ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

ĐÀO MẠNH HIỆP

NGHIÊN CỨU THIẾT KẾ BẢO MẬT THẺ RFID SỬ DỤNG MÃ HÓA ĐƯỜNG CONG ELLIPTIC

(Secure Lightweight RFID tag with Elliptic Curve Cryptography)

LUẬN ÁN TIẾN SĨ KỸ THUẬT ĐIỆN TỬ

HÀ NỘI – 2025

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

ĐÀO MẠNH HIỆP

NGHIÊN CỨU THIẾT KẾ BẢO MẬT THẺ RFID SỬ DỤNG MÃ HÓA ĐƯỜNG CONG ELLIPTIC

(Secure Lightweight RFID tag with Elliptic Curve Cryptography)

Chuyên ngành: Kỹ thuật điện tử
Mã số: 9520203

LUẬN ÁN TIẾN SĨ KỸ THUẬT ĐIỆN TỬ

NGƯỜI HƯỚNG DẪN KHOA HỌC:
1. GS. TS. Trần Xuân Tú
2. GS. TS. Vincent Beroulle

HÀ NỘI – 2025

DECLARATION OF AUTHORSHIP

I hereby declare that this thesis and the work presented in it are solely my own and have been generated as the result of my original research under the supervision of Prof. Xuan-Tu Tran, Prof. Vincent Beroulle, and Prof. Yann Kieffer. This work has not been submitted for any other degree or professional qualification except as specified.

ACKNOWLEDGMENTS

I extend my profound gratitude to Prof. Xuan-Tu Tran, Prof. Vincent Beroulle, and Prof. Yann Kieffer for their unwavering support, encouragement, and guidance throughout the project. I would like to express my gratitude to Dr. Duy-Hieu Bui for his indispensable guidance in creating the software and hardware simulation platform, as well as his insightful opinions during the peer-review process of my publications.

I acknowledge the collaborative support of colleagues at the Information Technology Institute, Vietnam National University, Hanoi (VNU-ITI), and researchers at the Center for Integrated Circuit and Application (CICA), for their support. I cannot fail to mention my gratitude to the staff and researchers at Laboratoire de Conception et d'Intégration des Systèmes (Valence, France) during the most difficult period of the pandemic. I also appreciate the assistance provided by the staff of the Faculty of Electronics and Telecommunications at UET, VNU.

Finally, I express my heartfelt gratitude to my family, Mô, and my friends for their unconditional support and encouragement emotional fortitude, and enduring patience, which sustained me through both academic and personal challenges.

Hanoi, August 2025

Manh-Hiep Dao

Contents

List of Abbreviations

Abbreviation	Description
ACB	Arithmetic Computation Block
AES	Advanced Encryption Standard
AI	Artificial Intelligence
ANN	Artificial Neural Network
ASIC	Application Specific Integrated Circuits
B2VP	Bottom-up Verification and Validation Process
BEC	Binary Edwards Curves
BWC	Binary Weierstrass Curves
CAGR	Compound Annual Growth Rate
CMOS	Complementary Metal-Oxide-Semiconductor
CPA	Correlation Power Analysis
DC	Direct Current
DLP	Discrete Logarithm Problem
DoS	Denial-of-Service
DRC	Design Rule Check
DSE	Design Space Explore
DSP	Digital Signal Processors
ECC	Elliptic Curve Cryptography
ECDH	Elliptic Curve Diffie-Hellman
ECDSA	Elliptic Curve Digital Signature Algorithm
EDA	Electronic Design Automation
EM	ElectroMagnetic

ESL	Electrical System Level
FPGA	Field Programmable Gate Array
FSM	Finite State Machine
GDSII	Graphic Design System II
GF	Galois Field
GHC	Generalized Hessian Curve
GNB	Gaussian Noram Basis
HDL	Hardware Description Language
IC	Integrated Circuit
ISO/IEC	International Organization for Standardization/International Electrotechnical Commission
LVS	Layout vs Schematic Check
MCU	Microcontroller Unit
MIMA	Man In The Middle Attack
MitM	Meet-in-the-Middle Design Methodology
MOSFET	Metal-Oxide-Semiconductor Field-Effect Transistor
PA	Power Analysis
RBD	Reverse Bias Diode Current
RFID	Radio Frequency Identification
RNG	Random Number Generator
RTL	Register-Transfer Level
SCA	Side-channel Attack
SoC	System on Chip
TVLA	Test Vector Leakage Analysis
VCD	Value Change Dumped format
VLSI	Very Large Scale Integration

List of Symbols

Symbols	Description
$\{C_1, C_2, \cdots, C_n\}$	Token Messages communicated between Tags and a Reader
r, k	Random generated scalar number
X, Y, K	Public points over the defined curve
$h(a, b)$	Hash function of concatenating message of a and b
$P(x_P, y_P)$	Affine Coordinates (x_P, y_P) of point P
x_i, y_i	Affine Coordinates
X_A, X_B	Mean of power traces A and B
S_A, S_B	Standard deviation of power traces A and B
N_A, N_B	Number of sampling point in the corresponding power traces
T	Welch's T-statistic
a_i, b_i	Individual sample points indexed i of the power traces A and B
ω_i	Differential coordinate of affine coordinates x_i and y_i
$\mathbb{K}, GF(2^m)$	Extended Binary Finite Field with m-order of field
m	Number of representation bits of the secret key
(W_i, Z_i)	Projective form of Differential Coordinate ω_i
$f(x)$	Irreducible Polynomial of the Galois Field
$f(a, b, \texttt{cond})$	Mathematical expression of proposed Arithmetic Computation Block (ACB) with control signal `cond`
$a(x), b(x)$	Polynomial expression of scalar a, b over the Galois Field
$T_{parallel}, T_{sequential}$ T_{total}	The cumulative latency of parallel, sequential, and top-level circuits measured in the number of clock cycles

$T^{ECC}, T^{hash}, T^{RNG}$	Latency of security primitives, quantified as the number of clock cycles
T_{norm}	Normalization of the latency for an independence circuit
A_{RFID}	Total physical area of the passive RFID tag
$A^{ECC}, A^{hash}, A^{RNG}$	Individual area of security primitives, quantified as the number of equivalent gates
E_{RFID}	Total energy consumed by the passive RFID tag
$E^{ECC}, E^{hash}, E^{RNG}$	Individual energy consumed by security primitives, quantified as micro joules
$P_{parallel}, P_{sequential}$, P_{total}	Power consumption of parallel, sequential, and the top-level circuit
$P^{ECC}, P^{hash}, P^{RNG}$	Individual power consumption of security primitives
$P_{dynamic}, P_{static}$	Dynamic and static power dissipation of the circuit
$P_{switching}$, $P_{short\_circuit}$	Power consumed by the switching activities and short circuit phenomenon
P_{norm}	Normalization power consumption of the independence circuit
$n_{ECC}, n_{hash}, n_{RNG}$	Number of sub-blocks dedicated to implementing specific field operator
α	Ratio of switching activity, by default, is 0.5
V_{DD}, C_L, f	Supply voltage, total load capacitance, and the operational frequency of the circuit
I_{leak}	Leakage current of the circuit
$I_{sub}, I_{gate}, I_{RBD}$	Leakage current caused by intrinsic phenomena of sub-threshold, gate leakage, and reverse bias diode
p_{ij}	Power consumption of operator i at a specific time instance $t = j$
$h(t)$	Impulse response of the digital circuit
$f_i(t)$	Data of the digital circuit at state i
$f_{in}(t), f_{out}(t)$	Input and Output data of the top-level circuit
`HD(a,b)`	Hamming distance of a and b
`HW(a)`	Hamming Weight of a
T_{limit}	Limitation of latency

List of Figures

List of Tables

Introduction

Radio Frequency Identification (RFID) has been the breakthrough of wireless authentication technology for the past few decades. Initializing in World War II (WWII) by the British, American, and German Air forces, RFID technology has been utilized for identifying and authenticating whether an aircraft belonged to the Allies or the Nazis. After the victory of WWII, RFID technology was developed for widespread use in several civil applications, such as supply chain management and access control, due to its benefits of low cost and high reliability.

Generally, in the RFID system, there are two main parties: tags and a reader for authentication. Depending on the particular application, the tags could utilize the internal battery (active RFID tags) or energy harvesting (passive RFID tags). Reducing the internal battery helps the device become smaller and cheaper. Besides, utilizing the passive RFID tags helps us remove the need for battery replacement. As a result, passive RFID tags significantly contribute to the development of hand-held applications such as e-passports, civil cards, contactless payment cards, and inventory tracking. Therefore, the market for total RFID, particularly passive RFID tags, is expected to grow significantly in the next few years.

Although the passive RFID technology, which is deployed in several critical applications, offers significant advantages, it still exhibits several inherent limitations. Their dependence on wireless communication and physical hardware renders them vulnerable to exploitation through both electromagnetic interception and direct tampering with embedded components. To mitigate these vulnerabilities, integrating embedded cryptographies within the secured protocols into passive RFID tags is critical to secure user data that is transmitted over the communication channels. Asymmetric cryptography has emerged as a robust solution, leveraging its unique properties. Notably, in this authentication scheme that utilizes asymmetric cryptography, each party of the authentication communication (e.g., tags and reader) will maintain a pair of matching keys: a private key ($k_{private}$), stored securely on the device and a public key (k_{public})

shared network-wide. Among asymmetric cryptographic methods, Elliptic Curve Cryptography (ECC) is notable for achieving equivalent security levels to other algorithms while requiring significantly lower computational complexity. The deployment of Elliptic Curve Cryptography (ECC) on passive RFID tags presents significant challenges, primarily due to the following constraints:

1. **Physical resource limitations:** Passive RFID tags face strict restrictions in power consumption, area cost, and latency. As the passive RFID tag harvests energy from the electromagnetic field over the rectifier antenna (rectenna), its operational power is limited by the efficiency of energy conversion. Besides, the communication time between the tag and the reader is quite short and standardized, which caps interaction time at 20 milliseconds, according to the ISO/IEC-14443, severely restricting available energy. Therefore, the harvested energy of the tag is significantly constrained. Additionally, minimizing the design's physical footprint is critical to reducing manufacturing expenses.

2. **Security requirements:** Protection approaches for the passive RFID tags to be robust against wireless and hardware attacks present a complementary complexity in processing. In particular, at the protocol layer, addressing the wireless attacks, countermeasures for ECC-based authentication protocols tend to implement additional security primitives, such as random generators, hash functions, and supplementary scalar multiplication steps. Consequently, the implementation costs of the system dramatically increase. Regarding hardware challenges, hiding or masking techniques are designed to prevent side-channel attacks that could expose secret keys. These techniques require additional calculations during processing to minimize the leakage of the secret key through side channels.

This thesis aims to propose an efficient hardware design of an ECC-based authentication protocol, which is the solution to the above-mentioned challenges. The proposed hardware design balances the physical resource limitations and also the security requirements. Firstly, we propose a novel algorithm with the corresponding architecture for Binary Edwards Curves (BEC). The area footprint and power consumption must be measured and compared to the state-of-the-art to demonstrate the efficiency of the proposed method.

Nevertheless, employing the ECC-based authentication protocol is processed over multiple design layers, from the protocol to the security primitives; there is a wide range of choices, such as communication schemes, algorithms, or architectures in each layer. Therefore, the design space of the ECC-based authentication protocol is complicated for designers to explore. Besides, there are issues of multi-object optimization in a

wide design space, as the ECC-based authentication protocol raises the time-to-market, which also impacts production expenses. Addressing this problem, we propose an innovative design methodology named Early Evaluation Meet-in-the-Middle (EEMitM), which enables designers to quickly evaluate and choose the most compatible design with the input requirements. The dissertation covers the following key contributions:

1. We proposed and implemented a low-cost, low-power Elliptic Curve Cryptography hardware architecture using Binary Edwards Curve (BEC), targeting the passive RFID tags by utilizing the conventional Top-down design methodology. The synthesized BEC that uses the TSMC CMOS 65nm technology shows the benefits of minimizing power consumption and physical area footprint. Besides, the proposed design is also secured against the Side-Channel Attack by passing the Test Vector Leakage Analysis (TVLA). This work is published in [C1].

2. To address the issue of time-to-product, an innovative EEMitM design methodology is proposed. The goal of this proposal is to provide a framework that helps designers quickly choose the most compatible design that meets multiple objectives of the constraints. To further reduce the effort for evaluating the security, we also propose a new approach to generate the pseudo power trace. This research is published in [C2, C3, J1, P1].

The thesis is divided into four major chapters to discuss these contributions in more detail. In chapter 1, the dissertation gives a brief introduction to the RFID system and the related security issues. This also demonstrates the motivation and objective of the thesis.

Chapter 2 presents an overview of the state-of-the-art design methodologies for hardware implementation. Conventional approaches, such as Top-down and Bottom-up, are quickly reviewed before an analysis of emerging AI-driven design methodologies. In the following, a particular design flow for hardware security design is presented to understand the security evaluation approaches that are supplemented into the mentioned conventional approaches.

Chapter 3 demonstrates our proposal of a low-cost, low-power point multiplication over Binary Edwards Curve (BEC), targeting the passive RFID tags. This proposal is employed according to the conventional design methodology, the Top-down approach. This design is synthesized by using the CMOS TSMC 65nm technology. The synthesized results show the compatibility of the proposed design with the resource constraints of passive RFID tags.

The last chapter focuses on the innovation of EEMitM design methodology for ECC-based authentication protocol for passive RFID tag applications. Firstly, the general

illustration of the proposed EEMitM is presented to provide systematic insights into the proposal. In the following, the complementarities of the early estimation and evaluation approaches are demonstrated in detail.

Finally, the thesis is concluded with a summary of our main contributions. Future works include the packaging plan and further development of the proposed EEMitM design methodology as an Electronic Design Automation framework.

Chapter 1

RFID System and Security Issues

1.1 General Concepts of Secured Passive RFID Systems

1.1.1 Authentication System by Radio Frequency

Radio Frequency Identification (RFID) is a technology that leverages radio frequency signals to transmit data, enabling the identification and authentication of objects within a system. In a general system, there are two common elements: a Reader and Tags, as demonstrated in Figure 1.1. The tag uses radio waves to broadcast its

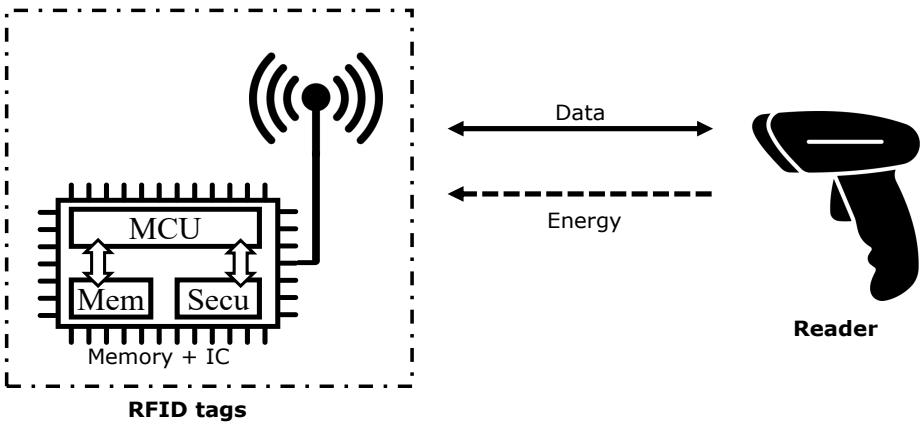


Figure 1.1: General Authentication System by Radio Frequency.

stored information to a dedicated reader device. Upon receiving the transmitted data, the Reader employs a predefined authentication protocol to verify the object's identity. If the information aligns with authorized parameters, the system grants the object access privileges or triggers subsequent actions, such as logging its presence or enabling entry. This process ensures secure, efficient interaction between physical objects and digital systems, with applications ranging from inventory management to secure access control, covering both military and civil applications.

1.1.2 Radio Frequency Identification Tags

RFID systems are primarily divided into passive and active configurations, each tailored for different applications. Active RFID tags, equipped with an integrated battery, can transmit signals across extended distances, spanning hundreds of meters, making them indispensable in logistics for real-time tracking of assets like shipping containers. Both active and passive systems adhere to standardized protocols (e.g., ISO/IEC 14443 for proximity-based cards or EPCglobal for supply chain management) to guarantee seamless compatibility between tags and readers worldwide.

In contrast, passive RFID tags, as depicted in Figure 1.1, operate without an internal power source. Instead, they draw energy from the radio waves emitted by the Reader to activate and transmit data. These low-cost, maintenance-free tags are ideal for short-range uses. As a consequence, they are popular in a variety of applications, such as retail stock management or contactless payment systems like credit cards.

Driven by these technological advantages, the global RFID market has experienced remarkable growth. The efficiency of active tags in long-range logistics and the cost-effectiveness of passive tags in retail and payment systems have fueled widespread adoption. According to a 2024 report by Market Data Forecast [1], the RFID sector is projected to expand at a compound annual growth rate (CAGR) of 7.08% through 2030, reaching a valuation of USD 25.93 billion in 2033. The Asia-Pacific region, strengthened by its large youth population and rapid technological integration, is anticipated to dominate this growth, with its market value expected to surge to USD 4.5 billion by 2028 at an impressive CAGR of 9.02%. This development highlights how RFID's versatility, from inventory tracking to secure authentication, is reshaping industries globally in the following decades.

1.2 Challenges of Passive RFID Tags

Although RFID systems have significantly advanced due to their benefits, they also face several drawbacks. This section provides a brief overview of passive RFID tags in terms of implementation costs and security.

1.2.1 Limitations related to physical implementation costs

As previously outlined, passive RFID tags operate by harvesting energy from electromagnetic signals emitted by the reader. This energy is captured by the tag's antenna, converted into direct current (DC) via a rectifier circuit, and then distributed to critical components like the microcontroller unit (MCU), which controls the processor of security primitives and memory, as described in Figure 1.1. Consequently, the overall efficacy of passive RFID systems hinges on the efficiency of this energy-harvesting process, particularly the rectifier's ability to convert radio frequency (RF) signals into usable power.

According to a report by Xu *et al.* [2], the state-of-the-art rectifier is reported to achieve peak conversion efficiencies of 45%, with a maximum output power of $-10dBm$. Given that passive RFID tags typically consume no more than 100 μW during operation, optimizing rectifier performance is paramount. However, designing

high-efficiency rectifiers involves overcoming material and structural challenges, such as impedance matching and minimizing parasitic losses. A key optimization strategy consists of reducing the system's power consumption, particularly in energy-intensive modules like encryption engines. To meet stringent power budgets, designers often simplify cryptographic algorithms or employ low-power hardware architectures.

Further complicating design efforts are industry requirements for rapid authentication. The ISO/IEC 14443 standard [3] specifies that authentication must be completed within 20 milliseconds, prompting the use of parallel processing architectures in encryption modules. While parallelism reduces latency, it significantly increases power consumption, forcing designers to strike a delicate balance between speed and energy efficiency. Additionally, cost constraints drive efforts to minimize the physical footprint of the MCU and other integrated circuits. These multiple objectives focusing on power efficiency and compact design underscore the multifaceted challenges in developing next-generation passive RFID systems tailored for applications like contactless payments and IoT devices.

1.2.2 Security Challenges for Passive RFID Tags

In other aspects, such as wireless communication, passive RFID tags are vulnerable to various threats. In the recent research, Juel *et al.* [4] described several vulnerabilities of passive RFID tags. They are categorized into two groups based on the objects to be attacked: hardware and wireless attacks. These vulnerabilities commonly exploit the weakness in the authentication algorithms or cryptographic mechanisms to compromise the security features, which are outlined in Table 1.1.

1.2.2.1 Wireless Attacks

Wireless vulnerabilities target the private information exchanged over wireless communication channels. These attacks generally pursue two objectives: compromising privacy or undermining system security. Privacy-focused attacks involve eavesdropping on communications or tracking RFID tags to extract sensitive data. In contrast, security-focused attacks aim to impersonate legitimate tags or readers to execute fraudulent transactions, such as unauthorized access or data manipulation.

Based on these objectives, wireless attacks are categorized into passive and active attacks [4]. Passive attacks, such as eavesdropping, intercept transmitted messages without altering their content. Active attacks, however, directly modify or replace messages on the communication channel.

Table 1.1: Security Characteristics of Passive RFID Tags.

Characteristics	Definition
Mutual authentication	Requires both the tag and the reader to authenticate each other.
Confidentiality	Ensures the secret key is accessible only to authorized parties.
Anonymity	This protects against the discovery and misuse of identity.
Availability	Ensures the RFID system remains reliably available
Scalability	The ability to handle a large number of tags without overload. A scalable RFID protocol should avoid performance degradation as tag numbers increase.
Forward security	This characteristic ensures that all the previous secret keys cannot be recovered if the long-term key or current session key is compromised.
Location privacy	Prevents disclosure of users' location and movements to untrusted entities.
Data integrity	Ensures transmitted data is not modified.

Additionally, wireless attacks can be classified by their targets: tag-level attacks, channel-level attacks, and system-level RFID attacks. Tag-level attacks are particularly concerning for passive RFID systems. Due to the stringent cost and power constraints of passive tags, robust encryption algorithms cannot be implemented, leaving them vulnerable to attacks that exploit the secret data.

1.2.2.2 Hardware Attacks

In RFID systems, security algorithms and sensitive data are processed on shared hardware components such as integrated circuits (ICs), digital signal processors (DSPs), and registers. During operation, these components emit side-channel information—including power consumption, processing time, and electromagnetic emissions. This information directly correlates with cryptographic processing operations. Attackers exploit these leaks by collecting and analyzing the data to deduce cryptographic keys, thereby compromising the system's security.

Side-channel attacks target hardware leakage sources, such as power consumption or electromagnetic radiation, particularly in passive RFID tags using lightweight cryp-

tographic algorithms. Due to strict cost and power constraints, these tags often can not implement adequate countermeasures, making them vulnerable to threats that exploit correlations between secret keys and hardware behavior, such as execution time, temperature fluctuations, or power draw to reverse-engineer keys. Among these techniques, Power Analysis (PA) [5], which analyzes variation in power consumption during cryptographic execution, is one of the most potent and widely used methods. By isolating patterns in power traces, attackers can break encryption even in resource-constrained environments. It is, therefore, necessary to implement complementary countermeasures that mask or hide the secret key in the passive RFID tag.

1.3 Security Strategies for Passive RFID Tags

Addressing these vulnerabilities, the popular countermeasures for passive RFID tags are implemented at several levels of the system. The countermeasures applied in the Authentication protocols help secure communication against wireless attacks. Meanwhile, the implementations at the Cryptography primitives level enable the system to be robust against hardware attacks.

1.3.1 Authentication and Identification Protocols

For RFID systems, the security aspect is a comprehensive concern that araises from the combination of multiple mechanisms, such as cryptographic algorithms, random generators, and authentication protocols for multiple RFID tags, readers, servers, and end-user applications. Within the scope of this thesis, the concerned objectives are limited to exploring the security approaches for the hardware-constrained passive RFID tags against the wireless and hardware vulnerabilities.

The authentication protocol establishes a structured exchange of messages between a device seeking verification (RFID tag) and the verifying device (Reader). These interactions confirm the validity and ownership of encrypted tokens (Token Messages) managed by the reader. In some security setups, the protocol also ensures the RFID tag can authenticate the reader's identity, guaranteeing that the tag interacts only with authorized devices.

Figure 1.2 presents an example of an authentication protocol proposed by Chou *et al.* [6]. In this scheme, the RFID system communicates with five token messages $\{C_0, C_1, C_2, C_3, C_4\}$. On the RFID tag side, the device implements an Elliptic Curve Cryptography (ECC) with a hash function.

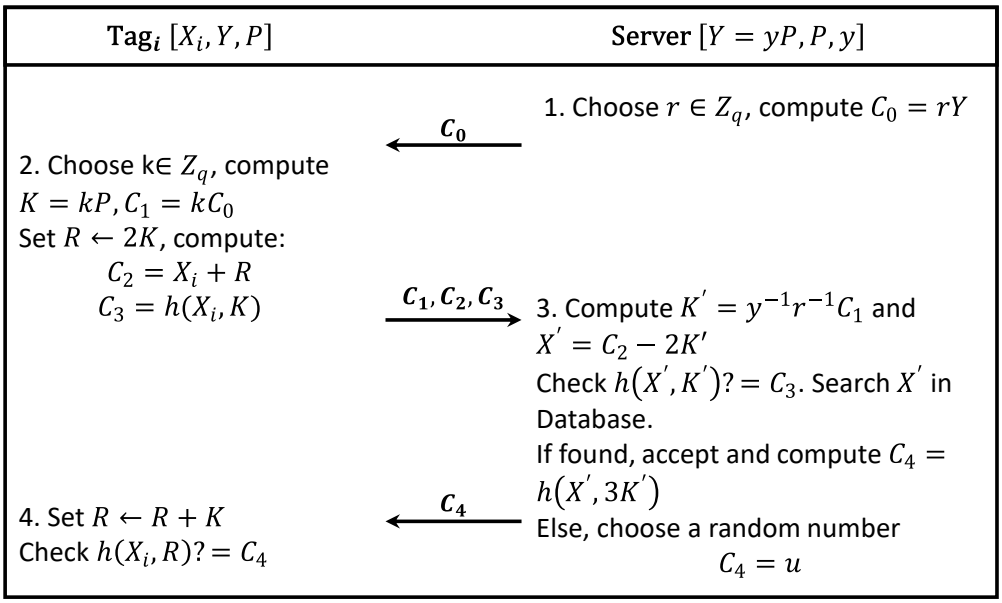


Figure 1.2: An example of an authentication protocol for a passive RFID tag.

In optimization strategies, the choice of computational steps depends on the system's security policies, operational constraints, and cost considerations. Thus, evaluating an authentication protocol's efficiency involves assessing the quantity and complexity of cryptographic operations and their relative resource constraints. Additionally, passive RFID tags operate under strict constraints on power consumption and hardware resources. This makes it critical to optimize both processing time and communication costs. Processing time refers to the minimum duration for a tag to complete authentication. The amount of data transmitted directly affects communication costs. These costs not only reflect the energy consumed during communication but also affect the storage capacity needed for passive tag hardware. Therefore, it is critical to design an authentication protocol that satisfies both security requirements and resource constraints.

1.3.2 Cryptography primitives

At another level of countermeasures implementation, cryptography primitives are used to encrypt data with a secret key. This processing ensures data integrity as described in Table 1.1. Based on the encryption mechanism, they are categorized into two families: symmetric and asymmetric, as presented in Table 1.2.

Symmetric Cryptography is the encryption/decryption algorithm that utilizes the same secret for both encryption and decryption. Table 1.2 shows that with the same security level, the symmetric algorithms generally require a shorter length of key com-

Table 1.2: Security Comparison of the General Cryptographic Mechanism with security level of 128 bits.

Family	**Cryptographic Algorithms**	**Resists Key Distribution**	**Key length (bits)**
Symmetric	AES [7]	No	128
	PRESENT [8]	No	128
	ASCON [9]	No	128
Asymmetric	ECC	Yes	256
	RSA	Yes	3072

pared to the asymmetric ones. As a result, the main benefit of symmetric cryptography, such as AES [7, 10], PRESENT [8], ASCON [9], is the low cost of implementation, compared to the asymmetric mechanism, as demonstrated in Table 1.3. Thus, symmetric algorithms are well-suited for resource-constrained applications, such as passive RFID Tags. However, sharing the same secret key introduces a risk that an attacker may derive the secret key from the token message, which is used to establish communication between parties the shared secret key between parties. As a consequence, subsequent communication may be exposed to the attackers.

Table 1.3: Hardware Cost Comparison of General Cryptographic Mechanism.

Designs	**Tech. node (nm)**	**Freq. (MHz)**	**Area cost (kGEs)**	**Power** (μW)	**Throughput** ($Mbps$)
AES [7]	65	10	8.6	20	28
PRESENT [8]	180	100	1.608	272.67	130.612
ASCON [9]	90	1	3.75	15	14
ECC [11]	65	1	10.95	-	0.917
RSA [12]	130	11.9	$1.283mm^2$	11.430	-

Addressing the drawbacks of symmetric algorithms, asymmetric cryptography utilizes encryption/decryption based on a private key ($k_{private}$) and a public key (k_{public}). The other devices use the public key k_{public} to encrypt the token message before broadcasting it. Only the legal reader, who owns the private key $k_{private}$, can decrypt the

received message to authenticate the tag. The key limitation of asymmetric cryptography is the high complexity of the computation compared to symmetric algorithms. As a result, the implementation cost of hardware utilizing asymmetric algorithms is significantly higher than that of lightweight symmetric cryptography, as shown in Table 1.3. Among the asymmetric algorithms, Elliptic Curve Cryptography (ECC) is considered the most lightweight cryptography.

1.4 Elliptic Curve Cryptography

1.4.1 Overview of Elliptic Curve Cryptography

The Elliptic Curve Cryptography (ECC) algorithm was first introduced in 1985 and relies on curves defined by Equation (1.1).

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{1.1}$$

where the set of coefficients $(a_1, a_2, a_3, a_4, a_6)$ characterizes the elliptic curve. The key operation in ECC is the point multiplication, which is shown in Equation (1.2).

$$Q = k \cdot P \tag{1.2}$$

The initialization point $P(x_P, y_P)$ defines the base point of the defined curve. The public key $Q(x_Q, y_Q)$ is generated by using the private key k. All the applications of the ECC are based on the Discrete Logarithm Problem (DLP). DLP is defined according to Equation (1.2) that even if the attackers know the points P and Q, the reverse derivation of the secret key k remains computationally infeasible.

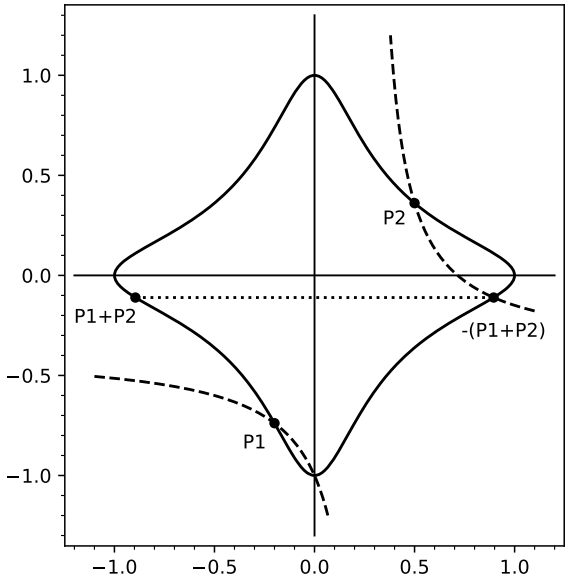
Among different curves, the Edwards Curve, which is represented as Equation (1.3), emerges as a secure curve due to its completeness characteristic.

$$x^2 + y^2 = 1 + d \cdot (xy)^2 \tag{1.3}$$

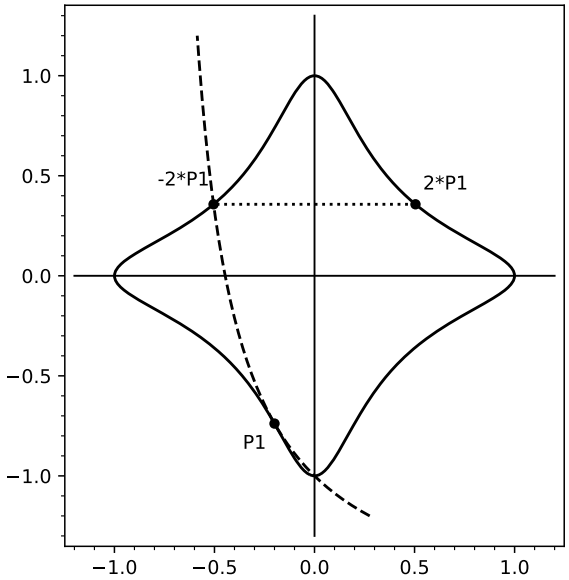
where d is not a square in the Finite Field. To compute the point multiplication in Equation (1.3) over the Edwards Curves, point addition and point doubling are processed iteratively. For two distinct points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$, which are defined within Edwards Curves, result of point addition yields $P_3(x_3, y_3)$. The addition law is presented in Equation (1.4):

$$\begin{aligned} x_3 &= \frac{x_1y_2 + y_2x_1}{1 + dx_1x_2y_1y_2} \\ y_3 &= \frac{y_1y_2 - x_1x_2}{1 - dx_1x_2y_1y_2} \end{aligned} \tag{1.4}$$

With the non-square value of d, the denominators of x_3 and y_3 are non-zero even in the case that the points P_1 and P_2 are identical. As a result of this property, the point doubling $P_4(x_4, y_4) = 2 \cdot P_1$ can be computed via the addition law, which is represented in Equation (1.4). Therefore, the addition law over Edwards Curves is complete. The visualization of the point addition and point doubling over the Edwards Curve $x^2 + y^2 = 1 - 19 \cdot x^2y^2$ is illustrated in Figure 1.3.



(a) Graph Expression of Point Addition in BEC.



(b) Graph Expression of Point Doubling in BEC.

Figure 1.3: Point Operators used in Edwards Curve of $x^2 + y^2 = 1 - 19 \cdot x^2y^2$.

Let $\mathbb{K}$ be a Binary Finite Field GF(2^m). A general Edwards curve over $\mathbb{K}$ is defined

as Equation (1.5):

$$E : d(x + y + x^2 + y^2) = xy + xy(x + y) + x^2y^2 \tag{1.5}$$

with d are the elements of $\mathbb{K}$ satisfying $d_1 \neq 0$ and no element $t \in \mathbb{K}$ that $t^2 + t + d = 0$. The addition law $P_3(x_3, y_3) = P_1(x_1, y_1) + P_2(x_2, y_2)$ is defined by Equation (3.6) with $A = x_1 + x_2, B = y_1 + y_2, C = x_1 + y_1, D = x_2 + y_2, E = x_1 + x_1^2, F = y_1 + y_1^2, G = x_1x_2, H = y_1y_2$:

$$\begin{aligned} x_3 &= \frac{d(A + C \cdot D) + E(x_2 \cdot (B + 1) + H)}{d + E \cdot D} \\ y_3 &= \frac{d(B + C \cdot D) + F(y_2 \cdot (A + 1) + G)}{d + F \cdot D} \end{aligned} \tag{1.6}$$

In the case of $P_1 = P_2$ or $(x_1, y_1) = (x_2, y_2)$, the denominator of the Equation (1.6) is non-zero. Consequently, point doubling could be performed via the addition law, ensuring point operations' homogeneity. As a result, the point multiplication based on BEC is secure against the Simple SCA and Differential SCA.

While Edwards Curves, especially Binary Edwards Curves (BEC), offer advantages in security properties compared to the other incompleteness curves in hardware implementation, they still struggle to meet the implementation cost and energy constraints of devices and incredibly passive RFID tags. Recent research prioritizes optimizing BEC hardware implementation costs. Additionally, efforts focus on enhancing security features to resist advanced side-channel attacks targeting ECC hardware by additional countermeasures such as masking or hiding strategies. However, the hardware of the supplementary countermeasures also increases the total cost of the system and possibly exceeds the physical resource constraints of passive RFID tags. Once again, this highlights a research gap in developing hardware-based passive RFID tags that balance the security level and the implementation costs.

1.4.2 Related works

1.4.2.1 Optimizing the implementation cost on ECC

In the literature, the selection of lightweight curves to minimize the computational complexity of the Elliptic Curve Cryptography has shown an attractive benefit in hardware implementation. General Weierstrass curves, also known as Binary Generic Curve, which require $6M + 4S$ for a point operator with M, S denoting the field multipliers and field squares, are widely used in various security standards for authentication and digital signature as the advantages in the least computational complexity according to Table 1.4.

To minimize the hardware implementation cost of the point multiplication block, Wenger [16] showed that the Binary Field $GF(2^m)$ is cost-effective compared to the prime field $GF(p)$. Among field operations (multiplier, squarer, inversion), Deschamps *et al.* [17] pointed out that field inversion is the most costly operation in terms of latency and power consumption. Thus, to avoid using the field inversion, the projective coordinates such as Lopez-Dahab [13, 18], Jacobian [19], and mixed projective coordinates [11, 13–15] are recommended, as described in Table 1.4. Besides, to reduce the number of operators, the Differential Addition Chain algorithm enables computing on the $\omega-$ coordinate, where $\omega_i = x_i + y_i$, instead of on two coordinates (x, y). However, recovering from the $\omega-$ coordinate to the affine coordinates (x, y) is much more complicated because of the use of the half-trace computation [11, 20].

Polynomial basis-based field operators are also widely used instead of the Gaussian Normal Basis (GNB) to optimize the area and energy consumption of the field operator modules. Meanwhile, the field square using the GNB is implemented as a shift register. The field multiplier over GNB is implemented as a matrix multiplication, which requires a large amount of storage and energy to calculate the multiplier in the finite field [21].

1.4.2.2 SCA Security Robustness Improving on ECC

Point multiplication in Elliptic Curve Cryptography (ECC) involves the repeated execution of either the point doubling followed by the point addition, or only point doubling, depending on the corresponding bit value of the cryptography key. This algorithm is known as the conventional Double-and-Add, which is presented as Algorithm 1. Consequently, the use of incomplete curves, such as the Weierstrass Curve or the Huff Curve, results in distinct representations or equations of point operations. Therefore, the point multiplication, which is executed by these point operations, is

Table 1.4: Implementation cost of different binary elliptic curves in alternative coordinate systems.

Curve	Coordinate System	Cost of Point Multiplication
Binary Generic Curve [13]	Mixed	$6M + 4S$
Binary Edward Curve [11]	Mixed	$6M + 4S$
Binary Edward Curve [14]	Affine	$I + 11M + 4S$
Binary Edward Curve [14]	Projective	$16M + S$
Generalized Hessian Curve [15]	Mixed	$9M + 4S$

* I, M, S denote the field inverter, multiplier, and squarer, respectively.

Table 1.5: Security completeness properties of different binary elliptic curves in alternative coordinate systems.

Curve	Binary Generic Curve [13]	Binary Edward Curve [11]	Binary Edward Curve [14]	Generalized Hessian Curve [15]
Completeness	×	✓	✓	✓
Implementation cost	$6M + 4S$	$6M + 4S$	$I + 11M + 4S$	$9M + 4S$

* I, M, S denote the field inverter, multiplier, and squarer, respectively.

Algorithm 1: Conventional Double and Add Algorithm of ECC Point Multiplication using Affine Coordinates [22].

Input: Generator point $P(x_P, y_P) \in GF(2^m)$ and scalar $k = (k_0 k_1 \cdots k_m)_2$
Output: Point $Q(x_Q, y_Q) = kP$

```
Initialize Q ← P;
for i ← m − 2 to 0 do
    Q ← 2 · Q;
    if k_i == 1 then
        Q ← Q + P;
return ω_kP = ω_1 and ω_kP+1 = ω_2;
```

vulnerable to Side-Channel Attacks (SCA) [22].

To mitigate such vulnerabilities, Edwards *et al.* [23] and Bernstein *et al.* [24] initially proposed complete curves as Binary Edwards Curves. The completeness property requires the same representations or equations of point operations. This characteristic enables the hardware design to eliminate the distinctions in the process that leak the data in the side channel. Nevertheless, using the complete curves also requires the overhead of the implementation cost, as indicated in Table 1.5, which leads to many challenges for low-cost and low-power devices.

In addition, the conventional *Double-and-Add* algorithm of ECC also leaks the side-channel information, according to the analysis by Abarzua *et al.* [25]. The timing and power traces of the encryption are directly dependent on the processed bit of the secret key k. By observing the variety of the power trace or timing, an adversary could determine between point doubling and point addition, and thus, the secret key $[k]$ is recovered. Therefore, the additional countermeasure techniques are proposed to hide the dependence by modifying the algorithm of point multiplication that performs repeated both point addition and point doubling, *Double-and-Always-Add*. Nevertheless,

these approaches increase computational complexity by approximately 33% and remain insecure against Correlation Collision Attacks [25].

Based on the idea of dumping an additional operation, the Montgomery Ladder algorithm [26–28] is proposed to provide fast point multiplication on elliptic curves and binary elliptic curves. By using this point multiplication algorithm, the devices are also being secured against Differential SCA, as analyzed in [25]. Similar to the *Double-and-Always-Add*, the Montgomery Ladder also requires higher computational complexity, and therefore, optimization of the hardware design to meet the requirements of the constrained devices, such as passive RFID tags, is one of the most challenging issues.

1.5 Summary

This chapter has provided a systematic overview of passive RFID technology and its critical research challenges. The main objective of the research lies in the hardware implementation of Elliptic Curve Cryptography (ECC) based passive RFID tags. The device requires balancing physical constraints such as cost, energy efficiency, and hardware footprint with security level requirements.

Existing solutions of ECC-based passive RFID tags mainly prioritize security at the expense of minimizing the implementation costs. Additionally, some approaches reduce the implementation cost by overlooking the hardware vulnerabilities in resource-constrained conditions. It is a challenge for designers to find optimal architectures that balance both the security level and physical constraints. Addressing this challenge, the design methodologies enable designers to systematically consider the possibilities of the deployment architectures, cryptographic algorithms, and hardware configurations. The following chapter reviews the drawbacks and benefits of the conventional design approaches to highlight the motivation of this thesis.

Chapter 2

Design Methodology for Secured RFID Tags

Following the detailed introduction of RFID technology and its associated security concerns, this chapter highlights a key challenge in conventional design approaches: achieving an effective trade-off between implementation cost and the security of passive RFID tags. The analysis presented in this chapter primarily focuses on the strengths and limitations of traditional methodologies used in hardware designs of secure passive RFID systems.

Section 2.1 outlines the advantages and drawbacks of existing design approaches, which serve as the motivation for developing a novel design methodology. Section 2.2 provides a concise overview of standard Top-down and Bottom-up design flows, followed by a brief review of emerging AI-assisted design methodologies. In Section 2.3, the discussion is concentrated on general methodologies specifically applied in hardware security design.

Through this analysis, it becomes evident that selecting an optimal design methodology involves a complex trade-off between minimizing implementation costs and ensuring robust security. Therefore, a new design flow is necessary to systematically incorporate security evaluation metrics and implementation cost estimations into every stage of the Design Space Exploration (DSE) process.

2.1 Background and Motivation

2.1.1 Definition of the Design Methodology

The rapid advancement in hardware design applications has led to a significant increase in the complexity of hardware systems. To address this, the Electrical System Level (ESL) [29] approach aims to model and analyze these systems functionally and architecturally at a high level of abstraction. For example, in an ECC-based authentication device, the ESL defines four primary levels of design, as illustrated in Figure 2.1.

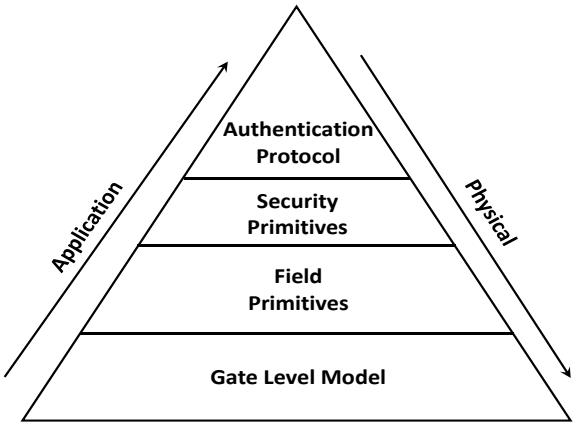


Figure 2.1: Pyramid showing the implementation levels of ECC-based authentication device.

The ESL models are particularly valuable during the initial design stages, enabling designers to conduct Design Space Exploration (DSE) or design methodology. Design methodology is a critical process that involves identifying and assessing various design alternatives to optimize the performance and cost of the hardware system.

Design Methodology or DSE offers several key advantages [30]:

- **Fast Prototyping:** Generates prototypes for simulation and profiling, enhancing the understanding of design impacts.
- **Optimization:** By evaluating comparable metrics, designers quickly eliminate inferior designs and focus on optimal solutions.
- **System Integration:** Using the knowledge of hierarchy and systematic architecture, designers identify feasible configurations that meet global design constraints.

2.1.2 Motivation of Research

Understanding the key benefits of design methodologies in hardware design reveals a gap that challenges the designers to quickly design a system that balances multiple objectives. For example, in security hardware design, the general design flow exhibits a variety of limitations that increase the design time of the final products. These limitations stem from the increased trial-and-error required as design complexity rises. As a result, the design time to complete an optimal design is significant.

Recognizing the need, several works [31–47] have been proposed to provide efficient design methodologies to help designers quickly identify solutions. In the beginning, these works focused on the balance between the implementation cost and performance in simple hardware designs. Some of them integrated artificial intelligence (AI) [43–47] to boost the efficiency of the design process of analog circuits. They show a significant improvement in the time-to-market cost. Additionally, some of the framework tools have recently proposed AI assistance for security hardware design. However, these improvements do not concern the application of security hardware design, which is constrained by physical resources. This thesis proposes a novel design methodology that integrates both security evaluation and implementation cost estimation into the DSE loop to address this research problem.

2.2 Traditional Design Methodology for Hardware Implementation

2.2.1 Top-down Design Flow

In digital circuit design, several design methodologies have been developed for designers to optimize and balance multiple objectives and constraints. Among these, the top-down design approach has emerged as the most popular and widely adopted methodology. This approach, extensively documented in the literature [31–35], provides a structured framework that guides designers from the initial functional specification to the final implementation details. The top-down methodology is particularly advantageous because it allows designers to systematically address balancing other critical design constraints, such as power consumption, area, and performance.

The top-down design process begins with formulating specifications for the digital circuit. These specifications define the operational requirements and constraints, including power budgets, area costs, and expected throughput. These parameters serve as the foundation for the entire design process, ensuring that the final implementation

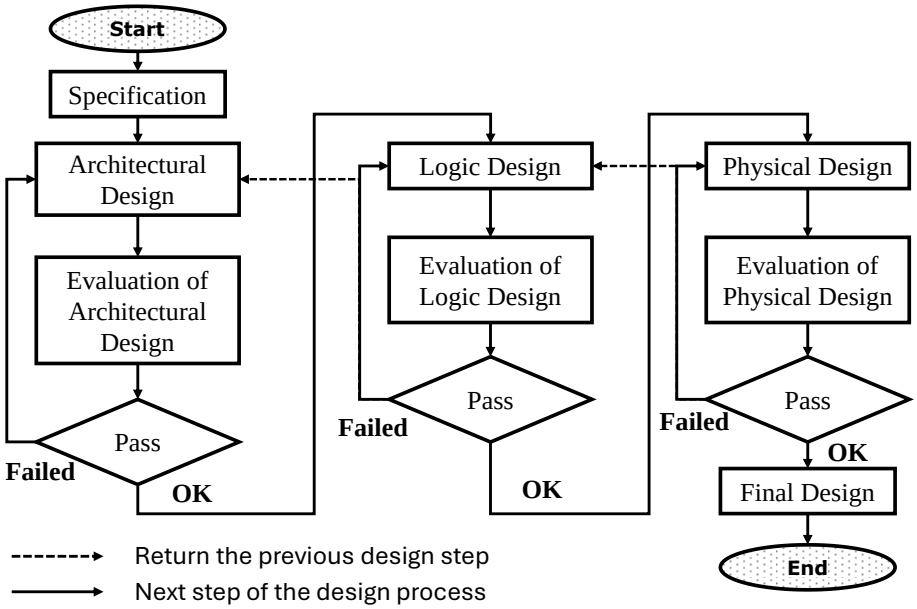


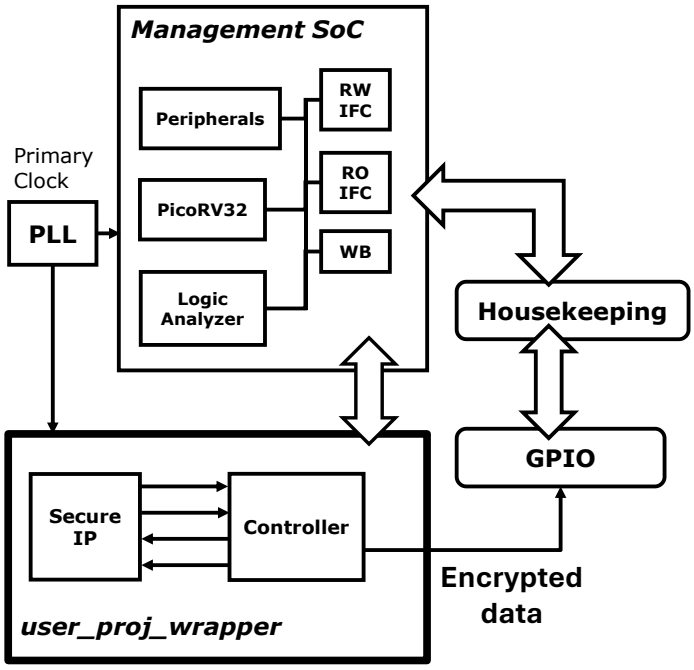
Figure 2.2: Regular Top-down Design Methodology for Hardware Design.

meets both functional and non-functional requirements. As depicted in Figure 2.2, the designer starts by translating these specifications into a high-level architectural design. This involves breaking down the overall functionality into smaller, manageable blocks, each responsible for a specific subset of operations. The architectural layout is typically represented as a block diagram, where each block operates within its predefined limits.

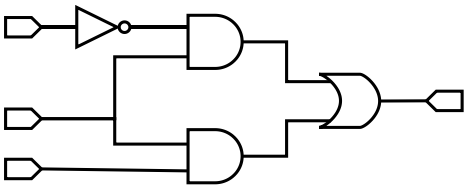
The main task of the architectural design phase is to define the data transmission protocol and the interconnections between blocks. Figure 2.3(a) shows an example that illustrates the block interconnections in the system implementing the cryptography-based authentication protocol. The implemented '`user_proj_wrapper`' implements the encryption algorithm, for example, the point multiplication over the Binary Edwards Curves. The encrypted results are controlled and authenticated by the processor named '`Management SoC`', which executes the authentication protocol.

The choice of protocol and the way blocks communicate significantly influence the system's overall performance. Especially in the context of the FPGA (Field Programmable Gate Array), the architectural design substantially impacts the systematic performance [36]. Therefore, designers must carefully evaluate the data transmission protocol and ensure that it aligns with the constraints of each block. This evaluation is a crucial step for the designer to predict the system's performance and identify potential issues early in the design process.

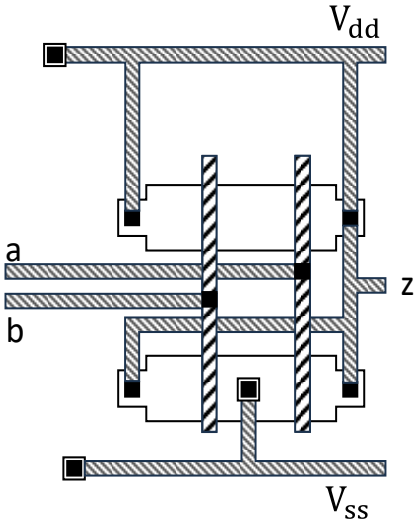
In the case of a successful evaluation, the next step is to implement each block at the logic level. This phase involves creating a detailed behavioral model of each block using



(a) Architectural design of the implemented system.



(b) Gate-level description of a particular block in the implemented system.



(c) Physical design of the NOR gate in the implemented system.

Figure 2.3: Design abstraction levels of design in top-down design methodology.

a Hardware Design Language (HDL), such as Verilog or VHDL, at the Register-Transfer Level (RTL) or Gate Level, as illustrated in Figure 2.3(b). The HDL description must adhere to the constraints defined during the architectural design phase, including the input/output (I/O) specifications and the data transmission protocol. Additionally, the functionality of each block must satisfy the requirements established in the previous stage, ensuring consistency across the design. The logic design phase is iterative and requires rigorous verification to ensure correctness. Designers regularly use simulation tools to evaluate the model behaviorally against various scenarios, verifying that the block operates as intended under different conditions. If the evaluation reveals discrepancies or failures, designers may need to revisit the logic design, refine the HDL description, or even reconsider the architectural layout. This iterative process ensures that the design is robust and meets all specified requirements before proceeding to the next stage.

The final phase of the top-down design process is the physical layout, where the logic design is translated into a geometric representation, and that is ready for fabrication. At this stage, all components of the design are instantiated with their precise geometric descriptions, as defined by Andrew *et al.* [37]. This includes the placement of blocks, cells, and gates, as well as the routing of connections between them. The physical placement must comply with the design rules of the target fabrication technology, which dictate parameters such as minimum feature sizes, spacing, and layer usage. On the other hand, routing is a critical aspect of the physical design process, as it directly impacts the circuit's performance, power consumption, and the possibility of successful manufacturing. Figure 2.3(c) shows the completed layout of a NOR Gate within the implemented system. Once the physical layout is complete, the design is exported in a standard format such as GDSII (Graphic Design System II), which is used for manufacturing ASIC technology, or Bitstream for FPGA technology. The final design needs to be evaluated to warrant that the final result meets all the series of checks, including Design Rule Checking (DRC), Layout Versus Schematic (LVS) verification, and timing analysis. These checks help identify and resolve any discrepancies or violations that could affect the circuit's functionality or the possibility of successful fabrication. If the evaluation reveals issues, designers must revisit the physical layout and make the necessary adjustments. In some cases, this may involve modifying the placement of components or rerouting connections. The final design needs to be evaluated to warrant that the final result meets all required checks, including Design Rule Checking (DRC), Layout Versus Schematic (LVS) verification, and timing analysis. This iterative process ensures that the final design is both functional and manufacturable.

The integration of the top-down design methodology with Hardware Description

Languages (HDLs) offers notable strengths, chief among them being the standardization of the design process into a unified framework. This harmonization allows designers to seamlessly transition between different Electronic Design Automation (EDA) tools within a shared environment, facilitating end-to-end development—from abstract conceptualization to hardware modeling [38–40]. By employing abstract block diagrams, designers gain the ability to visualize interconnections and assess individual block performance, enabling targeted optimizations for system components.

In contrast, the main disadvantage of the top-down design approach is insufficient granularity in sub-block architecture at the beginning design stages, which often delays verification of compliance with critical constraints, such as power efficiency, area utilization, performance, and hardware security, until later stages. Furthermore, incorporating security measures (e.g., safeguarding secret keys, randomization of operators, hiding the secret, or masking the operation) risks inflating implementation costs, potentially disrupting the equilibrium between physical resource allocation and security robustness. Alternatively, redesigning logic or architecture to meet security demands may extend development timelines, elevating time-to-market pressures. Thus, while the methodology fosters efficiency and optimization, it necessitates careful trade-offs to balance security, cost, and project timelines.

2.2.2 Bottom-up Design Flow

The bottom-up design methodology offers a compelling alternative to traditional top-down approaches, particularly in addressing localized optimization challenges within complex systems such as elliptic curve cryptography (ECC) modules. Unlike the top-down strategy, which begins with high-level abstractions and iteratively refines them, the bottom-up approach prioritizes the customization of fundamental circuit components—transistors, standard cells, and memory blocks—to achieve granular performance improvements. This methodology has gained traction in ECC design, as evidenced by works such as Pirotte *et al.* [41] and Lutz *et al.* [42], which highlight its effectiveness in balancing power, area, and security constraints. Below, Figure 2.4 demonstrates the three-stage bottom-up design process, its advantages, and its limitations, with a focus on ECC implementations.

The initial phase of the bottom-up methodology revolves around decomposing the system into discrete functional primitives. Designers begin by defining explicit constraints, including power budgets, area limits, performance targets, and other requirements. For example, with ECC modules, this might involve isolating critical operations such as modular arithmetic units, finite field multipliers, or point multiplication

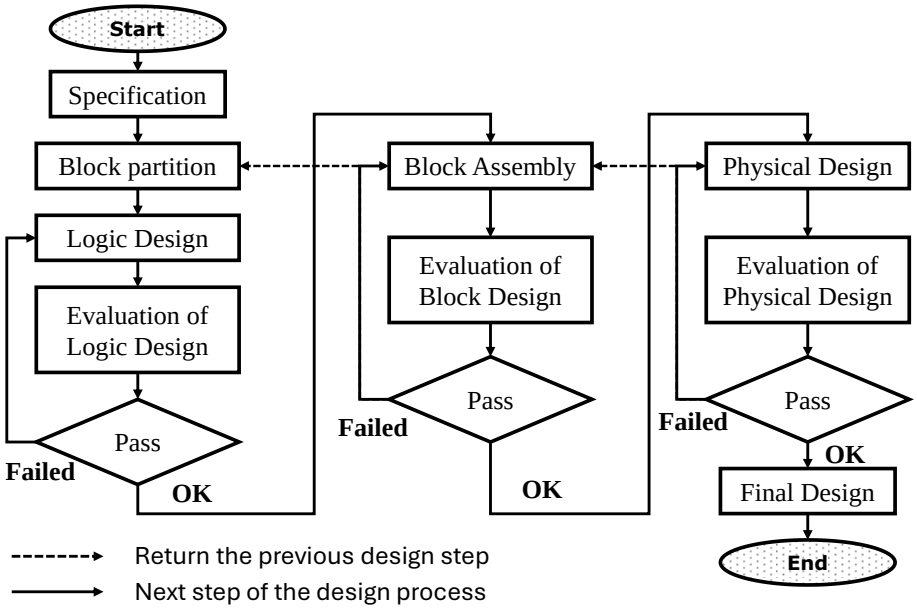


Figure 2.4: Conventional bottom-up design methodology for hardware design.

engines. Each primitive is treated as an independent entity, enabling designers to optimize its internal logic without external interference. For example, transistor-level tuning might reduce leakage power in a finite field adder, or custom standard cells could enhance the speed of a modular inversion block.

A key advantage of this stage is the simplicity of localized evaluations. By focusing on a single electrical element, designers rapidly iterate on logic designs, testing alternatives if evaluations fail. For instance, an optimal shift register with minimal depth is proposed by Pirotte *et al.* [41] to reduce significantly the total latency and the area cost of the ECC encryption. This compartmentalization accelerates optimization cycles, as simulations need only account for independent blocks. However, this isolation also introduces risks, as optimizations may inadvertently overlook global system dependencies, such as inter-block signal propagation delays or shared resource contention.

Once primitives are optimized, the second stage involves assembling them into a cohesive system. Abstract models of each sub-block—detailing input/output (I/O) behavior, area, and power consumption—are interconnected to form a high-level architectural blueprint. For ECC modules, this might involve integrating a field multiplication component with a field square and a field inversion unit to execute the point multiplication.

During assembly, designers evaluate functional correctness and constraint adherence through rapid simulations. At this step, a misalignment between one component's out-

put bandwidth and another's input requirements could necessitate rerouting interconnects or revisiting primitive designs. However, the abstraction's simplicity—omitting detailed electrical parameters—limits evaluation accuracy. A sub-block meeting local power targets might exceed global budgets when combined with others, a common pitfall in ECC designs where cryptographic operations are highly interdependent. In extreme cases, designers must reconstruct primitives, underscoring the iterative nature of this phase.

The final stage translates the assembled architecture into a physical layout. Designers floorplan sub-blocks, route metal layers, and plan power distribution networks, leveraging geometric data from earlier stages. In the context of an ECC system, this phase ensures that all components, such as field operators, are connected to the power supply and clock networks. Additionally, the electrical parameters, including parasitic capacitances and resistances, are incorporated into simulations to predict timing, power, and thermal behavior.

This stage also assesses manufacturability by evaluating design rule compliance (DRC) and lithography constraints. For instance, dense transistor arrays in ECC modules might require adjustments to meet foundry-specific spacing rules. Failed evaluations trigger layout replanning or even architectural overhauls. However, the precision of this phase comes at a computational cost: simulating nanometer-scale effects in large designs, such as multi-core ECC accelerators, can be prohibitively time-consuming.

Despite its strengths, the bottom-up approach faces significant issues. First, independently optimized sub-blocks often result in irregular shapes, leading to die-core fragmentation and underutilized silicon area. Schurmann *et al.* [48] demonstrated that such inefficiencies escalate with design scale, directly impacting cost, a critical concern for high-volume ECC hardware. Second, the reliance on abstract models during assembly introduces inaccuracies; a sub-block local power estimate might neglect global voltage drop effects, compromising system-level predictions.

Moreover, the exhaustive simulations required in the last design stage become impractical for complex ECC systems. For example, verifying side-channel resistance in a masked ECC implementation demands cycle-accurate power trace analyses, which scale exponentially with design size. This complexity often renders bottom-up methodologies incompatible with modern ECC primitives, which demand tight integration of security and performance.

A recurring critique of both top-down and bottom-up methodologies is their treatment of security as a peripheral constraint. In particular, for ECC designs, security levels that dedicate the resistance metric of the system against threats, such as side-channel attacks or wireless vulnerabilities, are frequently supplemented rather than

being embedded into early-stage primitives. This reactive approach complicates the pursuit of a balance between the implementation cost and security level.

2.2.3 Emerging AI-driven Design Methodologies

In the realm of Artificial Intelligence, the generative AI model plays a critical role as an assistant to the designer in optimizing the implemented hardware systems. Li *et al.* [43] presented an effective optimization automation tool, which is based on Artificial Neural Network (ANN) for local minimization of the implementation costs. Besides, for the global optimization, the automation tool utilized the genetic algorithm [44, 45].

Figure 2.5 demonstrates the proposed design flow for the analog design. The proposed work by Li *et al.* [43] shows the benefit of reducing four times speed enhancement and comparable results with the conventional approaches. However, Zhang *et al.* [46] explained the critical disadvantages of Li's ANN-based design methodology. By using the model-based approach, the previous work of Li *et al.* [43] requires a large set of simulation data to approximate. As a consequence, the implementation cost of the generated model is inaccurate and deviates from the fabricated circuit performance. Addressing this issue, Zhang *et al.* [46] proposed to use the Bayesian optimization framework with the combination of model-based and simulation-based approaches to improve the efficiency and the effectiveness of the design exploration. Besides, a fully automated tool named AnGel was recently proposed by Fayazi et [47], which enables maintaining the same accuracy with a reduction of 4.7x-1090x of labeled data and 2.9x-75x faster. With the significant advantages of utilizing AI to assist in hardware design, it is becoming more and more critical to implement AI to support the design process for secure hardware applications.

In the context of the innovation design methodologies for secured hardware applications, Aerabi *et al.* [49] systematically proposed a design methodology that enables designers to manually choose the appropriate cryptographic algorithms for the ultra-low-energy application. By using the Aerabi approach, the designers significantly reduce the trial-and-error of alternatives and save time-to-product costs. This proposal is the starting point for the AI-assistance design methodology targeting the secure hardware implementation. A few years later, the X-dfs tool was proposed by Mahfuz *et al.* [50] to systematically explain the security countermeasure for the AI optimization model. The proposal helps to guide the AI model for suggesting a suitable countermeasure against reverse engineering threats, hardware trojans, fault attacks, and side-channel attacks. It is explicit that AI assistance or design methodologies, which help designers rapidly and precisely choose the balancing hardware implementation between multiple objec-

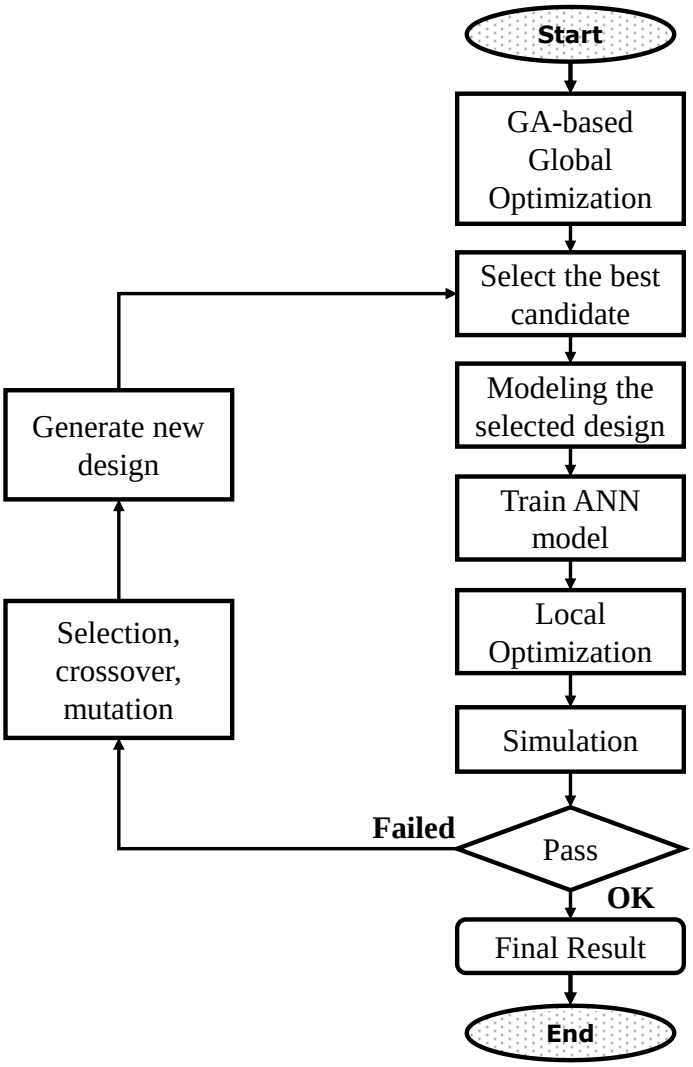


Figure 2.5: ANN-based design methodology for analog design [43].

tives, are increasingly emerging, especially for secure resource-constrained hardware applications.

2.3 Design Methodology for Hardware Security Design

In a particular application of implemented design, the complementary stages are supplemented into the original design flow, enabling the designer to quickly evaluate and modify the system to fit the design constraints. This section provides an overview of the complementary security evaluation for the ECC-based authentication system, which is constructed by utilizing the Top-down design methodology.

2.3.1 Complementary Security Evaluation for Design Methodologies

2.3.1.1 Post-Fabrication Security Evaluation

Adapting to the security hardware design, the designer complements the security evaluation of all necessary vulnerabilities at the end of the original Top-down design approach. After fabricating the hardware design, the designers take into account experiments to measure the security level of the implemented system. The security tests cover all the verifications from the wireless attacks to the hardware vulnerabilities, such as Side-Channel Attacks (SCA). Dao *et al.* [51] illustrated the approaches for measuring the security level of ECC-based RFID tags in devices, as depicted in Figure 2.6.

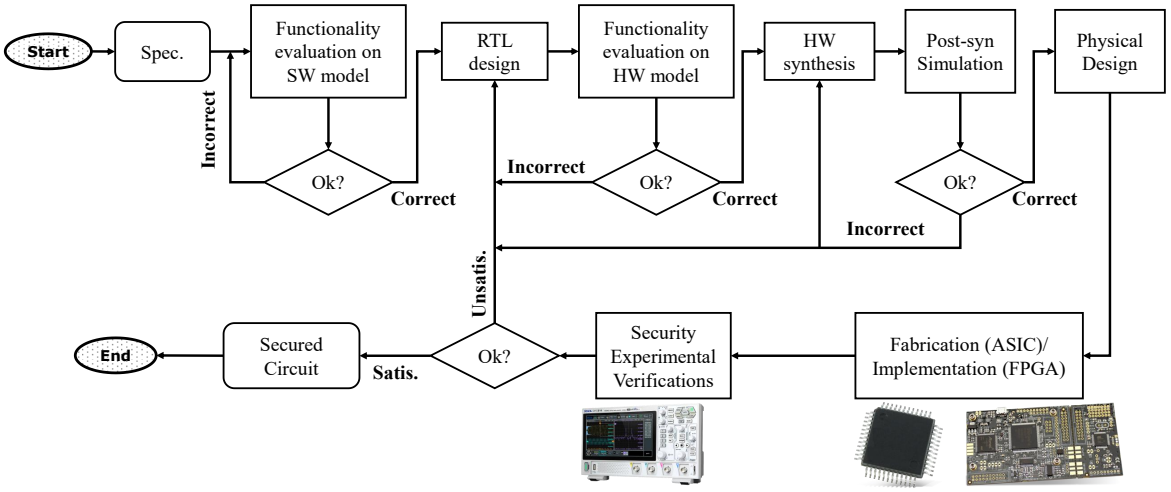


Figure 2.6: Security evaluation for post-fabrication methodology.

The complementary security assessments are performed after the circuit has been fabricated using ASIC technology or deployed on an FPGA. Engineers use oscilloscopes equipped with electromagnetic (EM) probes or shunt resistors to capture electromagnetic emissions or voltage fluctuations, which correspond to the circuit's power consumption during operation. These power traces are then analyzed to gauge the system's resilience against hardware-based vulnerabilities, such as side-channel attacks. The resulting power traces provide granular insights into the device's behavior but are inherently polluted by ambient electrical noise from the circuit and its environment. A key advantage of this method lies in its precision: for example, side-channel analysis leverages these traces to dissect correlations between power consumption patterns, operational data, and background noise.

However, extracting actionable intelligence from the raw data demands rigorous

preprocessing. This includes noise suppression and temporal alignment of traces to isolate coherent signals, ensuring that transient power spikes or cryptographic operations are accurately mapped. The complexity of these steps—coupled with the need for meticulous signal calibration and synchronization—renders the security assessment process both labor-intensive and technically demanding. Achieving reliable results requires substantial computational resources and domain expertise, particularly when balancing the trade-off between evaluation depth and operational efficiency.

If a security evaluation fails, designers must go back to the beginning, where the alternative RTL design with countermeasures will be proposed. This loopback of reconstructing the alternative secured system causes a critical problem. Designing alternative solutions and prototypes takes more time and costs more. Time-to-product and design costs rise dramatically when the system becomes more complicated.

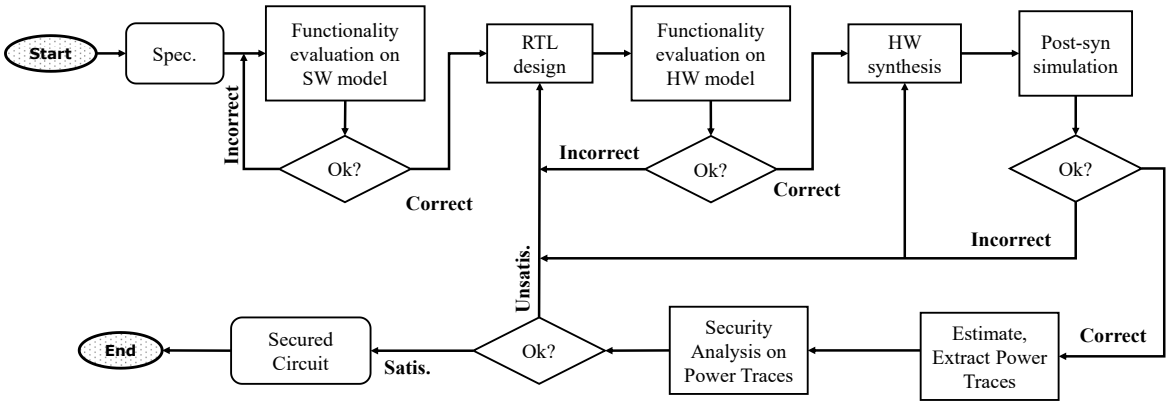


Figure 2.7: Security evaluation for post-synthesis methodology.

2.3.1.2 Post-Synthesis Security Evaluation

To address the challenges associated with post-fabrication security testing, modern design methodologies increasingly rely on simulated power traces generated during the prototyping phase prior to physical manufacturing [52]. These traces enable early-stage vulnerability assessments, reducing dependency on costly and time-consuming post-silicon evaluations. As shown in Figure 2.7, this approach streamlines the design cycle, significantly shortening time-to-market and lowering development costs. The process begins with synthesizing the Register-Transfer Level (RTL) design into a gate-level netlist using Electronic Design Automation (EDA) tools such as Synopsys Design Compiler or open-source alternatives like Yosys. These tools leverage technology libraries, which define transistor behavior, timing, and power characteristics, to simulate the system's functionality and estimate both dynamic and static power consumption.

By aggregating instantaneous power values during cryptographic operations (e.g., encryption or key generation), the EDA environment constructs simulated power traces that approximate the device's behavior. While these traces lack the fine-grained details of physical measurements (e.g., electromagnetic noise or analog effects), they capture gate-level dynamic power fluctuations, which are critical for identifying side-channel vulnerabilities such as data-dependent power leaks.

However, the utility of these simulated traces hinges on their accuracy and relevance to real-world scenarios. Post-synthesis simulations model switching activity and capacitive load effects at the gate level, providing insights into how logic cells interact during operation. Early detection allows designers to possibly implement countermeasures, such as adding masking logic or balancing power paths, before committing to a physical design. This proactive approach minimizes late-stage redesigns and ensures that security is built into the architecture rather than supplemented. Nevertheless, the limitations of simulation must be acknowledged: factors like process variations, interconnect parasitics, and environmental noise are abstracted away, leading to discrepancies between simulated and physical traces. Besides, power traces harvested at the gate level also cause a storage problem as the file size contains power traces. This leads the evaluation engine to arrange a large storage and processing strong enough for analysis.

Despite these gaps, the ability to perform early security evaluations during the first design phases allows designers to balance the implementation cost and the security level, especially for secured resource-constrained designs. Furthermore, utilizing the post-synthesis security evaluation design methodology results in a secured design at the gate level. This benefits the designer by concluding the security issues of the implemented design without several fabrications.

2.3.2 Security Evaluations for Secured RFID Tags

2.3.2.1 Protocol Security Evaluation Approaches

Distinguished the design methodology for the general security design, in the context of secured RFID tags, emphasizes distinct approaches for evaluating security at both the protocol and hardware levels. For the protocol level, the evaluation focuses on assessing the security of RFID (Radio Frequency Identification) tags that are resistant to unauthorized access. These evaluations analyze the ability of security protocols to counter wireless threats and maintain optimal system performance, as outlined in Table 1.1. By testing resilience against vulnerabilities like eavesdropping or cloning (detailed in Table 2.1), the study ensures protocols balance security with functional efficiency.

Together, these assessments provide a comprehensive framework for understanding the strengths and weaknesses of RFID security systems.

Table 2.1: General Vulnerabilities that are evaluated on Protocol Security Evaluation.

Vulnerabilities	**Definition**
Man-in-the-middle Attacks (MITMA)	Attackers may receive the messages and attempt to relay or alter communication.
Eavesdropping	Passive adversaries may attempt to intercept communication between tags and readers.
Replay	Attackers may receive the messages and attempt to resend them to the other parties.
Forward Security	Adversaries possibly recover the current or previous confidential information if they can determine the secret information in a tag.
Impersonation	In this attack, an adversary provides situations in which the legitimate reader accepts the adversary as a legitimate tag
Key Compromise	In the case a particular tag is vulnerable, the network should withstand the situation.
Location Tracking	Determining the current and visited location of specific tags by obtaining the same response between the tag and the reader.
Denial-of-Service (DoS)	Attackers create a large number of tags or transponders to disrupt the service of the system.
Cloning	Adversaries target the integrity characteristics of the RFID system when they can succeed in capturing a tag's identifying information to create clones of tags. The clones can overcome the counterfeit protection and act as a preparatory step in a theft scheme.
Desynchronization	Adversaries desynchronize the endpoints by sending the invalid message to destroy

In general, to speed up the improvement and assessment of the authentication protocols, automation validation tools, such as AVISPA [53].

2.3.2.2 Hardware Security Evaluations

Hardware attack, such as Side-Channel Attacks, as mentioned in Chapter 1, is one of the critical vulnerabilities of the passive RFID system. Therefore, measuring and

evaluating the hardware security, especially Side-Channel Attacks, is necessary for designers to prove the safety of the implemented hardware system. Power analysis is the most popular threat to passive RFID tags, as the attacker could execute these attacks in resource-limited environments.

a. Power Consumption Analysis In the realm of integrated circuit (IC) design, the overall power dissipation is primarily composed of two fundamental components: static power consumption and dynamic power consumption, as delineated in Equation (2.1). Static power dissipation, denoted as P_{static}, occurs due to the leakage currents when the circuits remain in an idle state. In this situation, the circuits are powered but not actively engaged in data processing. Conversely, dynamic power dissipation, represented as $P_{dynamic}$, occurs during active circuit operation and is dominated by switching power. This switching power is a consequence of the charging and discharging of capacitive loads during logic transitions. Mathematically, the total power dissipation of the system is expressed as below:

$$P_{total} = P_{dynamic} + P_{static} \tag{2.1}$$

Dynamic Power Dissipation Dynamic power dissipation constitutes a significant portion of the total power consumption. This power leakage is primarily caused by the charging and discharging of the total load, as well as by the short-circuit current. Generally, the dynamic power $P_{dynamic}$ is calculated by using Equation (2.2):

$$P_{dynamic} = P_{switching} + P_{short\_circuit} \tag{2.2}$$

where $P_{switching}$ represents the power consumed due to the switching activities within the circuit. $P_{short\_circuit}$ denotes the power consumed due to the direct current path that momentarily exists between the supply voltage and ground. This short-circuit phenomenon is transient, and, particularly in low-power applications operating at low frequencies, its impact is minimal compared to the clock period. Consequently, the short-circuit power dissipation is often deemed negligible.

In contrast, the amount of power consumed by the switching activities of the circuit is estimated by Equation (2.4).

$$P_{switching} = \alpha \cdot V_{DD}^2 \cdot C_L \cdot f \tag{2.3}$$

where α denotes the ratio of switching activity, which by default is 0.5. Besides, V_{DD}, C_L, and f are the supply voltage, total load capacitance, and the frequency of operation, respectively.

In other words, the production of $\alpha \cdot f$ is the number of the switching activities of the circuits, which directly reflects the dependency of the data. Therefore, by modeling the switching activity power dissipation, the designer possibly understands the leakage power model of the system. As a consequence, they can evaluate the side-channel analysis (SCA) vulnerabilities.

Static Power Dissipation This dissipation is predominantly caused by several intrinsic phenomena, such as the reverse bias diode (RBD), sub-threshold, and gate leakage current. These factors play a significant role in static power dissipation, and their impacts are illustrated in Equation ((2.4)):

$$\begin{aligned} P_{static} &= V_{DD} \cdot I_{leak} \\ &\simeq P_{sub} + P_{gate} + P_{RBD} \\ &\simeq V_{DD} \cdot (I_{sub} + I_{gate}) + V_{bs} \cdot I_{RBD} \end{aligned} \tag{2.4}$$

where P_{sub}, P_{gate}, and P_{RBD} represent the leakage power dissipated by sub-threshold, gate, and reverse bias diode phenomena. Similarly, I_{sub}, I_{gate}, and I_{RBD} denote the leakage current caused by the corresponding phenomena. V_{DD} and V_{bs} are the supply and body bias voltages, respectively.

Depending on the standard cells' technology, the static power either depends on or is independent of the processing data. For example, in the digital CMOS (Complementary Metal Oxide Semiconductor) standard cells, the leakage current is highly determined by the composition and state of the transistors, which in turn directly depend on the processing data. Consequently, the static power consumption of the CMOS logic is substantially data-dependent. However, compared to the dynamic power dissipation, the static leakage power of the CMOS consumes a negligible amount, which is possibly neglected in modeling.

In the other technology, MOSFET (Metal Oxide Semiconductor Field Effect Transistors), although there is a significant off-current flow between the terminals, which consumes a significant amount of static power, this amount is proportional to the number of powered logic cells in the circuit. It turns out that the static power of the MOSFET is independent of whether those standard cells are actively fed with input data or not. Consequently, in modeling the power leakage for the SCA evaluation, this amount of power is also possibly neglected.

Power consumption traces generated by electronic design automation (EDA) tools such as Synopsys PrimeTime are critical for evaluating information leakage in cryptographic hardware systems. These traces, derived from power analysis simulations, enable engineers to quantify the susceptibility of a design to side-channel attacks (SCAs),

which exploit unintended physical emissions, such as power fluctuations or electromagnetic radiation, to extract sensitive data like encryption keys. Multiple methodologies exist, such as Test Vector Leakage Assessment (TVLA) or Power Analyses to assess the vulnerabilities of a proposed hardware design against SCAs.

b. Security Evaluation Framework

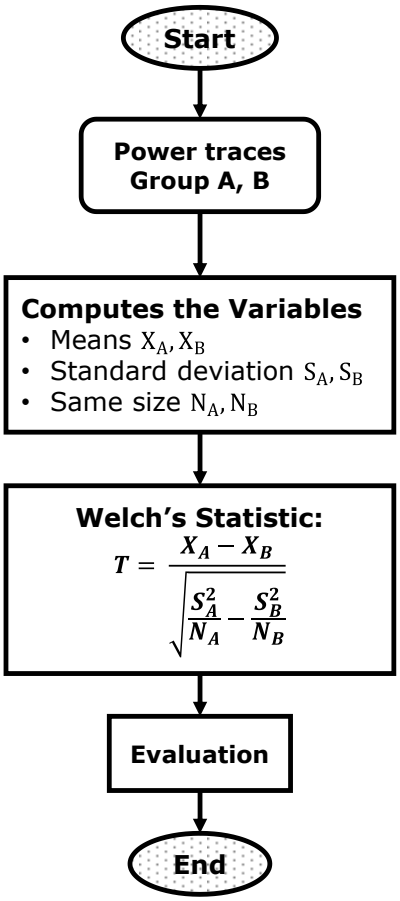


Figure 2.8: Methodology for evaluating the Test Vector Leakage Assessment (TVLA).

Test Vector Leakage Assessment One of the most widely adopted and comprehensive approaches is the TVLA. This approach provides a statistical framework for detecting subtle correlations between processed data (e.g., cryptographic operations) and observable side-channel signals, such as power consumption or electromagnetic emanations. The TVLA evaluation framework operates by statistically analyzing side-channel traces to detect data-dependent leakage. In practice, this involves feeding a cryptographic module with two distinct sets of test vectors: one fixed (e.g., a constant plaintext) and one random (e.g., varying plaintexts). Power or electromagnetic traces

are collected during the module's operation, and a statistical test—often a Welch's t-test—is applied to determine whether the traces differ significantly between the two sets, as depicted in Figure 2.8.

Firstly, the average of all the traces in each power trace set is computed and denoted as X_A and X_B, respectively. After that, the sample standard deviations S_A and S_B are also derived by the statistical tools. The T-statistic trace value is computed point-wise via Equation (2.5):

$$T = \frac{X_A - X_B}{\sqrt{\frac{S_A^2}{N_A} + \frac{S_B^2}{N_B}}} \tag{2.5}$$

Using Welch's t-test, a threshold of 4.5 standard deviations, corresponding to 99.999% confidence that a difference shown is not due to random chance, will be set. Further, for each test, the t-test should be run twice on independent data sets. If, for a given test, the same time point in both data sets exceeds ±4.5, the device is leaking data-related information through the power or electromagnetic traces.

The advantage of the TVLA is its generality. This leakage evaluation does not require prior knowledge of the cryptographic algorithm or attack methodology. As a consequence, TVLA is suitable for evaluating black-box systems. However, TVLA's effectiveness depends on the quality of the traces. Noisy or misaligned measurements can obscure leakage signals, necessitating preprocessing steps such as filtering, trace averaging, or alignment using reference patterns.

Power Analyses Evaluation Approaches Another approach for evaluating security is performing an experimental hardware attack, such as power analysis, to assess the security level of the implemented design. One of the most powerful engines for power analysis is the CPA.

CPA is a sophisticated side-channel attack technique designed to extract secret cryptographic keys by analyzing the power consumption patterns of a device during its operation. Unlike brute-force methods, CPA exploits the correlation between the power traces—recordings of a device's power usage over time—and hypothetical power consumption models derived from the cryptographic algorithm being executed. By statistically comparing measured power traces with predicted models, attackers can systematically deduce portions of the key, significantly reducing the computational effort required to compromise the system.

The methodology of CPA involves several key steps. First, an attacker collects many power traces generated by the target device while it performs cryptographic operations on known input data. Next, a hypothetical model of the cryptography model is simulated with the corresponding plaintext for each possible key guess. In

the following step, a leakage model of the hypothesis, which is based on the Hamming weight or Hamming distance, generates the reference power traces of intermediate data.

After successfully collecting both sets of traces, Pearson's correlation coefficient is then computed between the hypothetical power values and the actual measured traces at each point in time, as illustrated in the equation.

$$\rho_{AB} = \frac{\sum_i a_i b_i - n\overline{a}.\overline{b}}{\sqrt{N\sum_i a_i^2 - (\sum_i a_i)^2}.\sqrt{N\sum_i b_i^2 - (\sum_i b_i)^2}} \tag{2.6}$$

where, N denotes the number of samples in the power traces, a_i, b_i represent the individual sample points indexed with i of the power traces and hypothetical model.

A high correlation at specific intervals indicates that the hypothesized key portion is likely correct, as the model aligns with the observed power fluctuations. This process is repeated iteratively for all key segments until the full key is reconstructed.

The effectiveness of CPA underscores critical vulnerabilities in unprotected hardware and emphasizes the need for robust countermeasures. However, the effectiveness of the executing CPA depends on the amount of noise and the alignment of the measured power traces. To address these issues, the designers also need to perform additional preprocessing steps such as filtering, trace averaging, or alignment using reference patterns.

2.4 Summary

The conventional methodologies for hardware designs, such as top-down or bottom-up, are popular and effective in implementing a complete design from the assumption specifications. The biggest advantage of these traditional approaches is the ease of application. However, with the rapid development of resource-constrained applications and increasing security requirements, conventional methodologies face a lot of limitations. One of them is the high time-to-product and designing time.

Advanced design methodologies have enabled engineers to develop secure systems systematically by integrating security evaluations into the design workflow. However, these evaluations are typically relegated to the final stages of the design process, creating a critical limitation: Designers are often forced to retroactively address security concerns, which restricts their ability to holistically balance security requirements with hardware implementation metrics such as power, area, or performance. This sequential approach risks suboptimal trade-offs, as late-stage security adjustments may necessitate costly redesigns or compromises in system efficiency.

Recent advancements in AI-based Automatic Design Space Exploration (DSE) offer promising solutions to this challenge. By automating the exploration of design config-

urations and optimizing for multiple objectives (e.g., security, resource utilization, and latency), these tools empower designers to identify optimal solutions that harmonize competing priorities. Yet, while existing AI-driven DSE frameworks excel at multi-objective optimization, they often treat security as a static constraint rather than a dynamic, co-equal parameter to be actively optimized during the design process.

This gap underscores the need for a paradigm shift: embedding security-aware co-optimization into conventional frameworks. This thesis realized this limitation by implementing the first proposal of a low-cost, low-power hardware design for ECC-based passive RFID. The experimental results of the first proposal achieve a comparative hardware design. However, during the design process, the drawbacks of utilizing the conventional design methodologies have been shown. In chapter 3, a detailed description of the first proposal is discussed before recognition of the motivation for a novel design methodology, which is presented in chapter 4.

Chapter 3

Hardware Implementation of Elliptic Curve Cryptography

In the previous chapters, we showed one of the most critical issues of the ECC-based passive RFID tags. A balanced design, which is a compromise between the implementation cost and the security level against wireless attack and hardware attacks, is a challenge for the designers to implement by utilizing the design methodologies presented in Chapter 2. In this chapter, we propose a low-cost, low-power hardware implementation of Elliptic Curve Cryptography using the Top-down design methodology. The proposed hardware design has been implemented using the TSMC CMOS 65nm technology. To evaluate the security of the implemented ECC-based RFID tag, we use the TVLA evaluation methodology. The experiments confirm that the proposed hardware architecture of ECC is resilient against widespread side-channel attacks and satisfies the physical requirements of the passive RFID tags.

3.1 The proposed algorithm of Binary Edwards Curve

This section presents the proposed Binary Edwards Curve (BEC) algorithm, which is secured against SCA and fits the passive RFID tag requirements. Firstly, an introduction to conventional BEC point multiplication [23, 24] is presented. However, the original BEC provides the completeness property that is secured against SCA. However, the hardware implementation of the BEC surpasses the passive RFID tag's constraints [11, 20] in terms of power consumption, area footprint, and latency. Addressing this problem, in the last part of this section, a proposed hardware architecture corresponding to the reduced point multiplication algorithm is demonstrated in more detail.

3.1.1 The Projective ω- coordinates in BEC

As analyzed in Table 1.4, the implementation cost of point multiplication on BEC using Affine Coordinates is higher than the Binary Generic Curve (Weierstrass Curve). Additionally, the original BEC's point multiplication also requires field inversion, which is the most complex operator among field operators. Addressing the resource constraints of the BEC's point multiplication, the Differential Addition Chain is proposed to replace the original algorithm using the affine coordinates. This approach requires the calculation of only one coordinate ω instead of two affine coordinates (x, y). In particular, $\omega_i = x_i + y_i$ is the difference of two coordinates of the point $P(x_i, y_i)$. The addition law of the $\omega-$coordinate is Equation (3.1):

$$\omega_3 = \frac{\omega_1^2 + \omega_2^2 + \frac{1}{\omega_0}(\omega_1^2 + \omega_2^2)}{\frac{1}{\omega_0}(\omega_1^2 + \omega_2^2) + 1} \tag{3.1}$$

n the case $\omega_1 = \omega_2$, the doubling law of the $\omega-$ coordinate is Equation (3.2):

$$\omega_4 = 1 + d\frac{1}{d + (\omega_2^2 + \omega_2)^2} \tag{3.2}$$

Assuming that at the end of the point multiplication, two values $\omega_Q = \omega(kP)$ and $\omega_{Q+1} = \omega(kP + 1)$ are the results of the Montgomery Ladder algorithm. par To retrieve the affine coordinate (x_2, y_2) or $(x_2 + 1, y_2 + 1)$, solving a equation (3.3) by using half-trace method is required:

$$x_2^2 + x_2 = A \tag{3.3}$$

where:

$$A = \frac{\omega_{Q+1}(d + I(1 + J) + I^2) + dJ + (y_1^2 + y_1)(\omega_0^2 + \omega_Q)}{\omega_0^2 + \omega_0} \tag{3.4}$$

with $I = \omega_0\omega_Q, J = \omega_0 + \omega_Q$. Thus, Equation (3.3) needs $1I + 4M + 4S$ with I, M, S being field inversion, multiplier, and squares, respectively. Besides, solving Equation (3.3) by using half-trace algorithm [20] requires $\lceil \frac{m}{2} \rceil$ field squares.

After retrieving the x-term, the remaining coordinate is calculated by solving the quadratic equation as below:

$$y_2^2 + y_2 = C \tag{3.5}$$

where:

$$C = \frac{d(x_2^2 + x_2)}{d + x_2 + x_2^2} \tag{3.6}$$

Similarly, to derive the y-coordinate from the Equation (3.6), there is a need for $1I+1M$ to compute C. Additionally, a repetition $\lceil \frac{m}{2} \rceil$ times of field squares is also required for solving the quadratic equation (3.6) by the half-trace. The complexity of a general method using $\omega-$ coordinate is $m(2I + 2M + 3S) + mS + (2I + 6M + 5S)$ where m represents the order of the binary field.

3.1.2 The Proposed Point Operators using the Projective ω-coordinates in BEC

3.1.2.1 Proposed Point multiplication algorithm neglecting the retrieving affine coordinates

To reduce the complexity of the BEC-based RFID tags, firstly, the thesis proposes to remove the retrieving step from the ω-coordinates to affine coordinates. In the convention by Koziel *et al.* [11], the procedure of Point Multiplication using the ω-coordinates is presented in Figure 3.1. The steps in white boxes show the procedure of point multiplication using the *omega*-coordinate; meanwhile, the red blocks represent the retrieval of the affine coordinate steps.

At the beginning, the result of the point multiplication is the ω-coordinates of two consecutive points $\omega_Q = \omega(kP)$ and $\omega_{Q+1} = \omega(kP + 1)$. The term A in Equation (3.4) is processed in the following step before repeating $\lceil \frac{m}{2} \rceil$ times and accumulating the field square of A. As a result of the process, the x-coordinate of the result point $Q = kP$ is derived. The following calculates the term C in Equation (3.6) by the computed x-coordinate. In the next state, other repetition and accumulation of squared C are performed in $\lceil \frac{m}{2} \rceil$ times to calculate the y-coordinate of the result point Q. At the end of the process, after deriving both (x_Q, y_Q), the resulting affine point is sent to the server for the authentication process.

As explained above, the steps for retrieving the ω coordinate to the affine coordinate, represented as the red blocks in Figure 3.1, are complicated and incompatible

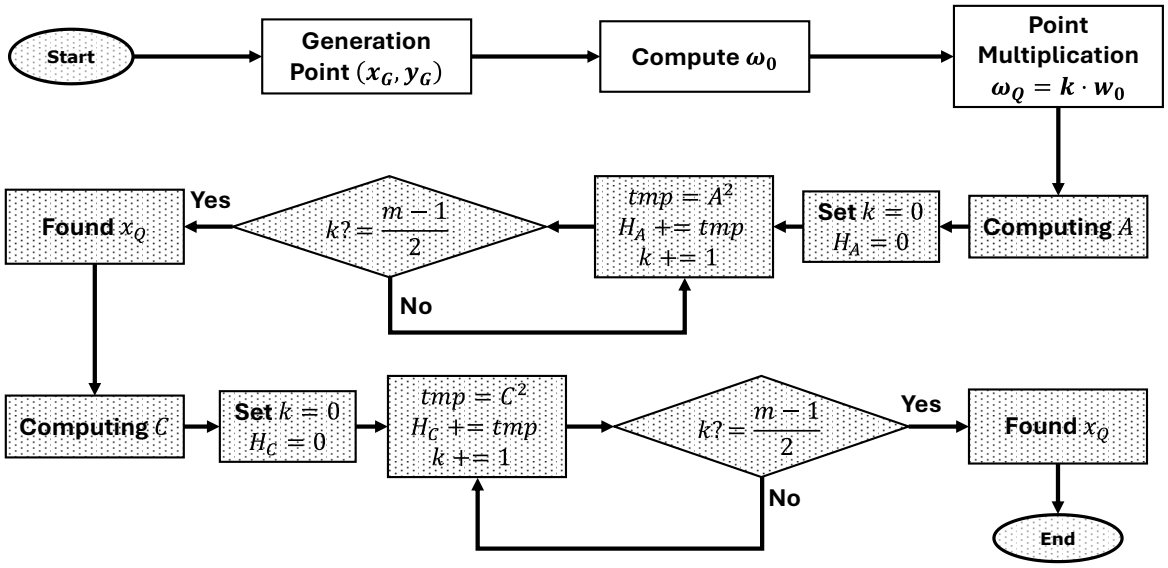


Figure 3.1: Conventional Computation of Point Multiplication $Q = kP$ using the ω-coordinates.

Algorithm 2: Montgomery Point Multiplication for BEC using ω-coordinates.

Input: Generator point $P(x_P, y_P) \in GF(2^m)$ and scalar k
Output: $\omega(Q) = \omega(kP)$ and $\omega(Q+1) = \omega(kP+1)$

Initialization:;
$\omega_0 \leftarrow x_P + y_P$;
Pre-compute $\frac{1}{\omega_0}$;
Set $\omega_1 \leftarrow \omega_0$;
Set $\omega_2 \leftarrow \texttt{DiffDouble}(\omega_1)$;
for $i \leftarrow m-2$ **to** 0 **do**
    if $k_i == 1$ **then**
        $\omega_1 \leftarrow \texttt{DiffAdd}(\omega_1, \omega_2, \frac{1}{\omega_0})$;
        $\omega_2 \leftarrow \texttt{DiffDouble}(\omega_2)$;
    else
        $\omega_1 \leftarrow \texttt{DiffDouble}(\omega_1)$;
        $\omega_2 \leftarrow \texttt{DiffAdd}(\omega_1, \omega_2, \frac{1}{\omega_0})$;
return $\omega_{kP} = \omega_1$ *and* $\omega_{kP+1} = \omega_2$;

with constrained devices. Consequently, the initial proposal involves relocating the recovery step to the server side, a non-constrained device potentially equipped with powerful accelerators.

The point multiplication is now executed in the passive RFID tags, as shown in Algorithm 2. Initially, the generator differential ω_0 is computed as $\omega_0 = x_P + y_Q$. In the next step, the point multiplication is executed in ω-coordinates, which results in

Table 3.1: Comparison of the complexity of various approaches for Point Multiplication in BEC.

Approach	Complexity
General	$m(2I + 2M + 3S) + \lceil \frac{m}{2} \rceil S + (1S + 4M + 4S)$
Neglecting retrieving step	$m(2I + 2M + 3S)$
Projective Coordinate with neglecting retrieving step	$m(6M + 5S)$

ω-coordinates of two consecutive points $\omega_Q = \omega(kP)$ and $\omega_{Q+1} = \omega(kP + 1)$. After that, the point multiplication is finished, as indicated in Algorithm 2. Consequently, the computational complexity is reduced by $mS + (2S + 5M + 4S)$.

3.1.2.2 Proposed the Point Operators

The computational complexity of the first proposal remains $m(2I + 2M + 3S)$. Nevertheless, according to the report by Deschamps *et al.* [17], the field inversion is the most complex operator over the binary finite field. The projective ω-coordinate is proposed to avoid the use of the field inversion. Instead of using the pure ω term, the differential coordinate is represented as $\omega = (W : Z)$, where $\omega = \frac{W}{Z}$. As a consequence, the Equations (3.1), (3.2) are modified. The modified point addition $(W_3, Z_3 = (W_1, Z_1 + (W_2, Z_2)$ and point doubling $(W_4, Z_4 = 2 \cdot (W_1, Z_1)$ are presented as Equation (3.7).

$$\begin{aligned} \frac{W_3}{Z_3} &= \frac{(W_1Z_2 + W_2Z_1)^2 + \frac{1}{\omega_0}(W_1Z_2 + W_2Z_1)^2}{Z_1^2Z_2^2 + \frac{1}{\omega_0}(W_1Z_2 + W_2Z_1)^2} \\ \frac{W_4}{Z_4} &= \frac{(W_1(W_1 + Z_1))^2}{d \cdot Z_1^4 + (W_1(W_1 + Z_1))^2} \end{aligned} \tag{3.7}$$

Considering the computational complexity, using the projective ω-coordinate, the point multiplication needs $m(6M + 5S)$, as described in Table 3.1. As discussed in Table 3.1, the projective ω-coordinate saves two field inversions, although it requires four field multipliers more and two field squares more. As a result, the hardware design implemented in this approach by Koziel *et al.* [11] requires five temporary registers. In hardware implementation, internal registers consume much more dynamic power than the combinational logic elements. Therefore, minimizing the number of internal registers reduces the power consumption of the device.

The second proposal in this thesis is optimizing the point operator to minimize the number of storage registers used. The quartic term of Z_1^4 in Equation (3.7) over

Table 3.2: Computation Order of the Proposed Point Operators.

	Register A	**Register B**	**Register C**	**Register D**
Step 0	-	-	$(W_2Z_1)^2$	-
Step 1	-	-	$C + (W_1Z_2)^2$	-
Step 2	-	-	-	$(Z_1Z_2)^2$
Step 3	-	-	$C + \frac{1}{\omega_0}C$	$D + \frac{1}{\omega_0}C$
Step 4	$(W_1(W_1 + Z_1))^2$	-	-	-
Step 5	-	$(Z_1Z_1)^2$	-	-
Step 6	-	$A + d \cdot B$	-	-

$GF(2^m)$ can be represented as below:

$$Z_1^4 = (Z_1 \cdot Z_1)^2 \tag{3.8}$$

Consequently, the alternative of the point addition and point doubling is presented as Equation (3.9)

$$\begin{aligned} \frac{W_3}{Z_3} &= \frac{C + \frac{1}{\omega_0} \cdot C}{(Z_1Z_2)^2 + \frac{1}{\omega_0} \cdot C} \\ \frac{W_4}{Z_4} &= \frac{A}{d \cdot (Z_1 \cdot Z_1)^2 + A} \end{aligned} \tag{3.9}$$

where common variables $C = (W_1Z_2)^2 + (W_2Z_1)^2$ and $A = (W_1(W_1 + Z_1))^2$.

Regarding the computational complexity, the quartic term Z_1^4 requires one field square and one field multiplier. The point multiplication using the proposed operators, which are presented in Equation (3.9), requires $m(6M + 4S)$. Consequently, using the proposed point operators, assuming that the hardware design includes one field multiplier and one field square the process of a point multiplication requires four m−bit registers instead of five.

In particular, in the first step, register C is assigned the results of the $(W_2Z1)^2$ followed by an accumulation with $(W_1Z_2)^2$ in second step. Third step processes the square of multiplication(Z_1Z_2) and assigning the results to register D. Fourth step appears the results of W_3, Z_3 at register C and D, respectively, by accumulating $1/\omega_0$. In step 5th, value of W_4 is computed as the result of $(W_1(W_1 + Z_1))^2$ and stored in register A. Step 6th computes the quadric term of Z_1 and stores in register B. In the last step, the value of Z_4 results from the XOR operation between register A and the multiplication of register B with the constant d. The order of the computing and storing of the temporary value for the point addition and point doubling is indicated in Table 3.2.

3.2 The proposed Hardware Architecture of Binary Edwards Curves

A systematic architecture of a low-cost, low-power Binary Edwards Curve is proposed to implement the proposed point multiplication algorithm that neglects the retrieval of affine coordinates for the constrained passive RFID tags. In particular, a top view of the proposed module is presented first before an insightful description of the proposed arithmetic computation block (ACB).

3.2.1 The Proposed Arithmetic Computation Block

An integrated block of arithmetic computation (ACB) is proposed to process immediately the field multiplier and the field square as depicted in Figure 3.2 for minimizing the number of registers. A detailed illustration of the proposed ACB block is presented in this section.

3.2.1.1 Integrated Arithmetic Computation Algorithm

To implement the second proposed point operator, which is explained in Equation (3.9), a general equation $f(a, b, \texttt{cond})$ is implemented as the integrated arithmetic computation in the following equation:

$$f(a, b, \texttt{cond}) = \begin{cases} (ab)^2, & \texttt{cond = 1} \\ ab, & \texttt{otherwise} \end{cases} \tag{3.10}$$

According to Equation (3.10), when `cond = 1`, the ACB processes the field square of

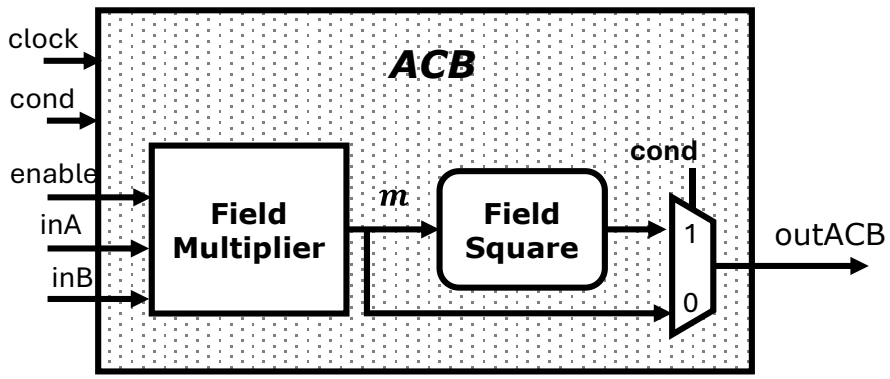


Figure 3.2: Proposed architecture of the arithmetic computation block.

the multiplication of inputs $f(a, b, 1) = (ab)^2$. Otherwise, when `cond = 0`, the ACB multiplies 163-bits inputs of a and b.

The input pins of the proposed ACB include synchronization signal (`clk`), controlling signals (`cond`, `enable`), and pins of 163 bits of `inA` and `inB`. The 163-bit output pin `outACB` is the result of the function $f(a, b, \texttt{cond})$.

3.2.1.2 Interleaved Field Multiplier

One field multiplier and one field square are integrated within the arithmetic computation algorithm. At the beginning, the field multiplier of $a(x)$ and $b(x)$ is processed as the following:

$$\begin{aligned} a(x) &= \sum_{i=0}^{m-1} a_i x^i \\ b(x) &= \sum_{i=0}^{m-1} b_i x^i \end{aligned} \tag{3.11}$$

The interleaved field multiplier of $a(x) \cdot b(x)$ `mod` $f(x)$, which is declared by Deschamps *et al.* [17], utilizes the bit-serial computation to minimize the implementation cost. Multiplication of two elements $a(x) \cdot b(x)$ `mod` $f(x)$ over the finite field $GF(2^m)$ with the irreducible polynomial $f(x) = f_0 + f_1 x + \cdots + f_m x^m$ is expressed as:

$$\begin{aligned} a(x) \cdot b(x) &= \sum_{i=0}^{m} a_i x^i \cdot b(x) \ \texttt{mod} \ f(x) \\ &= (a_0 \cdot b(x) + a_1 x \cdot b(x) + \cdots + a_{162} x^{162} \cdot b(x) \ \texttt{mod} \ f(x) \\ &= (a_0 \cdot b(x) \ \texttt{mod} \ f(x)) + (a_1 x \cdot b(x) \ \texttt{mod} \ f(x)) + \cdots + \\ &(a_m x^m \cdot b(x) \ \texttt{mod} \ f(x)) \end{aligned} \tag{3.12}$$

In Equation (3.12), the term $x \cdot b(x)$ `mod` $f(x)$ is expressed a one position shift-left register:

$$\begin{aligned} x \cdot b(x) &= x \cdot \sum_{i=0}^{m-1} b_i x^i \ \texttt{mod} \ f(x) \\ &= x \cdot (b_0 + b_1 \cdot x + \cdots + b_{m-1} \cdot x^{m-1}) \ \texttt{mod} \ f(x) \\ &= f_0 \cdot b_{m-1} + x \cdot (b_0 + f_1 \cdot b_{m-1}) + \cdots \\ &+ x^{m-1} \cdot (b_{m-2} + f_{m-1} \cdot b_{m-1}) \\ &= b_{m-1} \cdot f_0 + \sum_{i=0}^{m-2} (b_i + b_{m-1} \cdot f_{i+1}) \cdot x^{i+1} \end{aligned} \tag{3.13}$$

According to the expression depicted in Equation (3.13), the interleaved field multiplier value $a(x) \cdot b(x)$ `mod` $f(x)$ is utilized as Algorithm 3. Within $GF(2^m)$, the interleaved field multiplier requires $2m$ `#AND` $+(2m-1)$ `#XOR` and three $m-$ bit registers for storing $a(x), b(x)$, and $c(x)$.

Algorithm 3: Interleaved Field Multiplier of $a(x) \cdot b(x) \mod f(x)$.

Input: Polynomials $a(x)$ and $b(x)$ defined over $GF(2^m)$ with irreducible polynomial $f(x)$

Output: $c(x) = a(x) \cdot b(x) \mod f(x)$

```
Initialization:;
Set c(x) ← 0;
for i ← 0 to m − 2 do
    if a_i == 1 then
        c(x) ← c(x) + b(x);
    b(x) ← x · b(x);
return c(x);
```

3.2.1.3 Classic Field Square

In addition, the field square in this thesis is implemented as pure combinational logic to reduce the number of flip-flops. Thanks to this structure, which is explained in more detail by Deschamps *et al.* [17], the classic field square requires $(m^2 - 2m)$`#AND` $+(m^2 - 3m + 2)$`#XOR`. However, Paar [54] showed that the computational complexity of classic field square depends significantly on the selected irreducible polynomial. Therefore, by choosing the irreducible polynomial $f(x) = x^{163} + x^7 + x^6 + x^3 + 1$, the classic field square is implemented as 828 `#XOR`.

3.2.2 The Proposed System Architecture of BECs

3.2.2.1 Systematic Architecture of BECs

The Binary Edwards Curve (BEC) offers a range of key lengths correlating with different security levels. Depending on the application, security can vary from standard with 163-bit keys to robust with 571-bit keys. This key length is 80-bit to 256-bit security, similar to DSA [55]. The BEC architecture is versatile, accommodating various security needs.

This thesis focuses on optimizing the implementation costs of the BEC-163 over the Binary Field $GF(2^{163})$ for the passive RFID tag application. Several irreducible polynomials are available within the Finite Field $GF(2^{163})$. In an additional approach, the irreducible polynomial with the lowest Hamming weight is selected to minimize the design complexity, as expressed in Equation (3.14):

$$f(x) = x^{163} + x^7 + x^6 + x^3 + x + 1 \tag{3.14}$$

The proposed top-level module of BEC performs the point multiplication, represented as $Q = k \cdot P$, where k is a scalar integer playing the role of the private key of the passive

RFID tag and $P = (P_x, P_y)$ is the generation point. To safeguard the private key k, the Montgomery Point Multiplication is employed, which is presented in Algorithm 2. This approach computes iteratively using both point differential addition and point differential doubling. These processes are repeated over 163 loops.

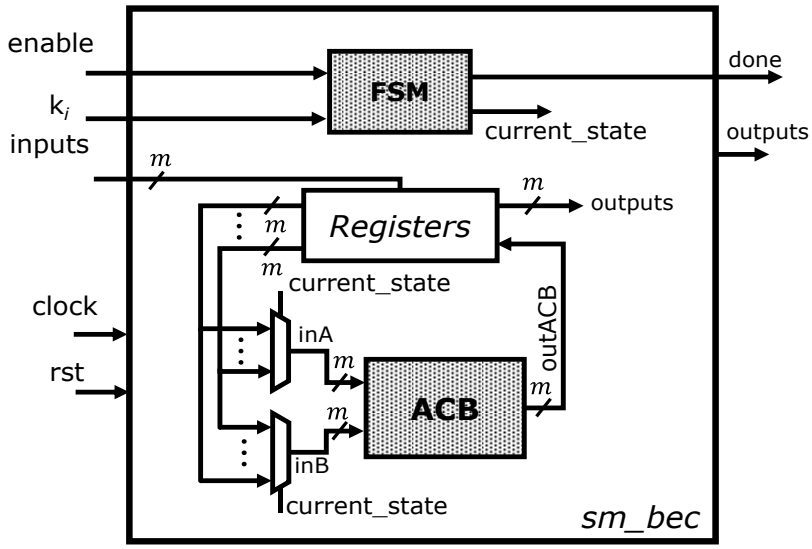


Figure 3.3: Proposed low-cost, low-power architecture of Binary Edwards Curve using bit-serial processing.

Furthermore, addressing issues regarding optimizing the implementation costs of the Binary Edwards Curve, the top module of the proposed hardware architecture applies bit-serial processing to minimize the number of internal registers and the number of arithmetic processing units, as shown in Figure 3.3. In particular, using the bit-serial architecture, the implemented hardware requires four internal registers (A, B, C, and D), each of 163 bits, and one block for field arithmetic computation. In addition, there is one controller component, which implements the finite-state machine and manipulates the computing activities of the module.

3.2.2.2 Interface of the Proposed Hardware Design of BECs

Table 3.3 presents the Interface of the proposed BEC. There are five input ports and two output ports. Among the inputs, signals `'clk'`, `'rst'`, and `'enable'` are used for controlling the module. In particular, the `'clk'` is the synchronization signal for aligning activities of the proposed BEC module. Besides, `'rst'` signal is the reset signal, which enables the initialization of the value of all internal registers. In addition, the signal of `'enable'` starts the proposed BEC module to compute the point multiplication when all the `inputs` data is fed fully into the proposed module. The `'inputs'` signal

includes six configuration values, each of 163 bits. They are $\omega_0, W_1, Z_1, W_2, Z_2$, which presents the generation points $P = (P_x, P_y)$; meanwhile, the pre-defined parameter d represents the particular Edwards Curve.

Table 3.3: The interface of the proposed BEC-163.

Name of port	Width (bits)	Group	Description
Input Ports			
`clk`	1	Synchronization Signal	Synchronous Clock Signal
`rst`	1	Control Signal	Asynchronous Positive Reset Signal
`enable`	1	Control Signal	Enabling the Top module to process
k_i	1	Data Signal	Bit of Private Key
`inputs`	163	Data Signal	Set of configuration signals including $\omega_0, W_1, Z_1, W_2, Z_2, d$
Output Ports			
`done`	1	Indicate Signal	Inform the Top Model completely computes the point multiplication
`outputs`	163	Data Signal	Set of resulted signals involving W_3, Z_3, W_4, Z_4

According to the illustration of Table 3.3, the top module of the proposed BEC owns one indicator signal `'done'` and one data out signal of 163 bits `'outputs'`. The `'done'` signal indicates the status of the top module that the design is, whether completely calculated or not. When this signal is not active, the point multiplication is in progress. In contrast, if the `'done'` signal is active, the point multiplication is completed, and the operation results are ready at the `'outputs'` port. The operation results involve four values, each of 163 bits, such as (W_Q, Z_Q) and (W_{Q+1}, Z_{Q+})

3.2.2.3 Finite State Machine

A model of the Mealy Finite State Machine (FSM) is presented in Figure 3.4 to control the datapath and the computational progress of the proposed BEC block. There are four states: `'IDLE'`, `'Load data'`, `'Compute'`, and `'Done'` to minimize the complexity of the controller block. In addition to the control signals explained above, the states

of the FSM are controlled by the loop counter value `'loop_cnt'`. Particularly, the illustrations of certain states in the FSM are listed below:

- `'IDLE'`: At this state, all the values of the internal registers are initialized to 0. The indicator signal `'done'` is also reset to 0. When the `'rst'` is not active, the current state is transformed to the next state `'Load data'`.
- `'Load data'`: At the `'Load data'` state, configuration parameters $\omega_0, W_1, Z_1, W_2, Z_2, d$ are assigned on the internal registers. After all the pre-defined parameters are loaded into the internal registers, the control signal `'enable'` activates and triggers the current state to transform to the state `'Compute'`.
- `'Compute'`: After loading all configuration variables to the internal registers, state `'Compute'` starts to compute and assign the results of the point operators to the internal registers. These operations are repeated over 163 loops. In the following, when the loop counter value `'loop_cnt'` = 163, the indicator signal `'done'` raises an active level and triggers the machine to transfer to the last state. Besides, the results of point multiplication (W_Q, Z_Q) and (W_{Q+1}, Z_{Q+}), which are stored in the internal registers, are fed into the `'output'` port.
- `'Done'`: In the last state of the proposed FSM, the indicator signal `'done'` keeps the same value before a transformation of state to the `'IDLE'` of the machine.

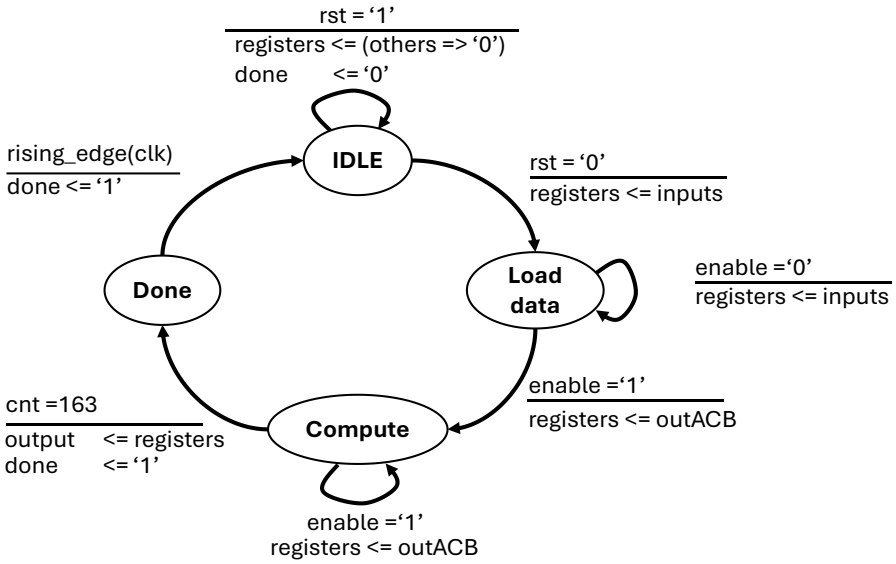


Figure 3.4: Block diagram of the proposed Mealy FSM.

In the `'Compute'` state, in addition to the register `'loop_cnt'`, another register, `'step_cnt'`, is utilized for manipulating the input data for the point operator. The

computation workflow of the point operator within this state is presented in Figure 3.5. When the FSM transforms to the `'Compute'` state, both `'loop_cnt'` and `'step_cnt'` are initialized to 0. Once `'enable'` signal is activated, the inputs of the points operator are assigned according to the `'step_cnt'`.

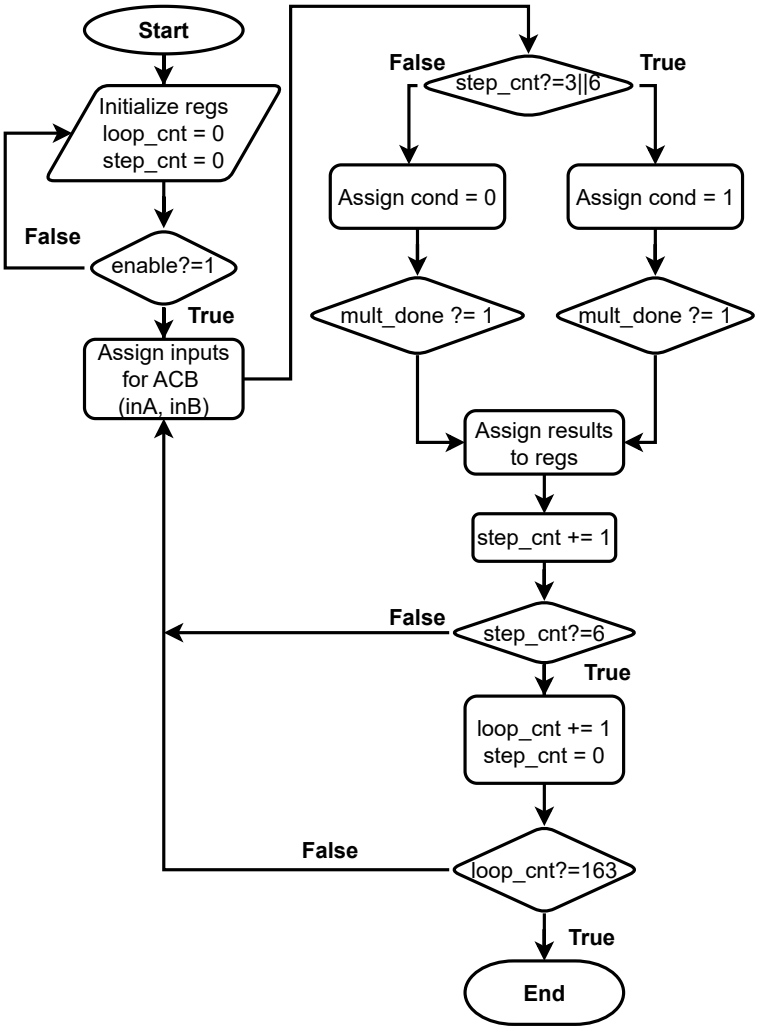


Figure 3.5: Flow chart of computation in `Compute` state.

After the field operator completes its process, the results are stored in the internal registers, and the `'step_cnt'` is incremented. If `'step_cnt'` reaches the last step, the register `'loop_cnt'` is incremented. Conversely, the inputs to the point operator are reassigned, and `'step_cnt'` is reset to 0. For example, in the proposed hardware of BEC, there are seven rounds of `'step_cnt'` ranging from 0 to 6. These processes are repeated for 163 iterations. During the final iteration, when `'loop_cnt'` equals 163,

computation concludes, prompting the FSM to transition to the next state.

3.3 Hardware Implementation Results

As standardized in the ISO/IEC-14443, the maximum protocol that executes the procedure of authenticating an object in an RFID application is no longer than $20ms$. Besides, the state-of-the-art rectifier antenna also shows that the budget of power consumption is $240\mu W$ [2]. By implementing the proposal architecture on hardware, the compatibilities of the proposal are validated.

3.3.1 Experimental Setup

In this section, the implementation of the proposed design of the BEC architecture will be presented. The implementation cost and the security properties of the proposed BEC design will also be provided and discussed. The top module is implemented with VHDL language, and behavioral simulation in QuestaSim is developed for system verification. A set 10,000 random keys (k_{rand}) with a generation point $P(\omega_P, z_P)$ and corresponding points $Q(\omega_Q, z_Q) = k \cdot P(\omega_P, z_P)$ is the golden model for the functional evaluation and post-synthesis simulation.

The validation of the system is processed by Synopsys Electronic Design Automation (EDA) tools such as Design Compiler and PrimeTime with the CMOS 65nm from TSMC at the operational frequency of $10MHz$. The implementation cost includes the area cost, which is measured by the number of equivalent gates of the design and reported by Synopsys Design Compiler. In addition, the latency of the point multiplication is measured by the duration of completing an operation of the point multiplication and reported from the post-synthesis simulation. The power consumption measured in (μW) is estimated and reported by the Static Timing Analysis process in the Synopsys PrimeTime. Regarding the security validation, there are several approaches to assessing the vulnerabilities of the proposed design against SCA. One of the most popular and general methods is measuring the leakage information from the side channel trace, such as power or electromagnetic, which is known as Test Vector Leakage Assessment (TVLA) [56] as illustrated in Figure 3.6.

In Figure 3.6, the power traces of two experiments, fixed key and random key, are used to compute the means and the standard deviation metrics. After that, by applying the *t-statistic* method using Welch's t-test equation in Equation (2.5), an assessment trace T is computed. The amount of leakage α less than the threshold of 4.5 is safe for the SCA vulnerabilities.

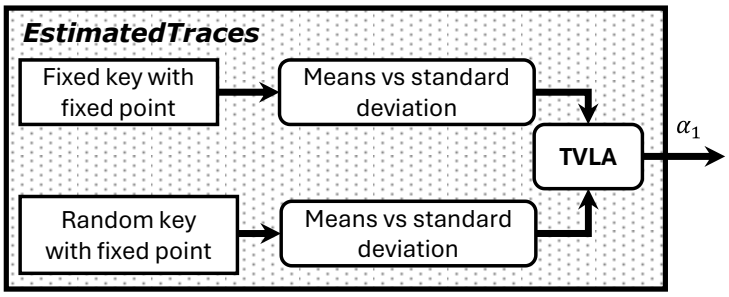


Figure 3.6: Evaluation flow of TVLA.

3.3.2 Hardware Implemntation Results

The equivalent area cost report of the proposed design is 21.68 $kGates$, approximately $0.031mm^2$, in which the ACB block takes 5.641 $kGates$ and 3.9 $kGates$ of the immediate register. The percentage of each component in the proposed top module is presented in Figure 3.7. It is also noticed that the proposed BEC point multiplication occupies only 12.7 $kGates$ and four internal registers of 163 bits for storing the data. The minimum average power consumption of the design is $126\mu W$, which consumes an energy of $2.4\mu J$ for each point multiplication. Regarding the latency of the proposed design, each point multiplication is executed in a delay of $19.05ms$.

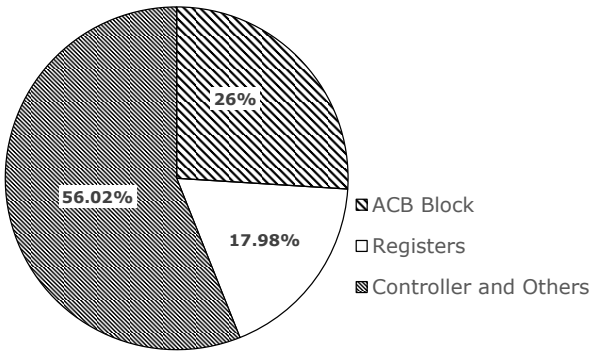


Figure 3.7: Percentage of the area cost of each block.

Table 3.4 compares the implementation cost between the related works and the proposed design of BEC. To fairly compare hardware implementations, the efficiency of the energy metric is defined as the number of processed bits in one microjoule over a clock cycle ($bits/(\mu J \cdot \texttt{cycle})$) as follows:

$$\texttt{Energy Eff.} = \frac{m}{E \cdot T} \tag{3.15}$$

where m denotes the number of representation bits of the secret key, E represents the energy consumption (μJ), and T is the latency of the point multiplication measured in number of clock cycles.

Table 3.4: Comparison of the related works.

Works	**Design**	**Tech.** (nm)	**Freq.** (MHz)	**Delay** (ms)	**No. Cycles**	**Area**		**Power** (μW)	**Energy** (μJ)	**Energy Eff.** ($b/(\mu J \cdot \texttt{cycle})$)
						Core (kGates)	**No. Flip-flops**			
Koziel [11]	BE163[(1)]	65	1	177.707	177,707	10.95	5·163	-	-	-
Kocabas [20]	BE163	130	0.4	547.87	1,369,675	11.72	6·163	7.27	**1.29**	$9.25 \cdot 10^{-5}$
Lara-Nino [57, 58]	BE251[(1)]	xc6s-lx16	100	8.24	824,284	-	-	83,000	680	$4.48 \cdot 10^{-7}$
Dan [59]	K163[(2)]	-	10	16.5	165,000	24,140	-	5,940	98.01	$1.01 \cdot 10^{-5}$
Rashidi[(3)] [60]	BE233[(1)]	180	**1,000**	**0.118**	126,983	29.52	$10 \cdot 233$	-	-	-
Rovzic [61]	K163	130	1.13	135.6	**86,000**	10,106	-	36.63	9.16	$2.07 \cdot 10^{-4}$
Bui [7]	AES-128	65	10	$4.3 \cdot 10^{-3}$	43	8.6	-	20	-	-
Gross [9]	ASCON	90	1	-	-	3.75	-	15	-	-
Bahnsawi [12]	RSA-1024	130	11.9	-	-	$1.283mm^2$	-	11,430	-	-
This work	BE163	65	10	19.05	190,547	12.74	4·163	126	2.4	$\mathbf{3.56 \cdot 10^{-4}}$

(1)BE163, BE233, BE251 implements the point multiplication over Binary Edwards Curve within $GF(2^{163})$, $GF(2^{233})$, and $GF(2^{251})$, respectively.
(2)K163 implements the point multiplication over Koblitz Curve within $GF(2^{163})$.
(3)This work implemented the digit-serial architecture of BEC with the size of word $\omega = 9$.

Compared to symmetric cryptography, such as AES [7] and ASCON [9], our proposal occupies a physical area 1.5 times and 3.5 times larger, respectively. Regarding the power consumption, our proposal BEC design takes $126\mu W$, which is 6 times and 8 times higher compared to the AES [7] and ASCON [9] respectively. However, these symmetric cryptography designs trade off the security issue related to the key distribution. Therefore, to deploy the practice system using symmetric cryptography, the key distribution mechanisms, which are based on the asymmetric algorithms, are supplemented to securely share the secret key with seperated parties. In another comparison, 1024-bit RSA [12] shows a higher demand for power consumption, which is approximately 100 times larger compared to our design. As a result of the comparison to the different cryptographies, such as AES [7], ASCON [9], and RSA [12], our proposal design shows an efficiency that balances between the hardware cost and security level.

Among the Elliptic Curve Cryptography, it has been noted that the fastest hardware implementation is proposed by Rovzic *et al.* [61], which executes a point multiplication over the Koblitz curve during 86,000 clock cycles, which is half of our design's demand. By minimizing the complexity of the point multiplication, hardware implementations of Dan *et al.* [59] and Rovzic *et al.* [61] are generally faster than those executing the BEC [11, 20, 57–60]. However, these curves are incomplete, which makes them vulnerable to power analysis. Among the hardware implementations of complete curves (BEC), Rashidi *et al.* [60] provided a digit-serial architecture with processed word $\omega = 9$, that requires the least cycles of clock for executing the point multiplication of 126,983 cycles. Besides, as implemented at $1GHz$ of operational frequency, the point multiplication proposed by Rashidi *et al.* [60] executes in $0.118ms$. The trade-off of this parallel architecture is a larger area, twice times compared to our proposed hardware design.

In terms of the physical area cost of the BEC implementation, Koziel *et al.* [11] proposed a low-cost architecture at 10.95 kGates of BEC and five registers of 163 bits, which are 14.05% smaller than our proposal. The main disadvantage of this design is disabling the completeness property [62]. By using the Co-Z trick, the BEC implementation of Koziel saves one field multiplier in each point operator. However, this leads to the variance of $(R_1 - R_0)$, where R_1, R_0 are the results of point addition and point doubling, respectively. As a consequence, this hardware architecture is vulnerable to power analysis attacks. Regarding energy consumption, our proposal offers the highest efficiency at $3.56 \cdot 10^{-4}(bits/(\mu J \cdot \texttt{cycle}))$, which is 1000 times better than the implementation on FPGA by Lara-Nino *et al.* [58].

According to the comparison analysis, this work proposed a low-cost, low-power, and energy-efficient architecture for BEC. Additionally, referring to the constraints of

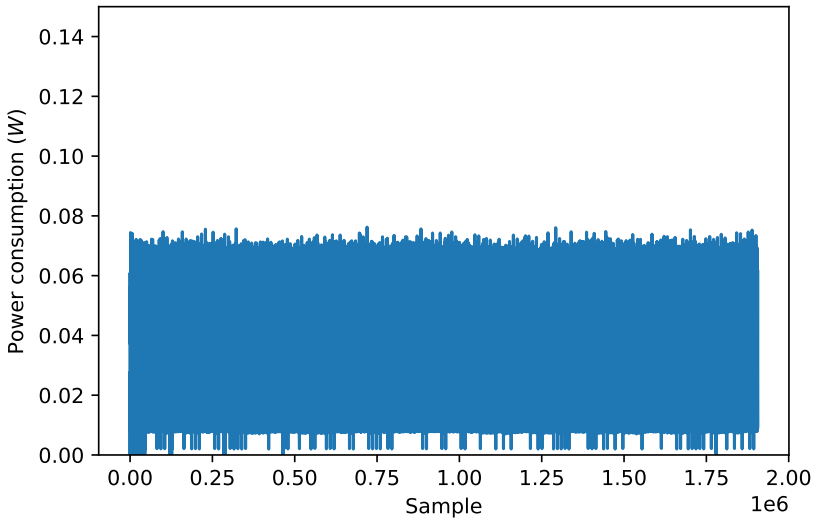


Figure 3.8: Power trace of the point multiplication recorded by PrimeTime.

the passive RFID tags, our proposed hardware design satisfies all the requirements of latency ($19.05 < 20(ms)$) and power consumption ($126 < 240(\mu W)$).

3.3.3 Security Evaluation

Thanks to the VCD file generated from the post-synthesis simulation, PrimeTime records the power traces of the point multiplications. The power traces describe the instance power consumption of the proposed design, depicted in Figure 3.8. They evaluated the system's robustness against SCA to ensure the security property of the system.

In this experiment, to evaluate the power leakage of the proposed BEC design, a set of 50 fixed keys and 50 random keys is implemented. The evaluation result is illustrated in Figure 3.9. The first 1000 samples of the $\alpha-$ traces show the most leakage source compared to the remaining part of the trace. Nevertheless, the maximum leakage, which occurs in the beginning of the encryption, is 3.36, smaller than the standard threshold of 4.5. In the remaining, the *tau*-value is smaller than 1, which provides a strong security against the power analysis techniques. Thus, the proposed design is secured against the SCA, especially the power analysis.

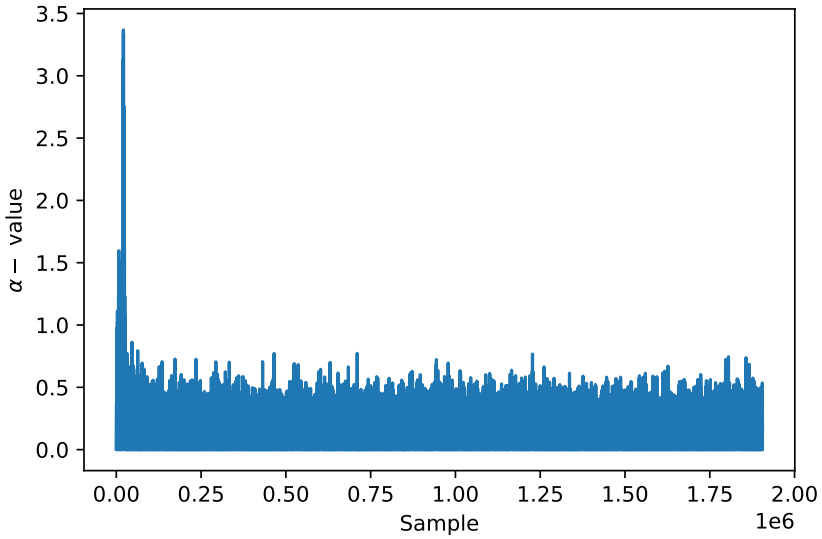


Figure 3.9: Experiment of leakage assessment of the proposed BEC design.

3.4 Summary

This section of the thesis proposes a low-cost, low-power architecture with a modified algorithm for point multiplication of BEC. This modification firstly removes the retrieving step from the affine coordinate to the $\omega-$ coordinate to reduce the complexity of the design. In addition, a minimized point operator of the BEC within the $\omega-$ coordinate is proposed to reduce the number of the internal registers. To deploy the minimized point operator, an arithmetic computation block (ACB) is lastly proposed to minimize the number of internal registers, decreasing the system's implementation cost. According to the synthesis results within the CMOS 65nm technology library from TSMC, the proposed BEC provides a low-cost, low-power design that satisfies the physical requirements of passive RFID tags. In addition, the TVLA evaluation of the system also validates the security property against the SCA threats. This proposal for the thesis is published in [C1].

However, the time-to-production is proportional to the repetition in selecting the design. Additionally, the late supplementing of the security countermeasure prevents the designers from considering both security and the implementation cost in the early design phase. Therefore, there is a critical need for a novel design methodology that enables the design consideration and evaluation of both security and implementation costs at the beginning.

Chapter 4

Proposed Early Evaluation Design Methodology

As stated in the summarization of the previous chapter, the time-to-production, as well as the time-to-market, is one of the most critical issues of the VLSI design. With the development of semiconductor applications, the complexity of electronic devices has risen dramatically. As a result, the integration of millions of sub-blocks with the heterogeneous constraints necessitates careful planning and optimization across various stages, including architectural design, simulation, verification, and physical layout. In addition, several device architectures, algorithms, and configuration parameters must be implemented and verified carefully before choosing the most compatible ones. Therefore, as significant as the repetition of these multiple stages is, it escalates the time-to-market cost.

This chapter presents the proposed EEMitM design methodology in Section 4.2. This combines the conventional design methodology and the early analysis mechanism, which helps designers consider both hardware security issues and the implementation cost of ECC-based RFID systems in the early stages of the design process. In addition, innovative approaches to generating the pseudo power traces are introduced in Section 4.3 before showing the results of implementing the proposed design methodology.

4.1 Introduction

In the context of security devices, such as passive RFID tags, complementing countermeasures at the end of the design process also involves a critical challenge. The additional computations increase the hardware implementation cost of the design. Consequently, this motivates designers to propose design flows incorporating the security evaluation in each design step.

Several established design methodologies, which are presented in more detail in Chapter 2, have been developed to include security considerations in digital design. Among these, the top-down design approach is the most prevalent for designing Elliptic Curve Cryptography (ECC) primitives, and it has been widely applied in numerous studies [31–35] within the literature. However, a significant limitation of these implementations is their lack of detailed information regarding the systematic architecture and implementation costs of the internal modules. Additionally, in the context of the ECC-based authentication protocol for passive RFID, the heterogeneous constraints of the alternative implementation levels challenge the designer to estimate the systematic costs. As a result, these approaches cannot substantiate their designs' compatibility with the specific constraints inherent to passive RFID tag systems, such as power limitations, computational capacity, and communication requirements.

In addition to the top-down approach, several works [41, 42] have explored using the bottom-up design methodology. A key limitation of this methodology is that it is time-consuming to develop, which becomes increasingly problematic when applied to complex systems. As the system's complexity grows, the simulation, verification, and validation become exponentially more difficult and resource-demanding. This results in a considerable increase in both the time and computational effort required to evaluate the correctness and functionality of the design. Consequently, the bottom-up methodology proves to be less compatible with the design of ECC primitives, particularly given the intricate nature of ECC implementations that require fine-tuning, optimization, implementation, and evaluation.

In summary, both the top-down and bottom-up approaches share a critical oversight: they often treat security as an afterthought or an additional feature to be integrated into the design at a later step of the design flow rather than considering it as a foundational requirement that must be embedded throughout the entire design process. This misjudgment undermines the ability to simultaneously achieve an efficient and secure solution, complicating the development of ECC-based authentication protocols. As a result, these methodologies struggle to strike an effective balance between meeting stringent security requirements and maintaining an acceptable implementation

cost, which is crucial for resource-constrained environments such as passive RFID systems. The inability to adequately address security at every stage of the design process leads to challenges in producing robust and cost-effective ECC-based solutions.

In order to solve the mentioned problem, this thesis proposes an Early Evaluation Meet-in-the-Middle design methodology. Specifically, the contribution regarding the design methodology in this chapter of the thesis is an innovative design methodology. The proposal enables designers to quickly choose the optimal design based on the database of the primitives. Thanks to using the primitive component database with the analyzing mechanism of the proposed design methodology, the time-to-production for finding the optimal design could be reduced by 480 times better than the conventional approaches.

4.2 Proposed Early Evaluation Meet-in-the-Middle (MitM) Design Methodology

This section presents the proposed Early Evaluation Meet-in-the-Middle (EEMitM) design methodology, as depicted in Figure 4.1. The proposed design framework integrates a bidirectional design approach, combining MitM architectural principles with a Bottom-Up Evaluation and Validation process. In the initial design phase, system architects define the functional specifications for the target system, including constraints on implementation costs and security requirements. Through iterative analysis of modular dependencies and design variations, the resultant design of optimized security primitive configuration and its associated authentication protocol architecture is achieved at the end of the process.

The Bottom-Up Evaluation and Validation phase constitutes a critical empirical verification stage within the proposed EEMitM designing framework, executed subsequent to the hardware integration of the derived system architecture. As a consequence of leveraging prototype implementations of the synthesized cryptographic primitives and authentication protocols, this process facilitates rigorous experimental validation of both functional correctness and security robustness under real-world operational constraints. Quantitative metrics—including security level, power consumption profiles, physical area cost, and timing analysis—are systematically extracted from the hardware-realized system to assess compliance with pre-established security benchmarks. Concurrently, structural integrity and protocol adherence are evaluated to ensure alignment with constraint specifications and threat mitigation objectives.

4.2.1 Early Evaluation Meet-in-the-Middle Design Process

The proposed Early Evaluation Meet-in-the-Middle (EEMitM) necessitates initial design specifications, including the expected physical area cost, power consumption, latency, and desired security level. These quantitative metrics would be utilized as the foundational criteria for the systematic evaluation, which enables an early analysis and optimization mechanism of the proposed design methodology. The methodology is structured into three sequential phases: Authentication, Scalar Multiplication, and Field Operators Design, as described in Figure 4.1.

The proposed EEMitM design framework begins by establishing the quantitative metrics of design specifications, followed by the authentication design phase. This initial phase focuses on selecting an optimal authentication protocol and its corresponding security primitives, leveraging existing knowledge from databases such as those detailing ECC (Elliptic Curve Cryptography) primitive blocks proposed in prior research.

The completion of the first phase within the EEMitM design methodology enables designers to gain critical insights into the constraints governing security primitives, ensuring the alignment with the operational limitations of resource-constrained devices. These constraints encompass quantitative metrics such as physical area cost, temporal latency, and equivalent energy consumption, which collectively define the implementation boundaries of cryptographic components like elliptic curve cryptography (ECC). By leveraging these specifications, a systematic selection process is initiated to identify optimal configurations for scalar multiplication.

In the last phase, the field operators phase, the design constraints of each field operator are determined as the output of the previous design phase. Algorithms and architectures of field operators are assessed to select the optimal design. At the end of the process, a detailed systematic design of the optimal ECC-based authentication protocol is determined before starting to implement it by using the Hardware Description Language (HDL).

Within each phase, the aforementioned analytical mechanisms are applied to expedite the assessment of distinct design alternatives. This evaluative process facilitates the rapid identification of optimal architectural configurations by comparing algorithmic implementations and parameter settings against predefined performance benchmarks. By integrating these assessment protocols, the framework enhances decision-making efficiency, allowing designers to prioritize solutions that align with the stipulated cost, efficiency, and security requirements.

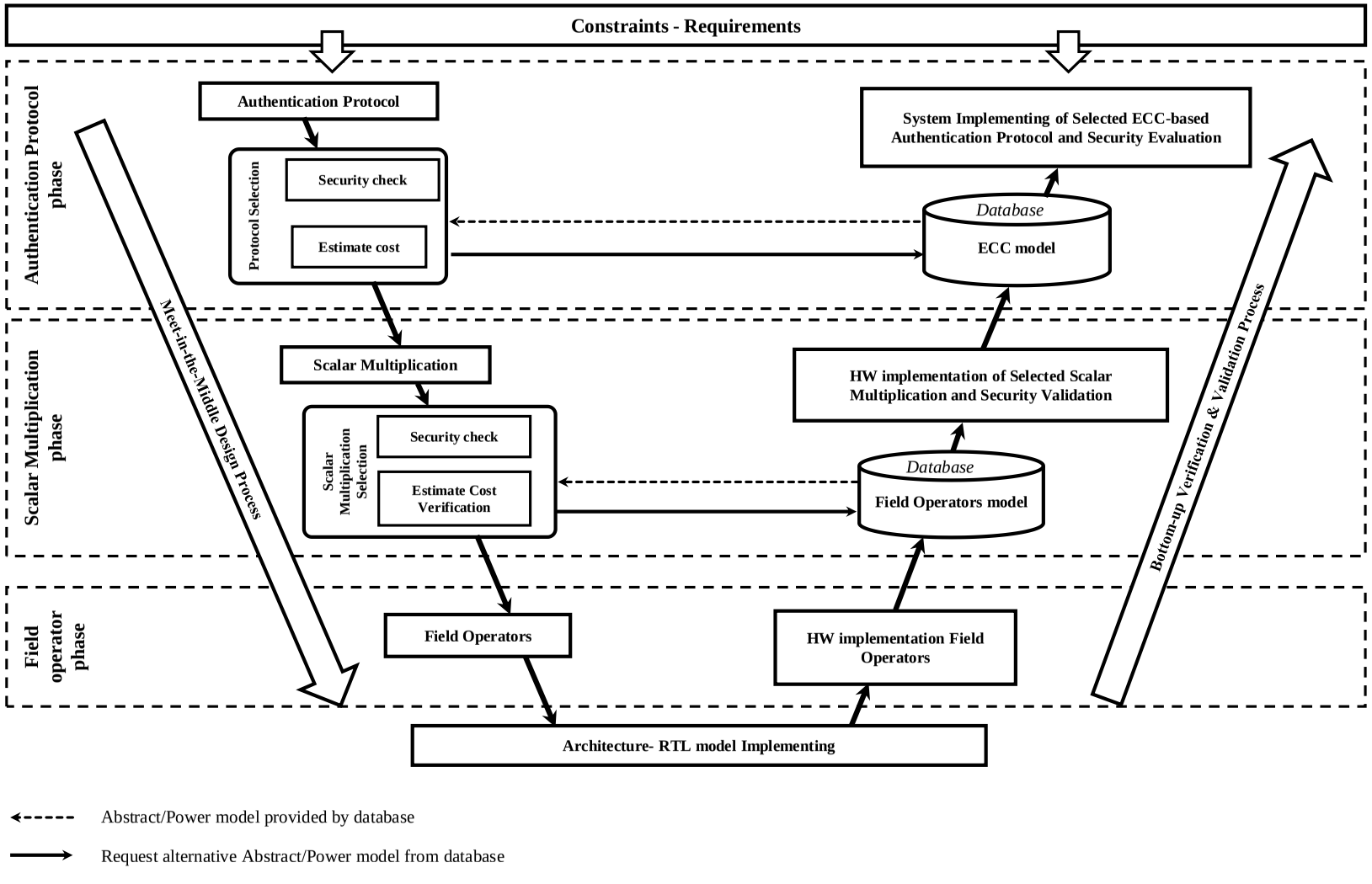


Figure 4.1: Proposed Early Evaluation Design Methodology.

4.2.1.1 Authentication Protocol Design Phase

In the first phase of the proposed design framework, designers evaluate the primitives based on implementation costs, including physical area, power consumption, and latency, as well as the security efficacy of their algorithms. By analyzing these factors, a model is constructed to balance security requirements with practical constraints, enabling informed decisions about protocol suitability. This systematic approach ensures that the chosen protocol aligns with both performance benchmarks and cryptographic robustness.

The database of hardware components for security primitives corresponds to the orange block illustrated in Figure 4.2. It includes mathematical representations of cryptographic elements such as elliptic curve cryptography (ECC), hash functions, random number generators (RNGs), and other essential operators, which are utilized during the functional assessment phase. During this step, these mathematical models are aligned with the authentication protocol's structure to verify its operational validity. If the protocol fails this evaluation, alternative protocols must be explored. If it succeeds, the process proceeds to calculate the implementation costs, such as resource usage and efficiency, of the selected authentication protocol, ensuring it meets both functional and practical criteria.

a. Estimating the execution time of the authentication protocol. Following successful functionality verification, the implementation cost estimation of the selected authentication protocol is conducted. During this phase, the database supplies critical metrics for security primitives, including execution time, area cost, energy consumption, and supplementary parameters. The total execution time of the protocol, quantified in clock cycles (`clk`), is computed by Equation (4.1):

$$\begin{aligned} T_{total}(\texttt{clk}) &= T_{total}^{ECC} + T_{total}^{hash} + T_{total}^{RNGs} \\ &= n_{ECC} \cdot t_{ECC} + n_{hash} \cdot t_{hash} + n_{RNGs} \cdot t_{RNGs} \end{aligned} \tag{4.1}$$

Equation (4.1), $T_{total}(\texttt{clk})$ denotes the cumulative latency, which is expressed in the number of clock cycles, aggregated from constituent cryptographic operations, such as ECC (T_{total}^{ECC}), hash function (T_{total}^{hash}), and RNGs (T_{total}^{RNGs}). For operations independent of a clock system, the latency of the computation is approximated using virtual clock cycles. The total latency for each subcomponent is derived from the product of the requisite instances of an operation ($n_{operator}$) and the latency of a single instance ($t_{operator}$), as formalized by using Equation (4.1).

The maximum frequency of the desired authentication protocol (F_{total}^{max}) is determined by the minimum value within the set $F_{total}^{max} = \texttt{min}\{F_0^{max}, F_1^{max}, \cdots, F_k^{max}\}$.

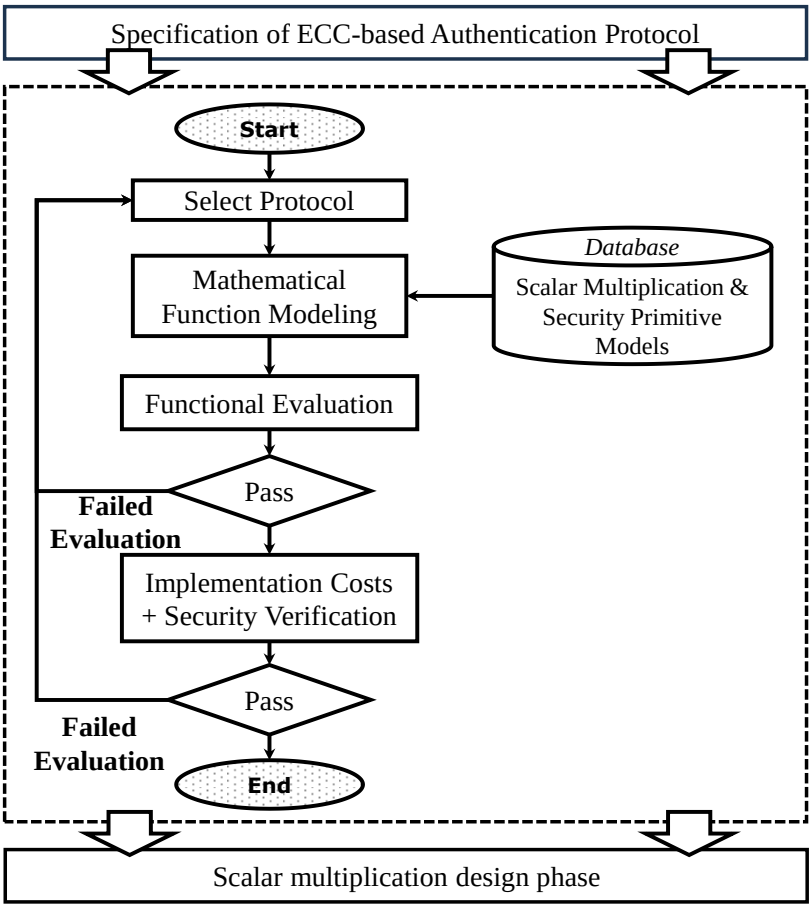


Figure 4.2: Authentication Protocol Design phase.

Equation (4.2) converts the total clock cycles into temporal units (seconds) :

$$T_{total}(s) = \frac{T_{total}(clk)}{F_{total}^{max}} \tag{4.2}$$

b. Estimating the Area Cost. The estimation of area cost constitutes a pivotal aspect in the design of the authentication protocol, which involves two entities: the passive RFID tag and the Receiver. Given that the passive RFID tag operates as a resource-constrained device, evaluating its physical area cost becomes imperative. This metric is quantified by Equation (4.3):

$$A_{RFID}(\texttt{kGates}) = A^{ECC} + A^{hash} + A^{RNGs} \tag{4.3}$$

where $A_{RFID}(\texttt{kGates})$ represents the total physical area cost of the passive RFID tag, expressed in `kGates`. Equation aggregates the individual area costs of cryptographic sub-blocks, specifically the ECC block (A^{ECC}), hash function block (A^{hash}), and RNGs block (A^{RNGs}).

Notably, this estimation model operates under the assumption that each cryptographic operation is executed by a dedicated sub-block. This premise arises from the absence of detailed architectural specifications for the RFID tag's implementation. Consequently, the model simplifies the design by assigning a single sub-block to each constituent cryptographic function—ECC, hashing, and random number generation—thereby enabling a baseline evaluation of area requirements. Such an approach facilitates preliminary resource allocation analysis while adhering to the constraints inherent in passive RFID systems.

c. Estimating the Energy Consumption. The estimation of energy consumption represents a critical consideration in designing passive RFID tags, which rely on energy harvested from the Receiver's emissions. Because of this dependency, the tag's operational energy expenditure is inherently constrained. Therefore, designers need to be precise in the evaluation of the system's energy consumption. The total energy consumed by the tag during the authentication protocol is quantified by computing Equation (4.4):

$$E_{RFID}(\mu J) = E^{ECC}_{equivalent} + E^{hash}_{equivalent} + E^{RNGs}_{equivalent} \tag{4.4}$$

where E_{RFID} denotes the cumulative energy consumption of the RFID tag, measured in microjoules (μJ). This formula consolidates the equivalent energy costs attributed to cryptographic sub-modules, such as the ECC block ($E^{ECC}_{equivalent}$), hash function block ($E^{hash}_{equivalent}$), and RNGs block ($E^{RNGs}_{equivalent}$).

The concept of equivalent energy consumption is employed, thereby neutralizing the influence of the specific technology node utilized in implementation to standardize the estimation across varying manufacturing technologies. This approach aligns with methodologies proposed by Still Maker *et al.* [63], whose conversion equations facilitate the derivation of equivalent energy values. Specifically, the equivalent energy consumption integrates the Energy Factor defined in [63] to compute these normalized metrics, ensuring a technology-agnostic assessment of energy requirements is derived as depicted by Equation (4.5):

$$E^{operator}_{equivalent} = \frac{E^{operator}}{\texttt{Energy Factor}} \tag{4.5}$$

The assessment of estimated implementation costs for the authentication protocol is a crucial step in evaluating its compliance with predefined device limitations and comparing it to other design options. This structured evaluation helps designers identify and discard authentication protocols, along with their corresponding hardware

security components, that do not meet the required performance or resource criteria. In addition, the filtering process provides valuable insights into possible improvements for the security components, directing designers toward focused adjustments in the following development stages. By aligning cost metrics with functional needs, this approach ensures that the authentication protocol associated with the cryptography primitives probability matches the operational constraints of resource-limited devices. The ongoing comparison of design alternatives supports well-informed decision-making, emphasizing solutions that balance both security effectiveness and practical feasibility.

4.2.1.2 Scalar Multiplication Design Phase

The completion of the first phase within the Early Evaluation Meet-in-the-Middle (EEMitM) design methodology enables designers to gain critical insights into the constraints governing security primitives, ensuring alignment with the operational limitations of resource-constrained devices. These constraints encompass quantitative metrics such as physical area cost, temporal latency, and equivalent energy consumption, which collectively define the implementation boundaries of cryptographic components like ECC. By leveraging these specifications, a systematic selection process is initiated to identify optimal configurations for scalar multiplication, including the choice of the finite field, algorithms, hardware architecture, projective coordinate systems, and other critical systematic configuration parameters, as illustrated in Figure 4.3.

The methodology's initialization phase involves deriving behavioral and abstraction models from the established database. Behavioral models formalize the mathematical operations underlying field operators, such as multiplication, inversion, and squaring, enabling functional verification of scalar multiplication processes. Concurrently, abstraction models encapsulate implementation-specific details, including hardware configurations, resource utilization metrics, power consumption profiles, and associated data traces. These models serve dual purposes: validating the security robustness of the proposed design and facilitating the precise estimation of implementation costs. By integrating these analytical tools, designers can iteratively refine configurations to balance computational efficiency with security requirements.

a. Physical area cost estimation of the ECC block. Following the successful functionality verification, the physical area cost estimation of the security primitives, particularly the ECC block, is conducted. By knowing the detailed hardware architecture, algorithm, and systematic configuration parameters of the security primitive, designers could know how many field operators are required to carry out the operation.

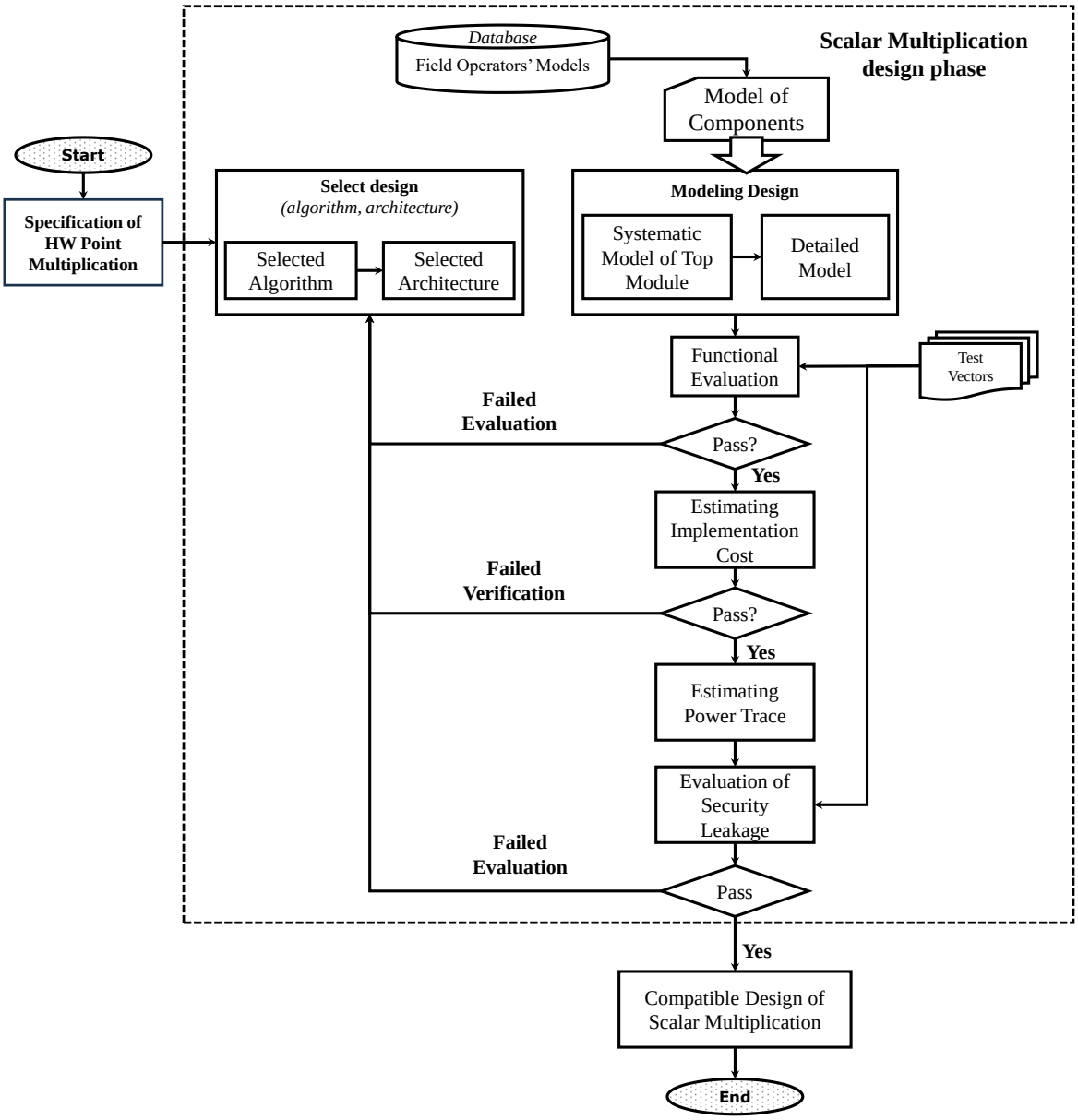


Figure 4.3: Proposed Scalar Multiplication Design Phase.

N_i denotes the number of required field operators with the index of $\{1; 2; 3\}$ being the field multiplier, field squarer, and field inverter, respectively. The architectural layout of the ECC system dictates the arrangement and interconnection of sub-blocks. Consequently, the total area of the ECC block can be computed using Equation (4.6), where n_i represents the number of sub-blocks dedicated to implementing a specific field operator i, and $A_{register}, A_i$ denote as the area occupied by the internal registers and each corresponding sub-block.

$$A_{total} = \sum_{i=1}^{3} n_i \cdot A_i + A_{register} \tag{4.6}$$

b. Power Consumption Estimating of the ECC Block. The estimation of power consumption within an ECC block constitutes a critical aspect of system design, particularly for optimizing energy efficiency in resource-constrained environments. This process relies on the architectural configuration of the ECC system, which determines whether cryptographic sub-blocks operate in parallel, sequentially, or in a hybrid mode. Each operational mode necessitates distinct computational models to quantify power consumption.

For parallel implementations, where field operators such as multipliers, squarers, and inverters function simultaneously, the total power consumption $P_{parallel}$ is calculated by using Equation (4.7):

$$P_{parallel} = \sum_{i=1}^{3} n_i \cdot P_i + P_{register} \tag{4.7}$$

where, n_i represents the number of sub-blocks implemented in the hardware that are assigned to a specific field operator i, and $P_{register}, P_i$ denotes the power consumed by internal registers and each sub-block carrying out its operation in parallel. This additive model accounts for the cumulative energy demand of all active sub-blocks operating concurrently.

In contrast, sequential operation — where sub-blocks execute tasks one after another — yields a power consumption profile dominated by the most power-intensive sub-block. Equation (4.8) formalizes this scenario:

$$P_{sequential} = \max\{P_i : i = (1, 2, 3)\} + P_{register} \tag{4.8}$$

Equation (4.8) dedicates the peak power requirement during sequential execution, as only one sub-block is active at any given time. The maximum value among $\{P_1, P_2, P_3\}$ defines the sequential system's power consumption.

In hybrid architectures, combining parallel and sequential modes requires a composite evaluation. Equation (4.9) aggregates the power contributions from both operational paradigms:

$$P_{total} = P_{paralle} + P_{sequential} \tag{4.9}$$

Equation (4.9) accommodates scenarios where certain sub-blocks operate simultaneously while others function in sequence, offering flexibility for designers to balance speed and energy efficiency.

c. Latency Estimating of the ECC block. Latency estimation for the ECC block represents a fundamental metric in evaluating the performance of hardware security primitives, particularly in applications requiring real-time cryptographic operations. This metric is derived from the architectural configuration of the ECC system, which dictates whether sub-blocks execute tasks in parallel, sequentially, or in hybrid architectures. Each sub-block i, responsible for cryptographic operations such as multiplication, squaring, or inversion, operates over N_{iter} iterations, with each iteration demanding N_i operations. The latency of an individual sub-block i is defined as $T_i(\texttt{clk})$, measured in clock cycles, reflecting the number of clock cycles required to complete its designated task.

For parallelized architectures, where sub-blocks operate concurrently, the total system latency $T_{parallel}$ is governed by the largest latency, adjusted for parallelism. The formula formalizing the equation to calculate $T_{parallel}$ is dedicated in Equation (4.10):

$$T_{parallel} = N_{iter} \cdot \max\{(\frac{N_i}{n_i} \cdot T_i) : i = (1, 2, 3)\} \tag{4.10}$$

Here, n_i denotes the number of parallel instances of sub-block i, effectively distributing the workload N_i across multiple units. The term $\frac{N_i}{n_i}$ represents the reduced operational burden per instance, and T_i denotes the intrinsic latency of the sub-block. The maximum value across all sub-blocks determines the critical path, as parallel execution cannot proceed faster than the slowest component.

In contrast, sequential architectures, where sub-blocks operate in series, accumulate latency across all components. Equation (4.11) quantifies the latency of the system in this scenario:

$$T_{sequential} = N_{iter} \cdot \sum_{i=1}^{3} \frac{N_i}{n_i} \cdot T_i \tag{4.11}$$

This generalized model assumes no overlap in sub-block operations, resulting in a linear aggregation of individual latencies. Sequential designs often prioritize resource efficiency over speed, as fewer parallel instances n_i are required, but at the cost of increased overall latency.

After estimating the implementation cost of the ECC block, the security evaluation of the design is performed. At this stage, the main vulnerability would be assessed as the hardware attack, especially the side-channel attack. The pseudo power traces of the ECC block are formalized based on the knowledge of the design and the information from the database for evaluating the security of the implemented circuit. The more detailed security evaluation of the ECC block is discussed in Section 4.3.

Based on the estimated metrics, including the physical area, power consumption, latency, and security level, the designers could consider several choices of architectures,

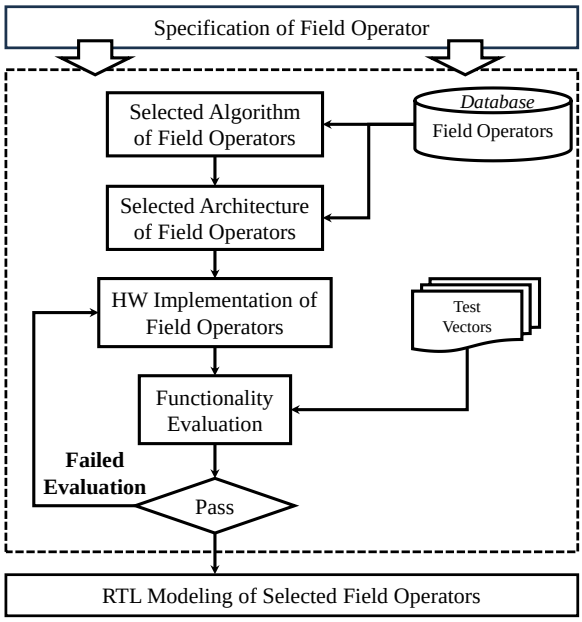


Figure 4.4: Field Operators Design phase.

algorithms, and systematic configurations. The consideration involves trade-offs between speed, area, and power consumption. High-performance choices that utilize the parallel configurations minimize latency by exploiting concurrency but demand greater hardware resources (e.g., additional multipliers or inverters). Sequential designs conserve area and power but may introduce bottlenecks in time-sensitive applications. Hybrid approaches, combining parallel and sequential elements, offer intermediate solutions tailored to specific performance constraints.

4.2.1.3 Field Operators Phase

In the last phase of the EEMitM design methodology, the specification of field operators such as field multiplier, squarer, and potentially the inversion is established as an outcome of the previous design phase. This phase, described in Figure 4.4, is critical in constructing a comprehensive architecture of field operators, which are subsequently interconnected to develop a fully functional system for ECC primitive design. The proper selection and integration of these field operators are fundamental to ensuring optimal computational efficiency and reliability in ECC implementations. At the end of this stage, designers transition from the conceptual framework to the implementation of these operators at the Register-Transfer Level (RTL), where their functional correctness and performance characteristics are rigorously evaluated in the Bottom-up Verification and Validation Process.

4.2.2 Bottom-up Verification and Validation Process

The bottom-up evaluation process is initiated with RTL models of field operators employed within the scalar multiplication component of ECC-based authentication protocols. The proposed Bottom-up Verification and Validation Process (B2VP) commences by rigorously validating the hardware implementations of field operators and foundational security primitives. Upon completion of the B2VP, the systematically implemented ECC-based authentication protocol undergoes comprehensive validation, encompassing both implementation cost efficiency and cryptographic security robustness.

At each iterative stage of the B2VP, quantitative implementation cost metrics (e.g., area, power, and timing) are rigorously compared against predefined estimated parameters to verify compliance with the hardware constraints established by the estimation framework. In instances where metric validation reveals nonconformity, designers may employ advanced optimization methodologies or re-evaluate and select alternative cryptographic primitives to ensure alignment with system constraints.

4.2.2.1 Field Operator Evaluation phase

The initial phase of the Bottom-up Verification and Validation Process (B2VP) focuses on the systematic evaluation and validation of cryptographic field operators, as delineated in Figure 4.5. This stage commences with the hardware implementation of RTL models describing the arithmetic and logical operations intrinsic to these operators. Functional verification is conducted via Electronic Design Automation (EDA) toolchains, employing combinatorial logic simulation and formal equivalence checking to ensure congruence with Galois field mathematical characteristics. Concurrently, implementation cost metrics of the hardware employed by field operators — encompassing physical area cost, critical path latency, and power dissipation — are quantitatively derived through technology-mapped synthesis, leveraging standard cell libraries specific to the target fabrication process.

The quantitative metrics derived from the hardware-implemented field operators are meticulously validated against the estimated computational costs obtained through the Early Evaluation Meet-in-the-Middle (EEMitM) design methodology. This validation process is crucial for ensuring the precision and efficiency of the implemented operators, as any significant discrepancies may indicate suboptimal performance or excessive resource consumption.

Failure to meet the expected cost estimations poses a critical risk to subsequent design phases, potentially leading to violations of physical design constraints such as

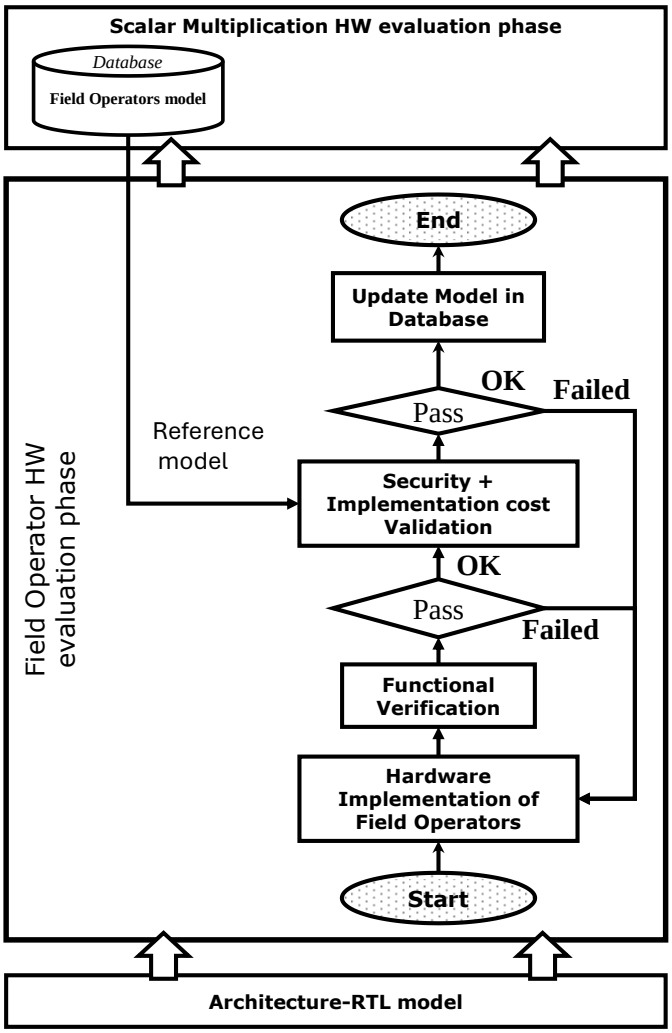


Figure 4.5: Field Operators Evaluation Phase.

area, power consumption, and timing requirements. Consequently, accurate verification is required to identify and address any inefficiencies before progressing to the next stage of development. This validation framework not only enhances the reliability of the hardware architecture but also facilitates informed design decisions for optimizing computational efficiency.

4.2.2.2 Scalar Multiplication Verification and Validation Phase

Following the evaluation and validation of field operators, these sub-components are systematically integrated into the ECC system to facilitate scalar multiplication. The primary input for this evaluation phase is the hardware implementation of the ECC system, ensuring that all design elements function as intended. At this stage, the design undergoes evaluation and validation processes to verify its correctness, efficiency,

and suitability. The evaluation process involves analyzing both functional accuracy and computational overhead, while the validation operations ensure that the hardware design adheres to predefined architectural constraints.

At the beginning of the Scalar Multiplication Verification and Validation Phase, as illustrated in Figure 4.6, power traces generated during scalar multiplication are subjected to post-silicon analysis using specialized Side-Channel Attack (SCA) assessment tools. This step is crucial for evaluating the system's resilience against power analysis attacks, which can reveal cryptographic keys through unintended power leakage. If the system fails this evaluation, indicating detectable leakage of sensitive information, designers must revisit the selection of scalar multiplication algorithms or strengthen countermeasures (e.g., adding blinding or masking techniques) to enhance resistance against SCAs.

Following the successful completion of the SCA evaluation, designers proceed with an in-depth assessment of the implementation costs associated with ECC primitives. This evaluation ensures that the realized ECC primitives maintain an optimal balance between computational efficiency and hardware resource constraints, as the estimation is performed in the EEMitM design methodology. If the implemented ECC primitive exceeds the estimated costs, designers iterate back to the Scalar Multiplication Design Phase in the proposed EEMitM design methodology to explore alternative architectures with the updated implementation cost of the field operators or optimize the configuration parameters. Conversely, if the ECC primitive meets cost benchmarks, the optimized design is cataloged in the security primitives database. This update enriches the repository for future projects, fostering scalability and reuse.

After evaluating and validating the field operators, these sub-blocks are assembled into the ECC system, carrying out the scalar multiplication. The input of this evaluation phase is the hardware implementation of the ECC system. At the beginning of the Scalar Multiplication Evaluation Phase, as demonstrated in Figure 4.6, the power traces of scalar multiplication are provided to the post-silicon evaluation tool for SCA assessment with the corresponding input data. If the evaluation fails, designers have to go back to re-select the algorithm of Scalar Multiplication or countermeasures for ECC.

Finally, the validated ECC primitive is embedded into the selected authentication protocol, such as Elliptic Curve Digital Signature Algorithm (ECDSA) or Elliptic Curve Diffie-Hellman (ECDH), enabling progression to the final implementation stage. Here, system-level integration is performed, verifying interoperability with peripheral components such as hash functions, random number generators, and communication interfaces. This end-to-end methodology balances security robustness with practical fea-

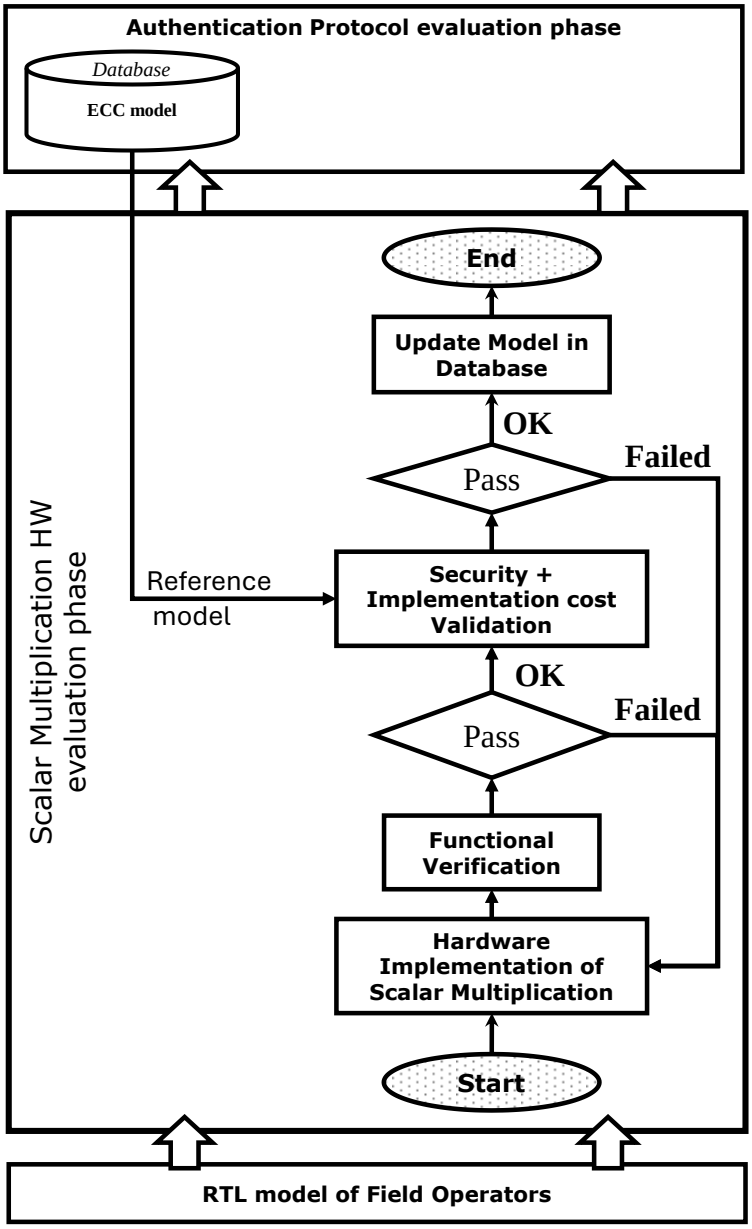


Figure 4.6: Scalar Multiplication Evaluation Phase.

sibility, ensuring that the resulting ECC-based authentication protocol adheres to both functional specifications and operational efficiency requirements. The cyclical nature of this process—iterating between evaluation, cost analysis, and redesign—underscores the importance of modularity and adaptability in cryptographic hardware development, particularly in applications demanding high assurance against physical and algorithmic threats.

4.3 The Proposed Pseudo Power Traces Generation Methodology

After successfully estimating the hardware implementation costs for the system, the next critical step is conducting a systematic security evaluation, mainly focusing on vulnerabilities to SCAs. SCAs exploit physical leakage information, such as power consumption, electromagnetic emissions, or timing variations, to extract sensitive cryptographic data. Traditional security evaluation frameworks rely on harvesting power traces from hardware designs or performing gate-level power analysis. While these methods provide valuable insights into leakage patterns, they are inherently time-consuming and too late to evaluate in the design process, as highlighted in [64]. This inefficiency significantly increases the time-to-product cost, posing a challenge for rapid development cycles.

A novel security evaluation methodology is introduced, utilizing the EEMitM design framework to handle these constraints. This approach facilitates the estimation of pseudo-power traces for cryptographic primitives without necessitating extensive physical measurements. By leveraging comprehensive insights into the algorithm, system architecture, and abstract representations of hardware components, a mathematical function is formulated to model the systematic power consumption behavior of the design. This function enables the generation of pseudo power traces that closely approximate the leakage characteristics observed in actual hardware implementations. Subsequently, these traces undergo analysis using established SCA evaluation techniques, such as TVLA and CPA, to detect potential security vulnerabilities.

This methodology significantly enhances the efficiency of security assessments by reducing the dependency on time-consuming physical testing while providing valuable insights into the resilience of cryptographic implementations against side-channel threats. By integrating this approach into the proposed EEMitM framework, the proposed methodology bridges the gap between theoretical security analysis and practical implementation, ensuring robust protection against physical attacks while maintaining efficiency in the design process.

Before performing the SCA evaluation using the proposed design methodology, a systematic power model is created based on knowledge of the implemented architecture and algorithm. There are two proposed approaches to forming the systematic power of the implemented circuits:

- Forming based on the sub-traces of the components that are utilized in the implemented circuits

- Resulting from the function of the switching activity of the implemented system.

4.3.1 Forming the Sub-traces of the Components

In the first approach to forming the systematic power trace, the reference traces, which are derived from primitive cryptographic operators, are utilized to estimate the power traces. This approach is designed to provide a detailed and accurate approximation of power consumption, enabling designers to optimize cryptographic implementations for energy efficiency, especially in resource-constrained environments such as IoT devices or hardware security modules. For example, in the context of the ECC system, the power trace of the implemented circuit is generated based on the reference power traces of field operators, such as field multipliers, field inversions, and field squares.

The first step in this process, which is depicted in Figure 4.7, involves the collection of power traces from the fundamental field operators that constitute the ECC system. These operators typically include field multipliers, field squarers, and field inverters, which are the building blocks of scalar multiplication. The power traces can be obtained through two primary methods: direct measurement from silicon devices or simulation using advanced power analysis tools. Direct measurement involves deploying the cryptographic operations on physical hardware and capturing the power consumption using specialized equipment, such as oscilloscopes or power analyzers. This method provides highly accurate traces but requires access to the actual hardware and a controlled testing environment. Alternatively, power simulation tools can be used to estimate the power traces in a virtual environment.

Once the power traces are collected, they are processed and organized into structured arrays for further analysis. Each power trace is represented as an array $P_i = \{p_{i0}, p_{i1}, \cdots, p_{ik}\}$, where p_{ij} denotes the power consumption of operator i at a specific time instance $t = j$. The variable k represents the length of the trace, which corresponds to the duration of the operation being analyzed. This structured representation allows for a systematic and mathematical approach to analyzing power consumption patterns. By breaking down the power traces into discrete time-based data points, we can more easily identify trends, anomalies, and correlations between different operators and their energy usage.

The next phase of the methodology involves using these arrays to estimate the power trace of the entire ECC system, specifically focusing on scalar multiplication. As illustrated in Figure 4.7, the power consumption of this operation is influenced by the sequence and combination of field operators used during its execution. By leveraging the reference power traces of individual operators, we can construct an approximate

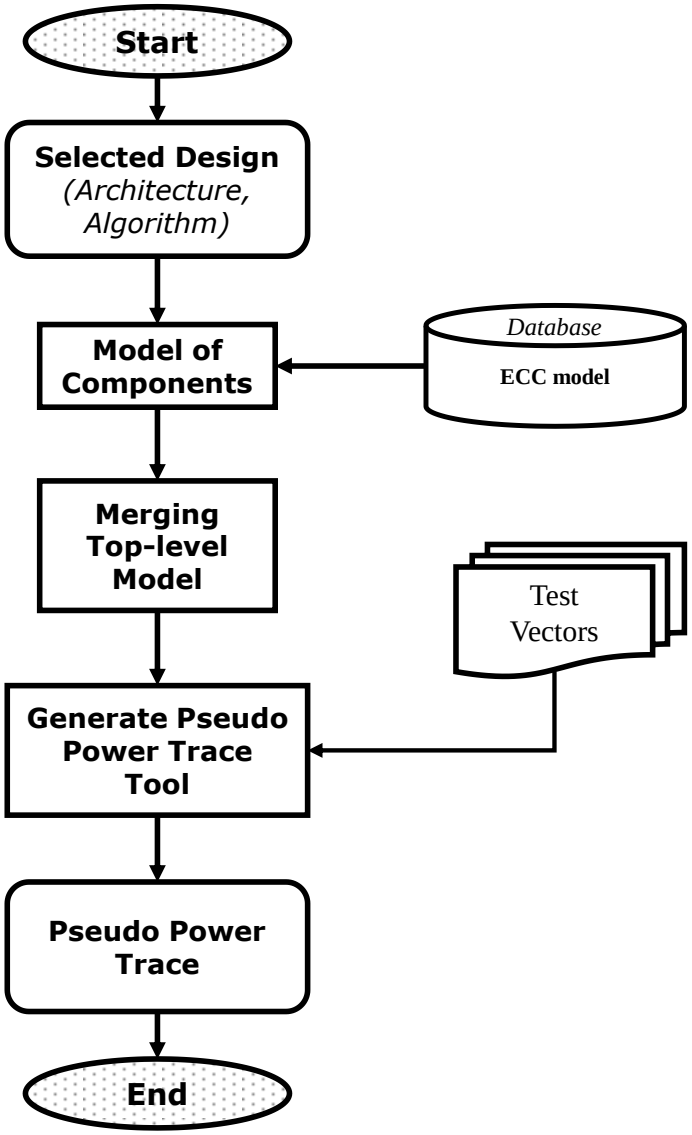


Figure 4.7: Process of estimating power trace of the cryptosystem.

power trace for scalar multiplication. This is achieved by combining the power traces of the constituent operators in a manner that reflects their sequence and frequency of use during the scalar multiplication process. For example, if a scalar multiplication operation involves multiple field multiplications followed by a field inversion, the power trace of the overall operation can be approximated by aggregating the traces of these individual operators in the appropriate sequence.

In particular, these operators work in parallel; their power traces are combined through accumulation. For instance, if two operators with power traces P_1 and P_2 operate simultaneously, the total power trace P_{total} is computed as the sum of their

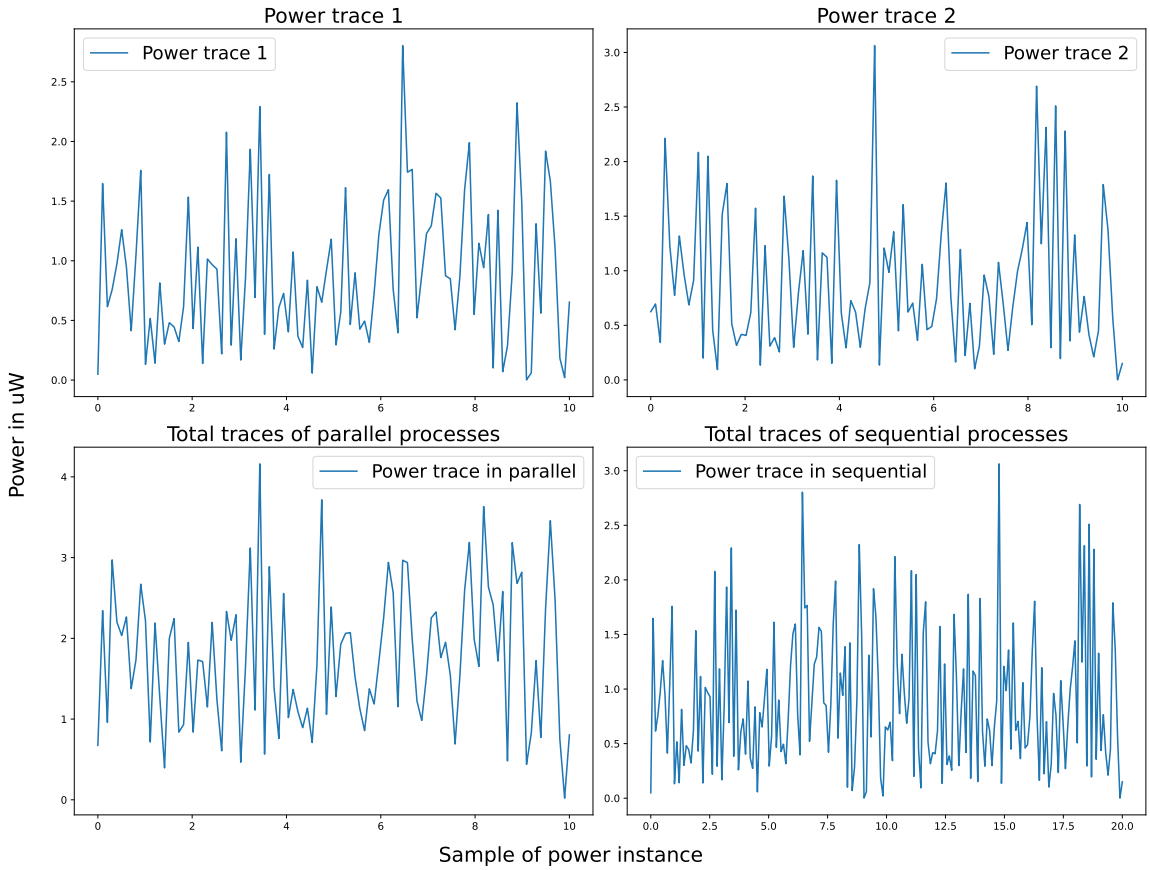


Figure 4.8: Examples of Forming the Trace from Sub-traces of Components.

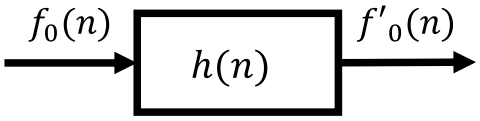


Figure 4.9: Modeling Mathematical Function of an Electrical Component.

individual power traces at each time instance. This is formalized as Equation ((4.12)):

$$\begin{aligned} P_{total} &= P_1 + P_2 \\ &= \sum_{i=0}^{k} (p_{1i} + p_{2i}) \end{aligned} \tag{4.12}$$

Here, p_{1i} and p_{2i} represent the power consumption of the first and second operators, respectively, at time instance i, and k represents the length of the power trace. The resulting P_{total} reflects the cumulative power consumption of the system during parallel execution.

In contrast, when operators work sequentially, their power traces are concatenated rather than accumulated. For example. if two operators with power traces P_1 and P_2 operate one by one, the total power trace P_{total} is formalized by appending the arrays as Equation (4.13):

$$\begin{aligned} P_{total} &= \{P_1|P_2\} \\ &= \{p_{10}, p_{11}, \cdots, p_{1k}, p_{20}, p_{21}, \cdots, p_{2k}\} \end{aligned} \tag{4.13}$$

In this case, the length of the total power trace is the sum of the lengths of the individual traces, reflecting the extended duration of sequential execution. By concatenating or accumulating the power traces of sub-modules 1 and 2, as illustrated in Figure 4.8, which is harvested from the physical deployment or hardware simulation by EDA tools, designers can accurately model the energy profile of systems that execute cryptographic operations in a step-by-step manner. This approach is particularly compatible with low-complexity designs, which typically generate shorter power traces. Furthermore, the simplicity of such systems enhances the designer's ability to analyze and interpret the system's behavior in each clock cycle with greater clarity.

4.3.2 Systematic Formalization of Switching Activity

In complicated systems, designers might not understand the architecture and functional behavior of the internal components well. Therefore, utilizing the first approach to generate the pseudo power trace is a challenge for designers to approximate the systematic power trace precisely. Therefore, the second proposed approach, which is based on the analysis of the systematic algorithm, is proposed to replace the first proposal.

In the second proposed approach to formalize the systematic switching activity, the designer models the impulse response of the digital circuit $h(n)$, which defines the mathematical dependence of the output on the input data. In contrast to the previous approach, the second proposed approach helps design a method to generate the pseudo power traces by analyzing the implementation algorithm. Figure 4.9 illustrates an example of a certain electrical component, which executes a function $h(n)$ by consuming the input data $f_0(n)$ and producing the output $f_0'(n)$. The mathematical expression of impulse response $h(n)$ is formalized as Equation (4.14):

$$f_0'(n) = h(n) * f_0(n) \tag{4.14}$$

As analyzed in Section 2.3.1, dynamic power dissipation in digital circuits is primarily caused by switching activities, which occur when logic gates transition between states, such as from 0 to 1 or vice versa. These transitions lead to transient current

flows, resulting in energy consumption. Understanding and modeling these switching activities is crucial for evaluating the security characteristics of a system. For instance, variations in power consumption can be exploited in side-channel attacks, where attackers infer sensitive information by analyzing power traces. The Hamming distance (HD) model is widely used for quantifying the number of switching activities of the circuit. This model measures the number of bit-level transitions in a digital component over a specific time interval, providing a reliable estimate of dynamic power dissipation.

The Hamming distance between two binary states, denoted as $f_0(n)$ and $f_0'(n)$, is formally defined as the number of bits that differ between them. Mathematically, this is calculated by first computing the bitwise XOR ($\oplus$) of $f_0(n)$ and $f_0'(n)$, which isolates the differing bits between the input and output data. In the following step, the Hamming weight (HW) is utilized for counting the number of bits that can be set to $'1'$ of the result, effectively quantifying the total transitions. In brief, according to the Mangard *et al.* [5] the Hamming distance between two values $f_0(n)$ and $f_0'(n)$ are encapsulated by the Equation (4.15) as below:

$$\mathtt{HD}(f_0(n), f_0'(n)) = \mathtt{HW}(f_0(n) \oplus f_0'(n)) \tag{4.15}$$

Substituting Equation (4.14) to Equation (4.15), the number of switching activities of the implemented system is expressed as Equation (4.16):

$$\begin{aligned} \mathtt{HD}(f_0(n), f_0'(n)) &= \mathtt{HW}(f_0(n) \oplus f_0'(n)) \\ &= \mathtt{HW}(f_0(n) \oplus \{h(n) * f_0(n)\}) \end{aligned} \tag{4.16}$$

Depending on the hierarchy of architecture, the total impulse response of the cryptography system is influenced. If these operators execute in parallel, their power traces are aggregated to form a combined representation through the accumulator. For instance, if two operators, in which the impulse responses are $h_0(n)$ and $h_1(n)$, respectively, operate simultaneously as depicted in Figure 4.10, the total systematical impulse response $h_{total}(n)$, which is formalized as the sum of the individual responses, as the expression (4.17):

$$h_{total}(n) = h_0(n) + h_1(n) \tag{4.17}$$

Equation reflects the cumulative effect of parallel execution, where the contributions of each operator are accumulated to produce the total system behavior. By substituting this result into Equation (4.14), the output data of the implemented system $f_{out}(n)$ is modeled as the expression:

$$f_{out}(n) = h_{total}(n) * f_{in}(n) \tag{4.18}$$

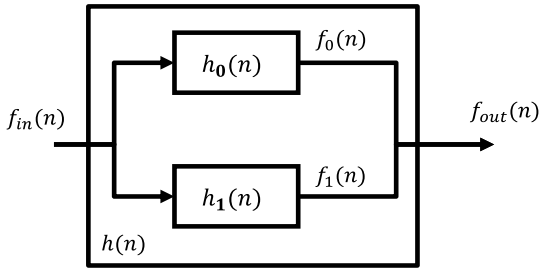


Figure 4.10: Modeling Mathematical Functions of Parallelism Electrical Components.

The output response of the parallelism electrical components is derived from the convolution of the input data $f_{in}(n)$ with the combined impulse responses of the operators dedicated as Equation (4.19) by expanding Equation (4.18).

$$\begin{aligned} f_{out}(n) &= (h_0(n) + h_1(n)) * f_{in}(n) \\ &= h_0(n) * f_{in}(n) + h_1(n) * f_{in}(n) \end{aligned} \tag{4.19}$$

In the following step, by utilizing the Hamming distance model, which is illustrated in Equations (4.15) - (4.16), designers can quantify the number of state changes occurring over a specific time interval, providing valuable insights into the system's power consumption profile. The process involves calculating the Hamming distance between two states, which represents the number of differing bits, and then using this metric to estimate the overall switching activity by using Equation (4.20):

$$\begin{aligned} \texttt{HD}(f_{in}(n), f_{out}(n)) &= \texttt{HW}(f_{in}(n) \oplus \{h(n) * f_{in}(n)\}) \\ &= \texttt{HW}(f_{in}(n) \oplus \{h_0(n) * f_{in}(n) + h_1(n) * f_{in}(n)\}) \\ &= \texttt{HW}(f_{in}(n) \oplus \{h_0(n) * f_{in}(n)\}) + \texttt{HW}(f_{in}(n) \oplus \{h_1(n) * f_{in}(n)\}) \\ &= \texttt{HD}(f_{in}(n), f_0(n)) + \texttt{HD}(f_{in}(n), f_1(n)) \end{aligned} \tag{4.20}$$

Equation (4.20) shows the methodology for combining the systematic dynamic power traces based on analyzing the algorithm of the sub-modules, which operate in parallel. By leveraging this methodology, designers could estimate the dynamic power traces of the implemented parallelism digital circuit by analyzing the algorithm corresponding to the architecture and accumulating the Hamming distance model of the sub-components.

In the case of sequential operation, the total systematic power traces are concatenated by ordering the Hamming Distance model of each sub-block. For example, two operations, in which the impulse responses are $h_0(n)$ and $h_1(n)$, respectively, execute

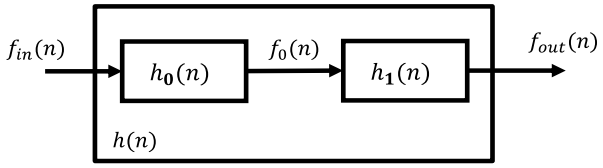


Figure 4.11: Modeling Mathematical Functions of Sequential Electrical Components.

block-by-block as depicted in Figure 4.11. The total systematical impulse response $h_{total}(n)$ is formalized as the Equation (4.21):

$$h_{total}(n) = \begin{cases} h_0(n) & 0 \leq n \leq n_1 \\ h_1(n) & n_1 < n \leq N \end{cases} \tag{4.21}$$

This formulation reflects a time-partitioned architecture where:

- Block 0 ($h_0(n)$) is exclusively active during the initial interval $[0, n_1]$.
- Block 1 ($h_1(n)$) operates subsequently in the interval $(n_1, N]$.

The sequential processing of the implemented circuit depicted in Figure 4.11 proceeds as follows:

1. The input signal $f_{\text{in}}(n)$ is first processed by Block 0, generating an intermediate output $f_0(n)$ via convolution with $h_0(n)$:

$$f_0(n) = h_0(n) * f_{\text{in}}(n).$$

2. The intermediate output $f_0(n)$ serves as the input for Block 1. The final output $f_{\text{out}}(n)$ is derived by convolving $f_0(n)$ with $h_1(n)$:

$$f_{\text{out}}(n) = h_1(n) * f_0(n).$$

Thus, the final output $f_{\text{out}}(n)$ is entirely governed by the impulse response of Block 1, following the sequential execution of both blocks. This hierarchical design ensures modular processing, with each block contributing to the system's behavior within its designated temporal domain. In other words, the power dissipation of the sequential digital circuit is caused by either Block 0 or Block 1 at a particular time slot. As a consequence, the Hamming distance model of the digital circuit is expressed as Equation (4.22):

$$\texttt{HD}(f_{in}(n), f_{out}(n)) = \begin{cases} \texttt{HD}(f_{in}(n), f_0(n)) & 0 \leq n \leq n_1 \\ \texttt{HD}(f_0(n), f_{out}(n)) & n_1 < n \leq N \end{cases} \tag{4.22}$$

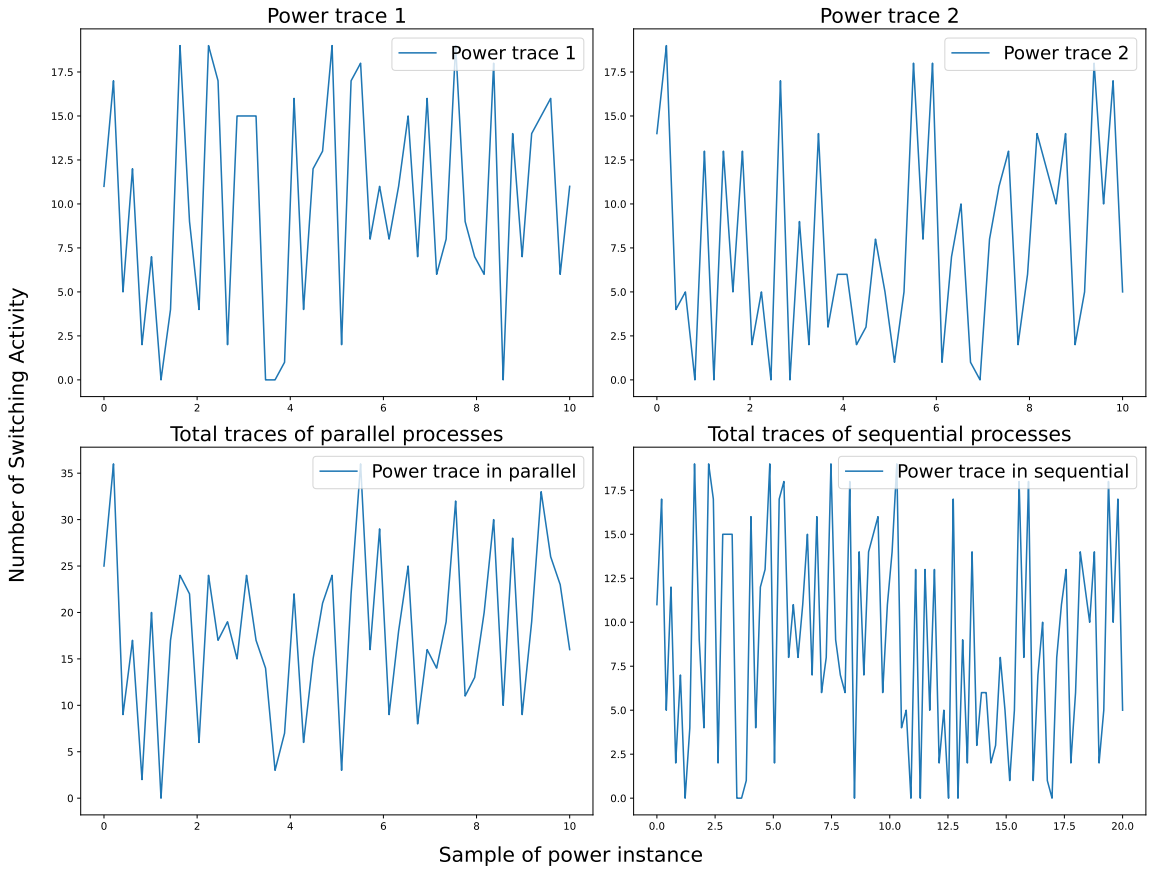


Figure 4.12: Examples of Forming the Trace from Analyzing the Architecture Corresponding to the Algorithm of Components.

In an alternative representation, the total Hamming distance of the sequential digital circuit is the Concatenation of the sub-Hamming distance functions of the sub-modules, as Equation (4.23).

$$\texttt{HD}(f_{in}(n), f_{out}(n)) = \{\texttt{HD}(f_{in}(n), f_0(n))|\texttt{HD}(f_0(n), f_{out}(n))\} \tag{4.23}$$

The total Hamming distance in digital circuits is modeled via Concatenation or accumulation, depending on sequential or parallel architectures:

- **Concatenation** applies to sequential circuits, where sub-module Hamming distances are combined chronologically to reflect stepwise execution, which is represented in Equation (4.23).
- **Accumulation** aggregates parallel sub-module Hamming distances, capturing simultaneous switching activities, which is presented in Equation (4.19)

These models enable the estimation of switching activity traces, which are critical for approximating dynamic power dissipation caused by logic gate transitions. An

example of the system's architectural dependence is shown in Figure 4.12. Pseudo power traces of modules 1 and 2 are approximated by estimating the behavior of the Hamming distance functions of blocks 1 and 2. Based on the system architecture of either sequential or parallel, the pseudo power traces of the top design are approximated by either concatenating or accumulating, respectively, as illustrated in Figure 4.12. By simulating power profiles early based on an understanding of the architecture and the corresponding algorithm, designers identify inefficiencies and vulnerabilities, such as side-channel attack (SCA) risks from predictable power patterns.

4.4 Evaluation Results

After demonstrating the proposed design methodology in Sections 4.2 and 4.3, this section carries out the experiments for evaluating and measuring the efficiency of the proposed methodology. The first executed experiment is the evaluation of the protocol stage for choosing the optimal authentication protocol with desired Elliptic Curve Cryptography modules stored in the database. In the following, according to the information derived from the protocol stage, the security primitive design begins with the determined constraints of the hardware implementation. The results of utilizing the proposed design methodology are compared to the conventional design flows to indicate the efficiency of the proposal.

4.4.1 Experimental Setups

The experiment was designed to evaluate the efficiency of a novel design methodology for Elliptic Curve Cryptography (ECC) based RFID tags by comparing it with a conventional design approach. The comparisons focused on analyzing implementation costs (e.g., area, power consumption, timing), performance metrics, security levels, and time-to-product efficiency of the proposed design methodology and the conventional design approach.

The experimental setup involved applying both the proposed and conventional methodologies to design ECC-based authentication protocols under identical constraints. These comparisons are executed under the same conditions of the machine. The server machine utilizes a single core of Intel(R) Xeon(R) CPU E5-2640 v4 2.40GHz with the availability of 128 GB RAM. In addition, the constraints of the design include:

- **Implementation costs:**
  - Maximum duration of one tag-server authentication: $T_{auth} = 20\ ms$.

– Maximum available power: $P_{max} = 240\ \mu W$.

- **Security properties:**

  – Robust against the popular Wireless Attacks.

The maximum timing of communication T_{auth} refers to the standard ISO/IEC-14443; meanwhile, the maximum power consumption is the peak harvesting at $-3dBm$ incident power by rectifier antenna, as proposed in prior work of Xu *et al.* [2].

The final designs, which are the results of applying two design flows, are compared in terms of similarity, performance, implementation cost, and security level for the passive RFID tags. This demonstrated the feasibility of the proposed methodology in optimizing resource utilization while maintaining compliance with critical standards and security benchmarks.

4.4.2 Results of Experiments for Protocol Stages: Design and Evaluation Stage

4.4.2.1 Constructing the database of the Security Primitives

As regards the ECC-based authentication protocols, the database includes the hardware models of ECC on FPGA or ASIC, according to work by Lara-Nino *et al.* [65]. Table 4.1 lists recent hardware designs of ECC on both FPGA and ASIC.

Table 4.1: Database of the low-cost, low-energy hardware implementations of ECC primitives.

Works	**Size field (bits)**	**Technology**	**Frequency (MHz)**	**No. Clocks**	**Power (μW)**	T_{norm} **(ms)**	P_{norm} **(μW)**
Salman [66]	192	FPGA Artix 7	187	824,212	-	69.97	-
Azarderakhsh [67]	163	CMOS 65nm	13.56	106,700	77.2	10.67	56.93
Wenger [68]	163	CMOS 130nm	1	54,376	181.7	5.44	1,817

In Table 4.1, synthesizes key performance parameters—operational frequencies ($F_{operate}$), clock cycles (N_{clk}), latency (T_{delay}) per independent execution, and power consumption (P)—derived from prior investigations [66–68]. Besides, the table further introduces two normalized metrics for the database:

- **Normalized latency (T_{norm})**: Standardized delay for executing scalar multiplication over $GF(2^{163})$, measured in equivalent clock cycles.
- **Normalized power consumption (P_{norm})**: Standardized power expenditure for the same operation, quantified in microWatts (mW).

These metrics are calibrated for a fixed operational frequency of 10 MHz, enabling cross-platform comparison of cryptographic efficiency. The normalization process eliminates frequency-dependent variability, ensuring equitable evaluation of latency and power consumption across heterogeneous hardware implementations. In particular, the normalization of the latency for an independence point multiplication of the design is demonstrated as Equation (4.24) below:

$$\begin{aligned} T_{norm} &= \frac{T_{delay} \cdot F_{operate}}{m} \cdot 10 \cdot 163 \\ &= \frac{N_{clk}}{m \cdot 10^{-5}} \cdot 163 \end{aligned} \tag{4.24}$$

where m is the size of the implemented finite field $GF(2^m)$. In addition, a normalization of the power consumption is also expressed in Equation (4.25):

$$P_{norm} = \frac{P}{m \cdot F_{operate}} \cdot 10 \cdot 163 \tag{4.25}$$

4.4.2.2 Evaluating the Authentication Protocols

The assessment of lightweight authentication protocols for passive RFID tags involves a dual focus on security robustness and practical implementation costs, particularly in resource-constrained systems. Security evaluations prioritize resilience against prevalent wireless attacks, such as eavesdropping, spoofing, and man-in-the-middle intrusions. Simultaneously, implementation costs are analyzed through metrics like power efficiency, computational latency, and hardware complexity of the ECC-based systems. In the following, the evaluation of the compatibility of the analyzed metrics with the passive RFID tags' constraints is presented.

a. Security Evaluations for the Authentication Protocols Recent studies, including a comprehensive survey by Gabsi *et al.* [69], highlight the vulnerability analysis of the recent lightweight authentication protocol based on ECC [70–73] for the constrained RFID tags. In addition, a novel low-cost authentication proposed by Gabsi *et al.* [74] is also considered. Security validation extends to benchmarking against established standards, ensuring alignment with requirements like mutual authentication,

Table 4.2: Security Analysis of the Lightweight Authentication Protocols for Passive RFID tags.

Vulnerabilities	Liao [70]	Zhao [71]	Alamr [72]	Zheng [73]	Gabsi [74]
MITMA	✓	✓	✓	✓	✓
Replay	✓	✓	✓	✓	✓
Anonymity	×	✓	✓	✓	×
Forward Security	×	✓	✓	✓	×
Impersonation	×	✓	✓	✓	×
Key Compromise	×	✓	✓	✓	✓
Location tracking	✓	✓	✓	✓	×
DoS	✓	✓	×	✓	✓
Cloning	✓	✓	✓	✓	✓
Server spoofing	✓	✓	✓	✓	✓
Desynchnization	✓	✓	×	✓	✓

✓ denotes that the protocol is robust against such vulnerability.
× denotes that the protocol is vulnerable to such an attack.

forward secrecy, and resistance to replay attacks. Additionally, attributes such as data confidentiality and user privacy are rigorously examined, as illustrated in Table 4.2.

With the authentication protocol proposed by Gabsi *et al.* [74], Asrlan *et al.* [75] analyzed and claimed that this protocol is robust against anonymity, position tracking, impersonation, and forward security.

b. Implementation Cost Evaluation with the Predefined Database Based on the survey by Gabsi *et al.* [69], the analyzed authentication protocols rely on cryptographic primitives such as scalar multiplication, hash functions, random number generators, and XOR operations. Scalar multiplication is emphasized as the most computationally significant component in these protocols. As a result, this evaluation focuses specifically on scalar multiplication's impact, while the contributions of other operations are excluded from the analysis. Estimating the cost of the considered authentication protocols with the entire database of lightweight ECC primitives is executed in the following, with the assumption that the server side has unlimited processing ability. Therefore, we assume that all the computations that are performed on the server side

are processed immediately.

By utilizing Equation 4.1 with the assumption that $t_{RNGs} \ll t_{hash} \ll t_{ECC}$, the total execution time of the protocol is estimated by Equation (4.26):

$$T_{total} = n_{ECC} \cdot t_{ECC} \tag{4.26}$$

The execution times for a dependent scalar multiplication t_{ECC} are listed in the database, represented in Table 4.1. The results of the total execution time of the protocols are demonstrated in Figure 4.13. The red dashed line in Figure 4.13 presents the limitation in the minimum total time of the authentication protocol at 20 ms.

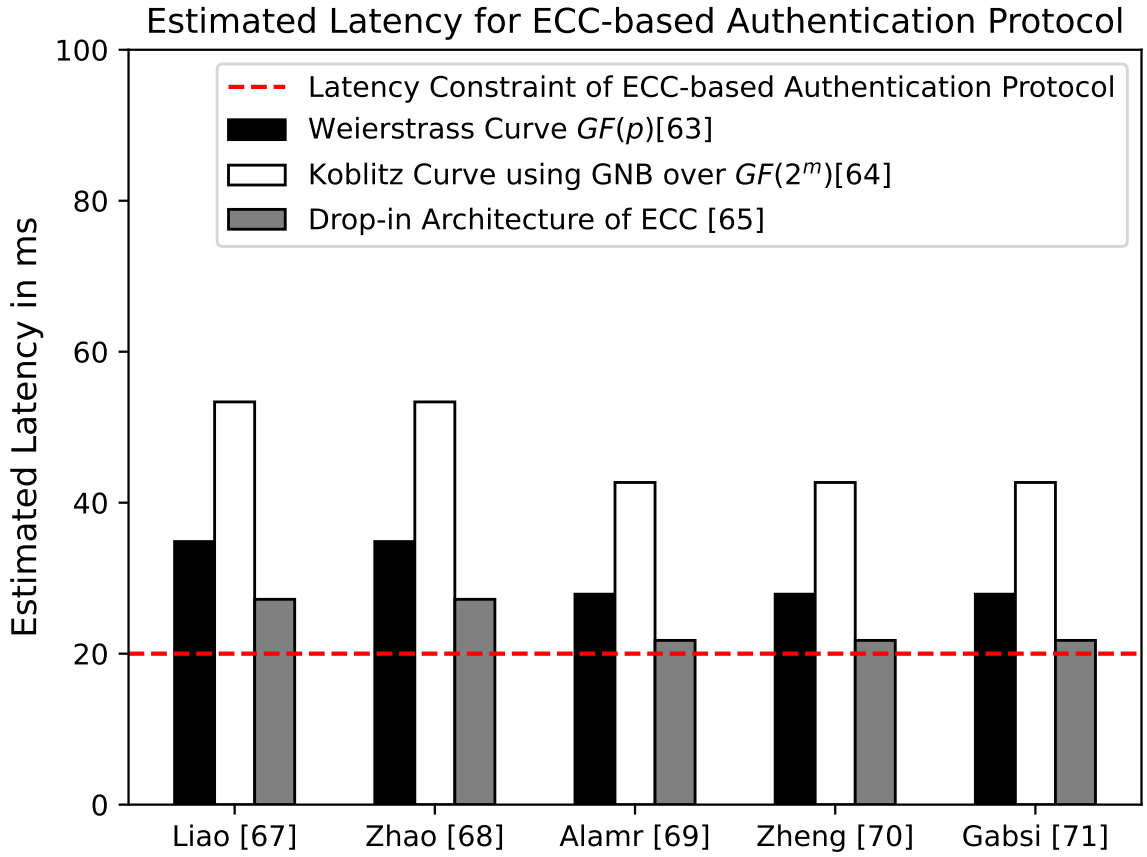


Figure 4.13: Latency estimations of the analyzed authentication protocols

As represented in Figure 4.13, all the existing lightweight authentication protocols with the existing ECC primitives module violate the time constraints of the passive RFID tags application. The most attractive authentication protocols are Alamr [72], Zheng [73], and Gabsi [74] with the ECC module provided by Wenger [68]. The estimated total latency of them is approximately 21.76 ms.

Regarding the security evaluation of these authentication protocols, which is dedicated in Table 4.2, the most secure protocols are proposed by Zhao [71] and Zheng

[73]. These protocols are secured against all the listed wireless vulnerabilities. Besides, Alamr's protocol [72] also deserves to be considered by providing almost all the characteristics except the DoS and Desynchronization attacks.

In general, when evaluating both security and processing efficiency, the procedure proposed by Zheng [73], which incorporates Wenger's ECC primitives [68], emerges as the optimal candidate for implementing the passive RFID tags. However, the ECC primitives from Wenger need to be improved. In particular, a key limitation arises in Zheng's authentication scheme, where the RFID tag performs four sequential scalar multiplications. Therefore, the maximum latency for an independent scalar multiplication of ECC imposes 5 ms, as expected by Equation (4.27). This improvement of the ECC will be executed in the following phase. Besides, the power consumption of the ECC module provided by Wenger is also significantly higher than the power budget. Therefore, the power consumption of the ECC is optimized by utilizing the subsequent phases.

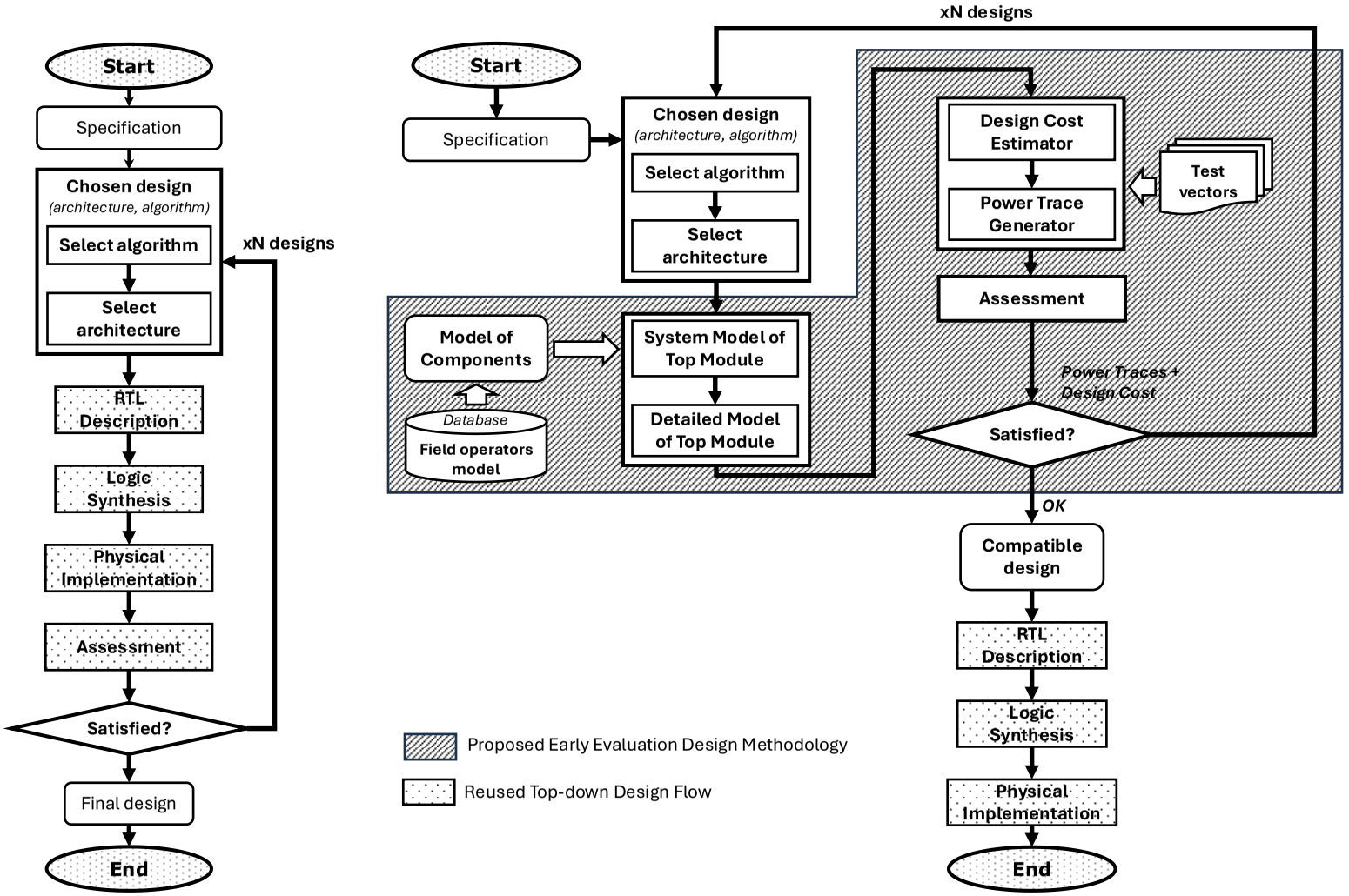
$$\begin{aligned} T_{expected} &= \frac{T_{limit}}{n_{ECC}} \\ &= \frac{20}{4} = 5(ms) \end{aligned} \tag{4.27}$$

4.4.3 Results of Experiment for Security Primitives: Design and Evaluations Stage

As a result of the authentication protocol phase, the ECC scalar multiplication needs to be refined. The expected constraints of the scalar multiplication are listed below:

- **Implementation costs:**
  - Maximum duration of one point multiplication: $5ms$.
  - Maximum power consumption: $240\mu W$ [2].
  - Area footprint: as low as possible.
- **Security properties:**
  - Robust against SCA.

These constraints are fed into both the conventional top-down design flow and the second phase of the proposed design methodology, as demonstrated in Figures 4.14(a) and 4.14(b), respectively. At the end of each phase, the metrics regarding the implementation cost and security of designs are compared to measure the accuracy of the estimation. Besides, the time-to-products is measured to prove the efficiency of utilizing the proposed design methodology over the conventional top-down design flow.



(a) Conventional Top-down Design Flows for Hardware Implementation.

(b) Proposed Early Evaluation Design Methodology for Hardware Implementation.

Figure 4.14: Description level of design in Top-down Design methodology.

4.4.3.1 Constructing the database for ECC Scalar Multiplication

Before starting to construct the database of field operators for the ECC, Wenger *et al.* [16] recommended choosing the Binary Field $GF(2^m)$ for hardware implementation instead of Prime Field $GF(p)$. Comparing the hardware costs of the primitives, the mathematical operators of the finite field implemented over the $GF(2^m)$ consumed fewer resources than over the $GF(p)$. The field operators include field multiplier, field squarer, and field inverter. In this experiment, the length of the key is 163 bits, which is referred to as standard FIPS-186-4.

The reference field operators are reused from work by Deschamps *et al.* [17] to construct the database of the field operators. The synthesized results are executed within the CMOS 45nm technology library from NANGATE at the operational frequency of 20 MHz. All the metrics of hardware implementation of these field operators are listed in Table 4.3.

The Classical Square module is purely a combinational logic block [17]. Therefore, it executes immediately ($2.3{\cdot}10^{-8}ms$) and requires 497 equivalent gates. Other modules are implemented as sequential blocks that use storage for computing. Consequently, they require more footprint area and more execution time. Field Multiplier and Field Square are executed by two algorithms:

- **School-book**: Classical Square and Interleaved Multiplier.
- **Montgomery**: Montgomery Square and Montgomery Multiplier.

Table 4.3: Database of hardware implementation of field operators over $GF(2^{163})$.

Works		**Latency** (μs)	**Power Consumption** (μW)	**Area (GEs)**
Field Multiplier	Interleaved Mult.	8.25	170	4,650
	Montgomery Mult.	8.15	3,561	6,700
Field Inversion		16.4	612	8,600
Field Square	Classical Sqr.	$2.3{\cdot}10^{-5}$	35.8	497
	Montgomery Sqr.	8.15	612	3,291.33
D Flip-Flop		-	0.344	8.09

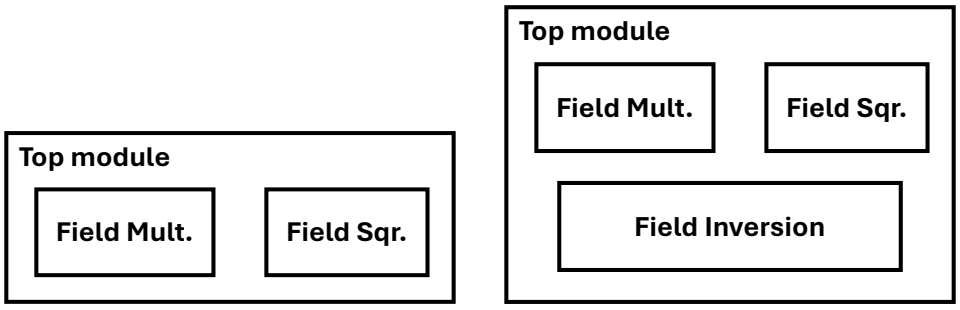
* GEs denotes the number of equivalent gates.

4.4.3.2 Evaluating the ECC's hardware architectures

The evaluation of the ECC's hardware architectures for passive RFID tags considers not only the implementation cost but also the security against hardware attacks, which is demonstrated in Section 2.3.2.2.

a. Evaluating the Elliptic Curves The comparative analysis of lightweight Elliptic Curve Cryptography (ECC) primitives, as outlined in Table 1.4, focuses on evaluating their efficiency under standardized hardware configurations. These primitives, referred to from cryptographic literature, employ fundamental field operators such as field inverters, multipliers, and squares. Two distinct hardware architectures are utilized in Figures 4.15 to ensure an equitable comparison.

The first assumed hardware model, which is presented in Figure 4.15(a), excludes a dedicated field inversion block, while the second architecture, which is demonstrated in Figure 4.15(b), incorporates one. The Binary Edwards Curve in Affine Coordinates (BEC) proposed by Kim *et al.* [14] represents a notable exception, as its implementation inherently requires a field inversion operation. This necessitates the use of the inversion-inclusive hardware model, Figure 4.15(b), for its evaluation, whereas other curves are assessed using the simpler architecture, Figure 4.15(a).



(a) Without field inversion. (b) With field inversion..

Figure 4.15: Assumed hardware models for evaluating the elliptic curves.

The efficiency of each ECC primitive is quantified using metrics involving latency in milliseconds, power consumption in microWatt, and area utilization measured in the number of equivalent gates. These metrics are derived from simulations using the hardware models in Figure 4.15 to ensure consistency across evaluations.

As these assumed hardware models process in sequence, applying Equations (4.8) and (4.11) for estimating the power consumption and the total latency of the point multiplication. The metrics of field operators are mapped by Table 4.3 for measuring

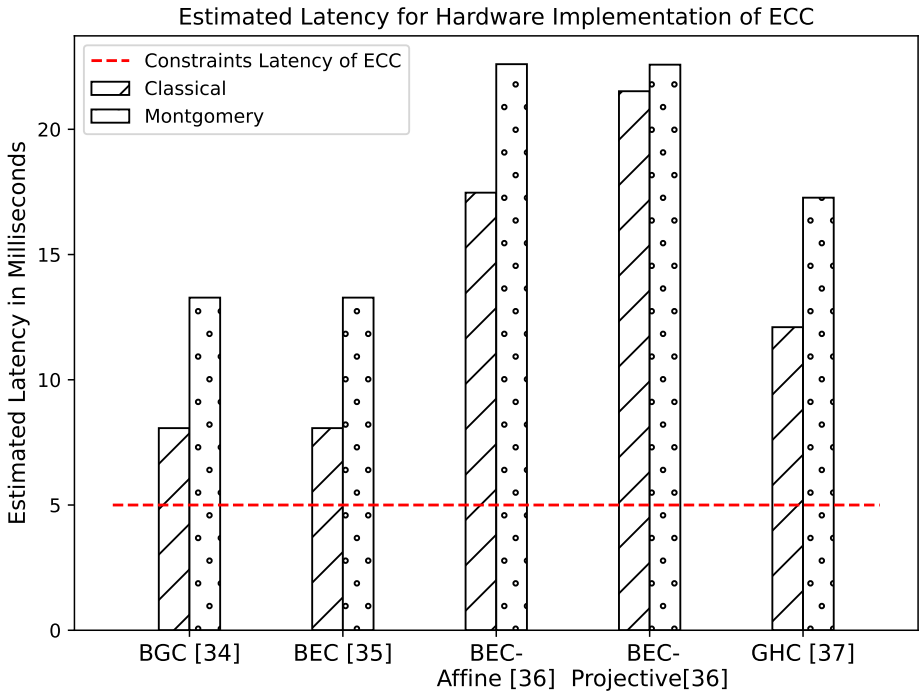


Figure 4.16: Estimated Latency of considered ECC designs.

the implementation costs, such as power consumption and delay, of the considered elliptic curves.

The measured latencies for an independent scalar multiplication of the considered ECC designs are demonstrated in Figure 4.16. As depicted in Figure 4.16, existing hardware architectures containing single-field multipliers and squares and employing each considering elliptic curves fail to meet the latency requirements for passive RFID tag applications. Specifically, elliptic curve cryptography (ECC) implementations using Binary Edwards Curves (BEC) in affine coordinates exhibit a worst-case latency of $53ms$, surpassing the permissible operational budget for passive RFID systems by 10 times the magnitude. Meanwhile, the Binary Weierstrass Curve (BWC) [13] and BEC [11], which adopt classical field operators, demonstrate latencies of $16ms$—fourfold higher than the predefined constraints. These results underscore the incompatibility of conventional hardware configurations with the stringent timing demands of passive RFID technologies, necessitating architectural optimizations to align with resource limitations.

Power consumption serves as an additional critical metric for evaluating the efficiency of Elliptic Curve Cryptography (ECC) implementations. This metric is estimated by Equation (4.8). Under the assumption of sequential processing, where operations are executed in a non-overlapping manner, the total power consumption of the system's top module is governed by the maximum power consumed by its constituent

sub-blocks. The estimated power consumption values for the analyzed ECC designs are graphically summarized in Figure 4.17. This graphical demonstration illustrates the relative energy demands across different implementations. These results highlight the influence of sub-block power profiles on overall system efficiency, particularly in resource-constrained environments where energy optimization is paramount.

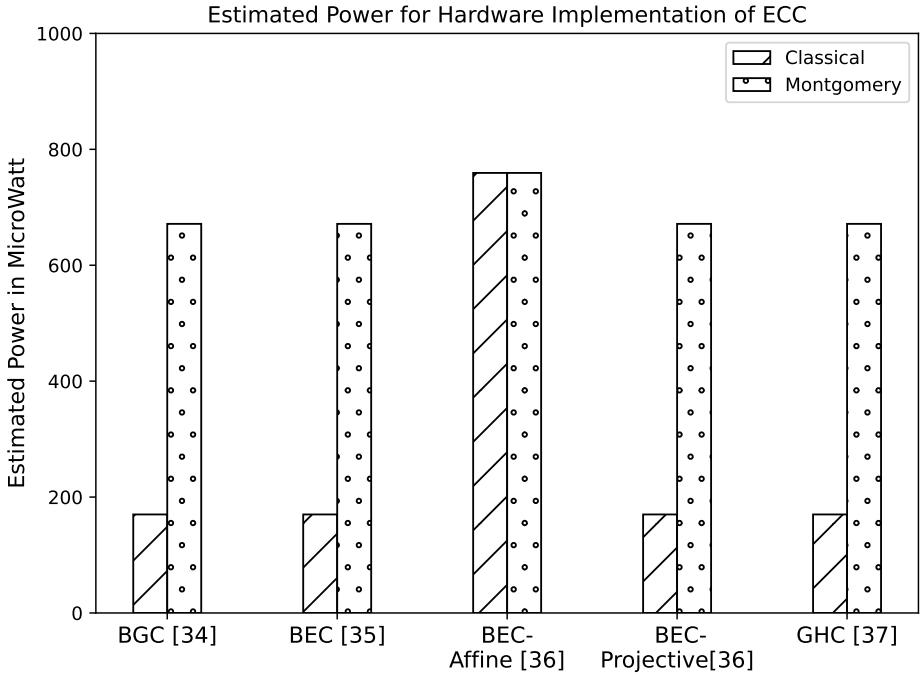


Figure 4.17: Estimated Power Consumption of considered ECC designs.

Figure 4.17 shows the Binary Edwards Curve [14] using the Affine Coordinate consumes significant power, which is 20 microWatt larger (25%) than other curves. This is caused by the field inversion, which is the most complicated operator over the finite field. Therefore, hardware implementation of field inversion is also more costly compared to the other field operators, which is also reflected in Table 4.3. Besides, the Montgomery algorithms in field multiplier and field squarer provide notable benefits in minimizing power consumption. By employing the components executing the Montgomery algorithms, the power consumption is saved by 10 microWatt (13%).

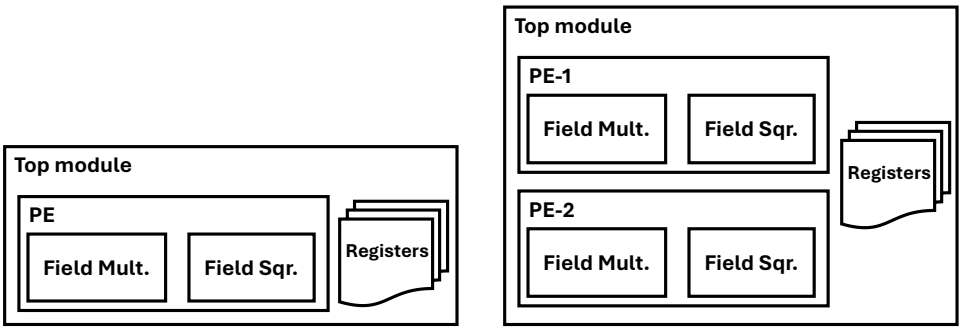
The general analysis further highlights the advantages of the Binary Edwards Curve proposed by Koziel *et al.* [11], which demonstrates superior implementation cost efficiency, demonstrated in Figures 4.16 and 4.17, and enhanced security properties. Specifically, its Mixed Coordinates implementation ensures cryptographic completeness, a critical feature for mitigating hardware-based side-channel attacks. However, despite these benefits, the curve's latency remains incompatible with the timing constraints of passive RFID systems. Consequently, architectural refinements are imper-

ative to align the BEC's performance with resource-constrained applications' requirements while preserving its robustness in security.

b. Evaluating the hardware architectures In the previous evaluation, the alternative ECC algorithms were assessed in terms of both efficiency and security. The evaluation results highlight the advantages of the algorithm implementing Binary Edwards Curves proposed by Koziel *et al.* [11]. This algorithm requires $6M + 4S$, where M and S denote the field multiplier and square, respectively.

To implement Koziel's algorithm on hardware, there are two architectures under consideration as listed below:

- **Single Processing Element (PE)**: Contains one field multiplier and one field square, which is illustrated in Figure 4.18(a), that requires nine internal registers of $m-$bits.
- **Dual Processing Elements**: Contains two couples of field multiplier and field square demonstrated in Figure 4.18(b), which requires 12 internal registers of $m-$bits.



(a) Single PE Architecture. (b) Dual PEs Architecture.

Figure 4.18: Two possible architectures under consideration.

By using the estimation equations, which are discussed in Section 4.2.1.2, in conjunction with the reference field operators listed in Table 4.3, the estimated hardware implementation costs of these architectures are given in Table 4.4. The benchmark metric $APT2$ is introduced, as defined in Equation (4.28) to facilitate a comprehensive evaluation of these implementations:

$$APT2 = A \cdot P \cdot T^2 \tag{4.28}$$

where A, P, T represents the area physical cost ($kGates$), power consumption (μW), and latency (μs), respectively.

Table 4.4: Estimation Hardware Implementation Cost by using proposed EEMitM Design Methodology.

Architectures	**Single PE**	**Dual PEs**
Area (kGates)	17.02	26.12
Power Consumption (mW)	710.45	1084.46
Latency (μs)	9.41	5.38
$APT2$	1,071,251	819,596

The benchmark is specifically designed to reflect the cost-effectiveness of each implementation within the framework of passive RFID tag applications. Given the stringent real-time constraints characteristic of such applications, latency emerges as the most critical design parameter. Accordingly, in the formulation of the $APT2$ benchmark, latency is assigned a second-order weighting, thereby emphasizing its significance relative to area and power in the overall evaluation of hardware efficiency.

Table 4.4 presents a comparative analysis of the hardware implementation metrics for Single and Dual Processing Element (PE) architectures. The Dual PE architecture exhibits higher area and power consumption but achieves significantly reduced latency compared to the Single PE design. As a result, the $APT2$ benchmark indicates a more favorable cost-performance trade-off for the Dual PE configuration, with a lower overall score of 819,596 versus 1,071,251 for the Single PE.

Both Single PE and Dual PEs are implemented on hardware within the CMOS 45nm NANGATE technology at the operational frequency of $20MHz$ by using the conventional design methodology with EDA tools to evaluate the accuracy of the estimation approach in the proposed EEMitM. Table 4.5 provides a detailed comparison of key performance indicators for the Single and Dual Processing Element (PE) architectures, including area, power consumption, latency, and the integrated APT2 metric. Notably, while the Dual PE configuration demands a higher silicon area and consumes more power, it offers substantially improved latency performance, which is critical in time-sensitive applications. Consequently, the enhanced latency of the Dual PE architecture yields a better APT2 score, demonstrating a more optimal balance between cost and performance despite the increased resource utilization.

c. Experimental Results of the proposed EEMitM Design Methodology

Implementation costs The benchmark precise metrics of the implementation costs, such as area physical cost, power consumption, and latency, are defined to facili-

Table 4.5: Estimation Hardware Implementation Cost by using conventional Design Methodology.

Architectures	**Single PE**	**Dual PEs**
Area (kGates)	20.93	29.88
Power Consumption (mW)	798.4	1,402
Latency (μs)	9.52	5.44
$APT2$	1,514,480	1,239,728

tate the accuracy of the estimation mechanism in the proposal. These metrics compare the corresponding variables between using the conventional design methodology and the proposed estimation approach as determined in Equation (4.29):

$$\begin{aligned} \texttt{Error}_{Area}(\%) &= 100\% - \frac{|A_{EDA} - A_{Estimation}|}{A_{EDA}} \\ \texttt{Error}_{Power}(\%) &= 100\% - \frac{|P_{EDA} - P_{Estimation}|}{P_{EDA}} \\ \texttt{Error}_{Latency}(\%) &= 100\% - \frac{|T_{EDA} - T_{Estimation}|}{T_{EDA}} \end{aligned} \tag{4.29}$$

where $A_{EDA}, P_{EDA}, T_{EDA}$ are the implementation costs, area physical cost, power consumption, and latency metrics derived from the EDA tools using the conventional design methodology. Additionally, $A_{Estimation}, P_{Estimation}, T_{Estimation}$ are estimated using the proposed EEMitM design methodology.

Furthermore, the additional feature of the proposed EEMitM design methodology is comparing the relative proportions of alternative designs. Relative ratios of implementation costs between two designs, defined as Equation (4.30), are also estimated and measured by using both the proposed design methodology and the conventional design methodology, respectively. Figure 4.19 shows in more detail the precision of the estimation approach in the proposed EEMitM design methodology, compared to the EDA tools using the conventional design methodology.

$$\begin{aligned} \texttt{R}_{Area} &= \frac{A_{Single\_PE}}{A_{Dual\_PEs}} \\ \texttt{R}_{Power} &= \frac{P_{Single\_PE}}{P_{Dual\_PEs}} \\ \texttt{R}_{Latency} &= \frac{T_{Single\_PE}}{T_{Dual\_PEs}} \end{aligned} \tag{4.30}$$

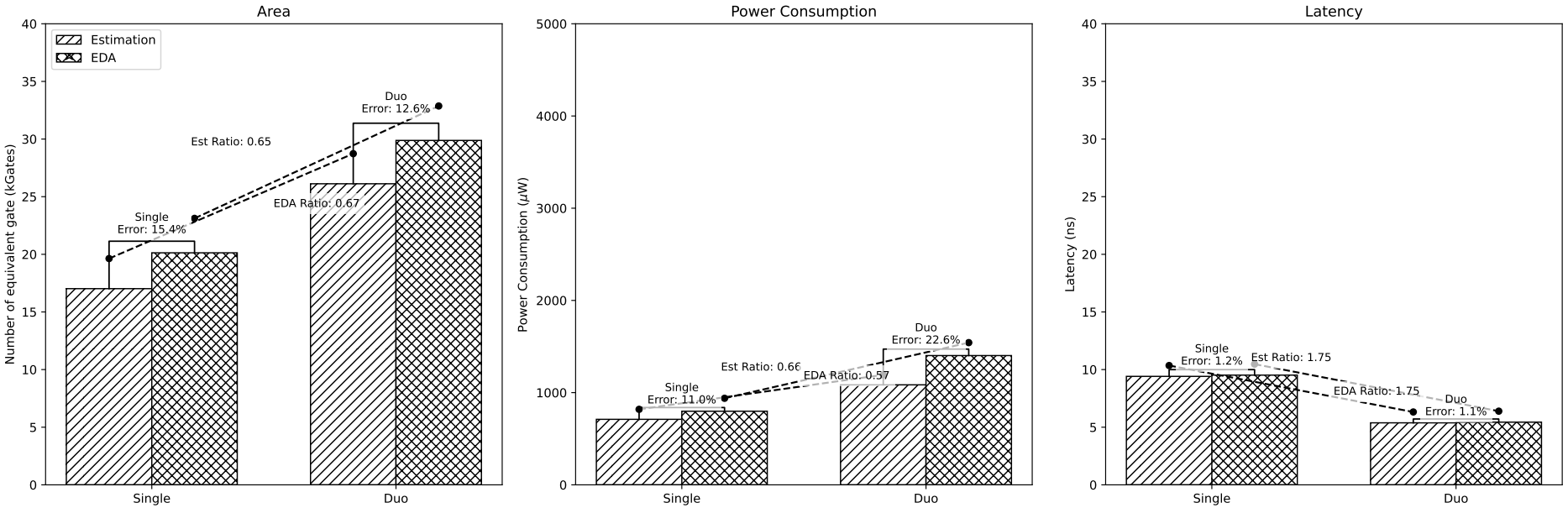


Figure 4.19: Proposed Early Evaluation Design Methodology.

In terms of power consumption, the estimated ratio of 0.66 shows reasonable correspondence with the EDA ratio of 0.57. This close correspondence suggests that the estimation model is effective in approximating real power values under the Single configuration. However, the Dual configuration exhibits a higher estimation error of 22.6%, indicating a noticeable deviation from actual power measurements. This larger discrepancy may stem from more complex interactions in the dual-core design that are not fully captured by the estimation model. Despite this, the power estimation still provides a useful baseline during early design stages.

The latency metric shows perfect alignment between estimation and EDA results, with identical ratios of 1.75 and minimal error margins of 1.2% (Single) and 1.1% (Dual). This demonstrates exceptional estimation precision for this parameter. This suggests that the estimation approach captures timing characteristics of the design with impressive precision. Such accuracy is particularly valuable in time-critical applications, where tight timing constraints must be met. It also validates the robustness of the latency estimation model across different architectural configurations.

For the area metric, the proposed estimation method performs well across both configurations. The estimated area for Single shows a ratio of 0.65 compared to the EDA value, with an error margin of 15.4%, which is acceptable for early design evaluation. In the Duo setup, the estimation ratio improves slightly to 0.67, and the error reduces to 12.6%. These results suggest that the estimation framework is fairly reliable in predicting gate count trends, especially as design complexity scales. Overall, the estimation results show a strong correlation with EDA outputs, particularly for area and latency metrics, validating the estimation methodology's effectiveness.

Duration of process Utilizing the proposed EEMitM reduces the time-to-market significantly. In our experiments, we consider four designs of ECC blocks, including the BGC [13], BEC [11], and the proposed BEC with single core and dual cores. Each design is sequentially synthesized, verified, and validated with the assumed design constraints with 1,000 reference test cases on the same server with a single core of the Intel(R) Xeon(R) CPU E5-2640 v4 2.40GHz.

In the conventional top-down design methodology, each design consumes approximately 1,440,000 seconds (2 weeks) to complete the design process. In contrast, by utilizing the proposed EEMitM, each design only needs 3,000 seconds (50 minutes) to estimate the implementation costs and generate pseudo power traces.

In the experimental scope of four designs, the total durations of the process of the conventional top-down and the proposed EEMitM design methodology are shown in Figure 4.20. Figure 4.20 shows the advantages of utilizing the proposed design

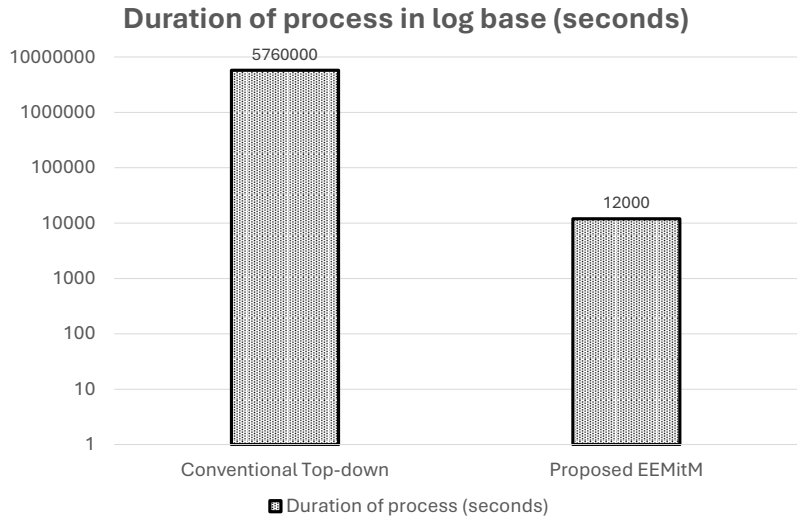


Figure 4.20: Comparison of the process's duration between the conventional top-down and the proposed EEMitM design methodologies in log base.

methodology. The proposed design methodology requires 12,000 seconds to find the optimal design of the ECC block, which saves 480 times compared to the conventional approach.

4.5 Summary

This section of the thesis addresses the effectiveness of balancing the security level and the implementation costs of the passive ECC-based Authentication Protocol RFID tags by proposing an innovative design process named Early Evaluation Meet-in-the-Middle (EEMitM). The proposed design methodology is firstly based on proposing a database of the primitive hardware components. The database provides the general information of these hardware modules, such as power consumption, physical area occupation, and mathematical model. This information is provided to the proposed estimation process, which is the second contribution of the proposed EEMitM, to quantify the approximated hardware implementation costs of the desired system. Lastly, the third proposal of the estimation power trace method appropriates the pseudo trace of the instantaneous power consumed by the hardware system. By reusing the information provided by the existing models in the database, the proposal EEMitM enables quick estimation and evaluation of the alternative designs. Thanks for the proposal, the time-to-product can be reduced by 480 times better than using the conventional top-down design approaches in the experimental condition. These results are published in [C3] and partly reported in [J1, C2]. Additionally, the proposed EEMitM is also proposed

in an accepted patent [P1].

In practice, there are N selections of hardware implementation. Therefore, if the designers carry out the conventional top-down design methodology, the time-to-products is approximately $N \cdot T_{HW}$, where T_{HW} is the minimum duration to implement a design on hardware. Consequently, the total time-to-product is significant, which increases the manufacturing price of the product.

By using the proposed innovative design methodology, the estimation and evaluation are executed over the software with the determined database in the duration of $T_{SW} << T_{HW}$. At the end of the assessment, the optimal design is implemented on hardware within T_{HW}. In total, the time-to-product of using the proposed EEMitM design methodology is $N \cdot T_{SW} + T_{HW}$.

Related to the precision of the estimation, the proposal provides an impressive accuracy of 1.2% of error in latency. In the worst estimation of the power consumption, the accuracy stays at an acceptable error of 22.6%. Additionally, the time-to-market in the condition of experiments can be reduced by 480 times better than using the conventional top-down design approaches.

Conclusion and Future Works

Conclusion

With the outstanding advantages of low cost and low power, passive RFID tags have become an essential part of our lives, ranging from civil to military applications, significantly impacting human daily life. However, the security vulnerabilities require the passive RFID tags to employ complementary countermeasures, such as ECC. These solutions help passive ECC-based RFID tags avoid many threats, such as wireless and hardware attacks. The drawback of the approaches is carrying out additional operators, which permits the reduction of the possibility of leakage. These complementary operators enlarge the system's total implementation costs, possibly exceeding the physical constraints of the passive RFID tags. Therefore, the balance between the implementation cost and the security level of the passive RFID tags is a critical challenge.

In order to address this question, several research studies have tried to develop major trends. The first attempt is to optimize the complexity of the ECC-based authentication protocol to match the implementation costs. The second orientation is to minimize the implementation cost of security primitives that are utilized in the ECC-based authentication protocol. The last is a hardware-software co-design approach, where efficient hardware of the security primitives with a compatible protocol is explored. However, these recent attempts contain limitations as follows:

1. Because of a lack of systematic information, the conventional attempts often treat security as an afterthought or a complementary feature at the later design phase rather than considering it at the beginning. Consequently, the ability to produce insecure and costly ECC-based solutions.

2. Due to the ECC-based authentication protocol being complicated, which is possibly optimized over multiple layers, the design space of the system is massive. With the traditional mechanisms, implementing trial-and-error to find the optimal ECC-based authentication protocol that fits with the design constraints is a

critical challenge. Scanning all the huge space of designs with the conventional design methodology increases the time-to-market.

Addressing these mentioned drawbacks, it is necessary to invent an innovative design methodology that helps designers consider both the security and implementation costs of the ECC-based authentication protocol in each design step. As a consequence, designers reduce the number of iterations for trial-and-error prototyping and minimize the time-to-product expense.

The ultimate goal of this research was to propose a comprehensive set of design phases that enable the designer to quickly and early estimate, evaluate, and choose the best systematic configurations for hardware implementation. Consequently, the designers save on the time-to-product cost. The main research works can be summarized as follows:

1. A 21.68 $kGates$ of a low-cost, low-power hardware architecture of BEC has been proposed and synthesized by using the CMOS TSMC 65nm technology, according to the conventional Top-down design methodology. The proposed hardware implementation consumes $126\mu W@10MHz$ of power, estimated by the Synopsys PrimeTime tool. The hardware implementation results of the proposal show that our proposed system is efficient in terms of implementation cost and energy consumption when compared with other works. Besides, the TVLA evaluation proves that the proposal is secure against side-channel vulnerabilities. These results are published in [C1].

2. To solve the problem of time-to-product, we proposed the Early Evaluation MitM design methodology. By using the database of the reference security primitives, the designers can possibly model the chosen system without prototyping. The proposed framework enables quick modeling over the software environment, which significantly reduces the time to produce. The experiments show that by applying the proposal, the time-to-product reduces by 480 times compared to the conventional design flow. Regarding the accuracy of the estimation of implementation costs, the experiments show a precise estimation of 1.2% of error in latency. In the worst estimation of the power consumption, the accuracy stays at an acceptable error of 22.6%. These discussions of the proposal are published in [C2, C3, P1] and partly presented in [J1].

Future works

While this study has made strides in addressing key challenges within the exploration of the design space for ECC (Elliptic Curve Cryptography)- based authentication protocols tailored to passive RFID (Radio-Frequency Identification) tags, several avenues for future research remain open. The work has focused on identifying optimal design configurations compatible with the resource-constrained characteristics of passive RFID systems. However, to advance the field further, the following directions are proposed as critical next steps.

Evaluate the security properties Aiming to prove the security benefits of the first contribution in this thesis, a general TVLA evaluation is performed and presented in Chapter 3. Although the TVLA evaluation of the system validates the security property against the SCA threats, the experiments to measure the robustness of the proposed system against the particular vulnerabilities, such as Correlation Power Analysis (CPA) and Differential Power Analysis (DPA) are necessary to strengthen the security advance of our work.

Develop Commercial EDA Tools with the proposed Early Evaluation MitM design methodology A primary objective moving forward is the refinement and integration of the proposed Early Evaluation MitM (Meet-in-the-Middle) design methodology into existing commercial Electronic Design Automation (EDA) tools. By embedding the Early Evaluation MitM framework, the goal is to create a seamless interface that bridges theoretical design exploration and practical implementation. A key component of this effort involves developing an intuitive graphical user interface (GUI) for the platform. This GUI would enable designers to interact dynamically with estimation and evaluation modules, allowing real-time adjustments to the parameters of the design.

Extending Application-Specific Database To increase the versatility and adoption of the proposed tools, expanding the Early Evaluation MitM methodology's reference databases is essential. Currently, the datasets are limited to a narrow range of ECC-based authentication protocols for passive RFID use cases. The platform's applicability would grow significantly by diversifying these databases to include applications. This expansion would motivate the foster cross-industry collaboration.

List of Publications

List of publications relevant to the thesis

C1 **Manh-Hiep Dao**, Vincent Beroulle, Yann Kieffer, Xuan-Tu Tran, and Duy-Hieu Bui. "Low-cost Low-Power Implementation of Binary Edwards Curve for Secure Passive RFID Tags." In 2023 IEEE 16th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC), pp. 494-500. IEEE, 2023.

C2 **Manh-Hiep Dao**, Vincent Beroulle, Yann Kieffer, and Xuan-Tu Tran. "How to Develop ECC-Based Low-Cost RFID Tags Robust Against Side-Channel Attacks." In International Conference on Industrial Networks and Intelligent Systems, pp. 433-447. Cham: Springer International Publishing, 2021.

C3 **Manh-Hiep Dao**, Vincent Beroulle, Yann Kieffer, and Xuan-Tu Tran (2023). Secure-by-Design methodology using Meet-in-the-Middle design flow for hardware implementations of ECC-based passive RFID tags. In ICWMC 2023, The Nineteenth International Conference on Wireless and Mobile Communications. IARIA (pp. 14-19).

J1 Souhir Gabsi, Vincent Beroulle, Yann Kieffer, **Manh-Hiep Dao**, Yassin Kortli, and Belgacem Hamdi. "Survey: Vulnerability analysis of low-cost ECC-based RFID protocols against wireless and side-channel attacks." Sensors 21, no. 17 (2021): 5824. (**SCIE, Q2**).

P1 **Đào Mạnh Hiệp**(2023), "Quy trình thiết kế phần cứng bảo mật cân bằng giữa chi phí thực thi và mức độ bảo mật" (VN Patent No. 1-2023-06893) (*Accepted*)

List of publications published during the thesis

J2 **Manh-Hiep Dao**, Koichiro Ishibashi, The-Anh Nguyen, Duy-Hieu Bui, Hiroshi Hirayma, Tuan-Anh Tran, and Xuan-Tu Tran. "Low-cost, High Accuracy, and Long Communication Range Energy-Harvesting Beat Sensor with LoRa and Ω-Antenna for Water-Level Monitoring." IEEE Sensors Journal (2025). (**SCIE, Q1**).

Bibliography

[1] Market Data Forecast, "Global RFID market size, share, trends & growth forecast report by component, tag type, industry and region (North America, Europe, Asia-Pacific, Latin America, Middle East and Africa), industry analysis from 2025 to 2033," https://www.marketdataforecast.com/market-reports/global-rfid-market, 2024.

[2] P. Xu, D. Flandre, and D. Bol, "Analysis, modeling, and design of a 2.45-ghz rf energy harvester for SWIPT IoT smart sensors," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 10, pp. 2717–2729, 2019.

[3] International Organization for Standardization, "ISO/IEC 14443: Identification cards — Contactless integrated circuit cards — Proximity cards," Geneva, Switzerland, 2016, retrieved from ISO Standards repository.

[4] A. Juels, "Rfid security and privacy: A research survey," *IEEE journal on selected areas in communications*, vol. 24, no. 2, pp. 381–394, 2006.

[5] S. Mangard, E. Oswald, and T. Popp, *Power analysis attacks*, 2007th ed. New York, NY: Springer, Dec. 2007.

[6] J. S. Chou, "An efficient mutual authentication rfid scheme based on elliptic curve cryptography," *Journal of Supercomputing*, vol. 70, 2014.

[7] D.-H. Bui, D. Puschini, S. Bacles-Min, E. Beigné, and X.-T. Tran, "Aes datapath optimization strategies for low-power low-energy multisecurity-level internet-of-things applications," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 12, pp. 3281–3290, 2017.

[8] J. G. Pandey, T. Goel, M. Nayak, C. Mitharwal, S. Khan, S. K. Vishvakarma, A. Karmakar, and R. Singh, "A vlsi architecture for the present block cipher with fpga and asic implementations," in *International Symposium on VLSI Design and Test*. Springer, 2018, pp. 210–220.

[9] H. Gross, E. Wenger, C. Dobraunig, and C. Ehrenhöfer, "Ascon hardware implementations and side-channel evaluation," *Microprocessors and Microsystems*, vol. 52, pp. 470–479, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0141933116302721

[10] M.-H. Dao, V.-P. Hoang, V.-L. Dao, and X.-T. Tran, "An energy efficient aes encryption core for hardware security implementation in iot systems," in *2018 International Conference on Advanced Technologies for Communications (ATC)*. IEEE, 2018, pp. 301–304.

[11] B. Koziel, R. Azarderakhsh, and M. Mozaffari-Kermani, "Low-resource and fast binary edwards curves cryptography," vol. 9462, 2015.

[12] M. A. Bahnasawi, K. Ibrahim, A. Mohamed, M. K. Mohamed, A. Moustafa, K. Abdelmonem, Y. Ismail, and H. Mostafa, "Asic-oriented comparative review of hardware security algorithms for internet of things applications," in *2016 28th International Conference on Microelectronics (ICM)*. IEEE, 2016, pp. 285–288.

[13] J. López and R. Dahab, "Fast multiplication on elliptic curves over gf (2m) without precomputation," in *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, 1999, pp. 316–327.

[14] K. H. Kim, C. O. Lee, and C. Negre, "Binary edwards curves revisited," in *INDOCRYPT*. Springer, 2014, pp. 393–408.

[15] R. R. Farashahi and M. Joye, "Efficient arithmetic on hessian curves," vol. 6056 LNCS, 2010.

[16] E. Wenger and J. Grossschadl, "An 8-bit avr-based elliptic curve cryptographic risc processor for the internet of things," in *2012 45th Annual IEEE/ACM International Symposium on Microarchitecture Workshops*. IEEE, 2012, pp. 39–46.

[17] J. P. Deschamps, G. D. Sutter, and E. Cantó, "Guide to fpga implementation of arithmetic functions," *Lecture Notes in Electrical Engineering*, vol. 149 LNEE, 2012.

[18] H. Mahdizadeh and M. Masoumi, "Novel architecture for efficient fpga implementation of elliptic curve cryptographic processor over GF(2^m)," *IEEE transactions on very large scale integration (VLSI) systems*, vol. 21, no. 12, pp. 2330–2333, 2013.

[19] D. Hankerson, A. J. Menezes, and S. Vanstone, *Guide to Elliptic Curve Cryptography*. Berlin, Heidelberg: Springer-Verlag, 2003.

[20] Ü. Kocabaş, J. Fan, and I. Verbauwhede, "Implementation of binary edwards curves for very-constrained devices," in *ASAP 2010-21st IEEE International Conference on Application-specific Systems, Architectures and Processors*. IEEE, 2010, pp. 185–191.

[21] V. Tujillo-Olaya and J. Velasco-Medina, "Hardware architectures for elliptic curve cryptoprocessors using polynomial and gaussian normal basis over GF(2^m)," *Transactions on Computational Science XI: Special Issue on Security in Computing, Part II*, pp. 79–103, 2010.

[22] C. D. Walter, "Simple power analysis of unified code for ecc double and add," in *Cryptographic Hardware and Embedded Systems-CHES 2004: 6th International Workshop Cambridge, MA, USA, August 11-13, 2004. Proceedings 6*. Springer, 2004, pp. 191–204.

[23] H. M. Edwards, "A normal form for elliptic curves," *Bulletin of the American Mathematical Society*, vol. 44, 2007.

[24] D. J. Bernstein, T. Lange, and R. Rezaeian Farashahi, "Binary edwards curves," in *Cryptographic Hardware and Embedded Systems – CHES 2008*, E. Oswald and P. Rohatgi, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 244–265.

[25] R. Abarzúa, C. Valencia, and J. López, "Survey for performance & security problems of passive side-channel attacks countermeasures in ECC," *Cryptology ePrint Archive*, 2019.

[26] M. Joye and S.-M. Yen, "The montgomery powering ladder," in *International workshop on cryptographic hardware and embedded systems*. Springer, 2002, pp. 291–302.

[27] D. J. Bernstein and T. Lange, "Montgomery curves and the montgomery ladder," 2017.

[28] T. Oliveira, J. López, and F. Rodríguez-Henríquez, "The montgomery ladder on binary elliptic curves," *Journal of Cryptographic Engineering*, vol. 8, pp. 241–258, 2018.

[29] I. Grout, *Digital systems design with FPGAs and CPLDs*. Elsevier, 2011.

[30] E. Kang, E. Jackson, and W. Schulte, "An approach for effective design space exploration," in *Foundations of Computer Software. Modeling, Development, and Verification of Adaptive Systems: 16th Monterey Workshop 2010, Redmond, WA, USA, March 31-April 2, 2010, Revised Selected Papers 16*. Springer, 2011, pp. 33–54.

[31] T. Shahroodi, S. Bayat-Sarmadi, and H. Mosanaei-Boorani, "Low-latency double point multiplication architecture using differential addition chain over gf(2m)," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, 2019.

[32] R. Salarifard and S. Bayat-Sarmadi, "An efficient low-latency point-multiplication over curve25519," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, 2019.

[33] P. Choi, M. K. Lee, J. H. Kim, and D. K. Kim, "Low-complexity elliptic curve cryptography processor based on configurable partial modular reduction over nist prime fields," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, 2018.

[34] S. R. Pillutla and L. Boppana, "Low-complexity bit-serial sequential polynomial basis finite field gf(2m) montgomery multipliers," *Microprocessors and Microsystems*, vol. 84, 2021.

[35] Q. Shao, Z. Hu, S. N. Basha, Z. Zhang, Z. Wu, C. Y. Lee, and J. Xie, "Low complexity implementation of unified systolic multipliers for nist pentanomials and trinomials over gf(2m)," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, 2018.

[36] E. Monmasson and M. N. Cirstea, "Fpga design methodology for industrial control systems—a review," *IEEE transactions on industrial electronics*, vol. 54, no. 4, pp. 1824–1842, 2007.

[37] A. B. Kahng, J. Lienig, I. L. Markov, and J. Hu, *VLSI physical design: from graph partitioning to timing closure*. Springer, 2011, vol. 312.

[38] T. Riesgo, Y. Torroja, and E. De la Torre, "Design methodologies based on hardware description languages," *IEEE Transactions on Industrial electronics*, vol. 46, no. 1, pp. 3–12, 1999.

[39] H. Chang, *A top-down, constraint-driven design methodology for analog integrated circuits*. Springer Science & Business Media, 1997.

[40] P. Mohan, O. Atli, O. O. Kibar, and K. Mai, "A top-down design methodology for synthesizing fpga fabrics using standard asic flow," in *Proceedings of the 2020 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, 2020, pp. 313–313.

[41] N. Pirotte, J. Vliegen, L. Batina, and N. Mentens, "Balancing elliptic curve coprocessors from bottom to top," *Microprocessors and Microsystems*, vol. 71, p. 102866, 2019.

[42] J. Lutz and M. Anwarul Hasan, "High performance elliptic curve cryptographic coprocessor," in *Wireless Network Security*. Springer, 2007, pp. 3–42.

[43] Y. Li, Y. Wang, Y. Li, R. Zhou, and Z. Lin, "An artificial neural network assisted optimization system for analog design space exploration," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 39, no. 10, pp. 2640–2653, 2019.

[44] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002.

[45] B. Tang, "Orthogonal array-based latin hypercubes," *Journal of the American statistical association*, vol. 88, no. 424, pp. 1392–1397, 1993.

[46] S. Zhang, F. Yang, C. Yan, D. Zhou, and X. Zeng, "An efficient batch-constrained bayesian optimization approach for analog circuit synthesis via multiobjective acquisition ensemble," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 1, pp. 1–14, 2021.

[47] M. Fayazi, M. T. Taba, E. Afshari, and R. Dreslinski, "Angel: Fully-automated analog circuit generator using a neural network assisted semi-supervised learning approach," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 70, no. 11, pp. 4516–4529, 2023.

[48] B. Schürmann, J. Altmeyer, and M. Schütze, "On modeling top-down vlsi design," in *Proceedings of the 1994 IEEE/ACM international conference on Computer-aided design*. Citeseer, 1994, pp. 508–515.

[49] E. Aerabi, M. Bohlouli, M. H. A. Livany, M. Fazeli, A. Papadimitriou, and D. Hely, "Design space exploration for ultra-low-energy and secure iot mcus," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 19, no. 3, pp. 1–34, 2020.

[50] T. Mahfuz, S. Bhunia, and P. Chakraborty, "X-dfs: Explainable artificial intelligence guided design-for-security solution space exploration," *IEEE Transactions on Information Forensics and Security*, 2024.

[51] M.-H. Dao, V. Beroulle, Y. Kieffer, and X.-T. Tran, "How to develop ecc-based low cost rfid tags robust against side-channel attacks," in *International Conference on Industrial Networks and Intelligent Systems*. Springer, 2021, pp. 433–447.

[52] F. Regazzoni, A. Cevrero, F.-X. Standaert, S. Badel, T. Kluter, P. Brisk, Y. Leblebici, and P. Ienne, "A design flow and evaluation framework for dpa-resistant instruction set extensions," in *Cryptographic Hardware and Embedded Systems-CHES 2009: 11th International Workshop Lausanne, Switzerland, September 6-9, 2009 Proceedings*. Springer, 2009, pp. 205–219.

[53] L. Vigano, "Automated security protocol analysis with the avispa tool," *Electronic Notes in Theoretical Computer Science*, vol. 155, pp. 61–86, 2006.

[54] C. Paar, "Efficient vlsi architectures for bit parallel computation in galios [galois] fields," 1994. [Online]. Available: https://api.semanticscholar.org/CorpusID:2111417

[55] E. B. Barker, "Sp 800-57. recommendation for key management, part 1: General (rev.5)," Gaithersburg, MD, USA, Tech. Rep., 2020.

[56] M. Tunstall and G. Goodwill, "Applying tvla to public key cryptographic algorithms," *Cryptology ePrint Archive*, 2016.

[57] C. A. Lara-Nino, A. Diaz-Perez, and M. Morales-Sandoval, "Energy/area-efficient scalar multiplication with binary edwards curves for the iot," *Sensors*, vol. 19, no. 3, p. 720, 2019.

[58] C. A. Lara-Nino, A. Diaz-Perez, and M. Morales-Sandoval, "Lightweight elliptic curve cryptography accelerator for internet of things applications," *Ad Hoc Networks*, vol. 103, p. 102159, 2020.

[59] Y.-p. Dan and H.-l. He, "Tradeoff design of low-cost and low-energy elliptic curve cryptoprocessor for wireless sensor networks," in *2012 8th International Conference on Wireless Communications, Networking and Mobile Computing*. IEEE, 2012, pp. 1–5.

[60] B. Rashidi, S. M. Sayedi, and R. R. Farashahi, "Full-custom hardware implementation of point multiplication on binary edwards curves for application-specific integrated circuit elliptic curve cryptosystem applications," *IET Circuits, Devices & Systems*, vol. 11, no. 6, pp. 568–578, 2017.

[61] V. Rožić, O. Reparaz, and I. Verbauwhede, "A 5.1 μj per point-multiplication elliptic curve cryptographic processor," *International Journal of Circuit Theory and Applications*, vol. 45, no. 2, pp. 170–187, 2017.

[62] J. J. Fournier, A. Loiseau, and J. Fournier, "Binary edwards curves for intrinsically secure ecc implementations for the iot," in *International Conference on Security and Cryptography*. SCITEPRESS-Science and Technology Publications, 2018, pp. 625–631.

[63] A. Stillmaker and B. Baas, "Scaling equations for the accurate prediction of cmos device performance from 180 nm to 7 nm," *Integration*, vol. 58, pp. 74–81, 2017.

[64] Z. Lyu and J. Shen, "An efficient algorithm for estimating gate-level power consumption in large-scale integrated circuits," *Microelectronics Journal*, vol. 146, p. 106143, 2024.

[65] C. A. Lara-Nino, A. Diaz-Perez, and M. Morales-Sandoval, "Elliptic curve lightweight cryptography: A survey," *Ieee Access*, vol. 6, pp. 72 514–72 550, 2018.

[66] A. Salman, A. Ferozpuri, E. Homsirikamol, P. Yalla, J.-P. Kaps, and K. Gaj, "A scalable ecc processor implementation for high-speed and lightweight with side-channel countermeasures," in *2017 International Conference on ReConFigurable Computing and FPGAs (ReConFig)*, 2017, pp. 1–8.

[67] R. Azarderakhsh, K. U. Järvinen, and M. Mozaffari-Kermani, "Efficient algorithm and architecture for elliptic curve cryptography for extremely constrained secure applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, 2014.

[68] E. Wenger, "Hardware architectures for msp430-based wireless sensor nodes performing elliptic curve cryptography," in *Applied Cryptography and Network Security: 11th International Conference, ACNS 2013, Banff, AB, Canada, June 25-28, 2013. Proceedings 11*. Springer, 2013, pp. 290–306.

[69] S. Gabsi, V. Beroulle, Y. Kieffer, H. M. Dao, Y. Kortli, and B. Hamdi, "Survey: Vulnerability analysis of low-cost ecc-based rfid protocols against wireless and side-channel attacks," *Sensors*, vol. 21, no. 17, p. 5824, 2021.

[70] Y.-P. Liao and C.-M. Hsiao, "A secure ecc-based rfid authentication scheme integrated with id-verifier transfer protocol," *Ad hoc networks*, vol. 18, pp. 133–146, 2014.

[71] Z. Zhao, "A secure rfid authentication protocol for healthcare environments using elliptic curve cryptosystem," *Journal of medical systems*, vol. 38, pp. 1–7, 2014.

[72] A. A. Alamr, F. Kausar, J. Kim, and C. Seo, "A secure ecc-based rfid mutual authentication protocol for internet of things," *The Journal of supercomputing*, vol. 74, pp. 4281–4294, 2018.

[73] L. Zheng, Y. Xue, L. Zhang, and R. Zhang, "Mutual authentication protocol for rfid based on ecc," in *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, vol. 2. IEEE, 2017, pp. 320–323.

[74] S. Gabsi, Y. Kortli, V. Beroulle, Y. Kieffer, A. Alasiry, and B. Hamdi, "Novel ecc-based rfid mutual authentication protocol for emerging iot applications," *IEEE access*, vol. 9, pp. 130 895–130 913, 2021.

[75] A. Arslan and M. A. Bingöl, "Security and privacy analysis of recently proposed ecc-based rfid authentication schemes," *Cryptology ePrint Archive*, 2022.