Body biasing injection: analysis, modeling and simulation (MAX 14 PAGES)

Geoffrey Chancel

***Abstract*—In the past decades, several fault injection techniques have been studied, such as Laser Fault Injection (EMFI), Electromagnetic Fault Injection (LFI), or Body Biasing Injection. Each method employs the manipulation of different physical quantities to achieve their goal. EMFI takes advantage of the electromagnetic susceptibility of integrated circuits, LFI uses the photo-sensitivity of the silicon, and BBI the conductive nature of the silicon substrate composing the ICs. Among these methods, BBI, although it has gained interest in the past few years, have still a lot of things to disclose. Particularly, methods to set up repeatable BBI platforms and experiment still need to be explored. In this context, this work aims at presenting our work on BBI in the past few years in its entirety, such as the set up of better platforms, a differential fault attack using BBI, and a simulation flow dedicated to BBI.**

***Index Terms*—Article submission, IEEE, IEEEtran, journal, LaTeX, paper, template, typesetting.**

I. Introduction

NOWADAYS, electronic devices are found in every economic sector, and very often manipulate sensitive and confidential data, such as in bank transactions, Internet of Things (IoT) devices, smartcards, or smartphones. To ensure data authenticity and confidentiality, these devices embed cryptographic algorithms. While theoretically secure and robust, once implemented on actual devices, these algorithms become fallible by leaking parts of the manipulated data through various physical quantities such as electromagnetic waves, infrared emissions, or sound emissions, not to cite them all. In addition to this, they are sensitive to external disturbances.

Cybersecurity, more specifically hardware security, takes place in this context When comes hardware security often comes side-channel attacks and fault injection attacks. On the one hand, side-channel attacks take advantage of the circuit leakage by measuring the various physical quantities available. On the other hand, fault injection aims at inducing physical disturbances into circuits, with methods like Electromagnetic Fault Injection (EMFI) [1], [2], Laser Fault Injection (LFI) [3], or Body Biasing Injection (BBI) [4], not to cite them all. Among these methods, EMFI and LFI are widely studied and understood. However, despite a resurgence in the past few years, BBI knowledge is still less mature compared to the previously cited methods. Therefore, this article is dedicated in presenting our work on Body Biasing Injection.

A. Fault injection objectives

Before going further in the discussion about BBI, let us first outline the main objectives of fault injection methods. Most commonly, they are set up to perform various malicious manipulation on integrated circuits, such as:

- Denial of service (DoS) → Stop circuit operation and the related services;
- Verification bypass → Modify data on the fly to fake authenticity (e.g. to bypass bootloader security);
- Confidential data extraction → Modify data to perform differential fault analysis.

To perform these objectives, we can use various injection methods, such as EMFI, LFI or BBI. Before presenting our work on BBI further, let us analyze the available and existing BBI platforms in the state-of-the-art.

B. BBI in the state-of-the-art

When compared to EMFI, BBI has a smaller state-of-the-art, whether in the amount of scientific papers published or in the amount of industrial platforms proposed. Currently, there are ten main works lingering on BBI [4]–[13]. Each one of them made a unique contribution for a better understanding of BBI.

The first one [4] introduced the technique and presented a Bellcore attack on the targeted IC. Then, one year later, another work [5] further studied the method, followed by a third work three years later [6], introducing an advanced test bench to work and perform attacks with BBI. After that, another work presented a low-cost BBI platform, dedicated to WLCSP devices [7]. Then, quite interestingly researchers proposed a study of BBI using an ESD gun as a voltage surge generator [8], [9]. At the same time, the impact of substrate thickness on BBI efficiency has been studied in [10], in addition to a simulation flow and better practices [11], [12]. Eventually, one last work proposed a safety-focused low-cost and open-source design for the practice of BBI [13].

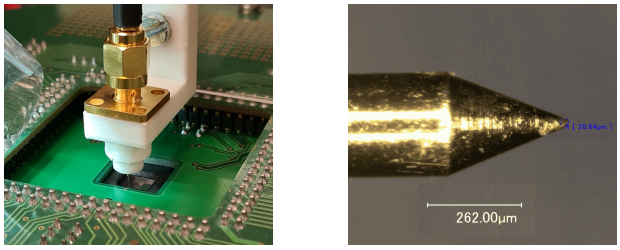
However, despite this extensive work, there are still unanswered questions, and the current works aims at bringing more answers thanks to a compilation of our work on Body Biasing Injection.

Before going any further, let us introduce the platform we use for the present work.

C. Our BBI platform

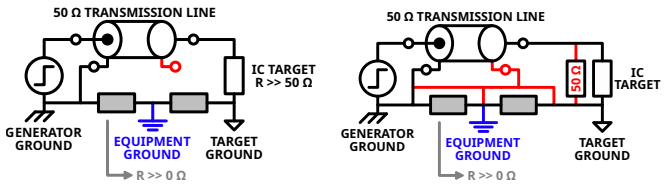
As the commercial platforms, our platform is focused on two main pieces of equipment: a voltage pulse generator and a metallic probe.

The generator model is the AVRK-4-B from the society Avtech Electrosystems Ltd. This model is commonly used for



(a) Global view (b) Zoomed view

Fig. 1: Our custom BBI probe. `bbi_probe`



(a) Typical (b) Improved

Fig. 2: A typical (a) and an improved (b) BBI setup. `bbi_setups`

EMFI, but is suitable for BBI, and its specifications are the following:

- Pulse amplitude: ± [150, 750] V;
- Pulse width: [6, 20] ns;
- First edge rise/fall time: 4 ns;
- Second edge rise/fall time: load dependent;
- Recovery time: $<$ 1 ms;
- Propagation delay (PD): 150 ns;
- Jitter: ± 100 ps ± 0.03 % of PD;
- DC-coupled output;
- Loaded with 50 Ω.

Probably the most distinctive piece of equipment when it comes top BBI is the probe. Some BBI probes can be active, others passive and less expensive. However, it is important to keep this piece of equipment relatively cheap as it endures most of the physical strain on a BBI platform, and should be easy to replace or repair. Fig. 1 shows two pictures of our probe from different angles. The one we use is custom made around three parts:

- A spring-loaded metallic tip, with a 20 µm head diameter;
- A SMA connector, where the tip is soldered;
- A custom 3D-printed enclosure holding the pieces together and cheap to replace.

The spring-loaded tip is 17 mm long and has a global diameter of 0.635 mm. It is specified for a 1.5 A nominal current, and its electrical resistance measures around 70 mΩ. The total cost of the probe is of 20 €.

D. BBI interrogations

With all

- What is the spatial resolution of BBI?
- What is the time resolution of BBI?
- Is thinning the substrate useful in any way?
- How BBI induced faults occur?
- How to properly model BBI?

II. How to perform BBI in a better way

A. BBI platforms in the state of the art

In the first place, we will analyze, from a theoretical perspective, a typical BBI platform. To do so, we created simple platform models allowing to highlight the major limiting factors of such platforms.

The typical platform model is described in Fig. 2.a and shows the main components making a BBI platform such as:

- The voltage pulse generator;
- The transmission line;
- The grounding installation;
- The IC target.

In addition to this, the schematic shows some important flaws we are going to address.

While this is not always the case, voltage pulse generators are typically specified to be loaded with a 50 Ω load, or more generally with a fixed load. When performing BBI, the backside of the IC is electrically connected to the generator output. Therefore, outside of luck alone, it is very rare that the impedance presented by the IC to the generator perfectly matches the required one. It implies that the generator will be, most of the time, out of specifications, and that the conditions will vary depending on the chosen IC and the location of the BBI probe. This can lead to issues such as errors in the set-point voltage and pulse width, and ringing in the transmission line. It then represents a first flaw to the typical approach.

Then, there is the grounding installation. The model presents a non-ideal but simple platform grounding. The reference, used by the oscilloscope and the main computer, is represented in blue and called ”equipment ground”. Ideally, every ground on the platform is connected to this reference with a very low impedance interconnection. However, depending on the hardware used, it may greatly vary from one platform to another. In the model, the secondaries generator and target grounds are connected to the reference thanks to vastly imperfect interconnections, whose impedance is significantly higher than zero. This mainly lead to set-point errors due to shifts in the voltage pulse amplitude. Therefore, it limits the inter-platform repeatability of BBI experiments.

B. Improvements proposed

To circumvent the previously introduced limitations, we propose two corrections to generalize the platforms and improve the repeatability.

First, let us talk about the improper grounding, Alleviating this issue is fairly straightforward. To do so, we propose to choose a reference, such as the equipment ground, and bypass all the grounds with low-impedance interconnections from this reference, as proposed in red in Fig. 2.b.

Then, concerning the impedance mismatch of the generator, multiple solutions can be approached. The best solution would be to implement an adaptive impedance matching system with active feedback, able to measure in real-time the impedance seen by the generator. However, adopting such a method is

costly and long to set up in comparison to the next solution. Therefore, we propose a much simpler approach. Since, most of the time, the impedance presented by the IC on its backside is in the order of 1 kΩ approaching the 50 Ω expected by the generator can be done by connecting in parallel to the IC a 50 Ω resistor, as it is shown in the schematic in Fig. 2.b.

C. Platform improvements in practice

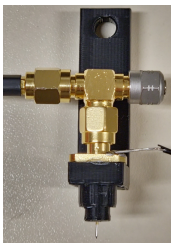


Fig. 3: Impedance matching in practice. `imp_match_real`

The proposed solution concerning the approximate impedance matching is shown in Fig. 3. The picture shows the BBI probe with a compensation load connected in parallel. To show the actual interests of these improvements, let us analyze signals from an actual platform.

We will compare before and after results and analyze the differences made by these improvements. To that end, we set up simple experiments consisting in injecting a voltage pulse into our IC target, measuring the voltage pulse at the probe and the current in the IC.

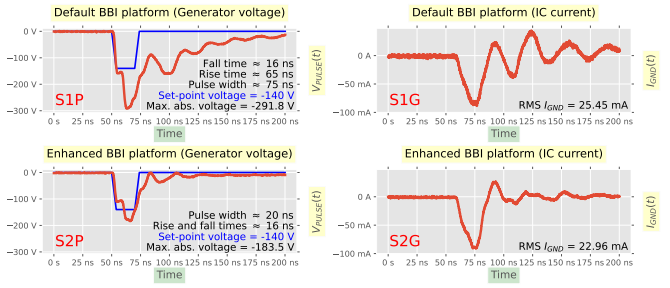


Fig. 4: Platform improvements in practice. `actual_imp`

Fig. 4 presents the waveform results of such experiments. The figure is split in two main parts, the top row shows the results before the improvements, and the bottom row shows the results after the improvements. The experimental conditions are the following:

- Voltage pulse amplitude = -140 V;
- Voltage pulse width = 20 ns;
- Rise and fall times = 4 ns.

The waveform S1P shows in blue the ideal waveform according to the generator settings and in red the measured waveform. In addition to this are annotated some noteworthy values. The first thing to notice here is the obvious undershoot of about -110 % under the set-point. It is far from being desirable when performing fault injection as the voltage amplitude is of great importance when considering the method effects on the IC. Furthermore, the pulse width is 275 & higher than the set-point, measuring 75 ns instead of 20 ns. It is an additional issue as it annihilates the accuracy needed in this context, and leads to longer pulses injected into the IC and potentially more energy than required. Additionally, the rise and fall times are also 4 to 16 times higher than expected. Eventually, we can notice damped oscillations, probably the proof of ringing in the transmission line.

Then, the waveform S1G, associated with the previous one, shows the IC ground current. Here, the damped oscillations are more clearly visible, in addition to the much longer than expected pulse duration. The RMS value of the injected current measures around 25 mA.

Afterwards, the waveform S2P shows the results with the proposed improvements The voltage pulse amplitude is much closer to the set-point, with an undershoot of -31 %. It is not perfect, but considering the simple nature of the impedance matching we propose, it was to be expected. On another note, the pulse width set-point is perfectly respected. However, the rise and fall times are still 4 times higher. Then, when looking at the S2G current waveform, we can remark the ringing reduction, while the amount of transferred energy remains approximately the same.

D. Further platform improvements comparison

To be able to illustrate further the actual interests of the proposed improvements, we did not only set up electrical measurements, but a complete differential fault attack (DFA). Indeed, performing fault injection is mainly used to perform attacks, therefore it makes sense to verify the soundness of the improvements in this context. We chose to perform a single bit DFA ou our IC target. The target embeds a dedicated cryptographic core, which we set up using a 128 bits AES. We then decided to perform the Giraud's DFA [14], originally described in 2002. This attack requires creating single bit faults in one or more bytes on the targeted AES. Our target was clocked at 40 MHz thanks to an external 8 MHz crystal, and powered with 3.3 V for the experiment.

1) Preliminary experiments: Before setting up the attack, we had to set up experiments allowing us to find the optimal locations in the AES sub-circuit where the attack could be performed. To do so, we created what we call Fault Analysis Mappings (FAM). These experiments consist in creating maps of a specific region of the IC, in that case the AES sub-circuit, and analyzing the IC behavior while performing BBI. The behavior is split into seven cases:

- Correct: the AES responds normally;
- Monobit Monobyte fault;
- Multibit Monobyte faults;
- Monobit Multibyte faults;
- Multibit Multibyte faults;
- Crash: the circuit did not respond correctly;
- Timeout: the circuit did not respond.

Therefore, as we need single bit faults, only two cases are valid for the Giraud's DFA. To be able to compare the two platforms, we performed a FAM on each one of them, using the following parameters:

- Pulse amplitude: from -150 V to -400 V with -5 V steps;

- Pulse width = 4.5 ns;
- Pulse delay of 150 ns + 553 ns targeting the penultimate AES round;
- A 40 mm displacement step.

Depending on the IC behavior, the experiments cant take up to 36 hours. The parameters were chosen to minimize the maximum energy transferred into the IC to avoid damaging it as much as possible.

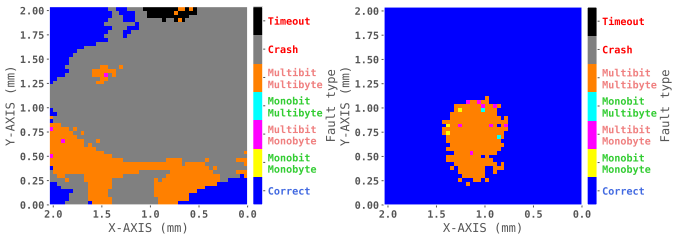


Fig. 5: Giraud's FAM `giraud_fam`

Fig. 5 presents the FAM results for both a typical and an improved platform. The mapped area encloses a little more than the actual AES, to be sure to map its entirety.

Let us look at Fig. 5.a first. What is interesting here is that we can spot numerous location where the circuit crashed. More specifically, they represent 70 % of the mapped area. This behavior is problematic in such an experiment as it cannot lead to any useful data for a DFA. Despite trying various experimental parameters, we could not observe single bit faults using this setup. However, multibit multibyte faults were easily observed, which was to be expected as they are easy to perform without much effort.

Then, let us discuss Fig. 5.b. The first interesting thing to remark is the total absence of IC crash. It is a desirable behavior as it indicates that we did not set a too high voltage pulse. Then, concerning monobit faults, we can spot five locations. It is a good sign for a preliminary experiment as it indicates the feasibility of such faults. It does not mean that we can perform the attack on one location. However, it means that we can use these locations as good starting points to perform the DFA.

2) DFA results: To perform the DFA, we focused on the five previously found monobit locations above the AES core. Then, for each location, we used the following parameters:

- Voltage ranging from -300 V to -600 V;
- Pulse width ranging from 4.5 ns to 5.5 ns;
- Injection delay ranging from ±10 ns around the penultimate AES round.

For each set of experimental settings, we had to set some limits when trying to inject faults. Indeed, it is required to create a finite experiment. The first limit consists in trying to retrieve a maximum of 100 single bit fault. Then, and because it cannot be achieved for every set of parameters, we set another limit of 10000 tries to achieve the previous limit.

Thanks to this, we retrieved, using the five locations, 14 out of 16 bytes of the AES secret key, as shown in Table I. The red cells indicate the two byes not retrieved thanks to the Giraud's DFA. To retrieve these last two bytes, we used a brute force method. Considering a slow laptop being able to compute approximately $10 \cdot 10^3$ AES encryptions per second, the 16 remaining bits representing 65536 combinations, we decided to blindly calculate every possibility. That represents around 6.5 seconds of total computation.

III. Modeling and Simulating BBI

SIMULATING a fault injection method behavior is an important part in understanding its mechanisms. Whether it is EMFI, LFI or BBI, it allows to predict and understand the underlying phenomena at work to set up reliable experiments. In this paper, we are focusing solely on BBI.

Ideally, we would want to directly observe signals inside integrated circuits, allowing for fine measurements of power supply voltages, logic levels and power current not to cite every physical quantity. However, embedding sensors into an already existing IC is not possible, and doing so on future IC is costly and takes time to fully implement. In addition to this, we do not have any guarantee that these sensors will not be disturbed too much by the fault injection. Therefore, we have decided to take the following approach:

$$\text{Simulation} \rightarrow \text{Conclusions} \rightarrow \text{Verification}$$

By doing so, we have freed ourselves from hardware limitations. However, other limitations remains. Indeed, modern ICs, even the smallest, embed millions of transistors, and with current technologies, it is impossible to evaluate with simulations entire circuits at a transistor level. Therefore, to tackle these limitations, we decided to adopt an hybrid approach, combining transistor-less models and local logic gates simulations. This approach is a compromise between accuracy and computational cost/time, and allows simulating relatively big circuits under BBI disturbances Overall, it is similar to what has been done for EMFI in [2]. The resulting simulation flow is divided in three consecutive steps:

- The simulation of an IC under BBI using a transistor-less model, allowing for a purely electrical analysis;
- The extraction of significant disturbed signals from the previous simulation;
- The simulation of functional logic gates under BBI thanks to the previously extracted signals.

A. An hybrid simulation flow: building the models

Building the correct models for the simulation flow pass through multiple steps. As the goal of the hybrid flow is to reduce the computational power required to evaluate an IC, it is still important to maintain a certain accuracy concerning the IC physical structure. To do so, the models are designed around actual IC implementations. The main building blocks of the models are the power supply network, the standard-cells, and the substrate structure. In this work, we are only focusing on bulk substrates: specifically dual-well and triple-well substrates.

#B	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
K10	0xFF	0x1F	0x42	0xE8	0xEF	0x44	0xA5	0x6A	0xCA	0xE7	0x55	0x3C	0xFD	0x65	0x39	0x26
KEY	0x01	0x23	0x45	0x67	0x89	0xAB	0xCD	0xEF	0xDE	0xAD	0xBE	0xEF	0x12	0x34	0x43	0x21

TABLE I: Giraud DFA

table_giraud

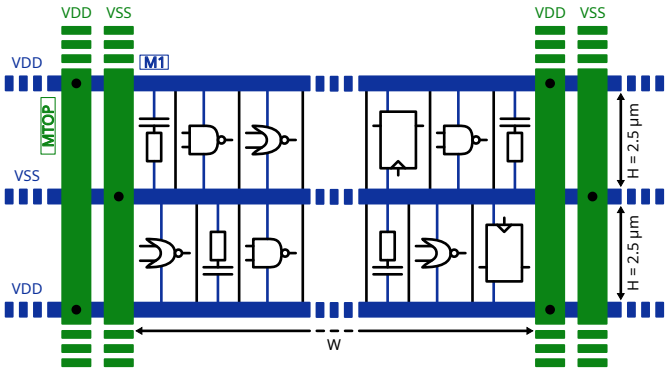


Fig. 6: A Standard-Cell Segment and its power delivery network.

fig_alim_std

1) Power supply rails and standard-cell segments: The power distribution inside an IC is typically made with a grid-like structure, composed of metal wires stacked on top of each other on planes. In each layer, the metal wires are equally spaced and have a dedicated width, which becomes thinner the deeper they are. The lowest layer brings the power directly to the transistors. Fig. 6 presents a common power delivery network, designed with two metal levels for simplicity.

Within the metal lines are located standard-cell segments (SCS), composed of decoupling, logic and sequential elements, and are pre-characterized by foundries and categorized depending on their performance (mainly but not exclusively power consumption and speed). As illustrated in Fig. 6, SCS have a constant height, in our case of 2.5 µm, and a variable width depending on how much logic gates each one of them embed. As we have stated previously, the hybrid simulation flow use transistor-less models as basic IC building blocks. Therefore, the transistors, hence the standard-cell segments, are modeled with passive elements such as resistors and capacitors.

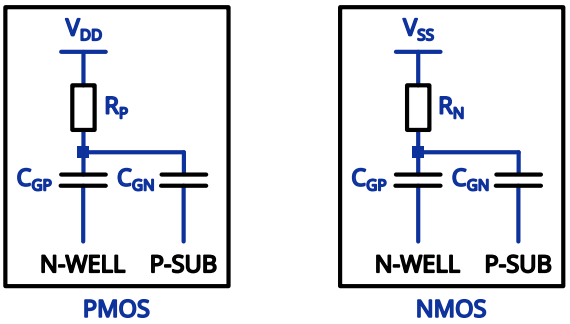


Fig. 7: aaa

mos_passive

To that end, the elementary SCS chosen measures 30 µm by 5 µm, representing two rows of logic cells. This represents about a hundred of logic gates, represented with four resistors and two capacitors, as shown in Fig. 7, with half of the transistors conducting, half not conducting. The conducting NMOS transistors, whose source is connected to V_{SS}, are equivalent to the passive resistor R_N. The conducting PMOS transistors, whose source is connected to V_{DD}, are equivalent to the passive resistor R_P. The resistors values depends on the considered technology, as well as the capacitors values, and can be adjusted and calculated according to one needs.

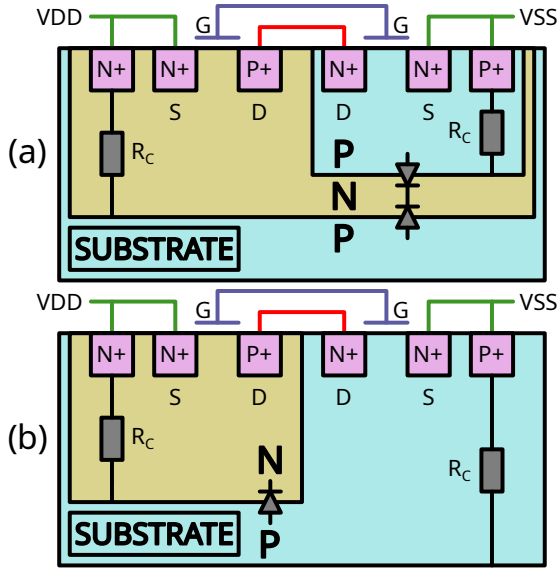


Fig. 8: Triple-well (a.) and Dual-well (b.) inverter cross-sectional view.

fig_sub

2) The substrate: Because BBI can be performed thanks to the silicon substrate as the main physical environment transferring energy from a generator to an IC, it is fundamental to elaborate a proper substrate model to precisely represent the various involved phenomena. As stated previously, our work focuses on bulk substrates, and in most cases, the substrate silicon is P-doped. There are two typical ways of lithographing the transistors in a bulk substrate, using dual-well or triple-well structures. Dual-well substrates are commonly found in moderately old circuits, while triple-well substrates are found in more recent circuits, while not bleeding-edge.

To properly understand how the differences between dual-well and triple-well substrates change the resulting model, let us analyze the cross-sectional schematics of an inverter created respectively in a triple-well and a dual-well substrate, as shown respectively in Fig. 8.a and Fig. 8.b:

- In the triple-well substrate, the NMOS transistors are lithographed into a P-doped silicon well, itself lithographed inside a N-doped well, buried inside the P-doped substrate. The PMOS transistors are located inside the N-doped well;

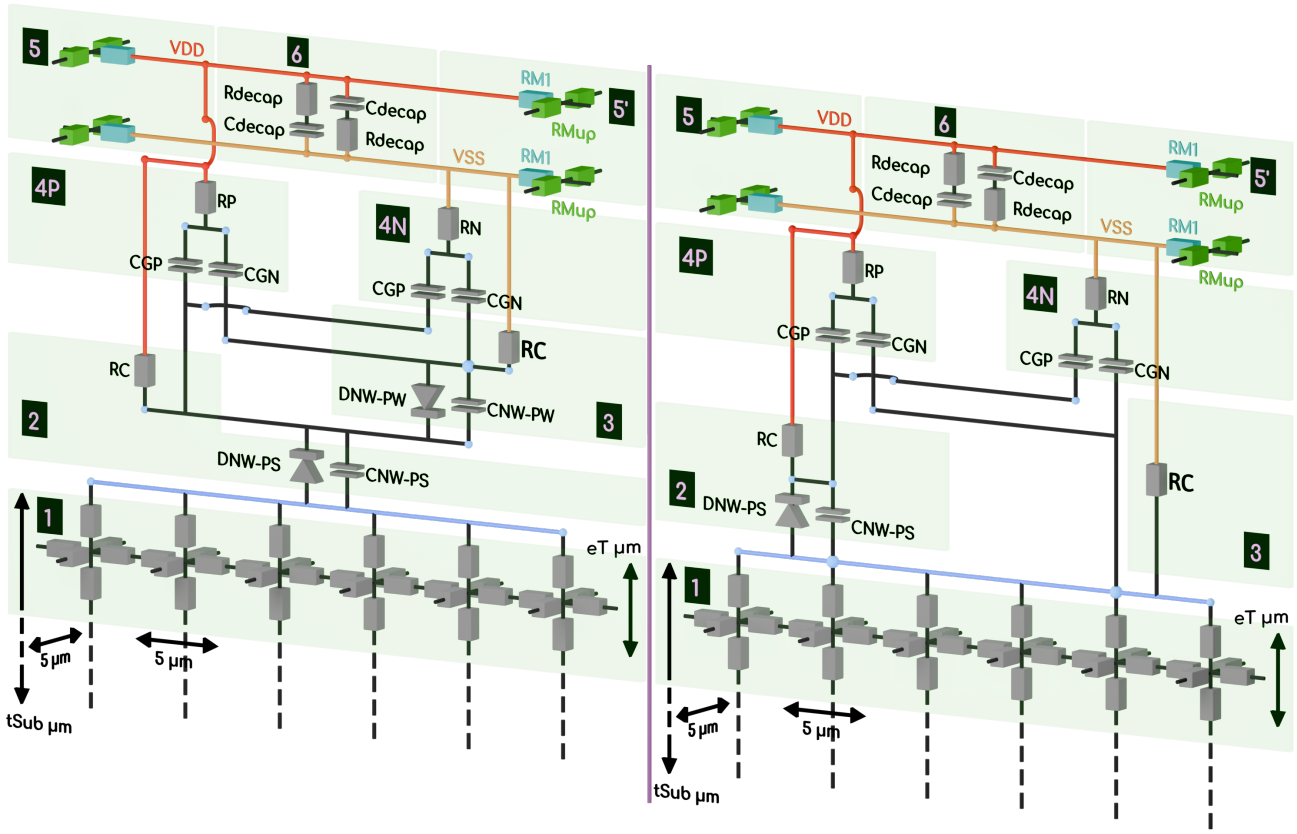


Fig. 9: Triple well (left) and dual well (right) std cell (PEUT ETRE FAIRE DES SOUS-FIGURES) fig_triplewellstdcell

- In the dual-well substrate, the PMOS transistors are still located inside the N-doped well, however, the NMOS are lithographed directly inside the P-doped substrate.

On the one hand, the triple-well substrate reveals two diodes:

- One formed between the P-well and the N-well;
- Another formed between the N-well and the P-substrate.

On the other hand, the dual-well substrate only reveals one diode between the N-well and the P-substrate.

3) The resulting model: Thanks to what we have introduced previously, we can now build the elementary building blocks for our hybrid simulation flow. It combines the power delivery network architecture, the equivalent logic gates models, and the substrate structure, all in an embedded model. This model represents an elementary section of the simulated IC, measuring 30 µm by 5 µm by t_{Sub} µm, the latter being the substrate thickness, a parameter which will vary depending on each considered IC.

As we consider both triple-well and dual-well substrate, there are two resulting elementary models, shown in Fig. 9. Each model is composed of various sub-regions, whose descriptions follow:

- 1 is the substrate network, divided into six sub-networks of six resistors for finer details;
- 2 is the first P-N silicon junction, common to both models;
- 3 is the access resistor (DW) or the second junction (TW);
- 4P is the PMOS equivalent section;
- 4N is the NMOS equivalent section;
- 5, 5' are the power supply metal layers (upper metal in green, first level in blue);
- 6 is the power supply decoupling.

As we have stated before, these models only represent a small portion of the modeled IC. To create an entire IC of a defined size, it is required to instantiate and interconnect as much as needed the elementary models. By doing so, we can create a bigger model of virtually any size. The language we have chosen to work with the simulation is the SPICE language. However, we created a custom Python script to interconnect the SCS together, place external power connections, and generate a SPICE file. For the current work, we decided to put the external power connections at the top and bottom of the IC (seen from above), and the BBI probe at the center of the IC (on the backside).

B. An hybrid simulation flow: performing simulations

Now that we set up the base models and their duplication, we can perform simulations with those models. To properly use these models, it is required, in the first place, to validate them through various steps to ensure their reliability. To that end, we generated an IC measuring 550 µm by 450 µm with a 140 µm substrate thickness, and performed an operating point to verify the correctness of the models for each substrate type.

Value	Triple-well	Dual-well
I_{GND}	2.88 nA	2.85 nA
I_{VDD}	-8.64 nA	-2.92 nA
$\overline{GND_{drop}}$	1.83 nV	1.76 nV
$\overline{V_{DD_{drop}}}$	1.2 nV	1 nV

TABLE II: op point

tab_op

We should expect almost no voltage drop and zero current consumption from such a model. Otherwise, it indicates an underlying issue with the model.

Table II shows the operating point results for both a triple-well and a dual-well circuit, and indicates a correct operating point, with idle currents and voltage drops close to zero. However, verifying the bias point alone is not sufficient to consider the model validated. As these models are dedicated to be mainly used in transient simulations, it is required to perform one and evaluate the soundness of its results.

Therefore, we performed transient simulations with a triple-well and dual-well IC, with the following parameters:

- A nominal power supply voltage of 1.2 V;
- A voltage pulse amplitude of ± 300 V;
- A voltage pulse width of 15 ns;
- Rise and fall times of 8 ns;
- A simulation duration of 80 ns;
- A simulation time step: of 50 ps.

C. An hybrid simulation flow: analyzing the results

Analyzing the simulation results involves observing various internal IC signals, for each substrate type, the ones presented in this section being:

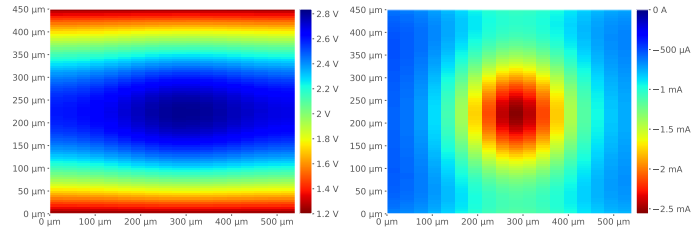
- The power supply voltage distribution;
- The epitaxial current;
- The substrate current distribution;
- The substrate pre-layer current density.

The observed signals are displayed in two dimensions and at the apex of the BBI disturbance. Each signal brings some insights on what happens inside the circuits during a BBI pulse. We will first analyze the dual-well results, then the triple-well ones, to finally conclude with a comparison of both.

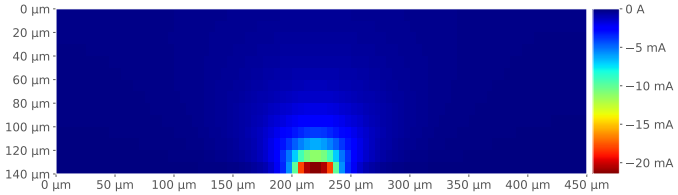
1) Negative dual-well simulation results: Fig. 10 shows the dual-well positive pulse results.

Sub-fig. left 10a represents the power delivery network (PDN) voltage across the entire IC as seen from above. In other words, it is the supply voltage of the transistors. Expectedly, far from the external power connections, we observe some deviation from the nominal 1.2 V power supply voltage. However, at the center of the circuit, in other words under the BBI probe, the voltage goes up to 2.8 V, being a 33 % increase from the nominal value.

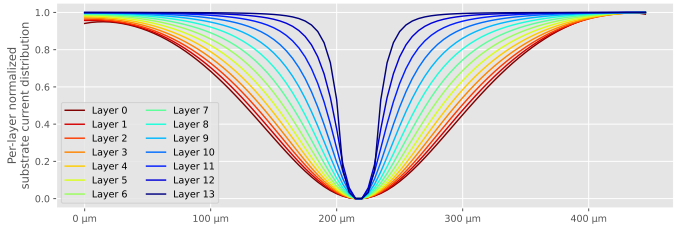
To put these values into perspective, let us look at sub-fig. right 10a, showing the epitaxial current distribution, representing the charges going from the substrate to the top of the SCS. According to the sub-figure, most of the charges are flowing at the center of the IC, under the BBI probe, as the current is the highest in that location. It is sound when comparing left and right sub-fig. 10a, as the voltage difference from the nominal value is higher where the epitaxial current is higher.



(a) Power delivery network (left), epitaxial current (right)



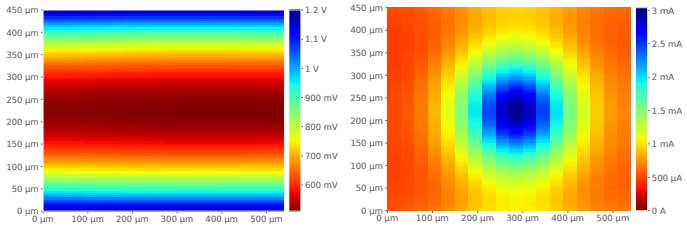
(b) Substrate cross-sectional view current



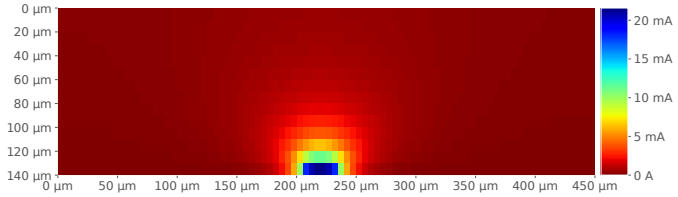
(c) Substrate per-layer normalized current density

Fig. 10: Dual-well IC negative pulse simulation results.

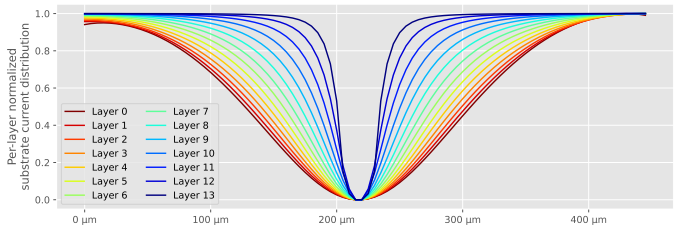
sim_res_dw_neg



(a) Power delivery network (left), epitaxial current (right)



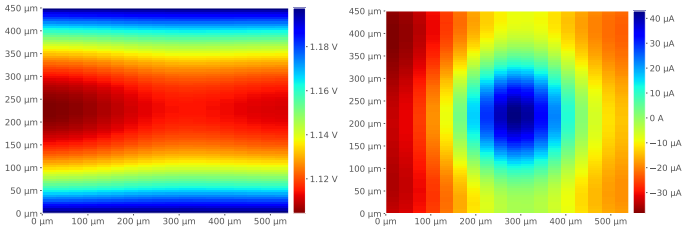
(b) Substrate cross-sectional view current



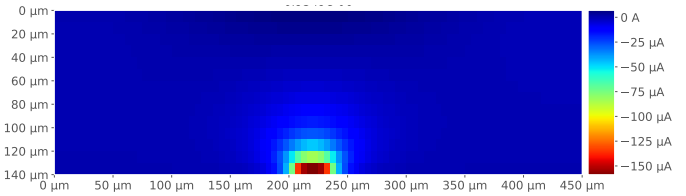
(c) Substrate per-layer normalized current density

Fig. 11: Dual-well IC positive pulse simulation results.

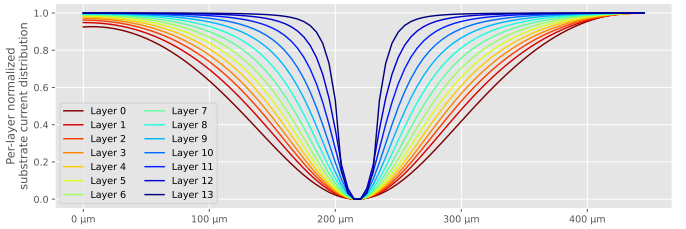
sim_res_dw_pos



(a) Power delivery network (left), epitaxial current (right)

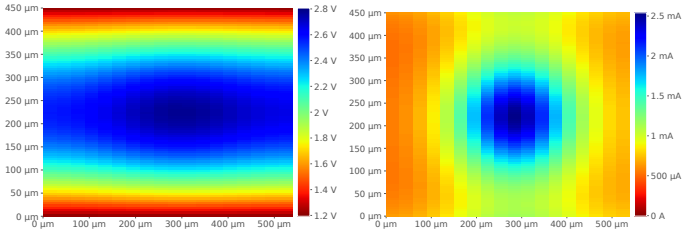


(b) Substrate cross-sectional view current

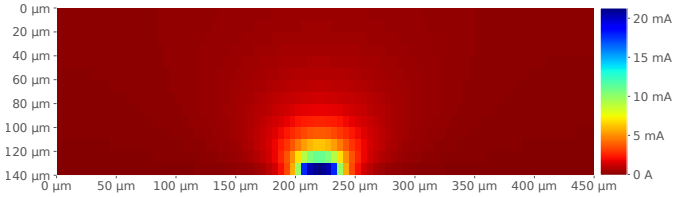


(c) Substrate per-layer normalized current density

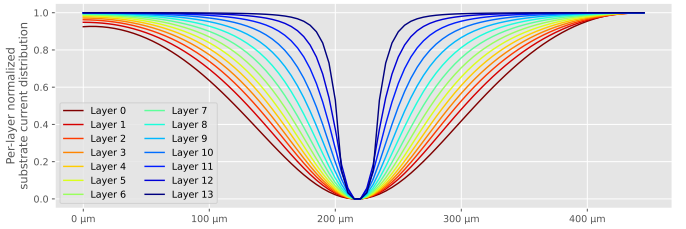
Fig. 12: Triple-well IC negative pulse simulation results. sim_res_tw_neg



(a) Power delivery network (left), epitaxial current (right)



(b) Substrate cross-sectional view current



(c) Substrate per-layer normalized current density

Fig. 13: Triple-well IC positive pulse simulation results. sim_res_tw_pos

Sub-fig. 10b and sub-fig. 10c both represent the same physical quantity in two different ways. We have chosen this approach to extract as much information as possible from these models and simulations. Sub-fig. 10b shows the cross-sectional view (from the Y-axis) of the current distribution inside the silicon substrate. The substrate being an isotropic environment, in other words, its resistivity is homogeneous in every spatial directions, we can observe a hemispheric current distribution in it. However, due to the large difference between the first layer (the farthest to the probe) and the last layer (the closest to the probe), it is difficult to do more observations. Therefore, we can look at sub-fig. 10c, which represents the same data in a different perspective. To better illustrate the inter-layer differences, we have chosen to normalize the data in a per-layer basis. Thus, it allows us to compare the current density between layers. It is important to note that the normalized values are calculated in a way that the closer they are to zero, the denser the current is, and vice-versa. The layer 0 is the closest to the logic gates, while the layer 13 is the closest to the backside (the probe). What is interesting to note here is that for each substrate layer, the current is focused where the probe is located. It is to be expected, as the substrate is isotropic. However, the deeper we are into the substrate, the less focused the current is. Once again, it is quite logical as the charges diffuse homogeneously inside the substrate.

2) Positive dual-well simulation results: Concerning the positive pulse dual-well results, let us look at Fig. 11. Compared to the previous results, sub-fig. left 11(a) shows that the PDN voltage exhibits not a voltage increase, but rather a voltage drop. Indeed, under the probe, the PDN voltage drops to 500 mV from 1.2 V. This is a substantial difference, which could lead, if applied to actual transistors, a significant change in behavior such as an incorrect biasing.

Concerning the epitaxial current, shown in sub-fig. right 11a, we can notice two key changes. First, the current polarity has changed, from a negative to a positive one. Once again, it was to be expected, as the voltage pulse polarity has changed. Then, in absolute value, the maximal current is 500 mV higher than previously, which indicates that more energy has been injected into the circuit. Eventually, regarding the substrate current, there are no major differences except the current polarity, both for sub-fig. 11b and sub-fig. 11c.

3) Negative triple-well simulation results: Let us take a closer look at Fig. 12. These results stand out all of the others, in many ways. First, if we take a look at sub-fig. left 12(a) regarding the PDN voltage, we can see that there are very little variations from the nominal voltage. Indeed, the voltage drops only to 1.1 V. Then, concerning the epitaxial current shown in sub-fig. right 12a, we can see that it is almost a hundred times lower than on other results. It is then confirmed in sub-fig. 12(c) with the substrate current distribution. However, the current density stays consistent with the previous results. Before analyzing further these results and explaining them, let us analyze the last case.

4) Positive triple-well simulation results: Quite interestingly, with a triple-well substrate and a positive voltage pulse, as displayed in Fig. 13, we observe results that are very similar to the dual-well negative case (Fig. 10), whether it is on the

Substrate	Polarity	Coupling NMOS	Coupling PMOS	Circuit	Danger
Dual-well	Negative	DC	AC	DC	☠☠☠
Dual-well	Positive	DC	DC	DC	☠☠☠☠
Triple-well	Negative	AC	AC	AC	☠
Triple-well	Positive	AC	DC	DC	☠☠☠

TABLE III: Caption dw_tw_table

PDN voltage or on the epitaxial current. Indeed, the PDN voltage disturbance is almost identical to sub-fig. left 10a, with an increase in voltage from 1.2 V to 2.8 V. Then, the epitaxial and substrate current maps are mirrors (in polarity) of sub-fig. left 10a and 10c. Eventually, the current density graph is very close to the other results.

5) Differences between dual-well and triple-well (negative and positive pulses): As we have seen through this section, we have four possible scenarios:

- A dual-well substrate and a negative voltage pulse;
- A dual-well substrate and a positive voltage pulse;
- A triple-well substrate and a negative voltage pulse;
- A triple-well substrate and a positive voltage pulse.

Each scenario behave differently than the others for one main reason: the electric coupling between the probe (substrate) and the SCS (logic). These differences in coupling are due to the substrate structure we encounter in dual-well and triple-well circuits.

As we have described before, the dual-well substrate embeds a P-N diode between the P-substrate and the N-well, and depending on the voltage pulse polarity, this diode is either blocking or conducting. This diode is interspersed between the substrate and the PMOS section. On the one hand, concerning the negative pulse scenario, the diode is blocking, thus creating an AC-coupling between the probe and the PMOS. On the other hand, the NMOS are DC-coupled to the probe as they are connected through a resistive path. Therefore, the circuit is globally DC-coupled to the probe, allowing the charges to flow all the time during the pulse. Then, concerning the positive pulse scenario, the diode conducts, creating another DC path to the transistors, reducing the effective circuit impedance seen by the probe. It explains the greater observed currents, as the charges have an additional DC path to follow.

On the triple-well side, the top of the SCS is barred with a first P-N diode (P-substrate N-well), and the NMOS are behind another diode. When using negative pulses, the first diode is blocking, therefore creating a pure AC-coupling between the probe and the circuit. It means that the charges are able to flow in and out of the SCS only on the pulse edges. Consequently, for a given voltage pulse, less energy is transferred into the IC in that case.

Then, regarding the triple-well positive scenario, the first diode become conducting, while the second stays blocking. Therefore, the PMOS are DC-coupled, while the NMOS stay AC-coupled. We come back to a scenario similar to dual-well negative.

Eventually, the main outcomes these simulation results show are in Table III, alongside a qualitative dangerousness appreciation of each scenario.

IV. Validating and Completing the Models

A. Validation the models

With the aim of verifying the soundness of the previous conclusions, we set up experiments using an actual IC composed of both triple-well and dual-well substrate on a monolithic die. These experiments consist in verifying if the difference in injected energy depending on the substrate type is actually significant or not.

The target used is a STM32F439 microcontroller, alongside the platform presented in the first chapter. The IC die measures approximately 5.5 mm × 4.5 mm. We call these experiments "IC ground current mapping", and quite naturally, they consist in measuring in specific conditions the current at the target circuit external ground connection. The entirety of the IC is mapped, and a voltage pulse is injected at each location. Then, we measure the current at the circuit ground and calculate its RMS value to represent it into a two-dimensional cartography.

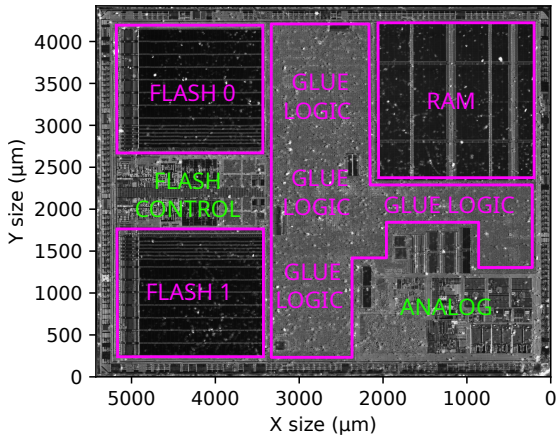


Fig. 14: Caption stm_ir_photo

Knowing the coarse structure of the considered IC, in addition to having insights on the substrate type, we could draw the coarse structure picture shown in Fig. 14. The "glue logic" regions are known to be made with triple-well substrates, while the "flash control" and "analog" regions are made with dual-well substrates. The memories, however, are made of a mix of both.

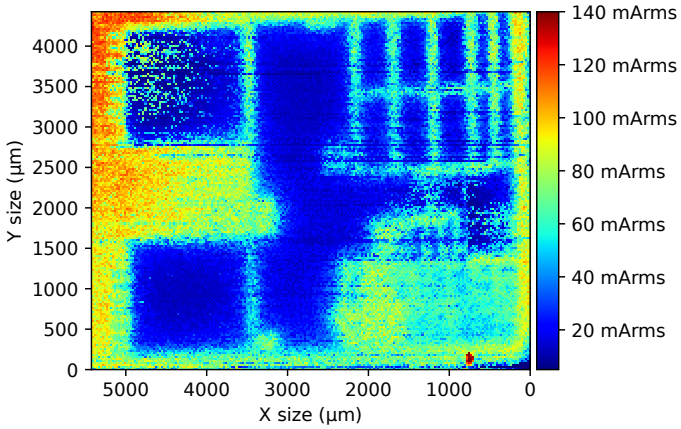


Fig. 15: Caption stm_ignd

The experimental results are shown in Fig. 15, and the experimental parameters are the following:

- Negative voltage pulse of 70 V amplitude;
- Pulse width of 20 ns;
- IC substrate thickness of 50 µm.

The voltage pulse used is of negative polarity as we have observed a very fast degradation of IC subjected to positive voltage pulses, therefore we decided to avoid them at all cost. When analyzing the results, we can notice significant differences in the measured current depending on various regions, and the IC floorplan seems to draw itself on the current map. The measured RMS current ranges from 10 mArms to 140 mArms, and as predicted by the simulation results, in the regions where the substrate is of dual-well type, the current is higher than on regions where the substrate is of triple-well type, such as the analog block or the flash control region.

These observations confirm the soundness of the proposed models. However, as we have seen previously, these models do not consider the functional nature of the considered ICs: their logic behavior. To circumvent this limitation, we decided to develop an addition to the initial simulation flow, thus the name ”hybrid simulation flow”.

B. Completing the models

As we have stated previously, it is required, to complete the models, to properly consider the logical behavior of the considered circuits, which allows for a better appreciation of BBI induced effects and their consequences. These additional steps consist in modeling actual logic and sequential elements in the same or in a close technology as the considered IC, while extracting the significant disturbed signals from the SCS simulation and injecting them into these logic devices. For this purpose, split this section into two subsections:

- A first section dedicated to studying a static logic gate: the classical inverter;
- A second section dedicated to studying a sequential element: the DFF.

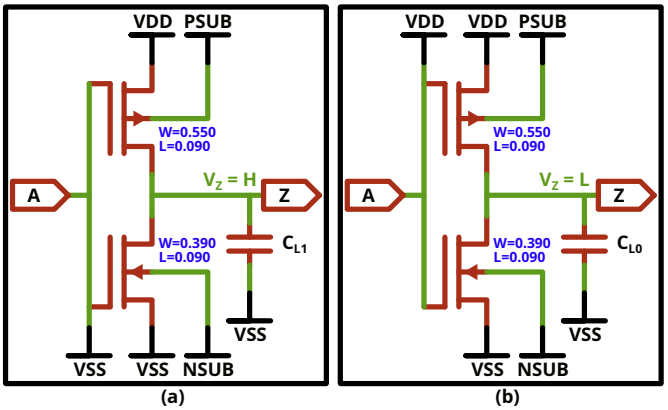


Fig. 16: Inverter schematic ivxbufmos

1) Static inverters under BBI: Because inverters can have two stable output states, we will consider two cases for each substrate scenario: a normally high inverter (Fig. 16.a) and a normally low inverter (Fig. 16.b). The inverters are connected to four external signals which are extracted from the previous SCS simulation:

- VDD: the power supply voltage;
- VSS: the power supply reference voltage;
- PSUB: the bulk voltage of the PMOS transistors;
- NSUB: the bulk voltage of the NMOS transistors.

The voltages PSUB and NSUB depend on the substrate type. On the one hand, in the dual-well scenario, NSUB is connected to the epitaxial layer, while PSUB is connected to the N-well. On the other hand, in the triple-well scenario, NSUB is connected to the P-well and PSUB to the N-well.

All of this gives us four scenarios to study. For clarity and because two of the four scenario are less noteworthy, we will only talk about two of them:

- The triple-well substrate and the normally high inverter;
- The dual-well substrate and the normally low inverter.

Then, for each scenario, we will analyze seven signals of interest:

- The backside voltage pulse, for reference purposes;
- The local differential power supply voltage;
- The current sum of the inverter;
- The inverter load current;
- The inverter output;
- The NSUB voltage;
- the PSUB voltage.

The signals extracted from the SCS simulations come from the standard-cell located directly below the BBI probe, a.k.a the cell targeted by the injection.

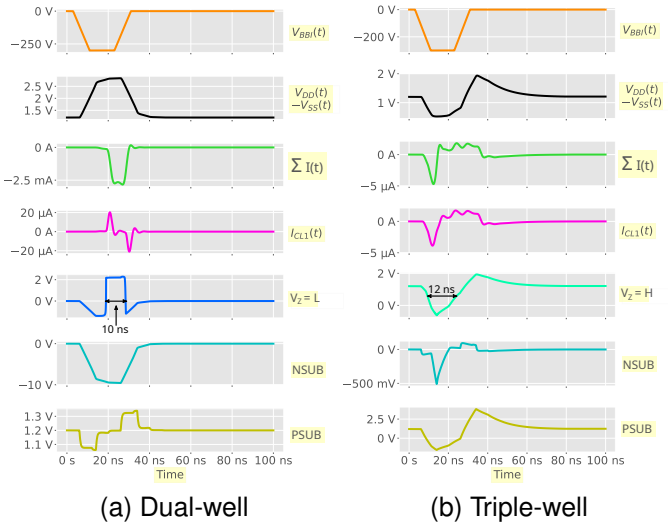


Fig. 17: Inverters simulation results ivxsimu

Fig. 17 presents the inverter simulation results for both considered scenarios.

First, let us focus on the dual-well inverter. The corresponding schematic is shown in Fig. 16b, and the simulation results are shown in Fig. 17a. In that case, as we have seen before, the global IC coupling is resistive, with a discrepancy

between VSS and VDD. Therefore, the inverter current sum (green) follows a DC-response, similar to the differential power delivery voltage (black). The inverter output follows the current sum curve, and its output goes from a low to a high logic value during 10 ns, then back to its original state. It is further corroborated by looking at the load current, which is charged on the first pulse edge, then discharged on the second one.

Second, concerning the triple-well inverter, where the results are shown in Fig. 17b and the schematic in Fig. 16a, the substrate is globally AC-coupled. It can be seen on the current sum curve, which follows almost exactly the capacitive load current curve. The inverter output, for its part, is discharged like the load, and goes from a high logic value to a low value during 12 ns before returning to its original value.

These observations are of great value because we can discuss a fault model for BBI, similar to what has been studied for EMFI and LFI. The previous results seem to indicate that the faults created using BBI are data-dependent. Indeed, if we lower the voltage of an inverter outputting a low value, or the opposite, it has theoretically no direct effect on the logic value. However, we have seen that it is possible, depending on the substrate type, to observe bit set or reset. Eventually, thanks to these results and the previous ones regarding current density in the substrate in Fig. 10, 11, 12 and 13, it seems that BBI effects are local.

REFERENCES

[1] 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.

[2] 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.

[3] Jean-Max Dutertre, Vincent Beroulle, Philippe Candelier, Stephan De Castro, Louis-Barthelemy Faber, Marie-Lise Flottes, Philippe Gendrier, David Hély, Regis Leveugle, Paolo Maistri, Giorgio Di Natale, Athanasios Papadimitriou, and Bruno Rouzeyre. Laser fault injection at the cmos 28 nm technology node: an analysis of the fault model. In *2018 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, pages 1–6, 2018.

[4] Philippe Maurine, Karim Tobich, Thomas Ordas, and Pierre-Yvan Liardet. Yet another fault injection technique : by forward body biasing injection. *"Yet Another Conference on Cryptography France (2012)"*, 09 2012.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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 2IRPS)*, pages 1–6, 2023.

[10] 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.

[11] 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.

[12] G. Chancel, J.-M. Galliere, and P. Maurine. A better practice for body biasing injection. In *2023 Workshop on Fault Detection and Tolerance in Cryptography (FDTC)*, pages 48–59, 2023.

[13] Colin O'Flynn.

[14] 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.