices," vol. 14s, pp. 1--31, 2023.

[10] M. Moradi, R. Mirzaee and S. Tao, "CMOS arbiter physical unclonable function with selecting modules," in *IEEE*, 2020.

[11] G. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," 2007.

[12] H. Zhuang, X. Xi, N. Sun and M. Orshansky, "A strong subthreshold current array PUF resilient to machine learning attacks," vol. 1, pp. 135--144, 2019.

[13] J. Guajardo, S. Kumar, G.-J. Schrijen and P. Tuyls, "FPGA intrinsic PUFs and their use for IP protection," in *Springer*, 2007.

[14] A. Xynos, V. Tenentes and Y. Tsiatouhas, "SiCBit-PUF: Strong in-Cache Bitflip PUF Computation for Trusted SoCs," in *IEEE*, 2023.

[15] F. Zhang, S. Yang, J. Plusquellic and S. Bhunia, "Current based PUF exploiting random variations in SRAM cells," in *IEEE*, 2016.

[16] Z. Chen, M. Wu, Y. Zhou, R. Li, J. Tan and D. Ding, "PUF-CIM: SRAM-Based Compute-In-Memory With Zero Bit-Error-Rate Physical Unclonable Function for Lightweight Secure Edge Computing," 2023.

[17] L. Lu, T. Yoo and T. Kim, "A 6T SRAM based two-dimensional configurable challenge-response PUF for portable devices," vol. 6, pp. 2542--2552, 2022.

[18] B. Razavi, Design of analog CMOS integrated circuits, 清华大学出版社有限公司, 2005.

[19] T. Blalock and R. Jaeger, "A high-speed clamped bit-line current-mode sense amplifier," vol. 4, pp. 542--548, 1991.

SHORT BIOGRAPHY

Dimosthenis Georgoulas was born in 1999 in Ioannina, Greece. He graduated from the High School of Astros Arcadias and continued and completed his undergraduate studies at the Department of Computer Science and Engineering at the University of Ioannina in 2022. At the same year he became a MSc student at the same department. His main interests lie in the field of digital circuit design, but he is also interested in the field of Machine Learning.