THESIS TO OBTAIN THE DEGREE OF DOCTOR OF THE UNIVERSITY OF MONTPELLIER

In SyAM - Automatic and Microelectronic Systems

Doctoral school: Information, Structures, and Systems sciences

Research Unit: LIRMM

Body biasing fault injection: modeling

Presented by Geoffrey Chancel

COMPILATION DATE: 2023-09-07 18:09:30+02:00

Under the supervision of TO BE COMPLETED

Thesis Committee:

Philippe Maurine ,	Associate Professor ?? ,	University of Montpellier	Thesis Director
Jean-Marc Gallière,	Associate Professor ?? ,	University of Montpellier	Thesis Supervisor

UNIVERSITÉ DE MONTPELLIER

Abstract 2023-09-07 18:09:30+02:00

Résumé de la thèse 2023-09-07 18:09:30+02:00

Acknowledgements 2023-09-07 18:09:30+02:00

The authors acknowledge the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-19-CE39-0008 (project ARCHI-SEC). They also acknowledge the French Ministère des Armées – Agence de l'innovation de défense (AID) under grant ID-UM-2019 65 0036.

Contents

List of Figures

List of Tables

List of Algorithms

List of Acronyms

AES	**A**dvanced **E**ncryption **S**tandard
BBI	**B**ody **B**iasing **I**njection
BSIM	**B**erkeley **S**hort-channel **I**GFET **M**odel
CPS	**C**yber-**P**hysical **S**ystem
DES	**D**ata **E**ncryption **S**tandard
DoM	**D**ifference **o**f **M**eans
DFA	**D**ifferential **F**ault **A**nalysis
DPA	**D**ifferential **P**ower **A**nalysis
ECC	**E**lliptic-**C**urve **C**ryptography
EMFI	**E**lectro-**M**agnetic **F**ault **I**njection
FAM	**F**ault **A**nalysis **M**apping
FIB	**F**ocused **I**on **B**eam
FSA	**F**ault **S**ensibility **A**nalysis
FSM	**F**ault **S**usceptibility **M**ap
GFI	**G**litch **F**ault **I**njection
HFI	**H**ardware **F**ault **I**njection
IoT	**I**nternet **o**f **T**hings
LFI	**L**aser **F**ault **I**njection
PCC	**P**earson **C**orrleation **C**oefficient
PLL	**P**hase **L**ocked **L**oop
RAM	**R**andom **A**ccess **M**emory
RSA	**R**ivest **S**hamir **A**dleman
SCA	**S**ide **C**hannel **A**ttack
SCS	**S**tandard **C**ell **S**egment
SPA	**S**imple **P**ower **A**nalysis
WLCSP	**W**afer-**L**evel **C**hip-**S**cale **P**ackaging

Publications 2023-09-07 18:09:30+02:00

- [1]

General introduction 2023-09-07 18:09:30+02:00

Over the past years, various fault injection methods, representing a significant threat for secure integrated circuits, have been extensively studied, like laser fault injection (LFI), or more recently electromagnetic fault injection (EMFI). The purpose of these studies is to propose efficient countermeasures to the right cost. They have had multiple objectives, such as understanding the various phenomena at the origin of fault creation, or being able to simulate fault propagation over multiple abstraction levels...

Voltage pulse substrate fault injection, commonly called Body Biasing Injection (BBI), while being contemporary to EMFI, led to very few researches and studies in comparison. Up to the best of our knowledge, three scientific papers existed at the beginning of my thesis, back in 2020.

The LIRMM (Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier: Computer Sciences, Robotics and Microelectronics Laboratory of Montpellier), inventor of this technique in 2011, proposed this thesis to answer various questions such as:

- What are the phenomena at work leading to fault injection?
- What kind of spatial resolution does BBI offer?
- What is the time resolution of this method?
- Is it relevant to thin the silicon substrate of BBI target ICs?
- Can constraining fault attacks be performed with this method?

These questions have guided my thesis work through the last three years. These works have led me to propose CMOS integrated circuits simulation models in a BBI context, in addition to proposing improvements for the practice of BBI. My thesis manuscript is structured in five chapters. Each one of them attempt to provide answers to the preceding interrogations.

The first chapter of this manuscript provides an overview of the existing fault injection techniques, with a particular emphasis on BBI.

The second chapter describes improvements for the practice of BBI. These improvements have been conceived and obtained through my studies concerning BBI resolution and accuracy, both in time and space. Additionally, this chapter describes the practical results of a differential fault attack performed thanks to BBI and requiring single-bit faults.

The third chapter is dedicated to CMOS integrated circuits modeling under BBI. It introduces the established simulation models, in addition the designed algorithms allowing to simulate circuits subjected to BBI. The models and methods introduced allow us to simulate circuit behavior in reasonable duration, which allows us to perform parametric analysis of BBI effects.

The fourth chapter discusses a common practice in fault injection methods: the thinning of integrated circuits' substrate. While this topic has been extensively addressed concerning LFI, it is not the case for BBI. It relates to studying IC behavioral differences and BBI efficiency on different substrate thicknesses circuits. Various models are introduced to get different approaches, allowing to predict differently electrical and physical phenomena at work. Mathematical models are also derived from the previous models, enabling the calculation of optimal experimental parameters, in addition to predicting circuit behavior.

The fifth is dedicated to the understanding of fault creation in circuits subjected to BBI. It allows deriving a fault model from the simulations.

Eventually, the last chapter presents a general conclusion of my thesis work. In addition to this, outlooks are provided. The latter are interrogations remaining unanswered by my thesis works, mostly concerning more specific BBI effects on integrated circuits.

I

Introduction and state of the art 2023-09-07 18:09:30+02:00

Contents

1.1 Summary 2023-09-07 18:09:30+02:00

This chapter reviews the state-of-the-art concerning fault injection methods. It first defines the interest of studying fault injection and its context. Then, various fault injection techniques are presented and their differences, advantages and disadvantages are analyzed. Specifically, platforms equipment across all methods is described alongside the different techniques employed to perform such fault injection. Eventually, body biasing injection is introduced, and we will study its interests in a fault injection context.

1.2 Introduction 2023-09-07 18:09:30+02:00

In our time, almost every business sector and every part of our surroundings, directly or indirectly, use integrated electronics circuits. It ranges from smart-cards to supercomputers, through military devices, cell-phones, Cyber-Physical Systems (CPS) and Internet-of-Things (IoT) objects to name but a few.

Traditionally, integrated circuits design mainly focused on performance upgrades over the generations. Performance was measured thanks to two factors: computation speed and silicon surface. Within this context, power consumption was not a design constraint, therefore, integrated circuits became more and more energy-consuming. However, with the advent of portable devices, power consumption became a predominant design factor over speed, and space and got included into the former design flows. Nevertheless, less space and more speed does not physically equate with less energy. Alongside, new systems have emerged and have massively grown these past decades: IoT and CPS. On one hand, CPS are often systems where hardware and software are interlaced and thought together, and can be drastically different from one application to another. On the other hand, IoT systems have often less coordination between hardware and software, but are commonly more flexible. Whatsoever, both of these systems have something strong in common: their security is fundamental. Therefore, in this context, as it has been proposed in [2], and because security had been adopted as a countermeasure after the design flow, it has to enter as a fourth design rule when creating integrated circuits. This is required because a secure system has to ensure that every data going in and out of it are subject to the following criteria:

- Authenticity: data received have to come from the sender
- Integrity: data cannot be altered in any way
- Confidentiality: data cannot be accessed (read or written) by third-parties

Therefore, it is imperative to study and comprehend the strategies for enhancing IC security in order to develop future integrated circuits that are designed with security in mind from the initial stages of development to its completion.

Currently, electronic devices implement security in two distinct ways, namely from a software or hardware standpoint. To accomplish this objective, encryption algorithms have been integrated. It is possible to distinguish two distinct categories of encryption algorithms, namely symmetric and asymmetric algorithms.

In short, symmetric cryptographic techniques use a unique key for encrypting and decrypting messages. The most popular algorithms are the AES (Advanced Encryption Standard), DES (Data Encryption Standard), IDEA (International Data Encryption Algorithm), RC5 (Rivest Cipher 5), and TDES (Triple DES) not to cite them all. The key must be kept confidential and only shared among parties in order to maintain a confidential connection between them. The requirement for a single key is the main drawback of symmetric encryption methods. As a result, every possible step must be taken to safeguard key secrecy, such as avoiding key exchanges on public networks. However, symmetric encryption has a clear advantage over asymmetric encryption. As a result of utilizing a single key, symmetric algorithms are typically simpler than asymmetric algorithms, resulting in a reduction in computing power required for encryption. It is therefore possible to encrypt a large amount of data in a short amount of time.

In contrast, when it comes to symmetric cryptographic techniques, commonly referred to as public key cryptography techniques, a pair of keys is employed. The keys are usually referred to as public-key and private-key. The public key is used to encrypt a message, and anyone can use it. The private-key is, however, kept confidential to ensure that only authorized parties can decrypt a message that has been encrypted with the public-key. The primary motivation behind having two keys is that it is impracticable to reconstruct the public-key from the private-key. The most commonly employed asymmetric algorithms include the RSA (Rivest–Shamir–Adleman) algorithm, the ElGamal encryption system, the ECC (Elliptic-curve cryptography), and the Cramer-Shoup system, to name a few. The main drawback of symmetrical algorithms is that they involve large mathematical calculations, which implies a higher time complexity. Hence, these techniques are capable of encrypting a limited quantity of data. Therefore, to achieve this objective, in the majority of systems, a hybrid approach is employed to employ both encryption methods, thereby ensuring optimal security and a brief calculation time.

On the one hand, if all the previously mentioned algorithms are mathematically reliable, their reliability will decrease when they are implemented on actual integrated circuits. Indeed, every integrated circuit uses electrical energy to function. Therefore, when an electric current appears in a conductor, there is inevitably an electromagnetic field associated with this current. Moreover, every measurable physical quantity concerning the IC could be a point of information leakage. This is particularly true when considering the fact that these quantities will exhibit varying variations based on the calculations performed by the IC. When evaluating these quantities, it is possible to retrieve confidential information. We described what is called a **"side-channel attack" (SCA)** when considering cybersecurity.

On the other hand, physical quantity measurement is not the only flaw in actual algorithm implementations. In fact, every physical IC has specifications under which it can execute its functions properly. It includes temperature, clock frequency, power supply voltage, and the electromagnetic environment. When pushed beyond its specifications, any integrated circuit will exhibit unpredictable behavior. However, it is still possible to control an IC's behavior outside its specifications with a certain degree of success. By doing so, it allows running the IC calculation incorrectly by finely controlling how much time and by which amount the IC is outside its specifications, thus enabling, with specific mathematical algorithms, to retrieve hidden data manipulated by the IC. This process is commonly referred to as a **"fault injection attack"**.

We have identified two potential attacks on robust algorithms that have been implemented into actual integrated circuits. However, it is customary to categorize cyberattacks into three distinct categories based on their execution methods.

Despite being technically advanced, noninvasive attacks are the most materially trivial. SCA are included in this set, which do not require any hardware modification to the targeted ICs, even if there is no physical contact. It is a delicate task to detect them; hence, they are deemed to be highly dangerous and are commonly considered in the initial stages of designing integrated circuits.

It is then possible to distinguish semi-invasive attacks. Systematically, they are accompanied by device physical preparation, which is entirely devoid of noninvasive attacks, but they are not accompanied by device physical modification. ICs integrity is therefore theoretically preserved. A typical IC modification involves the removal of the chip package. It enables access to either the front or back side of the integrated circuit, thereby facilitating micro-probing, laser injection, or substrate pulse injection. Furthermore, substrate thinning is also commonly considered and used, as it facilitates the fine-tuning of certain fault injection techniques, such as laser fault injection (LFI). These attacks necessitate specialized hardware, tools, and expertise and are frequently challenging to establish and execute.

Eventually, there are invasive attacks. They imply further physical modifications to integrated circuits. For instance, it is common to eliminate the layers of a chip, thereby enabling the photographing of the various layers and the reverse engineering of the target. A focused ion beam (FIB) can also be used to change the IC target internally by making electric connections that did not exist before. Contrary to semi-invasive attacks, invasive attacks frequently involve the definitive destruction of the target, primarily due to the absence of physical integrity during the process.

This doctoral thesis is dedicated to the study of a specific fault injection method: Body Biasing Injection (BBI). In this particular context, we will examine in this chapter the current state of the art in relation to side-channel attacks and fault injection techniques as outlined in the literature. This allows us to explain the interests of the current work regarding hardware security.

In the first place, we will briefly discuss side-channel attacks. We will then examine the various fault injection platforms commonly described. Eventually, we will ponder the interests of BBI in this context.

1.3 Side-channel attacks 2023-09-07 18:09:30+02:00

1.3.1 Timing attacks 2023-09-07 18:09:30+02:00

The most fundamental side-channel attack was initially introduced in 1996 [3]. This attack involves determining the duration required to execute cryptographic computations. By executing this method, the adversaries were able to obtain a variety of algorithmic keys, specifically for the RSA algorithm. The computation cost of this attack is low, thereby enabling it to execute swift attacks. Indeed, as per the RSA algorithm, as outlined in [4], the encryption of a message necessitates the calculation of the following relationship:

$$C \equiv E(M) \equiv M^e (\text{mod } n) \tag{1.1}$$

M denotes the message to be encrypted, while C is the ciphertext and (e, n) the encryption key pair. The objective of the attack outlined in [3] is to retrieve e. To achieve this objective, the integrated circuit must perform multiple computations of the equation 1.1 for varying values of M, while maintaining identical values of e. Subsequently, the attacker must evaluate the duration of each computation. If the value of e differs for each operation, the attack cannot be executed. After the demonstration of this attack, countermeasures were implemented, including the implementation of constant-time cryptographic algorithms allowing the elimination of leaks through the utilization of timing analysis. More recently, other, more advanced countermeasures have also been proposed [5].

1.3.2 Power analysis and electromagnetic analysis attacks 2023-09-07 18:09:30+02:00

Subsequently, more elaborated side-channel attacks were explained in 1999, as documented in [6]. This paper presents the concepts of simple power analysis (SPA) and differential power analysis (DPA).

On the one hand, SPA entails the measurement and direct interpretation of power consumption traces of a cryptographic integrated circuit. For instance, it enables the counting of DES or AES rounds to gain insights into the utilized implementation. Furthermore, it allows for the observation of power consumption variations depending on the executed instruction. A proposal has been made to prevent the utilization of secret keys or information during conditional branching logic, with the objective of preventing simple power analysis.

On the other hand, DPA is a more elaborate approach that aims to identify the effects and variations associated with data processed by ICs. The aforementioned variations are more subtle and frequently obscured by noise. Therefore, DPA proposes to use statistics tools to reveal hidden system information, specifically by computing the difference of means (DoM) between traces. Therefore, preventing DPA is more complicated than preventing SPA. One of the simplest methods is to add electrical noise. Another technique is to reduce measurable signal amplitude. It is done first by optimizing code execution, by finely choosing which operation is performed to reduce electromagnetic leakage. Second, it is also possible to shield the device, but it increases the IC's cost significantly.

In addition to these attacks, there is also another attack which is commonly studied: correlation power analysis (CPA) [7]. As well as DPA, CPA uses statistical tools. However, as opposite to computing the difference of means, it involves calculating the Pearson correlation coefficient (PCC), allowing to measure the linear correlation between different power consumption traces.

It is important to note that SPA, DPA and CPA are historically performed using traces directly measured from the ICs power consumption. However, these attacks can also be performed thanks to IC electromagnetic radiation analysis [8]. Because electric charges are circulating into the IC, they inevitably generate electromagnetic waves. Therefore, it is possible to pick up these waves, and similar to power consumption, their shape depends on the data being processed. There has been numerous active research concerning this method for twenty years. It can be explained thanks to its advantages compared to bare power consumption analysis. Indeed, when measuring the entire power consumption of an IC, it is not possible to target a specific area. It leads, especially with complex ICs and countermeasures, to an impossibility to perform such attacks. On the contrary, electromagnetic analysis attacks have multiple advantages over power consumption analysis attacks:

- No sample preparation required
- No physical contact with the target
- It requires only little equipment: probe and voltage amplifier

As we stated previously, power consumption analysis attacks target an entire IC, whereas electromagnetic analysis attacks allow having fine resolutions. Indeed, small probes with a size down to 50 µm have been proposed [9]. Such small probes allow focusing the measurement on the cryptographic area of the IC, while excluding from the measurement, with a certain amount, any undesirable electromagnetic emission which could potentially harm the attack efficiency. In addition to that, electromagnetic probes, depending on their design, can have very high cutoff frequency. Therefore, it allows analyzing ICs running at high frequencies, enabling attacks on recent devices such as smartphones [10].

1.4 Fault-injection attacks 2023-09-07 18:09:30+02:00

Fault injections are widely described in the literature and can be utilized for a variety of purposes. For instance, during integrated circuits testing, it is common to find fault injection susceptibility tests, allowing for engineers to test fault detection circuits, recovery capabilities and reconfiguration possibilities of ICs. In this work, we are going to take a closer look at hardware fault injections (HFI) techniques solely, which fall in two distinct categories, similar to side-channel attacks:

- HFI with physical contact
- Contactless HFI

For each kind of HFI, multiple outcomes are aimed. On the one hand, the HFI can produce, in the targeted IC, branching errors leading secret codes to be revealed or protected rights to be acquired by an attacker. On the other hand, HFI can produce incorrect behaviors, allowing to retrieve hidden and protected data thanks to mathematical tools. In that case, HFI targets are mostly cryptographic algorithms, and can be segmented in non-comprehensive set of categories.

One of the most performed HFI is called differential fault attack/analysis (DFA). The principle of DFA lies in inducing computation errors during the decryption process of cryptographic algorithm thanks to fault injection. Several DFA were proposed on different algorithms [11, 12, 13, 14, 15]. Every DFA implies that the attacker has access to at least two ciphertexts, a correct one, denoted C, and a faulty one, denoted C_F. In addition to that, the attacker must also know the characteristics of the induced faults, such as the amount of faulted bits, in which operation they are faulted, etc. Eventually, it is needed to be able to induce the expected faults depending on the fault model required for the DFA.

Another common HFI is the fault sensitivity analysis (FSA) [16]. As every HFI, it is still required to have physical access to the device. FSA usefulness comes from the fact that alongside fault characteristics, other information can be used by attackers, in that case: the IC sensitivity to faults. As defined in [16], fault sensitivity is a condition where the faulty output begins to show specific characteristics. Specifically, this work defines a critical condition, similar to the PLL capture ranges (lock-in, hold-in, pull-in, etc.), where the IC starts to exhibit a faulty behavior or when it stops this behavior. Then, to perform an attack with this information, the attacker has to know the relationship between the fault sensitivity and the computed data, without knowing the insights of the cryptographic algorithm at work. It states that the algorithm will inevitably exhibit data-dependency of fault sensitivity. Hence, it allows using the IC as an almost black box.

In the next paragraph, we are going to analyze deeper a specific fault attack and its implications

1.4.1 Giraud's differential fault attack 2023-09-07 18:09:30+02:00

1.5 Fault-injection techniques 2023-09-07 18:09:30+02:00

1.5.1 Glitch fault injection 2023-09-07 18:09:30+02:00

Glitch fault injection (GFI) are one of the first historical documented fault injection attacks. They are simple and require little equipment. For the most part, they are non-invasive, which means that they are reversible, physically speaking. Various physical quantities can be disturbed, but the power supply voltages (VDD or GND), and the IC clock are the most common. Each physical quantity can be modified at the attacker's discretion, with a certain amount. However, the disturbances have to be short enough to avoid IC shutdown concerning power supply glitches, but also not powerful enough to avoid the IC destruction. On the one hand, the main advantage of such attack is its easiness to set up compared to other methods. On the other hand, their main disadvantage is the complete lack of locality with the injection effects. Indeed, disturbing IC's macro-parameters interfere with the entire chip and does not guarantee a useful faulty behavior. In addition to that, every modern IC is prepared to detect such attacks and thus protect itself by resetting its electronics.

1.5.2 Laser fault injection 2023-09-07 18:09:30+02:00

Laser fault injection (LFI), sometimes called optical fault injection, has been introduced in 2002 [17] and is a more complex technique than GFI. However, its precision is immensely better, at the cost of being semi-invasive, and sometimes invasive. LFI consists in targeting specific regions of the IC with laser beams of specific wavelengths. Several other parameters are involved for this method to succeed, such as the light emission duration, the area/volume of the targeted region, the IC substrate thickness, etc. Although LFI requires chip preparation, it is often minimal. LFI works thanks to the fact that every silicon semiconductor device (diode, transistor...) is intrinsically sensitive to light, typically with wavelengths ranging from 400 nm to 1000 nm. Therefore, if the light conveys enough energy, it is possible to change the state of some transistors, thus affecting logical values. The main shortcoming of LFI is the platform price. Add more details.

1.5.3 Electromagnetic fault injection 2023-09-07 18:09:30+02:00

Electromagnetic fault injection (EMFI) is a more recent and more studied technique, introduced in 2002 [18]. Its principle is basic: an electric current in a wire (probe) near an IC creates a corresponding electric current in the IC power delivery network, similar to an electric transformer. Similar to GFI, the attack can be non-invasive, although this method yields better results while being semi-invasive. Indeed, the closer the probe to the IC, the better the coupling and the mutual inductance, which often required to remove the IC's plastic package. This injection technique efficiency greatly varies depending on the probe's characteristics, the IC transistors size, the targeted location, the field duration, etc. Over the time, electromagnetic probes were constantly improved, and it is common to find probes with a ferrite core, allowing for better injection locality. In 2020, a modeling workflow was proposed [19], allowing to explain how EM probe can couple to IC power delivery networks. Add more details.

1.5.4 Body biasing injection 2023-09-07 18:09:30+02:00

Eventually, there is another fault injection method, less studied and more recent than the others, commonly called Body Biasing Injection (BBI), which is the research topic of this thesis. This technique has been introduced in 2012 [20], and further studied in 2013 [21] and 2016 [22]. At the beginning of this thesis, a fourth article was published [23], studying the interests of BBI concerning Wafer-Level Chip-Scale Packaging (WLCSP). The principle behind BBI is fairly simple: applying voltage pulses directly onto the backside of IC targets, thanks to a metallic probe. On the one hand, despite this simple premise, in the vast majority of cases, BBI is a semi-invasive method. Indeed, as most IC are encapsulated in a ceramic or plastic package, it is required, to access to the substrate, to partially remove a piece of the package. On the other hand, building a BBI platform is not expensive and technically easier when compared to LFI or EMFI. Indeed, a metallic probe with a custom armature costs around 10 euros at worst, and is easy to build at hand, while manufacturing a precise EMFI probe requires more knowledge to achieve good results. Considering that EMFI and BBI both require similar voltage pulse generator, which is often the most expensive piece of equipment, the overall platform cost is lower concerning BBI.

II

Body Biasing Injection platforms and good practices 2023-09-07 18:09:30+02:00

Contents

2.1 Introduction 2023-09-07 18:09:30+02:00

The first part of this chapter introduces Body Biasing Injection platforms equipment,with a special focus on the metallic probe and the voltage pulse generator, two major tools in the practice of BBI. Afterward, we are going to study BBI in practice. In the first place, I introduce an electrical model describing a typical BBI platform and how to evaluate BBI platform characteristics. Then, we will analyze the results of such platform by highlighting its inherent flaws, such as:

- Poor control of the platform's characteristics
- Obvious ringing leading to poor temporal accuracy
- Platform dependent parameters such as the ground installation quality
- Main physical quantities, such as the voltage and the pulse width, set points not met

Thereafter, I propose enhancements to overcome BBI platforms shortcomings:

- Matching the output impedance of the generator to reduce the ringing and bring the measurements closer to the specifications and the set points
- Bypassing the grounding installation to minimize platform dependency

After that, we will evaluate the enhancements with further analysis, measuring ringing, set points accuracy, load and transmission line dependency. Then, we will study actual BBI platforms while discussing various techniques allowing to match the generator's output impedance, in addition to introducing practical grounding installation bypass. Next, I will introduce actual measurements of such platforms, illustrating the enhancements in practice. Eventually, I will introduce a constraining differential fault attack set-up. It includes the attack description, followed by a thorough description of the IC target, sustained with experiments allowing me to perform the attack with more ease, ended up by the attack results.

Parts of this work were published in FDTC 2023. (Add reference.)

2.2 BBI platforms 2023-09-07 18:09:30+02:00

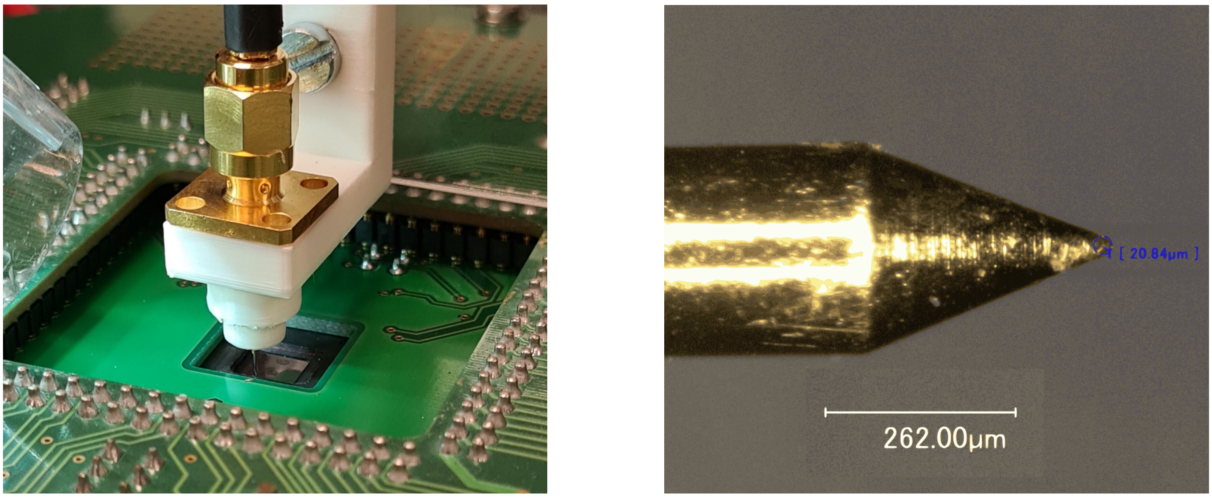
Body Biasing Injection platforms are constituted of many tools and equipment. It is quite similar of the one used in EMFI, and is composed of:

- A metallic probe (EM probe in EMFI), allowing high electric current and fast pulses
- A voltage pulse generator capable of generating very high, short and precise pulses

- A 3D positioning table, with a precision under 10 µm to place the probe precisely with correct pressure on the backside
- A vibration-proof table to minimize probe jitter due to vibration caused by other equipment or natural vibration
- A high precision oscilloscope to measure various physical quantities that might help to practice BBI

Among these tools, most are optional and play a role in improving the flexibility and accuracy of BBI. Therefore, we describe in details the probe and the generator as they represent the key equipment when performing BBI.

2.2.1 Probes 2023-09-07 18:09:30+02:00



(a) *Custom BBI probe in mechanical contact with the IC backside, seen through the perforated PCB*

(b) *Custom BBI probe microscopy physical measurements*

Figure 2.1: *Custom BBI probes photographs*

The most distinctive piece of equipment when working with BBI compared to other fault injection methods is the electrical probe. It is commonly made with a metal tip, a connector of any sort and a mechanical support to hold the structure together. Any size available on the market can be used depending on the needs. In the case of BBI, the probe is used to establish the electrical contact with the substrate of integrated circuits, the latter being poorly conductive, but not isolating. For this work, we designed a custom probe, allowing us to control its characteristics, around three simple parts:

- An SMA connector, to have a low-cost, small and standard interconnection available with almost every high-speed equipment
- A spring-loaded metallic tip soldered onto the SMA connector providing a better control over the applied pressure onto the backside

- A custom 3D-printed support holding the parts together, shaped to fit with our other tools

Fig. 2.1 shows detailed pictures of the probe we designed, with a photograph in operation in Fig. 2.1a, and a photograph under a microscope of the probe's tip-end in Fig. 2.1b, allowing to measure its actual size before its first usage. The metallic probe we had chosen has a 0.635 mm diameter and is 16.35 mm long. The specified maximum nominal current of the probe is of 1.5 A, and the electrical contact resistance measures approximately 70 mΩ. The tip has a diameter roughly equal to 20 µm, and it is important to note that this value tends to increase when the probe is utilized, due to the physical contact and the pressure with the IC backside. The bill-of-materials cost for our custom probe tool is roughly equal to 20 $, ignoring manual labor to assemble everything together.

2.2.2 Generator 2023-09-07 18:09:30+02:00

The other fundamental piece of equipment when practicing BBI is the voltage pulse generator. It is, generally, one of the most expensive platform tool, similar to EMFI. Indeed, because BBI relies on voltage pulses to disturb an IC, it is necessary to provide the researcher a precise control over the pulse parameters, such as the voltage set point, the pulse duration, the rise and fall times, etc. However, these pas few years have been introduced lower cost alternatives, at the cost of some drawbacks which, in some cases, can be a good compromise between cost and accuracy. During my work, I used a precise voltage pulse generator to be able to finely study the voltage pulse characteristics effects on BBI, but because other alternatives exist, I introduce one of them in this section, alongside the fast and precise generator available at our laboratory.

2.2.2.1 Low-cost generator 2023-09-07 18:09:30+02:00

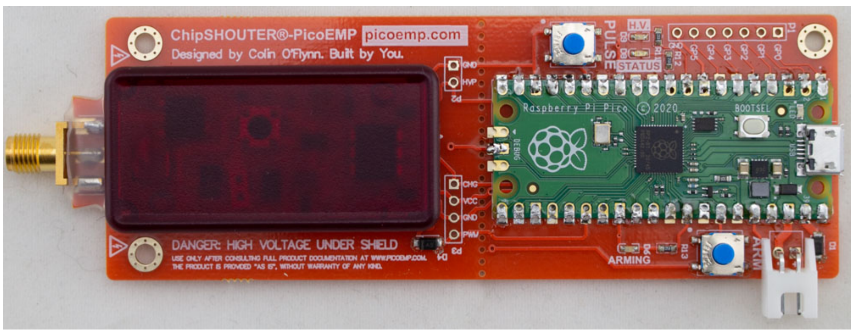


Figure 2.2: *ChipSHOUTER®-PicoEMP from NewAE Technology Inc.: a low-cost alternative to fast high voltage pulse generator typically used in fault injection platforms. It provides much less instantaneous power than a typical generator (around 99 % less power), has a long recovery time (from 1 to 4 seconds), has a lower maximum amplitude (250 V), is not calibrated, and has no controllable pulse width. However, it costs 94 % less than a typical generator, giving it a considerable advantage when building low-cost BBI platforms.*

The low-cost generator presented in this section is the NewAE Technology Inc. ChipSHOUTER®-PicoEMP, illustrated in Fig. 2.2. Despite being originally designed for EMFI, it is suitable as well for BBI. The bill-of-materials for this tool is roughly equal to 15 $, excluding the manual labor, which is less than our custom probe. It also has the advantage of being open-source, making it a future-proof community maintainable solution. Its main characteristics and drawbacks are the following:

- The output transformer is low-power, around up to 200 mW
- It uses a transformer, therefore no DC voltage option is available at its output
- The recovery time is slow, measured between 1 to 4 seconds depending on operating conditions
- The maximum voltage pulse is of approximately 250 V
- There is no pre-calibration
- It does not allow pulse width control by default. However, it is possible through drive signal control, even though being less accurate.

2.2.2.2 Fast, high voltage generator 2023-09-07 18:09:30+02:00

Figure 2.3: *Front side of the Avtech Electrosystems Ltd. AVRK-4-B High Voltage Pulser, used during all my thesis experiments.*

For my thesis, the generator used in all experiments is from the model AVRK-4-B from Avtech Electrosystems Ltd. It is illustrated in Fig. 2.3. This generator is high-speed and high-voltage. Similar to the low-cost generator described previously, it is commonly used for EMFI, but is also suitable for BBI. Its main specifications are the following:

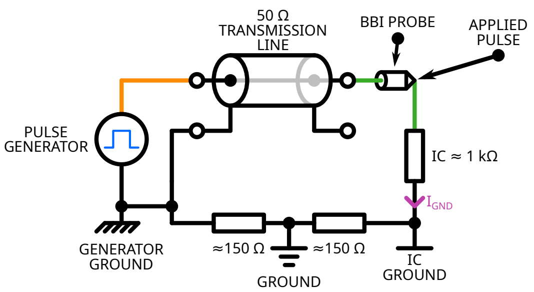
- The voltage pulse amplitude is specified between 150 V and 750 V with positive and negative polarities. The generator can go below and above these thresholds, there is no guarantee of the set point value correctness.
- The pulse width is specified between 6 ns and 20 ns. Similar to the voltage, the generator can go down from 4.5 ns, up to 22 ns, but is not specified out of the default range.
- Rise time (resp. fall time) for positive (resp. negative) pulses is specified to be precisely of 4 ns. Fall time(resp. rise time) for positive (resp. negative) pulses is not specified and depends on the generator load characteristics.

- The recovery time is inferior to 1 ms, allowing a pulse repetition frequency up to 1 kHz.
- The propagation delay measures 150 ns.
- The output is DC-coupled, allowing the generator to continue providing energy to the load (if resistive or inductive) during the pulse's plateau.
- All the specifications presuppose that the generator is loaded with precisely 50 Ω.

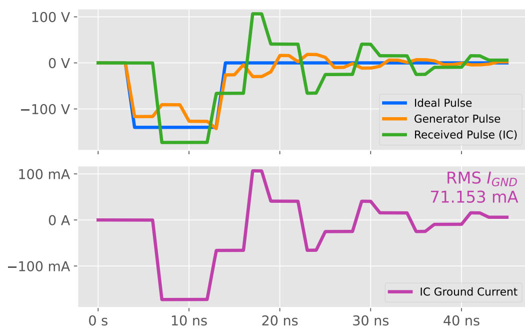
2.3 BBI in practice 2023-09-07 18:09:30+02:00

In this section, I am going to describe coarse electrical BBI platform models. They will allow us to concisely understand, evaluate and simulate BBI platforms behavior, limitations and room for improvement. We will then be able to analyze the platform's performance and point out its weaknesses.

2.3.1 BBI platform model 2023-09-07 18:09:30+02:00



(a) *Simplified electrical model of a BBI platform in the state-of-the-art. The model represents the major piece of equipment involved in a BBI platform, with the voltage pulse generator, the BBI metal probe, the transmission line used to interconnect the generator and the probe, the IC, modeled with a 1 kΩ resistor, and the platform grounding impedance, represented with two 150 Ω resistors. This model allows for fast evaluation and prediction of the BBI platform macro-behavior.*



(b) *Simulation results of the simplified model. The waveform colors match the schematic colors' for better clarity. In blue is shown the ideal voltage pulse one might want to generate (voltage set point of -140 V, pulse width set point of 10 ns). In green is the effective signal received by the IC on its backside. In orange is the pulse observed at the generator output, and in purple is the IC ground current measurement. These results highlight the platform limitations, such as the obvious ringing, causing incorrect voltage amplitude and pulse width set point.*

Figure 2.4: *BBI platform electrical model developed for my thesis to quickly evaluate various platform's parameters, alongside the model simulation results.*

To be able to quickly predict and analyze BBI platforms, I developed a very simple electrical model, illustrated in Fig. 2.4a. This model represents the key components of a BBI platform, which are:

- The voltage pulse generator

- The transmission line, used to connect the probe to the generator
- The BBI probe
- The targeted IC, modeled by an electrical resistance
- The grounding installation, consisting of electrical resistances connected between equipment grounding

2.3.2 Platforms evaluation criteria 2023-09-07 18:09:30+02:00

For the purpose of evaluating BBI platforms, we decided to focus on two important criteria, allowing to represent the platform quality:

- The voltage pulse measured at the generator's output characteristics, allowing us to observe how the generator behave when loaded with the transmission line and the IC
- The target ground current waveform characteristics, allowing to monitor exactly what is actually injected into the IC

2.3.3 Raw results 2023-09-07 18:09:30+02:00

To that end, we will deeply analyze the platform's simulation results shown in Fig. 2.4b. Four signals are displayed, with their colors matching the colors in Fig. 2.4a for greater clarity. There are three voltage waveforms and a current waveform. The blue waveform is the ideal voltage pulse an attacker want to apply to the backside of an IC during a body biasing injection. Its characteristics are the following: a voltage set point of -140 V and a pulse width set point of 10 ns. It is a steep, fast and precise pulse with controlled rise and fall times, pulse width and voltage. However, when performing real experiments, which the model allow us to evaluate, this ideal pulse falls apart. It can be seen thanks to the orange and green waveforms, representing respectively the pulse observer at the generator's output and the pulse effectively applied onto the backside of the IC target. There are multiple obvious observations that can be made concerning the received pulse (green) signal:

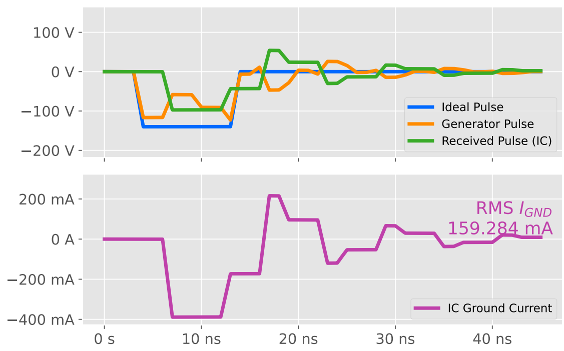
- The voltage set point is not respected, with a 23.5 % undershoot and a 107 V overshoot
- There is obvious ringing, causing the pulse width to be longer than expected in addition to damped oscillations

These effects can also be observed on the IC ground current waveform (purple), as it is a mirror of the applied pulse due to the pure resistive nature of the IC in that model.

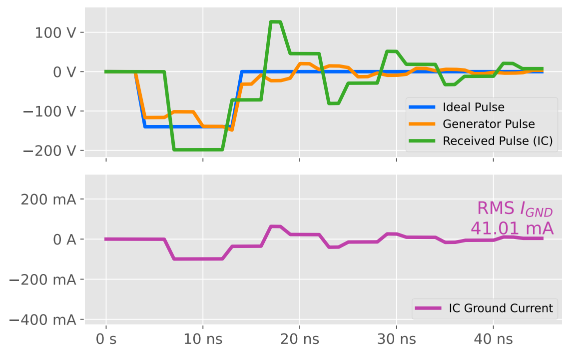
2.3.4 Analysis conclusions 2023-09-07 18:09:30+02:00

In order to understand the implications of such observations, let us analyze each one of them.

The first important thing to note is that the various model's numeric values are extracted from our actual platform. Therefore, the 150 Ω grounding and the 1 kΩ IC are average measured values of actual devices. Thus, these parameters, in addition to the transmission line characteristics, will inevitably vary with a certain amount from one platform to another.



(a) *Simulation result with an IC load equals to 250 Ω. The most obvious differences concern the received voltage pulse and the IC ground current. The voltage set point is not met, and the amplitude is -30 % lower than the requested value, with obvious ringing. However, due to the lower IC load value, there is more energy injected into the IC. This is mainly caused by the voltage divider formed by the generator output stage and the IC load.*



(b) *Simulation result with an IC load equals to 2 kΩ. The most obvious differences concern the received voltage pulse and the IC ground current. The voltage set point is not met once again, and the amplitude is 40 % higher than the requested value, with obvious ringing on all waveforms. Due to the higher IC load value, there is less energy injected into the IC. This is mainly caused by the voltage divider formed by the generator output stage and the IC load.*

Figure 2.5: *Platform simulation results with different IC load values.*

Indeed, the backside surface of an IC does not equal to a constant load. Therefore, depending on the probe location, the generator might not see the exact same load. To illustrate the induced effects of such differences, I changed the IC load value when running the simulations, and the results are shown in Fig. 2.5, both for a 250 Ω load (Fig. 2.5a), and a 2 kΩ load (Fig. 2.5b). In both cases, due to the non-zero generator output impedance, the latter forms a voltage divider with the IC load. On the one hand, with an IC load value one quarter lower, there is more current in it, while the applied pulse amplitude is 30% lower. On the other hand, with an IC load value two times higher, there is less current in it, while the applied pulse amplitude is 40% higher. It represents a 70 % range around the set point value, which is excessively high. However, in both cases, the ringing is still present with the same amount relative to the pulse amplitude.

Eventually, all of these observations allow us to spot three major flaws of such platform:

- The platform parameters are difficult to control, leading to unknown values concerning pulse width, voltage set point, etc.

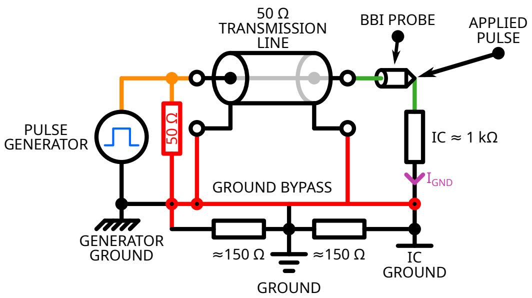
- It leads to a poor temporal accuracy, thus minimizing the chances to perform a good quality fault injection.
- At last, all parameters are platform dependent, leading to a low reproducibility rate, thus lowering the credibility of experiments performed on such platforms.

In this context, I present in the next section various simple improvements to the BBI state-of-the-art platform.

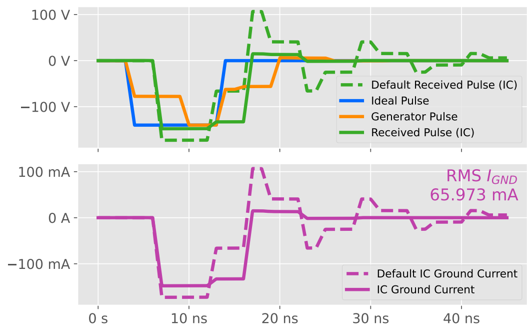
2.4 Enhanced BBI platform model and simulation 2023-09-07 18:09:30+02:00

In this section, I propose platform enhancements over the state-of-the-art BBI platform previously introduced. These improvements aim at being low-cost, fast and easy to set up, to represent an interesting addition without drastically increasing the platform's financial cost. We will then be able to draw conclusions on such improvements thanks to simulation results.

2.4.1 Matching the generator output impedance 2023-09-07 18:09:30+02:00



(a) *Enhanced simplified electrical model of a BBI platform. The model represents the major piece of equipment involved in a BBI platform, with improvements over state-of-the-art ones. There are the voltage pulse generator, the BBI probe, the transmission line, the IC (1 kω resistor), the original platform grounding (both 150 Ω resistors), and the proposed enhancements highlighted in red. The enhancements consist in first, creating an approximate impedance matching for the generator, second, to bypass the poor grounding with low impedance copper wires.*



(b) *Simulation results of the simplified enhanced BBI platform model. The waveforms colors match the schematic. In blue is the ideal voltage pulse (voltage set point of -150 V, pulse width of 10 ns). In green is the effective pulse received by the IC. In orange is the pulse observed at the generator output, before the transmission line. Eventually, in purple is the IC ground current. The dotted waveforms are the waveforms observed in Fig. 2.4b for the state-of-the-art platform, for comparison purposes. The most obvious observed improvements concern the set points, which are fully respected, in addition to the drastic ringing reduction, leading to better temporal control.*

Figure 2.6: *BBI platform enhanced electrical model developed for my thesis to quickly evaluate various platform's parameters, alongside the model simulation results.*

The first proposed improvement concerns the generated voltage pulse characteristics. As we observed previously, the various parameters set points were not met. In a fault injection context, it is an undesirable behavior, as it is required to finely control the generated pulse to produce controlled disturbances into ICs. Therefore, and because most high speed high voltage pulse generator are specified to be loaded with a precise impedance, I simply propose to connect a known load directly at the output of the generator model. In my model's case, a 50 Ω resistor is loaded at the generator's output, as illustrated in red in Fig. 2.6a. Thus, the generator will see the impedance network formed by the compensation load, the IC, the transmission line, and the grounding installation. However, because the grounding installation is platform dependent, it is required, in order to perform a better impedance matching of the generator output, to improve the grounding, which leads us to the following section.

2.4.2 Improving the grounding installation 2023-09-07 18:09:30+02:00

In many platforms, the grounding installation might be perfectly fine, and the following section may not apply to them. However, with our platform, we quickly observed that the grounding impedance was far from negligible. Indeed, with an average IC impedance around 1 kΩ, and inter-equipment ground impedance around 150 Ω, it represents a 15 % increase in the total impedance seen by the generator. Therefore, in order to transfer the maximum amount of energy into the IC, especially in areas where the IC impedance might be closer to the grounding impedance, it is required to cancel as much as possible the latter.

To that end, I propose a very simple setup modification. It consists in keeping the platform as is, and adding short copper wires between equipment grounds. Therefore, it shunts the platform ground and creates a low-impedance path for electric charges, thus allowing the previous section approximate impedance matching to perform better.

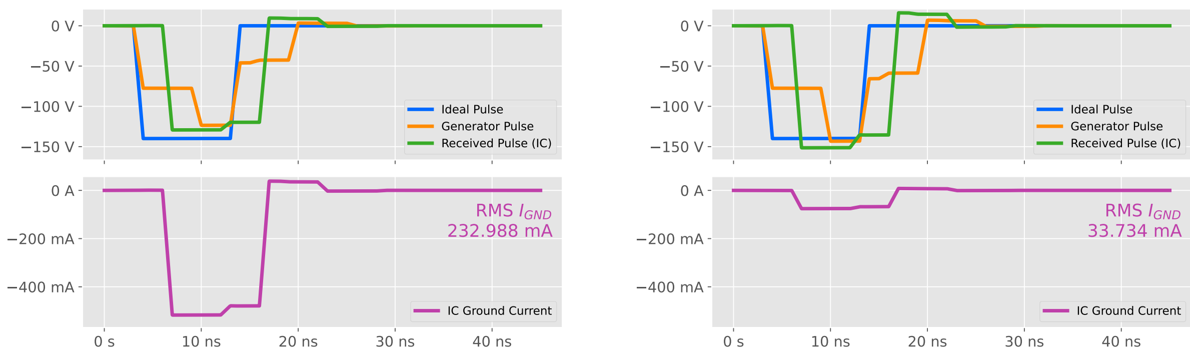
2.4.3 Simulation results 2023-09-07 18:09:30+02:00

To verify the soundness of the previously proposed enhancements, I performed simulations thanks to the model presented in Fig. 2.6a, and the simulation results are shown in Fig. 2.6b.

In that case, unlike in the state-of-the-art platform, the voltage set point is almost met concerning the received pulse (green waveform), with a slight undershoot of 6%. It is mirrored on the IC ground current waveform, where the ringing is drastically reduced, which leads to a steeper and more accurate pulse. It is especially noticeable when directly comparing the state-of-the-art waveforms in dotted lines. Concerning the generator pulse (orange waveform), it is still distorted as the ringing has not disappeared, but is less of a concern since the waveform of interest is the one effectively applied to the IC backside.

2.4.3.1 Load dependency 2023-09-07 18:09:30+02:00

To further analyze the benefits of the proposed improvements, I performed, as for the state-of-the-art platform, additional simulations with various loads. As before, 250 Ω and 2 kΩ were chosen to have a common point of comparison.



(a) **(b)**

Figure 2.7: *Simulation results of the enhanced platform with a 250 Ω IC load 2.7a and a 2 kΩ IC load 2.7b. The most obvious thing is the current increase in 2.7a, which is natural due to the load impedance reduction. However, the effective pulse amplitude relative to the set point has only a -7 % error, which is a drastic improvement over the previous -30 %. Therefore, we can say that the set point is met. Then, in 2.7b, there is an obvious current decrease, which once again, is logical given the higher impedance value. The effective pulse amplitude relative to the set point has only a 7 % error, which is a drastic improvement over the previous 40 %. In addition to this, the ringing almost disappeared in every case. It is caused by to the fact that the generator is not only loaded by the IC, but by the equivalent load composed of the IC and the compensation load, which reduces the effective load variation when changing the IC load value.*

Fig. 2.7 presents the simulation results for such loads. For both loads, the measured voltage moves away from the set point by around 7 % in each case. It represents a 14 % range around the -140 V set point, which is immensely better than in the previous platform. It is still not perfect, but the platform is overall less dependent to the IC load, which is desirable in order to have repeatable voltage pulse across the entire IC backside.

Then, quite naturally, for the 250 Ω load, the current is higher than for the 1 kΩ, and with the 2 kΩ, the current is lower. In addition to that, thanks to the 50 Ω resistor placed at the generator's output, it reduces the range in which the effective load (the compensation load in parallel with the IC) changes. Indeed, it goes from around 42 Ω to about 49 Ω, instead of going from 250 Ω to 2 kΩ in the previous case. Eventually, in addition to all of the above, these enhancements have also drastically reduced ringing, which contributes to the applied pulse amplitude being closer to the set point.

2.4.4 Simulation conclusions 2023-09-07 18:09:30+02:00

All of this leads to better control over the various platform parameters, allowing for more accurate and shorter pulses, closer to the expectations. In addition to that, the platform is less design dependent thanks to the minimization of impedance mismatch and poor grounding installation. It leads to a better time accuracy, leading to a potentially more controllable fault injection.

2.5 Actual enhanced BBI platform 2023-09-07 18:09:30+02:00

The previous models being a useful tool to draw quick conclusions and predictions, it does not represent the reality. To that end, I set up the various presented enhancements in an actual BBI platform in order to verify the soundness of all the outcomes. In the first place, we are going to discuss how to perform the approximate impedance matching. Then, I will explain how to set up an efficient grounding bypass. After that, we will take a look at actual measurements allowing to spot the improvements.

2.5.1 Generator impedance matching in practice 2023-09-07 18:09:30+02:00

Add pictures of real platform impedance matching. An ideal impedance matching implementation should be adaptive and vary the impedance seen by the generator to perfectly match 50 Ω in every case. It would require a system with feedback, capable of measuring in real time the impedance presented by the IC target, in addition to the transmission line characteristics, to be able to adapt the compensation load impedance value. However, this is not the approach I have chosen. Indeed, the goal here is to minimize financial cost and platform modification, while allowing for better control over the platform's parameters.

Another possibility would be to first measure the average IC backside impedance over its entire area or only the targeted area (such as the cryptographic core for instance). Then, thanks to the average value, the compensation load impedance could be chosen to approximately match the required 50 Ω.

Eventually, the selected solution is the simplest one. It consists in connecting a compensation load at the generator's output, consisting in a 50 Ω SMA terminator. It is far from ideal, as this solution does not consider the transmission line nor the IC effective load nor the capacitive and/or inductive nature of the IC in addition to its resistive nature. However, it is a solution requiring little to no changes to an existing platform and has proven to be good enough thanks to the previous models.

2.5.2 Grounding installation bypass in practice 2023-09-07 18:09:30+02:00

Add pictures of real GND bypass. As we discussed previously, the grounding installation can drastically vary from one platform to another. Its effective impedance can be very high, such as in our platform, where equipment is grounded thanks to the platform's earthing, with inter-equipment ground of around 150 Ω. To alleviate the effects of such ground impedance, I simply decided to shunt the existing earthing with short low-resistance copper wires. To that end, I chose an arbitrary piece of equipment as the reference, and connected every other piece of equipment local ground to the reference. It allowed reducing the effective platform ground impedance to a negligible value close to 0 Ω.

2.5.3 Practical analysis 2023-09-07 18:09:30+02:00

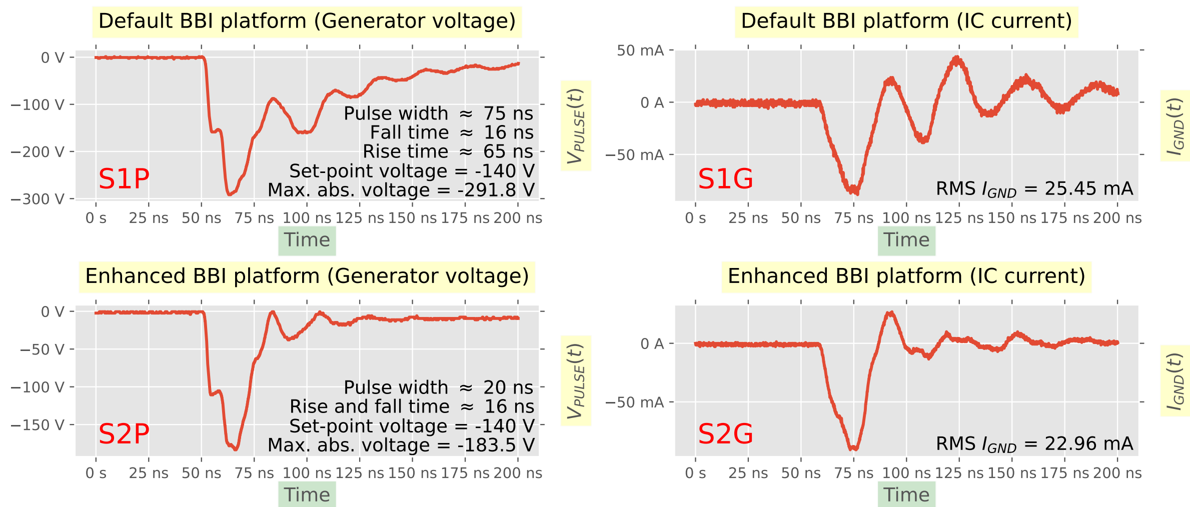


Figure 2.8: *BBI platforms comparison: state-of-the-art (S1P and S1G) versus the proposed enhanced platform (S2P and S2G). These waveforms code names are used to quickly indicate which signal from which platform is talked about. "S" stands for "setup", the number indicates whether it is the default (1) or the enhanced (2) platform, and the last letter indicates whether it is the voltage pulse (P) or the IC ground current (G) waveform. Time scales are identical on each waveform, and voltage/current scales are waveform dependent. These measurements illustrate in practice the benefits of the proposed enhancements. The ideal voltage pulse has a negative amplitude of 140 V and a pulse width of 20 ns, with rise and fall times of 4 ns. On S1P there is a -108 % voltage undershoot, which is too big not to be concerning. In addition to this, the pulse width is 275 % too high with a 75 ns value. The fall time is 4 times higher than requested, and the rise time is more than 15 times higher. S1G highlight specifically the ringing, which could already be seen on S1P, with energy going back and forth multiple times during the pulse. On S2P, the voltage set point error goes from -108 % to only -31 %,. It is still not negligible but is far better than on S1P. The pulse width now perfectly matches the set point value of 20 ns. However, rise and fall times are 4 times higher than they should be, despite both being consistent. S2G helps us spot the significant ringing reduction, while maintaining the same amount of transferred energy into the IC.*

Now that I presented how to practically set up the enhancements, let us analyze actual measurements on the platform. We will compare before and after results, allowing us to analyze each evaluation criterion. As it was done for the simulations, we will observe the voltage pulse and the IC ground current. Fig. 2.8 presents the various waveform results. The voltage pulse was measured at the IC backside during the injection, and the IC ground current was measured using a current probe thanks to the IC PCB ground interconnection. Therefore, the measured current is precisely the IC ground current, excluding any other equipment. The four waveforms displayed in Fig. 2.8 are code named using three characters for clarity. The first character is common to all waveforms, denoted "S" for "setup". Then, the number indicates which platform is concerned, "1" being the default platform, "2" being the enhanced one. Eventually, the last letter indicates which waveform is observed, "P" being the voltage pulse, "G" being the IC ground current. Therefore, the default platform contains S1P and S1G waveforms, while the enhanced one contains S2P and S2G signals. Fig. 2.8 also display the waveforms characteristics for more clarity. The ideal voltage pulse applied has a maximum negative amplitude of 140 V, a pulse width of 20 ns, and 4 ns rise and fall times.

S1P shows a clear undershoot of -108 % under the set point. It is a clearly non-negligible value, which is far from desirable when performing fault injection, as most of the time , the voltage value has a great importance concerning efficiency and repeatability. In addition to this, the pulse width is 275 % higher than its set point. It is an additional undesirable behavior, especially when one wants to inject precise disturbances into an IC under test. Then, fall time is four times higher than requested, and rise time is more than fifteen times higher. Put with the longer pulse width, it worsens the pulse accuracy.

S1G brings to light the obvious ringing issue, also observable to a lesser extent on S1P, which leads to longer than expected disturbance inside the IC. Considering that the ringing is mainly caused by impedance mismatch between the generator and the IC, it will drastically change from one location to another, further reducing repeatability.

S2P, on the other hand, shows a better voltage amplitude, with a -31 % undershoot. It is far from perfect, but given the approximate nature of the proposed impedance matching, it is expected. Concerning the pulse width, the set point value is perfectly respected, which is very important for precise disturbance duration. However, rise and fall times are now consistent, but still four times higher than requested. It can easily be explained by the fact that the transmission line, the probe, the IC and the power installation are not a purely resistive load. Therefore, any capacitive element in the chain will inevitably reduce the system's response time, thus elongating rise and fall times, leading to a shorter pulse plateau.

Concerning S2G, the approximate impedance matching shows a clear ringing reduction, with a steep current pulse, leading to a precise disturbance.

2.6 Enhanced BBI platform in a fault attack context 2023-09-07 18:09:30+02:00

Now that we have seen with simple actual experiments the benefits of the proposed enhanced BBI platform, let us linger on further experiments to verify more thoroughly the soundness of these enhancements. To that end, I performed a differential fault attack on our IC target. More specifically, a constraining fault attack requiring single bit faults on one or more bytes working on an AES cryptographic core, introduced by C. Giraud [15] in 2004.

2.7 Conclusion 2023-09-07 18:09:30+02:00

~~In this chapter, we first discussed the hardware and software commonly used for the practice of BBI, as well as presenting our hardware and software. We then proposed an enhanced platform for the practice of BBI, consisting of reducing the ground impedance and approximating the voltage pulse generator impedance matching. We first studied these enhancements using a coarse platform. Furthermore, we then performed analog experiments using real hardware. In order to further verify the soundness of the BBI platform improvements, we set up and conducted a Giraud's differential fault attack. We observed that it would be impossible to conduct the attack without the aforementioned BBI platform enhancements, thus confirming their usefulness.~~

III

Integrated circuits modeling 2023-09-07 18:09:30+02:00

Contents

3.1 Summary 2023-09-07 18:09:30+02:00

This chapter presents the work carried out concerning the modeling and simulation of integrated circuits and platforms in a body biasing fault injection context. The presented work focused on elaborating electrical models allowing to evaluate with simulations the behaviors of ICs subjected to BBI. The chapter introduces the elaborated models and the algorithms used to create them, and then goes on to present various validation steps to check the meaningfulness of the models. Parts of this work have been published both in [1] and [24].

3.2 Introduction 2023-09-07 18:09:30+02:00

When evaluating and studying ICs under BBI, it is important to be able to fully predict and understand the underlying mechanisms at work in order to set up reproducible and reliable experiments, as well as being able to set up efficient countermeasures. However, to model and simulate integrated circuit behavior subject to fault injection is not an easy task. Specifically, simulating an entire IC at a transistor level under fault injection is unrealistic with current resources and technology. It is especially true when considering time cost, as current digital ICs are composed of about a million of transistors for standard microcontrollers. Furthermore, no software nor algorithm is currently dedicated to simulate the functional, electrical behavior of millions of transistors at the same time while some of them are disrupted by strong and transient disturbances. In addition to that, to be able to set up a reliable model, one should have access to the detailed architecture of each considered IC, which is almost never the case, as most studied architectures are proprietary. Therefore, it is required to find alternative workarounds in order to be able to study IC behavior and their various responses to fault injection techniques.

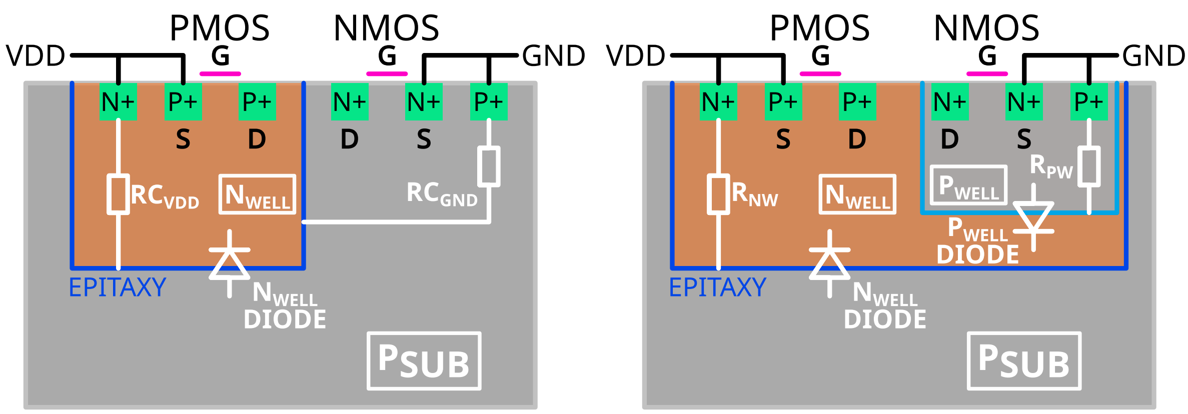
This has been first proposed in 2019 concerning Electromagnetic Fault Injection (EMFI) [25], and further extended in 2021 [19]. Especially in the latest work [19], the proposed solution consisted in establishing an equivalent non-logical model of the section of an IC. Instead of modeling each logic gate with as many transistors as required, in addition to the power delivery network and the silicon substrate, it was chosen to represent a hundred of logic gates in an average way, solely with a few resistors and capacitors. This results in a transistor-less model, achieved using manufacturing data for the studied IC. The authors assumed that the first half of the transistors are conducting while the other half are blocking. Then, two levels of power delivery network were added, simply modeled with electrical resistances. Eventually, and because the modeled IC was manufactured using a dual-well substrate type, the silicon substrate and the P-N junction respectively are modeled by six resistors going in every direction in addition to a diode and its capacitance respectively. This clever design allows to drastically reduce the computing work required to analyze and predict behaviors of ICs subject to EMFI. Indeed, simulating the average behavior of a hundred of logic gates only with four resistors and four capacitors is immensely lighter than simulating the equivalent with BSIM (Berkeley

Short-channel IGFET Model) transistors. However, the main shortcoming being the lack of functionality with the produced ICs, it is therefore impossible to evaluate their functional or logical behavior.

Body biasing injection being less documented than EMFI, no distributed model has yet been proposed to simulate ICs under BBI. In this context, our motivations were to set up and evaluate electrical models being able to reliably predict both in time and space IC behavior in order to understand how BBI induced disturbances propagate and create faults inside ICs. The current work main goal being to model and simulate BBI similarly to EMFI, we decided to start from the model proposed in [19], to improve and adapt it in order to be able to implement it in a BBI context.

This chapter begins with a general presentation of the enhanced models, followed by a closer look at each model and its specific features. Eventually, various model validation are studied in order to verify their soundness.

3.3 Electrical models 2023-09-07 18:09:30+02:00



(a) *Dual-Well* **(b)** *Triple-Well*

Figure 3.1: *Dual-well and triple-well inverter silicon sectional view*

On one hand, when performing EMFI (usually on the front side of the IC), air is the physical support to convey energy through electromagnetic waves. It is achieved by coupling the loop wire probe to the power delivery network loops. On the other hand, when working with BBI, the context is different. Indeed, the energy is conveyed through electrical charges through the silicon substrate. Therefore, the carriers have to go through the metallic probe and the whole substrate to reach the logic gates and the power delivery network in order to disturb the IC operation. Thus, the substrate type and design could have a significant impact on BBI efficiency. As a result, we explored and studied BBI in two specific scenarios depending on the substrate types: dual-well and triple-well. Fig. 3.1 shows the sectional views of two inverters manufactured in a dual-well and a triple-well substrate respectively. These simple schematics are helpful in understanding the reasoning behind the design of the electrical models.

Fig. 3.1a depicts the cross-sectional view of a dual-well CMOS inverter. The P-doped silicon substrate is colored in gray, with RC_{GND} being the access resistance from the epitaxy layer to the NMOS bulk. This physical environment is the conducting support of electrical charges which flow up to the NMOS transistor. The orange region is the N-doped silicon well, located inside the P-substrate to manufacture the PMOS transistors. RC_{VDD} is the access resistance from the epitaxy to the PMOS bulk inside the N_{WELL}. In addition to the P-substrate, the N-well is the last environment electrical charges have to go through before reaching the PMOS transistor.

Fig. 3.1b shows the cross-sectional view of a triple-well CMOS inverter. As before, gray areas represent P-doped silicon, and orange areas N-doped silicon. R_{NW} is the N_{WELL} access resistance from the epitaxy to the PMOS bulk, and R_{PW} is the P_{WELL} access resistance from the $N_{WELL} - P_{WELL}$ junction to the NMOS bulk. In this case, two silicon junctions are present, represented by two independent diodes. In order to reach the PMOS transistors, charges have to go through the exact same environments as before. However, concerning NMOS transistors, they have to pass through two silicon junctions instead of none. As discussed in Chapter 5, this has a significant impact on BBI induced effects. However, these schematics are incomplete and do not allow simulating ICs behaviors under BBI.

Therefore, as it has been done in [19], ICs are spatially split in elementary sections called standard-cells segments (SCS). However, in addition to the improvement of the dual-well proposed model proposed in [19], we also introduce a triple-well model in order to fully appreciate the behavioral differences of BBI applied to both substrate types.

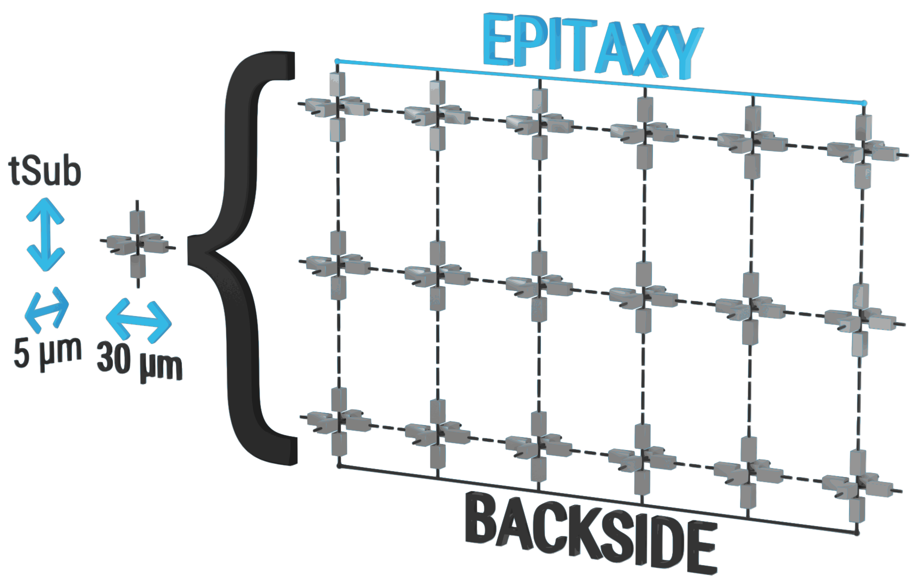
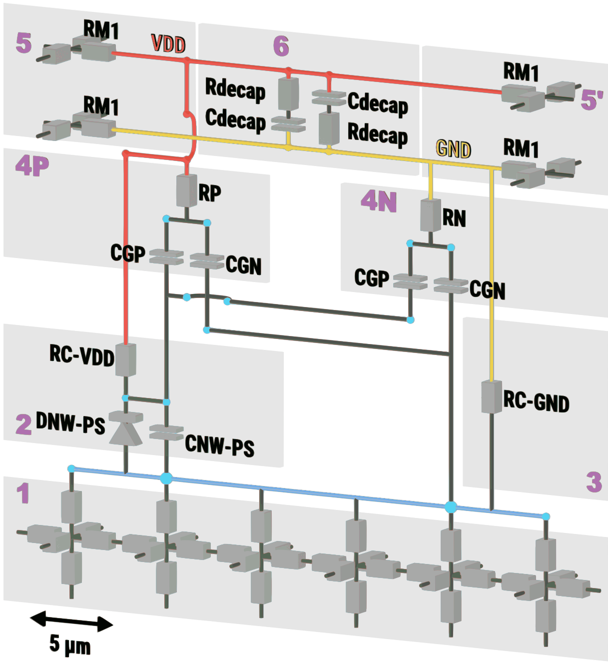
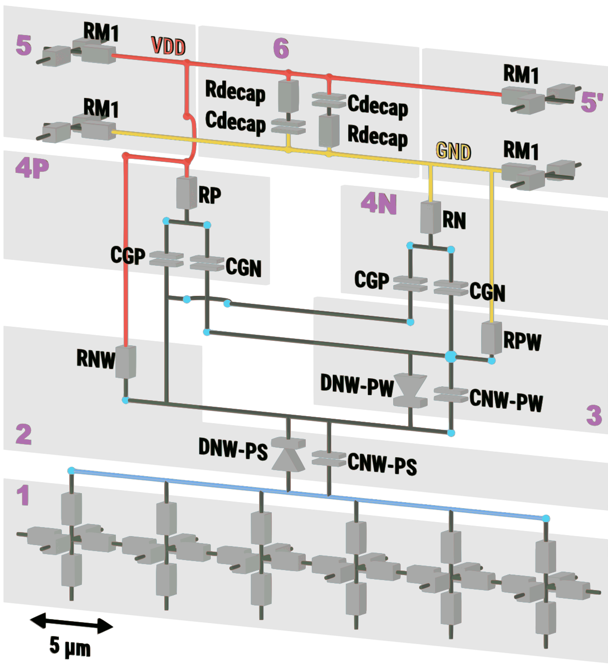


Figure 3.2: *Surface subdivision improvement.*

The main improvement over the dual-well model proposed in [19] concerns the substrate resistive network, as shown in Fig. 3.2. In [19], the substrate network is coarse and only consists of six electrical resistances for each SCS. It means that they represent the entire SCS substrate thickness, width, and height (on the left in Fig 3.2). Even though it is sufficient to appreciate



(a) *Dual-Well*



(b) *Triple-Well*

Figure 3.3: *Three-dimensional Dual-Well and Triple-Well IC comprehensive standard-cell electrical schematic.*

the injection method effects while studying EMFI, mainly because the substrate is almost transparent when it comes to electromagnetic waves, but also because EMFI is mostly performed at the IC front side, it is not precise enough to model the spreading of he voltage pulse from the IC backside to the transistors.

To that end, we decided to split as much as possible these resistors, as shown in Fig. 3.2, to provide a precise enough substrate sub-model while keeping realistic computational workload. For the final models, it was decided to use an editable elementary thickness of $10\ \mu m$, and fixed width and depth of $5\ \mu m$ for each elementary six-resistors substrate models, according to the footprint of an SCS on the XY plane ($5\ \mu m\ \times\ (6\ \mu m\ \times\ 5\ \mu m)$), resulting in a $30\ \mu m$ wide and $5\ \mu m$ deep SCS. One can remark that in Fig 3.3, no number is given concerning the substrate thickness, as similar to LFI, it is an important parameter which does not have a fixed value. Indeed, an attacker may want to thin the substrate or not before performing BBI.

Furthermore, as shown in Fig. 3.3, both SCS models contain various electrical components describing the IC structure, roughly composed of:

- Its substrate
- Its silicon junction(s)
- Its logic gates
- Its power supply rails

These two models, while being close to each other, allow, thanks to their subtle differences, to properly consider the different behaviors each substrate type exhibits. In the next section, dual-well SCS model and triple-well SCS model are consecutively considered and analyzed.

3.3.1 Standard-cell segment models 2023-09-07 18:09:30+02:00

Historically, IC substrate was manufactured using an exclusive dual-well structure. However, nowadays, it is common to find on relatively modern ICs a mix of dual-well and triple-well structures on a monolithic die. Triple-well substrate structures bring significant advantages over dual-well substrates. In digital ICs, it is mainly used to body bias transistors to optimize their performance under power constraints. When used in analog or mixed designs, it gives two main advantages: substrate cross-talk and noise reduction, in addition to power supply decoupling thanks to the additional P-N junction capacitance [26]. This is why we decided to cover dual-well and triple-well structures in our models.

Fig. 3.3a depicts an SCS dual-well model. Each significant section of the SCS is gray-framed and numbered:

- The section $\boxed{1}$ represents the substrate environment: resistive and isotropic.
- The section $\boxed{2}$ is the $P - N$ silicon junction between the P-substrate and the N-well, represented by a diode and its junction capacitance, in addition to an access resistance $RC - VDD$, being the N-well electrical resistance.

- The section $\boxed{3}$ is the substrate access resistance.
- The sections (4P) and (4N) contain the average non-logical model of a hundred of logic gates.
- The sections $\boxed{5}$ and $\boxed{5'}$ are the two levels of the power delivery network, which are low resistive metals.
- The section $\boxed{6}$ is the decoupling between both GND and V_{DD} power networks.

Fig. 3.3b depicts the SCS triple-well model as follows:

- The section $\boxed{2}$ is the $P-N$ silicon junction between the P-substrate and the N-well, represented by a diode and its junction capacitance, in addition to an access resistance R_{NW}, being the N-well electrical resistance.
- The section $\boxed{3}$ is the $N-P$ silicon junction between the N-well and the P-well, represented once again by a diode and its junction capacitance, in addition to an access resistance R_{PW}, being the P-well electrical resistance.
- The sections $\boxed{1}$, (4P), (4N), $\boxed{5'}$ and $\boxed{6}$ being the same as before.

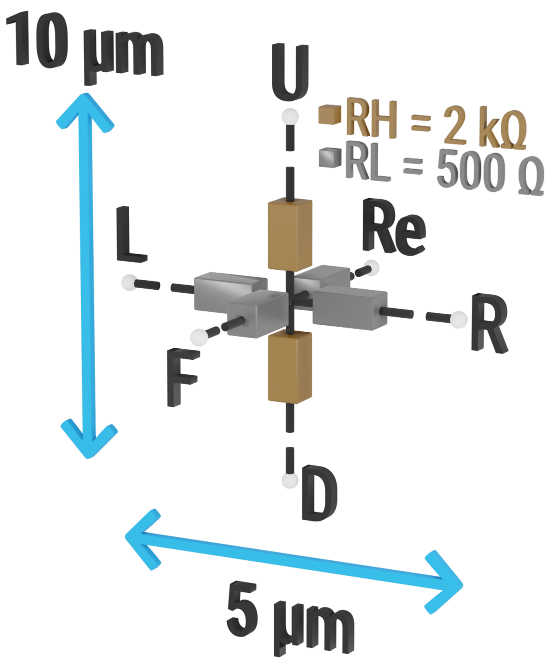


Figure 3.4: *Elementary substrate 3D netlist*

Each area of the elementary SCS models were automatically generated using a custom algorithm, shown in Alg. 1. It was mainly designed in order to reduce as much as possible any human intervention to limit difficult to debug errors and inconsistencies. Furthermore, it provides a degree of flexibility due to the ease of user modifications directly into the generation

```
.subckt elementary_bloc D F L R Re U
R1 U N001 RH
R2 N001 D RH
R3 Re N001 RL
R4 N001 F RL
R5 N001 L RL
R6 R N001 RL
.ends elementary_bloc
```

Figure 3.5: *Elementary substrate SPICE netlist*

```
.subckt elementary_blocx6 D1 D2 D3 D4 D5 D6
+F1 F2 F3 F4 F5 F6 L R RE1 RE2 RE3 RE4 RE5 RE6
+U1 U2 U3 U4 U5 U6 VSUBCintC
XX1 D1 F1 L VSUBCintL2 RE1 U1 elementary_bloc
XX2 D2 F2 VSUBCintL2 VSUBCintL1 RE2 U2 elementary_bloc
XX3 D3 F3 VSUBCintL1 VSUBCintC RE3 U3 elementary_bloc
XX4 D4 F4 VSUBCintC VSUBCintR1 RE4 U4 elementary_bloc
XX5 D5 F5 VSUBCintR1 VSUBCintR2 RE5 U5 elementary_bloc
XX6 D6 F6 VSUBCintR2 R RE6 U6 elementary_bloc
.ends elementary_blocx6
```

Figure 3.6: *SCS substrate layer SPICE netlist*

algorithm parameters, as opposed to netlist editing, thereby reducing errors further. These models only represent a section of an integrated circuit. For effective use and verification, it is necessary to replicate and interconnect these models spatially as much as possible. This was accomplished by utilizing customized Python scripts coupled with procedural generation. The IC generation algorithm enables the modification of multiple settings to produce the desired outcomes, albeit with certain inherent structural limitations. Two of the main limitations are the fixed width and depth of the elementary SCS models, and the fixed number of metal levels in the power delivery network. On the contrary, the following is a non-exhaustive list of user-modifiable settings:

- Global IC size.
- Probe position.
- IC global substrate thickness.
- IC elementary substrate thickness.
- Substrate type (dual-well, triple-well, or mixed).
- Voltage pulse amplitude.
- Voltage pulse width.

- Voltage pulse rise and fall times.
- Simulation time and step.

Eventually, the generator script incorporates a visual inspection tool in order to quickly verify the correctness of the generated netlist. Alg. 1 shows the IC generation algorithm main function, which is to create the coordinates for every net in the netlist.

3.4 Preliminary model validation 2023-09-07 18:09:30+02:00

Because validating such models is a complex task, we chose to trim validation into elementary steps. As these models aim at modeling and report back average IC behaviors, it is required to verify their soundness in trivial scenarios. Specifically, two class of measurements are going to be discussed in this section:

- Global quiescent leakage current evaluation
- Quiescent power network IR drop verification

These are important parameters to verify before going any further because any inconsistent or unrealistic value would result in meaningless models and simulations.

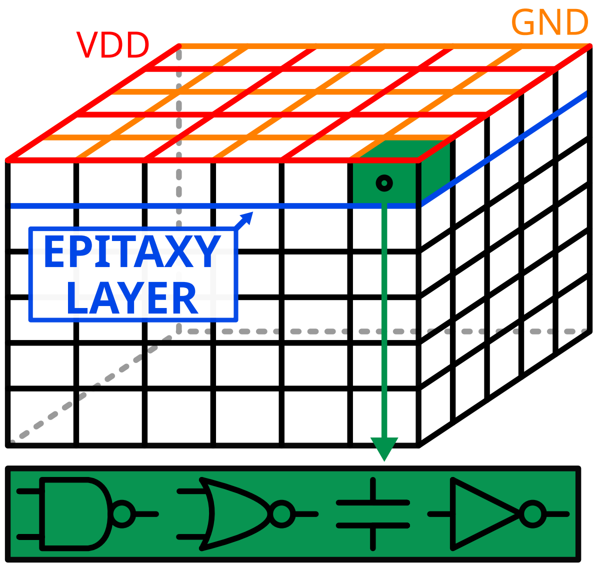


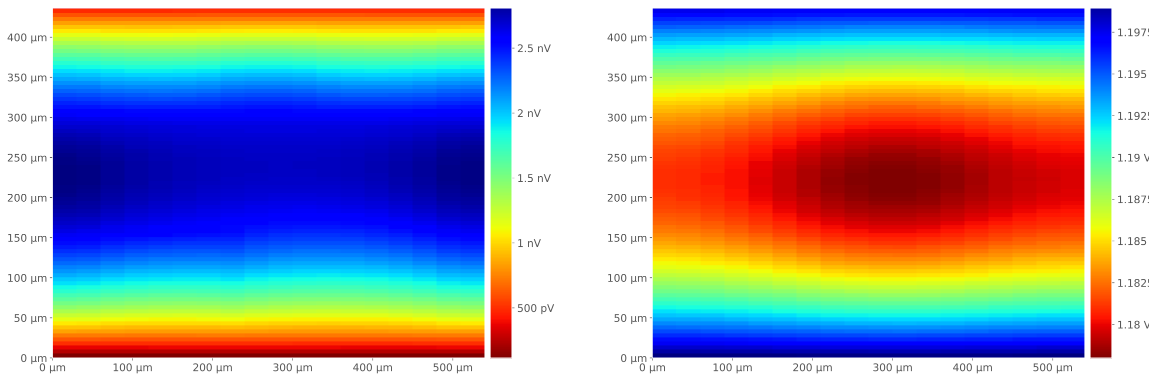
Figure 3.7: *Three-dimensional standard-cell segments interconnection example.*

To that end, we decided, as stated previously, to create an IC composed of several SCS. Fig. 3.7 depicts in a general way how the various SCS required are spatially connected to each other. In blue is indicated the epitaxy layer, which is the junction between the highest substrate level and the top of the SCS. All SCS share the power delivery network at their top and the silicon substrate at their bottom. As mentioned earlier, each SCS represent the average behavior of about a hundred of logic gates. The resulting IC measurements are the following: a width of $550\ \mu m$, a depth of $450\ \mu m$, and a thickness of $140\ \mu m$. First, we will present the global leakage current, then, we will analyze mappings of the simulated ICs power distribution networks. Dual-well, triple-well and mixed substrates models are analyzed, and most importantly, the simulated circuits do not include the voltage pulse generator nor any other external component required to work with BBI as what we present here is the first validation step. They are proposed as is, and Table 3.1 presents the operating point results for each substrate type.

Table 3.1: *Dual-well, triple-well and mixed substrates SCS operating point.*

Measurement	Description	Dual-well	Triple-well	Mixed substrates
I_{GND}	IC global ground current	1.92 nA	1.94 nA	3.4 nA
I_{VDD}	IC global VDD current	-1.96 nA	-5.8 nA	-3.5 nA
GND_{AVG}	Average GND voltage	1 nV	1 nV	1.75 nV
VDD_{AVG}	Average VDD Voltage	1.2 V	1.2 V	1.2 V

Looking at Table 3.1 indicates the absence of any significant leakage current and power supply voltage drop. However, to check the models relevance further and in a more reliable way, it is interesting to look at voltage mappings of the power delivery networks (VDD and GND), as shown in Fig. 3.8.



(a) *GND* **(b)** *VDD*

Figure 3.8: *Mixed substrates operating point.*

Concerning both GND and VDD operating point maps, there is no significant voltage drop across both maps, which indicates further the absence of significant leakage current in the simulated IC. With this in mind, we then introduced the generator into the model.

3.5 Voltage pulse generator model and further validation 2023-09-07 18:09:30+02:00

Introducing the generator did not come without major problems. Indeed, the latter inevitably interacts with the target IC, and depending on the real generator output stage architecture, this interaction can drastically vary from one to another.

For example, when using ESD guns as in [27, 28], their output stages are usually AC-coupled, while on our works, we mostly use DC-coupled generators. These subtle differences in practice become major issues in simulation when not treated correctly. Indeed, even considering the transmission line as it has been recommended in Chapter **??**, most DC-coupled high voltage generators use a high-impedance mode to disconnect the load from the generator before and after the generated pulses. Therefore, one has to consider this specific aspect when designing a proper BBI electrical model, as we will explain in this section.

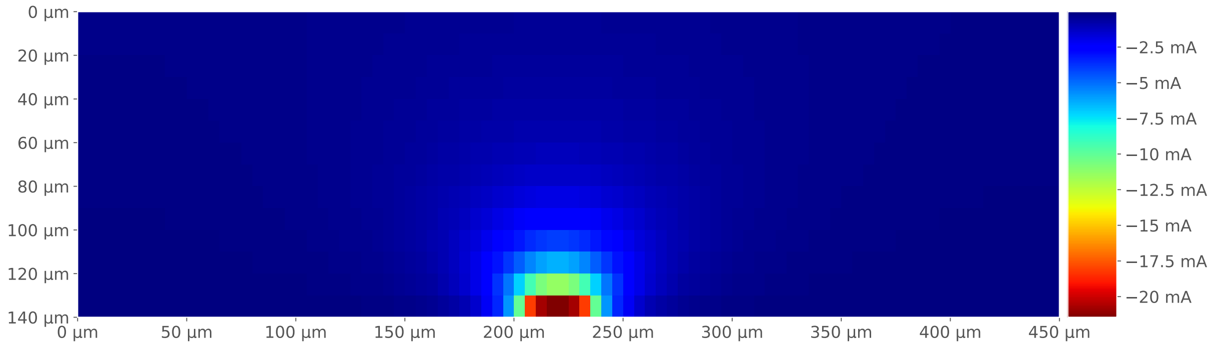
3.5.1 Early generator models 2023-09-07 18:09:30+02:00

The first models consisted in a PWL voltage source directly connected to the substrate of the IC, and we quickly observed abnormal operating point values. Je dois rajouter des valeurs chiffrées. Indeed, in this setup, at rest, the generator is equivalent a DC voltage source applying 0 V to the backside of the simulated IC. Therefore, it applies an undesired bias to the substrate and thus changes the operating point, inducing a high amount of charges flowing between power sources, thus disturbing the power delivery network. To circumvent this issue, we chose to mimic the behavior of an actual high voltage pulse generator and to switch between a high impedance mode and a voltage pulse mode as a function of the pulse time. This allowed to observe correct operating points with the generator connected, as it is the case in a real experiment. Je rajouterai les figures.

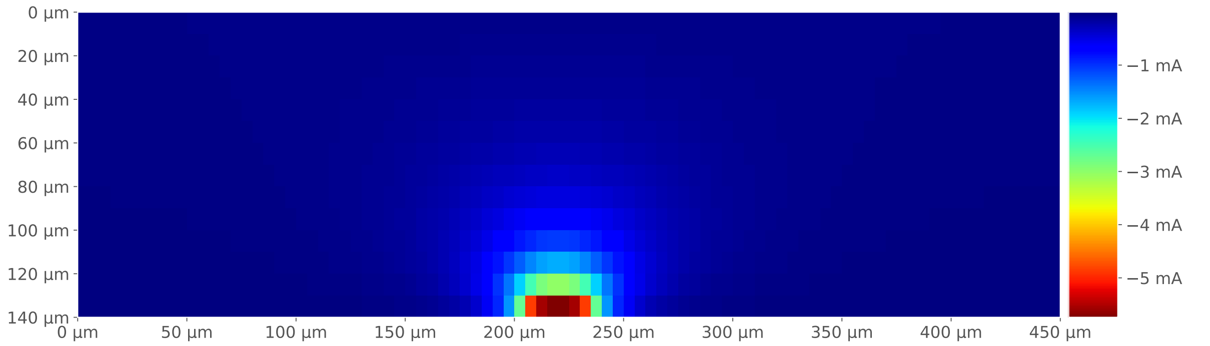
3.5.2 Further generator models and verification 2023-09-07 18:09:30+02:00

Because the previously explained generator model is electrically perfect and does not include any impedance mismatching effects, we extended the model to include the generator output impedance and the transmission line. Peut-être faire un schéma ? It allowed us to observe impedance mismatch effects, which are of great importance when performing BBI (Chapter **??**), as the injected pulses are very fast and of high amplitude. Thus, impedance mismatch greatly changes the effective applied voltage pulse and injected currents, while also modifying unpredictably the induced disturbances, as we will observe further in this manuscript.

In order to verify more thoroughly the soundness of the proposed models, a circuit under BBI is simulated in order to analyze the current distribution and amplitude, specifically at the peak of the voltage pulse. Fig. 3.9 presents the results for both dual-well and triple-well ICs. The substrate being a resistive environment, it is natural to observe isotropic hemispheric current distributions. However, it is interesting to notice that the results show a lower amount of current concerning the triple-well IC compared to the dual-well one. It can be explained thanks to the coupling between the probe/substrate and the logic gates. On one hand, as shown in Fig. 3.1, in the dual-well IC, the charges do not have to cross any silicon junction in order to reach the NMOS transistors, while there is one junction between the probe and the



(a) *Dual-well*



(b) *Triple-well*

Figure 3.9: *Dual-well and triple-well cross-sectional current distribution view at the apex of the voltage pulse*

PMOS transistors. On the other hand, concerning the triple-well IC, there is always at least one silicon junction to cross in order to reach the transistors. Because of this, and because the voltage pulse will inevitably bias the diode, it will change the coupling whether the diode is conducting or blocking. When the diode is conducting, the transistors are DC-coupled to the probe, whereas when the diode is blocking, the transistors are AC-coupled. In the second case, it means that charges can flow only on the edge of the pulse. Thus, during the pulse's plateau, there is no charge flow.

3.6 Experimental comparisons 2023-09-07 18:09:30+02:00

CREUSER PLUS EN DÉTAILS DANS LES SECTIONS PRÉCÉDENTES LES DIFFÉRENCES DUAL/TRIPLE, PARCE QUE C'EST IMPORTANT DANS LE MODÈLE ! In order to complete this chapter, we are going to analyze, in this last section, experimental results highlighting the differences between dual-well and triple-well substrates.

3.7 Conclusion 2023-09-07 18:09:30+02:00

In this chapter, we presented enhanced electrical models which can be utilized to simulate integrated circuits under body biasing fault injection. These models, supported by older ones originally designed for ICs under EMFI, cover two substrate types commonly found in commercial ICs: dual-well and triple-well substrates. The substrate type is of great importance when considering BBI as it is the only physical environment where charges can circulate. Each sub-models contain:

- The power delivery network
- The average electrical model of a hundred of logic gates
- The various silicon junctions
- The silicon substrate

Standard-cells segments models representing a portion of an IC, they need to be replicated and connected with each other in order to be meaningful. In addition to this, they propose refined substrate sub-models in order to improve the model spatial accuracy over their predecessors. The main advantage of these models is their relative lightness, computationally speaking. Indeed, they are only composed of passives components, in order to be able to simulate large resulting ICs. However, their main advantage is also their main shortcoming, they do not represent any function of the modeled IC, but its average electrical behavior.

Algorithm 1 Integrated circuit SPICE netlist generation algorithm.

```
Require: SUBTYPE      ▷ IC substrate type: Dual-well, Triple-well, Mixed
Require: TSUB         ▷ IC substrate thickness
Require: ESUB         ▷ Elementary substrate block thickness
Require: VPUU         ▷ Voltage pulse amplitude
Require: PW           ▷ Voltage pulse width
Require: TFR          ▷ Voltage pulse rise and fall times
Require: SIMTIME      ▷ Simulation duration
Require: SIMSTEP      ▷ Simulation time step
Require: TEX          ▷ Desired X size (µm)
Require: TEY          ▷ Desired Y size (µm)
Require: prbX         ▷ BBI probe X coordinate
Require: prbY         ▷ BBI probe Y coordinate
  RH ← 2000           ▷ Elementary substrate up-down/front-rear resistor value
  RL ← 500            ▷ Elementary substrate left-right resistor value
  WSEG ← 30
  HSEG ← 5
  W6SEG ← 30 ÷ 6
  nC ← TEX ÷ WSEG     ▷ Number of column
  nL ← TEY ÷ HSEG     ▷ Number of lines
  nH ← TSUB ÷ ESUB    ▷ Number of substrate layers
Ensure: nC, nL and nH are integers
  var SCS[nL, nH]     ▷ Array containing each standard-cell properties
  RH ← RH × (ESUB ÷ 10)   ▷ Adjust RH value according to user defined variable
  RL ← RL × (ESUB ÷ 10)   ▷ Adjust RL value according to user defined variable
  for all cY in ⟦0 ; nL⟦ do
    for all cX in ⟦0 ; nH⟦ do
```

$$\vec{X} \leftarrow \begin{bmatrix} cX \times WSEG \\ cX \times WSEG + 1 \times (W6SEG \div 2) \\ cX \times WSEG + 3 \times (W6SEG \div 2) \\ cX \times WSEG + 5 \times (W6SEG \div 2) \\ cX \times WSEG + 7 \times (W6SEG \div 2) \\ cX \times WSEG + 9 \times (W6SEG \div 2) \\ cX \times WSEG + 11 \times (W6SEG \div 2) \\ cX \times WSEG + 12 \times (W6SEG \div 2) \end{bmatrix} ; \vec{Y} \leftarrow \begin{bmatrix} cY \times HSEG \\ (cY + \frac{1}{2}) \times HSEG \\ (cY + 1) \times HSEG \end{bmatrix}$$

```
      if Y⃗[0] = 0 ∨ Y⃗[0] = TEY then     ▷ Determines if SCS has external power
        SCS[cY, cX].power = True
      else
        SCS[cY, cX].power = False
      end if
      if X⃗[0] = prbX ∧ X⃗[2] = prbX ∧ Y⃗[0] ⩽ (prbY + 15) ∧ Y⃗[0] ⩾ (prbY − 15) then
        SCS[cY, cX].probe = True
      else
        SCS[cY, cX].probe = False
      end if
    end for
  end for
```

IV

Substrate thinning analysis 2023-09-07 18:09:30+02:00

Contents

4.1 Summary 2023-09-07 18:09:30+02:00

This chapter proposes to study the interests of thinning the substrate of integrated circuits with the aim to enhance Body Biasing Injection efficiency. First, we are going to present a geometrical approach in order to appreciate with a certain abstraction from electronics the effects of substrate thinning on ICs behaviors. Second, thanks to the models presented in Chapter 3, in addition to the geometrical approach, we are going to theoretically analyze the effects of substrate thinning from an electrical point of view. Eventually, in order to verify the soundness of the geometric approach and the simulation results, experiments are going to be studied thanks to an actual analysis of substrate thinning on identical IC targets behavior.

4.2 Introduction 2023-09-07 18:09:30+02:00

When working with integrated circuits in a fault injection context, several physical parameters of the considered IC are of great importance. For example, as we have seen in the previous Chapter, the type of substrate used to manufacture the IC has a significant impact on BBI efficiency and behavior. In addition to this, the transistors' size, power supply voltage, the IC package or the IC substrate thickness can drastically change fault injections results. Among these examples, one of great interest for body biasing injection is the substrate thickness.

Indeed, as there are different manufacturing processes depending on the purpose of each manufactured IC, it is common to find various substrate thicknesses depending on the IC targeted application. On one hand, it is not rare to find 700 µm thick wafers with 300 mm diameters for generic applications. On the other hand, in other specific applications like SoCs, where vertical stacking is commonly used, or in Smart-cards and ID cards, the typical substrate thickness value is lower, around 200 µm. In addition to these differences one can find in commercial products, the practice of thinning the substrate of ICs is not uncommon in a context of fault injection. More specifically, substrate thinning has been thoroughly studied concerning Laser Fault Injection (LFI) [29, 30], and has proven to greatly enhance LFI efficiency, in addition to drastically reducing the power required to create faults. However, it had not been studied for Body Biasing Injection at the beginning of this work.

In this context, this work was first done in order to assess whether substrate thinning has similar effects on BBI as it has on LFI. Second, because thin ICs commonly found in smart-cards have unavoidable security constraints, third because BBI is performed using the silicon substrate as the physical environment to carry energy through electrical charges. Therefore, this Chapter will evaluate the interests of substrate thinning on BBI efficiency. In other words, we will analyze the electrical and behavioral differences between identical ICs with different substrate thicknesses. This analysis will take place using multiple approaches. In the first place, we will address the question using a geometric approach to appreciate the effects of substrate thinning on voltage propagation inside the substrate while taking a step back from electrical modeling. Then, the geometric approach will be completed with an electrical simulation analysis of two identical ICs with different substrate thicknesses, created thanks to the models proposed in Chapter 3. Eventually, experimental results will be analyzed in order to verify the correctness of the previous approaches, in addition to studying the actual effects of substrate thinning concerning faults creation.

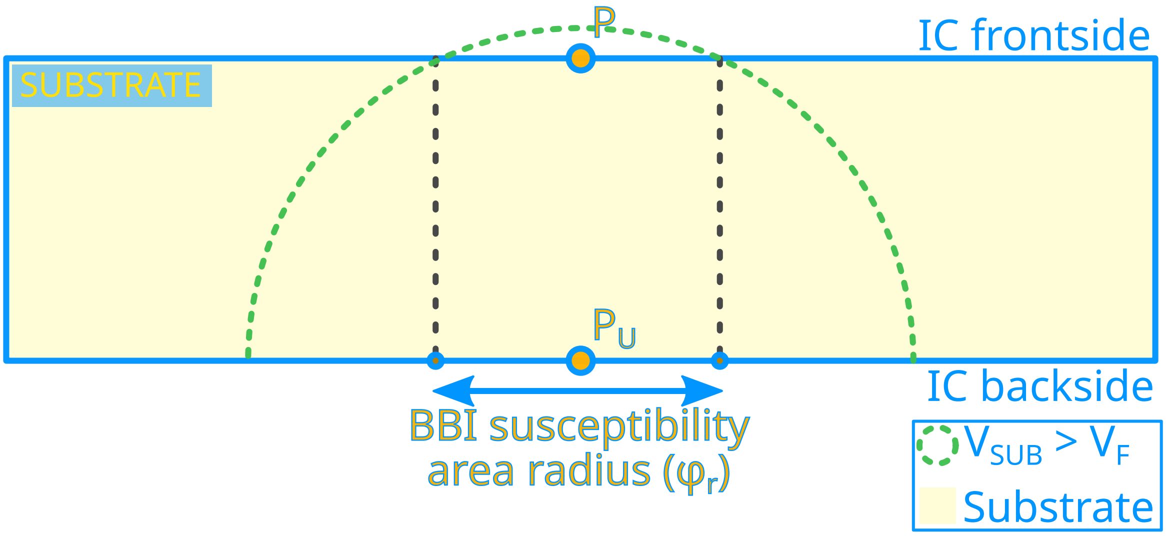
4.3 Geometric and electrical modeling 2023-09-07 18:09:30+02:00

To begin with, we will address the geometric approach. It has been chosen thanks to the advantages it brings forward, such as the abstraction from electronics it enables, thus allowing easier and faster modeling. However, because this approach alone is insufficient, we will then study an analogous electrical one.

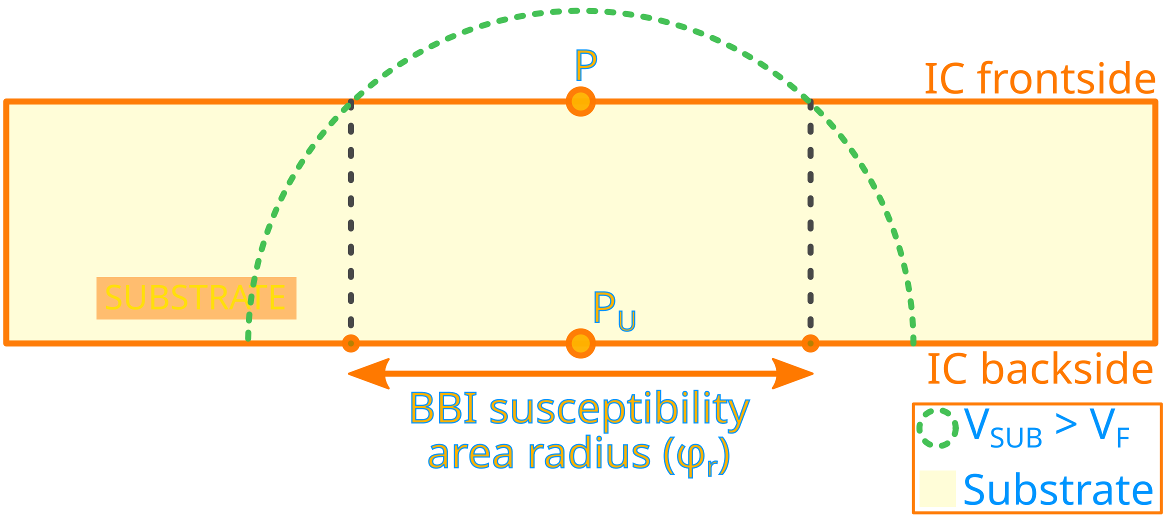
4.3.1 Geometric modeling 2023-09-07 18:09:30+02:00

For the purpose of geometric modeling, let us consider two identical ICs. A commercial one, with an arbitrary standard substrate thickness, and another one with its substrate thinned by a certain amount in order to perform fault injection. Fig. 4.1 illustrates the two-dimensional cross-sectional views of the considered ICs substrates during an arbitrary BBI voltage pulse. The silicon substrate being an isotropic resistive environment, it is quite natural to expect the electrical charges to flow and spread evenly when injected into it at any given time. Therefore, equipotentials form half-sphere surfaces inside the substrate volume. These surfaces are highlighted in two-dimensions as green half-circles in Fig. 4.1.

In this scenario, an attacker wants to induce a fault in the logic gates, located at the top of each IC. To that end, they need to change the voltage enough at point P, called V_P, in order to disturb the transistors and change the logic gates behavior. In addition to that, and for the sake of simplicity, let us assume that P is the only location in the considered IC where faults can be injected. However, in order to observe faults at point P, V_P needs to reach a minimal threshold voltage, called V_F. Because the attacker is working with BBI, a metallic probe is connected onto the backside of the IC, at point P_U, in order to inject energy into the IC. Depending on the amount of injected energy, in other words, the maximum amplitude of the voltage pulse because the substrate effective resistance is static, the voltage at P might never reach V_F, therefore, no faults will be observed. Let us consider that the attacker chose an amplitude V_{PU} big enough such that at a moment in the injection, V_P reaches V_F or more in each considered IC. In that scenario, the area on the IC front side where $V > V_F$ is a disk of radius ϕ, centered in P, called the BBI susceptibility area radius. It means that the attacker can position the probe anywhere on the backside within this disk to reach V_F at P, and therefore



(a) *Normal IC*



(b) *Thinned IC*

Figure 4.1: *BBI susceptibility area cross-sectional 2D view*

induce a fault at P.

The half-sphere equipotential radius relative to time can be determined thanks to the following formula:

$$r(t) = \frac{\rho_{SUB}}{\sqrt{2}} \cdot \frac{|I_G(t)|}{|V_{PU}(t) + V_F|} \tag{4.1}$$

with ρ_{SUB} the resistivity of the silicon substrate, $I_G(t)$ the instantaneous sum of the current distribution contained in the half-sphere, and $V_{PU}(t)$ the instantaneous voltage pulse applied on the backside of the IC. Then, logically, the BBI susceptibility area radius, denoted ϕ_r, is described by:

$$\phi_r(t) = 2 \cdot \sqrt{r(t)^2 - t_{SUB}^2} \tag{4.2}$$

with t_{SUB} being the IC substrate thickness.

As it is illustrated in Fig. 4.1, thinning the substrate inevitably increases the size of the susceptibility area if the experimental conditions are constant. It means that the susceptibility evolution ratio is always greater than 1 when thinning the substrate:

$$\frac{\phi_r^{THIN}}{\phi_r^{THICK}} = \sqrt{\frac{r^2 - t_{THIN}^2}{r^2 - t_{THICK}^2}} > 1 \tag{4.3}$$

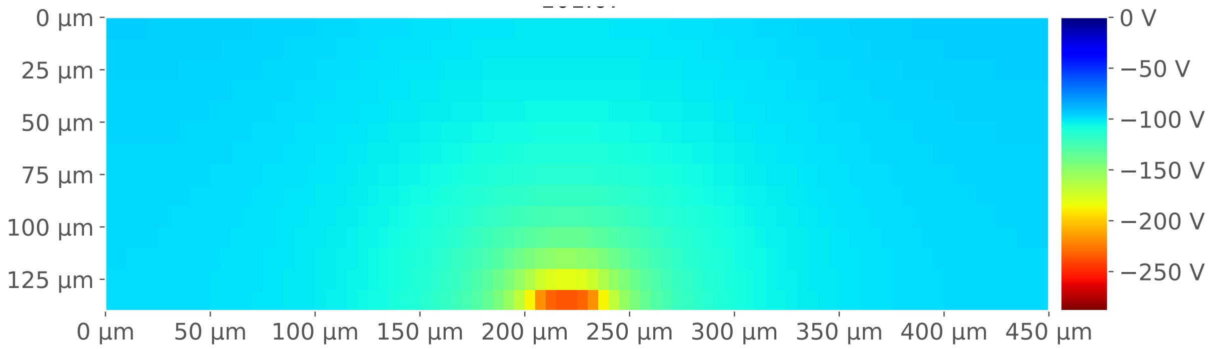
Therefore, in order to obtain the same susceptibility area with a thinner IC, it is required to reduce the voltage pulse amplitude, thanks to the following relation:

$$V_{PU}^* = \frac{t_{THIN}}{t_{THICK}} \cdot V_{PU} + V_F \cdot (1 - \frac{t_{THIN}}{t_{THICK}}) \tag{4.4}$$

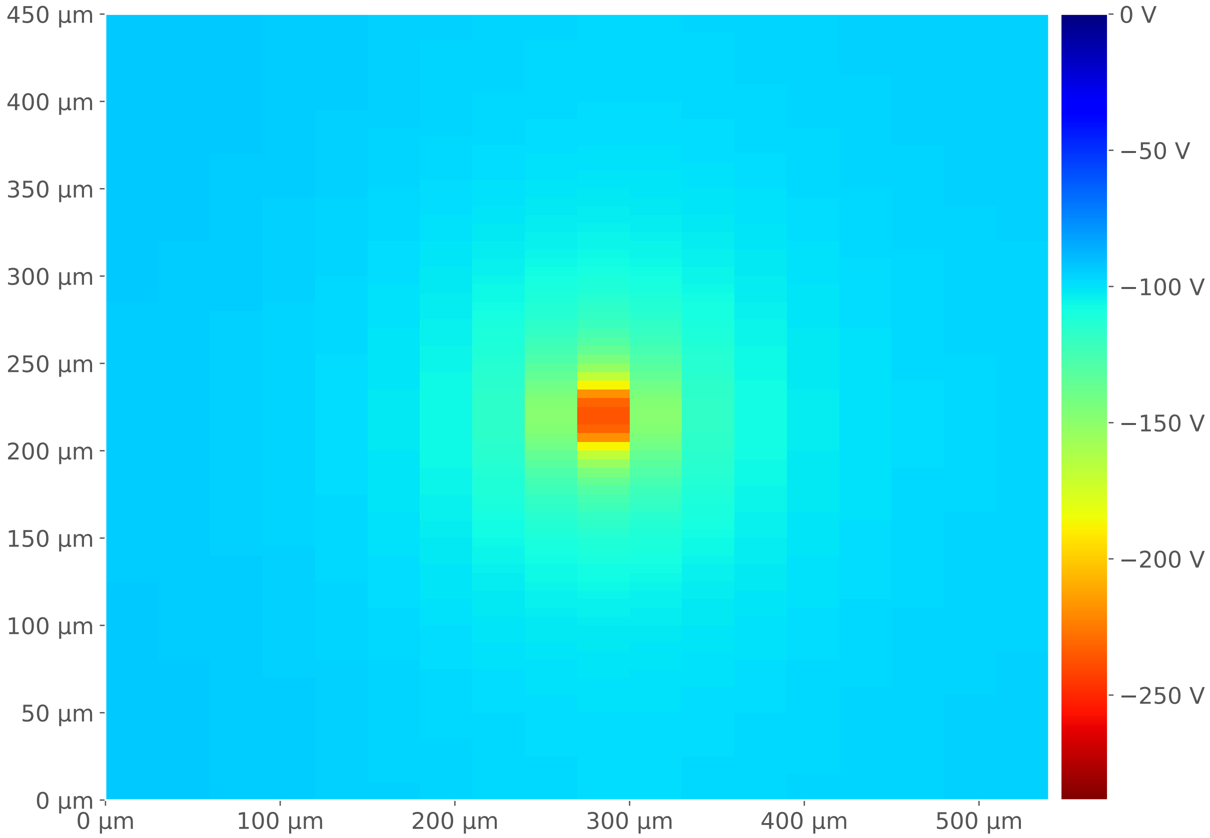
Eventually, this geometrical approach allows deducing three conclusions:

1. Thinning the substrate allows reducing the minimal voltage pulse amplitude required to induce a fault while keeping a constant susceptibility area.
2. The BBI susceptibility area increases while the substrate thickness decreases while working at a constant voltage pulse V_{PU}.
3. Thinning the substrate alone does not have an influence on BBI spatial resolution, as the susceptibility area depends on the couple (t_{SUB}, V_{PU}). Thus, similar spatial resolution could be obtained with different substrate thicknesses by changing V_{PU}.

4.3.2 Electrical approach 2023-09-07 18:09:30+02:00

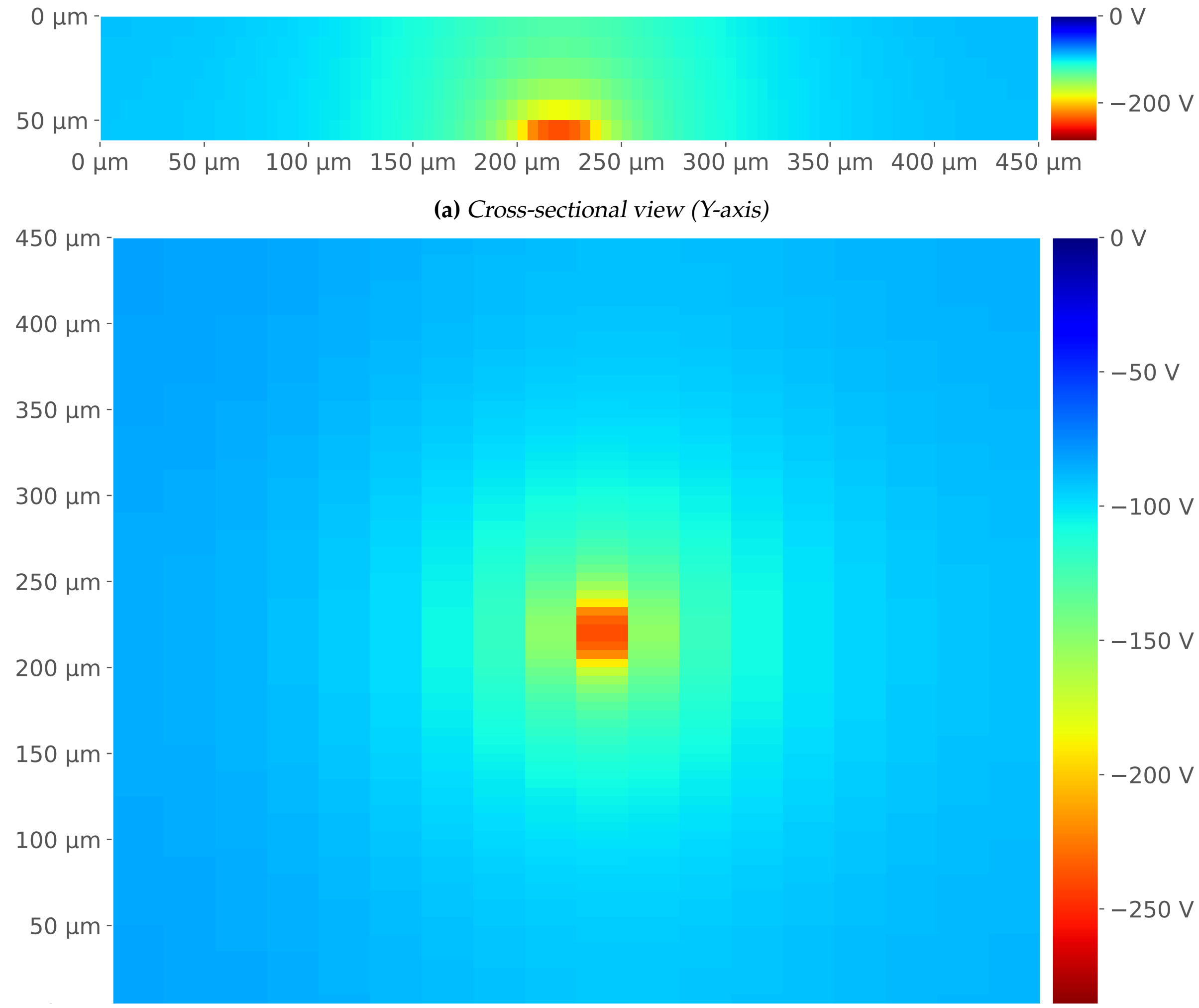


(a) *Cross-sectional view (Y-axis)*



(b) *Probe point of view*

Figure 4.2: *Simulated non-thinned IC (140 µm) substrate voltage distribution: peak of the first voltage pulse edge*



(a) *Cross-sectional view (Y-axis)*

(b) *Probe point of view*

Figure 4.3: *Simulated thinned IC (60 μm) substrate voltage distribution: peak of the first voltage pulse edge*

As stated previously, in order to verify the meaningfulness of the geometrical approach, we will complete it with an electrical modeling approach. For this purpose, the models introduced in Chapter 3 are reused. The electrical approach consists in generating ICs with different substrate thicknesses and simulating them during BBI. The considered ICs are 550 μm wide and 450 μm deep. Two substrate thicknesses are analyzed, 60 μm and 140 μm. The simulation parameters are the following:

- Triple-well substrate
- Required voltage pulse: -300 V
- Required pulse width: 20 ns
- Required rise and fall times: 8 ns

Fig. 4.2 and Fig. 4.3 show, for each simulated IC, the voltage bias across the substrate through different point of view at the apex of the voltage pulse first edge. For simplicity, results are shown in two dimensions and from two point of views: a cross-sectional view and a bottom view. The first interesting thing to note is that, as predicted thanks to the geometric model and as shown in Fig. 4.2 and 4.3, equipotentials effectively form half-circles into the substrate (half-spheres in 3D). They can be first observed from the bottom, where the voltage is spreading across the backside surface of the IC. Second, in the cross-sectional view, as it was illustrated previously with the geometrical model............... IL Y A BEAUCOUP DE CHOSES À DIRE MAIS JE MANQUE D'INSPIRATION POUR CETTE PARTIE, JE REVIENDRAI PLUS TARD DESSUS.

4.4 Models validation 2023-09-07 18:09:30+02:00

This section presents the conducted experiments allowing to validate the previously presented models.

4.4.1 IC substrate thinning quick look 2023-09-07 18:09:30+02:00

As substrate thinning is quite widespread when performing fault injection, let us have a quick look on how it is performed. Commonly, It is done using Selected Area Preparation (SAP) or Focused Ion Beams (FIB) milling. SAP milling consists in a very precise mechanical milling tool, generally able to remove material with a precision down to a few micrometers. However, it can often lead to uneven surfaces. FIB milling consists in a physical milling which does not imply a mechanical contact with the material to be removed, and allows nanometer-level precision. For that purpose, FIB is commonly used in combination with SAP [31] to produce even substrate surfaces. In addition to substrate thinning, SAP milling machines allow removing the plastic package and eventual internal metallic heat-sinks of ICs prior to substrate thinning. It has the advantage of providing low damage to thinned ICs, thanks to low spindle speed and low temperature rise compared to traditional high speed milling.

4.4.2 Experiments with thinned circuits 2023-09-07 18:09:30+02:00

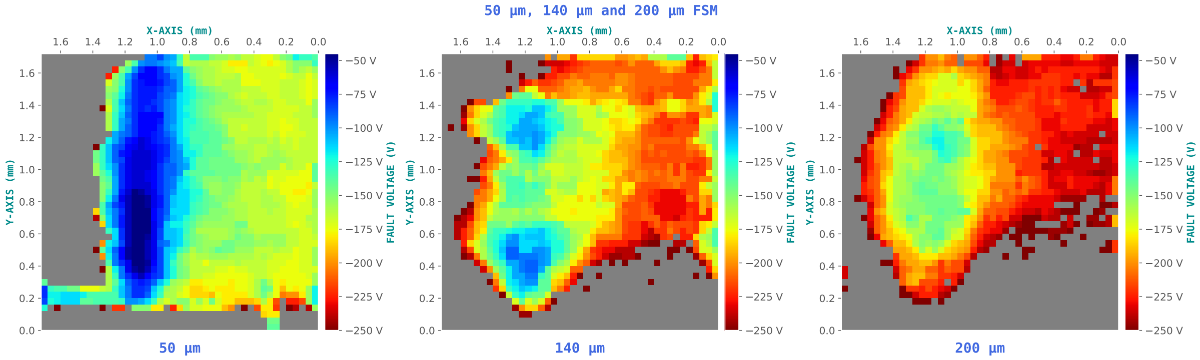


Figure 4.4: *Fault susceptibility maps*

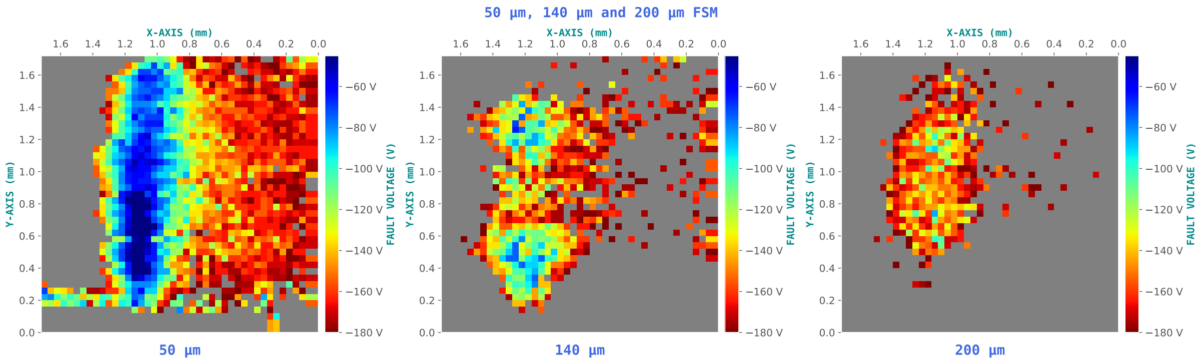


Figure 4.5: *Susceptibility area spreading*

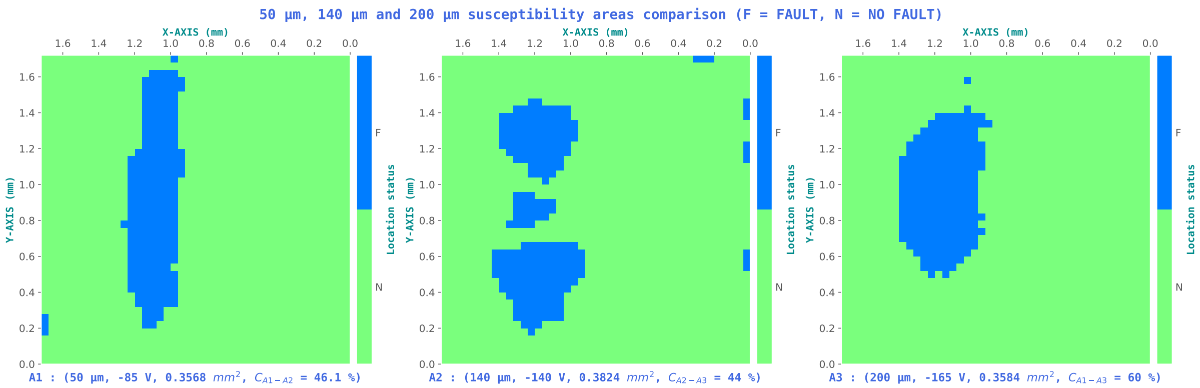


Figure 4.6: *Fault susceptibility maps couples*

With geometric and electrical modeling complete, it is now possible to conduct actual experiments in order to verify the meaningfulness of the previous approaches. In this context, three an STM32F439VIT6 LQFP100 identical targets were thinned to three different levels, from 750 µm to respectively 200 µm, 140 µm and 50 µm, respectively named ST200, ST140 and ST50 for the rest of this Chapter. In order to verify the three conclusions extracted from the modeling section, three experiments are conducted for each target.

The first experiment aims at measuring the minimal voltage pulse amplitude V_{PU}^{MIN} required to induce a faulty behavior on an IC performing computations. These experiments are called Fault Susceptibility Maps (FSM). They allow spotting the region where the IC is sensitive to BBI, no matter which type of induced fault. Therefore, when mapping an entire IC, it is common to spot various areas not directly involved in the targeted calculation, like the analog voltage regulator or the FLASH memory logic control logic not to cite them all. As a result, and because in a fault injection context the cryptographic core is very often targeted, it was decided to focus the maps above the STM32 AES core only. Fig. 4.4 presents the three performed FSM. From left to right, t_{SUB} goes from 50 µm, then to 140 µm, finally to 200 µm. As stated before, the maps are performed above the hardware AES core of the IC, temporally aiming the penultimate AES round. The scanned area measures 1.7 mm by 1.7 mm, with a displacement step of 40 µm between each point. V_{PU} was limited to the following range: [30 V ; 280 V], with 5 V steps and a negative polarity. The pulse width was fixed at 6 ns. The first important thing to note here is that, as predicted with the geometric and electrical modelings, a thinner substrate allows a lower fault induction threshold. It is mainly shown thanks to the measurement of the average voltage required to induce a fault across the entire map, annotated at the top of each map. All of this sustains the first conclusion made in section 4.3.

Then, the second experiment, whose results are shown in Fig. 4.5, consist in analyzing the spreading of the BBI susceptibility area. The core of the experiment is identical as before. However, in order to highlight the spreading effect, it was required to set a lower maximum voltage amplitude (in absolute value). The value of 180 V was chosen as it is the average voltage of the medium-thinned IC. What is interesting here is that, for the ST200 target, because the voltage at the epitaxy level cannot reach the threshold value V_F in most cases, the fault area is tiny compared to the other targets, and focused on the AES core. Then, concerning the ST140 target, thanks to the thinner substrate, the voltage at the epitaxy level can reach a higher value, and thus can cause more logic gates or further logic gates from the probe to have a faulty behavior. Eventually, the ST50 target shows the largest fault area. These experiments help to sustain the second conclusion of section 4.3.

Eventually, the last experiment consisted in finding, whenever possible, (t_{SUB}, V_{PU}) couples for which the susceptibility area is identical across all targets. The search for the couples of values was done by first choosing an arbitrary couple for ST200 target, and then calculating the correlation for each couple between the other two susceptibility areas and finding the highest correlation. Then, to confront the geometric modeling predictions, we calculated, thanks to equation 4.4, couples corresponding to

4.5 Conclusion 2023-09-07 18:09:30+02:00

This chapter introduced the interest of thinning the substrate of integrated circuits on Body Biasing Injection efficiency. In the first place, we studied thanks to a geometrical approach the potential benefits of this practice, further completed with electrical simulations. The geometric approach brought mathematical relations allowing to evaluate preliminary the effects of thinning the substrate of a target IC.

À FINIR.

V

Fault model 2023-09-07 18:09:30+02:00

Contents

5.1 Summary 2023-09-07 18:09:30+02:00

In this chapter, we present a fault model for BBI, extrapolated from a fault model used for EMFI. The objective of this chapter is to provide an explanation of the mechanisms and causes of faults in integrated circuits that are subjected to body biasing injection. Electrical models presented in the chapter 3 can be used to explain how electrical charge displacement in the IC during a BBI pulse allows changing some logic gates output values. Targeting an IC via BBI forces electric charges to be injected (resp. absorbed) with positive pulses (resp. negative). Therefore, it is possible to target a critical time in the IC calculation thanks to the ability to finely control the induced disturbances. Eventually, to verify the correctness of the proposed analysis, both substrate charge propagation and logic gate behavior studies will be conducted in parallel.

5.2 Introduction 2023-09-07 18:09:30+02:00

To further complete the understanding of BBI, in addition to having a reliable model to predict IC behavior, it is of great importance of having a precise fault model, in order to be able to set up countermeasures. Indeed, the main objective of studying fault injection techniques is to protect further secured ICs in order to consider during the design of new ICs, the implications and impacts of such countermeasures on the design. As it has been said in Chapter 3, simulating at a transistor level an entire IC is unrealistic, at least computationally speaking. Therefore, and because the previous models do not represent the logical functions of the considered ICs, we propose an additional step to the simulation workflow proposed in Chapter 3. This addition consists in extracting the propagated disturbances from standard-cell segments models, and injecting them into functioning logic gates. This method allows appreciating logic gates behavior under BBI in order to get a deeper and more precise understanding of both electrical and functional fault creation mechanisms. All of this is part of the required steps to set up efficient countermeasures, as we need to understand precisely the insights of fault creation.

5.3 Charge extortion 2023-09-07 18:09:30+02:00

This section explains the charge extortion mechanism at work during BBI which allows fault creation. The voltage pulse generator, at each edge of its pulse, injects and then extorts electrical charges into and out of the IC.

5.3.1 Sequential logic operation and simple fault model 2023-09-07 18:09:30+02:00

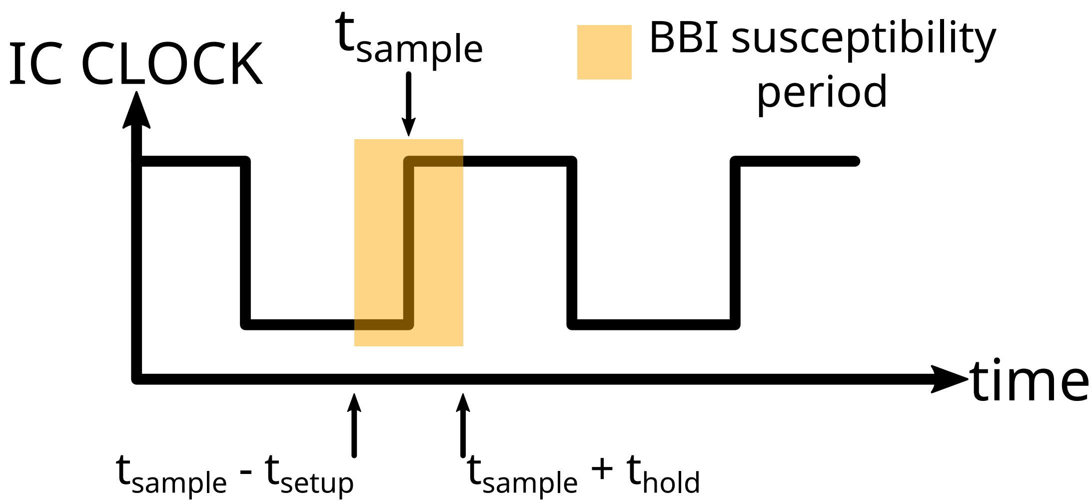


Figure 5.1: *Sequential logic operation and BBI sampling fault susceptibility*

As sequential logic is ubiquitous in contemporary integrated circuits, we shall examine its fundamental workings in greater detail. Sequential logic relies on a core element: the edge-triggered D flip-flop (DFF). They are a memory component that is governed by a clock. At each rising-edge or falling-edge (depending on the design) of the clock, DFFs sample their input and replicate it at their output. Between DFFs are placed the logic gates, which fulfill a specific logical function. Because of this, values in sequential logic circuits can only be changed at the clock edges. Hence, in the event that an adversary is able to alter a logical value for an extended period of time at the input of a DFF, the subsequent combinatorial logic will yield an incorrect value, which will propagate to the subsequent DFFs.

We are going to use this fault model, depicted in short in Fig. 5.1, for the rest of this chapter. It is important to note that this model was first introduced for EMFI [32].

5.4 Silicon substrate charges propagation 2023-09-07 18:09:30+02:00

On a déjà montré ces cartos dans des parties précédentes.

5.5 Logic gates simulation under BBI 2023-09-07 18:09:30+02:00

VI

Conclusion

Bibliography

[1] G. Chancel, J.-M. Galliere, and P. Maurine. Body biasing injection: To thin or not to thin the substrate? In Josep Balasch and Colin O'Flynn, editors, *Constructive Side-Channel Analysis and Secure Design*, pages 125–139, Cham, 2022. Springer International Publishing. xix, 28

[2] Prasanna Ravi, Zakaria Najm, Shivam Bhasin, Mustafa Khairallah, Sourav Sen Gupta, and Anupam Chattopadhyay. Security is an architectural design constraint. *Microprocessors and Microsystems*, 68:17–27, 2019. 2

[3] Paul C. Kocher. Timing attacks on implementations of diffie-hellman, rsa, dss, and other systems. In Neal Koblitz, editor, *Advances in Cryptology — CRYPTO '96*, pages 104–113, Berlin, Heidelberg, 1996. Springer Berlin Heidelberg. 5

[4] A. Shamir R.L. Rivest and L.Adleman. A method for obtaining digital signatures and public-key cryptosystems. In *Communications of the ACM*, volume 21, pages 120–126, 1978. 5

[5] Boris Köpf and Markus Dürmuth. A provably secure and efficient countermeasure against timing attacks. In *2009 22nd IEEE Computer Security Foundations Symposium*, pages 324–335, 2009. 5

[6] Paul Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In Michael Wiener, editor, *Advances in Cryptology — CRYPTO' 99*, pages 388–397, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg. 5

[7] Eric Brier, Christophe Clavier, and Francis Olivier. Correlation power analysis with a leakage model. In Marc Joye and Jean-Jacques Quisquater, editors, *Cryptographic Hardware and Embedded Systems - CHES 2004*, pages 16–29, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg. 6

[8] Vincent Carlier, Hervé Chabanne, Emmanuelle Dottax, and Hervé Pelletier. Electromagnetic side channels of an fpga implementation of aes. Cryptology ePrint Archive, Paper 2004/145, 2004. https://eprint.iacr.org/2004/145. 6

[9] Thomas Ordas, Mathieu Lisart, Etienne Sicard, Philippe Maurine, and Lionel Torres. Near-field mapping system to scan in time domain the magnetic emissions of integrated circuits. In Lars Svensson and José Monteiro, editors, *Integrated Circuit and System Design. Power*

and Timing Modeling, Optimization and Simulation, pages 229–236, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. 6

[10] Aurélien Vasselle, Philippe Maurine, and Maxime Cozzi. Breaking mobile firmware encryption through near-field side-channel analysis. In *Proceedings of the 3rd ACM Workshop on Attacks and Solutions in Hardware Security Workshop*, ASHES'19, page 23–32, New York, NY, USA, 2019. Association for Computing Machinery. 6

[11] Eli Biham and Adi Shamir. Differential fault analysis of secret key cryptosystems. In *Annual International Cryptology Conference*, 1997. 7

[12] Dan Boneh, Richard A. DeMillo, and Richard J. Lipton. On the importance of checking cryptographic protocols for faults. In Walter Fumy, editor, *Advances in Cryptology — EUROCRYPT '97*, pages 37–51, Berlin, Heidelberg, 1997. Springer Berlin Heidelberg. 7

[13] Mathieu Ciet and Marc Joye. Elliptic curve cryptosystems in the presence of permanent and transient faults. *Designs, Codes and Cryptography*, 36(1):33–43, July 2005. 7

[14] Ingrid Biehl, Bernd Meyer, and Volker Müller. Differential fault attacks on elliptic curve cryptosystems. In Mihir Bellare, editor, *Advances in Cryptology — CRYPTO 2000*, pages 131–146, Berlin, Heidelberg, 2000. Springer Berlin Heidelberg. 7

[15] Christophe Giraud. Dfa on aes. In Hans Dobbertin, Vincent Rijmen, and Aleksandra Sowa, editors, *Advanced Encryption Standard – AES*, pages 27–41, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg. 7, 25

[16] Yang Li, Kazuo Sakiyama, Shigeto Gomisawa, Toshinori Fukunaga, Junko Takahashi, and Kazuo Ohta. Fault sensitivity analysis. In Stefan Mangard and François-Xavier Standaert, editors, *Cryptographic Hardware and Embedded Systems, CHES 2010*, pages 320–334, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. 7

[17] Sergei Skorobogatov and Ross Anderson. Optical fault induction attacks. volume 2523, pages 2–12, 08 2002. 8

[18] David Samyde, Sergei P. Skorobogatov, Ross J. Anderson, and Jean-Jacques Quisquater. On a new way to read data from memory. *First International IEEE Security in Storage Workshop, 2002. Proceedings.*, pages 65–69, 2002. 8

[19] M. Lisart M. Dumont and P. Maurine. Modeling and simulating electromagnetic fault injection. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 40(4):680–693, 2021. 9, 28, 29, 30

[20] Philippe Maurine, Karim Tobich, Thomas Ordas, and Pierre-Yvan Liardet. Yet another fault injection technique : by forward body biasing injection. 09 2012. 9

[21] K. Tobich, P. Maurine, P.-Y. Liardet, M. Lisart, and T. Ordas. Voltage spikes on the substrate to obtain timing faults. In *2013 Euromicro Conference on Digital System Design*, pages 483–486, 2013. 9

[22] Noemie Beringuier-Boher, Marc Lacruche, David El-Baze, Jean-Max Dutertre, Jean-Baptiste Rigaud, and Philippe Maurine. Body biasing injection attacks in practice. In *Proceedings of the Third Workshop on Cryptography and Security in Computing Systems*, CS2 ’16, page 49–54, New York, NY, USA, 2016. Association for Computing Machinery. 9

[23] Colin O’Flynn. Low-cost body biasing injection (BBI) attacks on WLCSP devices. In Pierre-Yvan Liardet and Nele Mentens, editors, *Smart Card Research and Advanced Applications*, pages 166–180, Cham, 2021. Springer International Publishing. 9

[24] G. Chancel, Jean-Marc Gallière, and P. Maurine. Body biasing injection: Impact of substrate types on the induced disturbances*f*. In *2022 Workshop on Fault Detection and Tolerance in Cryptography (FDTC)*, pages 50–60, 2022. 28

[25] Mathieu Dumont, Philippe Maurine, and Mathieu Lisart. Modeling of electromagnetic fault injection. In *2019 12th International Workshop on the Electromagnetic Compatibility of Integrated Circuits (EMC Compo)*, pages 246–248, 2019. 28

[26] Yasuhiro Ogasahara, Masanori Hashimoto, Toshiki Kanamoto, and Takao Onoye. Supply noise suppression by triple-well structure. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 21(4):781–785, 2013. 32

[27] Takuya Wadatsumi, Kohei Kawai, Rikuu Hasegawa, Takuji Miki, Makoto Nagata, Kikuo Muramatsu, Hiromu Hasegawa, Takuya Sawada, Takahito Fukushima, and Hisashi Kondo. Voltage surges by backside esd impacts on ic chip in flip chip packaging. In *2022 IEEE International Reliability Physics Symposium (IRPS)*, pages P14–1–P14–6, 2022. 37

[28] Takuya Wadatsumi, Kohei Kawai, Rikuu Hasegawa, Kazuki Monta, Takuji Miki, and Makoto Nagata. Characterization of backside esd impacts on integrated circuits. In *2023 IEEE International Reliability Physics Symposium (IRPS)*, pages 1–6, 2023. 37

[29] Breier et al. Extensive laser fault injection profiling of 65 nm fpga. *J Hardw Syst Secur 1*, pages 237–251, 2017. 42

[30] Jakub Breier and Chien-Ning Chen. On determining optimal parameters for testing devices against laser fault attacks. In *2016 International Symposium on Integrated Circuits (ISIC)*, pages 1–4, 2016. 42

[31] C. Boit, R. Schlangen, A. Glowacki, U. Kindereit, T. Kiyan, U. Kerst, T. Lundquist, S. Kasapi, and H. Suzuki. Physical ic debug and - backside approach and nanoscale challenge. *Advances in Radio Science*, 6:265–272, 2008. 48

[32] S. Ordas, L. Guillaume-Sage, and P. Maurine. Electromagnetic fault injection: the curse of flip-flops. *Journal of Cryptographic Engineering*, 7(3):183–197, Sep 2017. 55