A Tale of Two Models: Discussing the Timing and Sampling EM Fault Injection Models

Roukoz Nabhan, Jean-Max Dutertre, Jean-Baptiste Rigaud, Jean-Luc Danger, Laurent Sauvage

A Tale of Two Models: Discussing the Timing and Sampling EM Fault Injection Models

Roukoz Nabhan*, Jean-Max Dutertre*, Jean-Baptiste Rigaud*, Jean-Luc Danger† and Laurent Sauvage†
*Mines Saint-Etienne, CEA, Leti, Centre CMP, F-13541 Gardanne, France
†LTCI, Télécom Paris, Institut Mines-Télécom, 91120 Palaiseau, France
*{roukoz.nabhan, dutertre, rigaud}@emse.fr †{jean-luc.danger, laurent.sauvage}@telecom-paris.fr

***Abstract*—Investigating the dynamics and mechanisms of Electromagnetic Fault Injection (EMFI) attacks, which expose an active circuit to electromagnetic disturbances, presents a persisting challenge due to the diverse and complex fault mechanisms involved. An improved understanding of EMFI modeling is paramount for developing proficient on-chip detection sensors, serving as countermeasures to these attacks. In light of this, our research evaluated the effectiveness of EMFI detection sensors, introduced by Elbaze et al., which rest on the premise that the sampling fault model accounts for EMFI. To assess the functionality of these sensors, we integrated them into an Advanced Encryption Standard (AES) accelerator of a Field-Programmable Gate Array (FPGA) and performed a series of experiments. The resulting evidence suggests that the explanation for EMFI is not a singular fault model but rather, two underlying mechanisms are implicated. At high frequencies, which corresponds to low slack, electromagnetic disturbances, in tandem with the target's Power Distribution Network (PDN), initiated timing constraint violations. This violation subsequently increased the logic propagation times, surpassing the clock period. Contrarily, at low to moderate frequencies, the induced faults generally aligned with the sampling fault model. However, certain deviations from the theoretical framework called into question the model's validity. Upon a deeper examination of the results, we determined that these faults, rather than being sampling faults, were tied to a different mechanism. Electromagnetic disturbances, when coupled with a target's Clock Distribution Network (CDN), can cause timing constraint violations due to EMFI-induced voltage glitches within the target's clock tree. By integrating the mechanisms of EMFI-induced clock glitches and timing faults into the timing violations fault model, we attain a holistic comprehension of EMFI mechanisms. It encapsulates both mechanisms induced by EMFI, spanning the full-frequency spectrum of the target.**

***Index Terms*—EMFI, timing violations fault model, EMFI-induced clock glitches, timing faults, sampling fault model, fully digital sensor.**

I. Introduction

The protection of IoT devices remains a challenge. The confidentiality and integrity of their data is continually at risk, predominantly due to hardware attacks, the most notable being Fault Injection Attacks (FIA). These attacks specifically aim to force faults during the computational processes of these devices. They allow attackers to employ efficient secret extraction methods such as a differential fault analysis to retrieve, for instance, the cryptographic key of an embedded AES [1]. Several FIA techniques exist in the literature. The focus of our research is on the ElectroMagnetic Fault Injection (EMFI) attacks [2], [3]. From the attacker's point of view, EMFI offers enticing advantages. As it is efficient and local [4], faults can be injected into a selected part of a target, it does not necessarily require the chip package to be opened, and it is more affordable than laser FIA [5]. One way to prevent such attacks is to design sensors that detect abnormal phenomena leading to fault creation. Developing effective on-chip detection sensors as countermeasures against EMFI attacks means that it is crucial to study the mechanisms involved in injecting faults due to EM disturbances. Numerous efforts have been carried out in the state-of-the-art [2], [3], [6], [7] to understand and elucidate the origins of EMFI-induced faults. However, this field is still relatively young, requiring further investigation and experimentation to achieve a deeper understanding of the impact of EM disturbances on a chip.

In a recent study, Dumont et al. [3] offered an electrical-level analysis of the so-called sampling fault model [8], [9], attributing it as the fundamental cause of EMFI. In the pursuit of an efficient embedded digital sensor that provides robust defense against EMFI, the fully digital detector developed by Elbaze et al. [10] emerged as a viable candidate. This sensor's design is based on a sampling fault model which has been evaluated in [10] and more recently expanded upon in [3]. Through our current research, we evaluate the performance of this sensor [10] within the AES accelerator of an FPGA, testing its functionality across the full-frequency range of the target.

The sensor turned out to be effective at low or moderate frequencies. Despite initial expectations, the observed faults did not align with the sampling fault model but instead, they were found to be consistent with a different mechanism: EMFI induced voltage glitches in the clock network. Upon experimentation, we determined that the coupling of EM disturbances with a target's Clock Distribution Network (CDN) induced clock glitch in the clock tree. This ultimately leads to faults prompted by timing constraint violations. Furthermore, when the target's clock frequency was increased to reach the limit of operating conditions, the sensor was found powerless. The undetected faults uncovered another EMFI mechanism; the timing faults resulting in an increase in the logic propagation times exceeding the clock period due to the coupling of the EM disturbances with the target's Power Distribution Network.

This paper represents a significant advancement in the understanding of EMFI models, integrating two mechanisms. It provides useful insights in the field of secure circuitry for

This work was supported by the French National Research Agency (ANR) under grand ANR-20-CYAL-0007

designing sensors based on a complete fault model.

Our contributions are outlined as follows:

- Discovered that electromagnetic-induced faults can occur through two distinct mechanisms within the timing violations fault model: timing faults, which result from EM coupling with the PDN, and EMFI-induced clock glitches within the clock network.
- Identified the injected faults did not align with the sampling fault model.
- Detailed the necessary conditions for the injection of timing faults, investigated the EMFI effects induced by clock glitches on the target's CDN and conducted an in-depth analysis of EMFI-induced clock glitches.
- Underscored the potential risks of using an EMFI detection sensor based on an incomplete fault model, while providing an enhanced explanation of the EMFI models to aid designers in developing more effective detection sensors.

The remainder of this paper is organized as follows: section II provides a reminder of the principle of EMFI and the related works which consider various mechanisms of EMFI models. Section III details our experimental setup: architecture of the embedded digital detection sensor, targeted test vehicle, and EM injection equipment. The experimental results are reported and analyzed in section IV with a focus on the corresponding injection mechanisms. The results obtained are further discussed in section V. Section VI reveals the effects of EMFI, specifically those induced by clock glitches, on the CDN. An extensive examination of the clock glitches triggered by EMFI is undertaken in section VII. Finally, section VIII provides a concise summary of our research findings.

II. Background

A. Electromagnetic fault injection principle

EMFI attacks are based on the generation of an EM disturbance close to an Integrated Circuit (IC). This is achieved by sending a voltage pulse with sharp transitions into an EM probe (made of a few copper wire loops around a ferrite core) located over a chip. EMFI has a local effect [2], [4]. These localized EM disturbances induce a transient voltage within the chip, corrupting its normal operation and causing digital faults.

B. EMFI models

The EM disturbance output from the EM probe induces a fault through its coupling with the target's on-chip main networks. Three on-chip networks have been accounted for in the literature [6] linked to different mechanisms and fault models: the Power Distribution Network (PDN), the clock distribution network, and the reset (or set) network. The following sections provide further elaboration.

1) Power Distribution Network - Timing constraint violations fault model: This was the initial hypothesis put forward to explain EMFI [2].

a) Timing constraints: Synchronous digital circuits are required to fulfill the setup timing constraint expressed in eq. 1:

$$T_{clk} > D_{clk2q} + t_{p_{max}} + t_{setup} - t_{skew} \quad (1)$$

where T_{clk} represents the clock period, D_{clk2q} is the delay to update the output data of a D Flip-Flop (DFF) after a clock rising edge. $t_{p_{max}}$ is the maximum propagation time through the target's logic gates. t_{setup} is the required time for a DFF's input data to be stable before a clock rising edge, and t_{skew} is the slight phase difference between clock signal inputs of all DFFs.

According to [4], any violation of this constraint results in the injection of faults, which aligns with the so-called timing fault model.

b) Timing Slack: The time margin related to (1) is called the timing slack, i.e., the difference between required arrival time and actual arrival time at the input data of a register [4]. It should be positive in order to meet timing requirements and to avoid timing violations.

c) Timing fault model mechanism: Dehbaoui et al. [2] suggested that EMFI could induce timing faults. It provokes a transient decrease of the supply voltage, thus inducing an increase of the target's logic gates propagation delay t_p. At a given level of increasing t_p, a timing constraint violation occurs and a fault is injected [11]. The timing fault model has four main characteristics:

- Faults are injected into the target's critical paths.
- Fault incidents gradually escalate with rising EM stress.
- Faults are contingent on input data, given that logic propagation times are data-dependent.
- The faults that arise are 100% reproducible.

2) Power Distribution Network - Sampling fault models: In contrast to the timing fault model, Ordas et al. [8], [9] introduced the sampling fault model by observing the effect of EM disturbances on the DFF's sampling operation around the clock rising edge which is at a vulnerable moment during an IC operation. They illustrated the existence of temporal windows around the clock rising edge called EMFI susceptibility windows. Within these windows, the probability of injecting faults is maximal and limited elsewhere. The width of these windows is constant and independent of the clock frequency.

Later, Dumont et al. [3] proposed a theoretical explanation for the sampling faults thanks to an electrical modeling of the effect of EM Pulses on the PDN of a generic IC model. The EM disturbances induced a temporary and local inversion in the polarity of the DFF power supply, freezing the IC operation. When DFF returned to its normal operating conditions near a clock rising edge, a bit-set or bit-reset fault occurred depending on the polarity of the pulse.

3) Clock Distribution Network - Timing constraint violations fault model: Ghodrati et al. [6] provided on an experimental basis another explanation of EMFI: that the coupling was made with the target's clock distribution network and that it induced clock glitches. This was illustrated based on experimental results on a RISC microprocessor (LEON3-design, 180 nm TSMC).

4) Reset Network - Reset fault model: Ordas et al. [7] showed the possibility to induce faults in a logic circuit at rest (no clock). This research illustrates the possibility of an EM coupling between the EM probe and the asynchronous set and reset signals of the target's DFFs that then produced bit-set or bit-reset faults.

III. Detection sensor and experimental setup

A. Sampling fault-based digital detection sensor

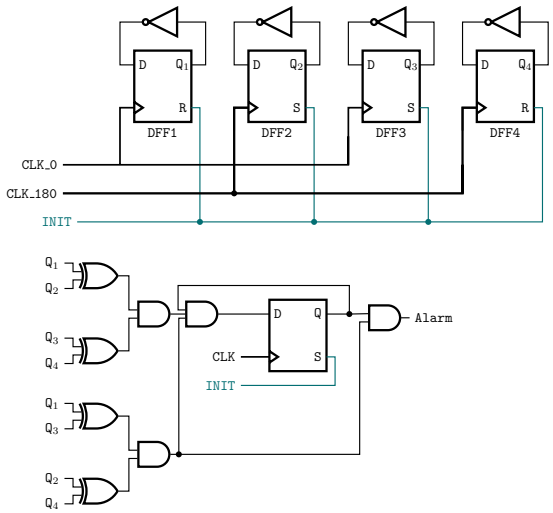


Fig. 1: Digital detection sensor architecture [10].

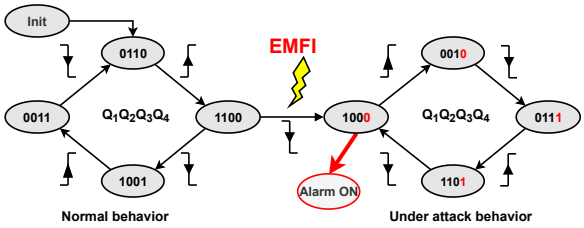


Fig. 2: Digital detection sensor behavior under normal conditions and under attacks.

Our research was centered on a digital detection sensor. By design, this sensor is finely tuned to the sampling fault model, enabling it to detect EMFI attacks [3]. It is based on the principle that EMFI occurs according to the sampling fault model [8], [9] and therefore that a sensor taking advantage of this mechanism is expected to detect any EMFI attack. According to this model, faults are injected in a target DFFs when they change state (i.e., either a $0 \rightarrow 1$ or $1 \rightarrow 0$ transition) at the clock rising edge. Hence, the EMFI sensitivity windows span periodically around clock rising edges. Fig. 1 (top part) provides the sensor's architecture. Two DFFs (DFF1 and DFF3) are clocked by the target's main clock signal and their output is inverted and fed back to their input. As their states are initialized at 0 for DFF1 and at 1 for DFF3, both $0 \rightarrow 1$ and $1 \rightarrow 0$ transitions occur at each clock rising edge. In case of an EMFI attack, the course of these transitions will be disturbed raising an alarm. Two additional DFFs (DFF2 and DFF4) are added according to the same principle but clocked by an inverted clock signal adding additional detection windows centered on the main clock falling edges.

The left-hand side of Fig. 2 depicts the course of the transitions of the four DFFs from an initial state $Q_1Q_2Q_3Q_4 = 0110$ under normal operation conditions. The right-hand side of Fig. 2 illustrates the effect of EMFI on the sensor transitions. The pathway deviates from its normal course, a divergence that is detected by a specialized logic block depicted in the bottom of Fig. 1. This block is responsible for generating an alarm signal when such a deviation occurs.

B. Experimental setup

1) Target FPGA: We used the Nexys Video 7 board [12], which embeds an Artix-7, XC7A200T FPGA manufactured in a 28 nm CMOS technology.

2) Pulse generator: An AV-Tech voltage pulse generator was selected for this study. This device can generate pulses with amplitudes up to +/- 750 V and pulse widths ranging from 4.5 ns to 20 ns. The amplitude, width, and delay of the pulse were adjusted using its Ethernet connection.

3) EM injection probe: We used a homemade EM probe. The probe was composed of a 0.2 mm diameter enameled copper wire, wound 4 times around a cylindrical ferrite core with a diameter of 2 mm (the target's chip area is 12*11mm).

4) Mixed-Mode Clock Manager - MMCM: This module was used to generate multiple clocks with a defined phase and frequency. It allowed the remote control of the target by dynamically changing the target's clock frequency without modifying the bitstream file. Additionally, we employed the MMCM to generate the sensor's clock signals. This ensured perfect synchronization between the primary clock signal and the 180° phase-shifted clock signal.

5) Block diagram: To investigate the efficiency of the sensors at detecting EMFI, as well as to study further the related mechanisms, the sensors were embedded in an AES accelerator. The full design consisted of:

- A hardware 128-bit AES accelerator that execute a full encryption in 11 clock cycles.
- A serial data link (UART) for communication purposes (it uses a distinct and fixed clock signal set to 100 MHz).
- A finite state machine (FSM) that controlled the execution flow of the target. Its clock frequency was fixed at 100 MHz.
- A MMCM block to generate the various clock signals of the design.
- A block of 16 EMFI detection sensors evenly placed within the AES accelerator.
- A Key calculation block delivering the keys of the AES computations rounds.

6) Maximum DUT clock frequency: The max. DUT clock frequency was measured above 200 MHz ($t_{critical} \approx 4.5$ ns as clock period).

IV. Experimental results and fault model analysis

A. Floorplan consideration

Fig. 3 provides the floorplan of the design, extracted from the Vivado tool, and its EMFI sensitive areas. On the right part

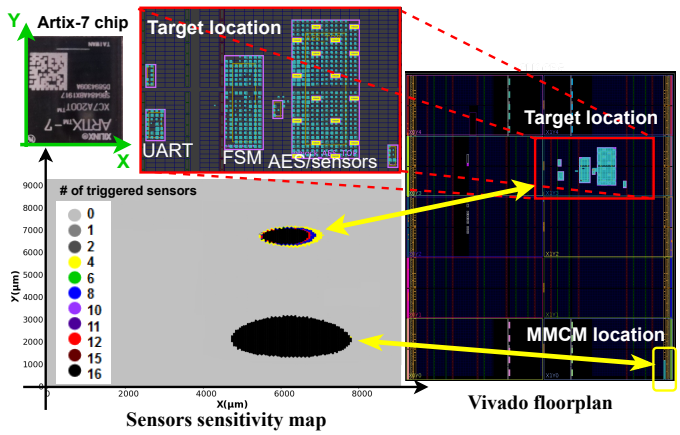


Fig. 3: Target's floorplan from Vivado, on-chip location of the design's building blocks and its EMFI sensitive areas.

of Fig. 3, the AES blocks are placed away from the MMCM to differentiate the EM disturbance effects on it from that on the clock generation block. A close-up (top center) shows the AES encryption block and how the 16 sensors are regularly distributed in it (sensor locations are highlighted in yellow). The sensors triggered an alarm when exposed to EM disturbances as the injection probe was located above two different areas shown in the EM sensitivity map depicted on the bottom left of Fig. 3. A color code shows the number of triggered sensors for each location of the EM probe. The bottom EM sensitive area corresponds to the MMCM location. The fact that it is drawn in black indicates that all 16 sensors were triggered simultaneously. The top EM sensitive area matches the location of the AES accelerator. Its color progressively changes from yellow, to blue, red and black, as the EM injection probe is moved from its edges to its inner part (from 4, to 8, 12 and 16 triggered sensors). Note that we rigorously ascertained these correspondences between the design logic blocks and the EM sensitive areas by testing several different locations of the logic (AES and/or MMCM) on the FPGA floorplan and observing the effect it had on the sensitive areas location.

B. Experimental methodology

The EMFI sensitivity map of Fig. 3 was drawn spatially for different locations of the EM injection probe above the target. Prior research, specifically [10] and [4], have already thoroughly explored the ability of embedded sensors to cover the physical area of a target from EMFI. Specifically, these studies addressed how detection sensors should be positioned and spread within the protected logic to avoid weak areas that could allow undetected fault injections to occur. The research objectives of our work differed from these studies. Our aim was to test the intrinsic detection ability of a sensor constructed based on the sampling fault model when used over the complete frequency range of a target. Therefore, the EM injection probe was positioned in the center of the AES accelerator sensitive area (a place where it should be at its best efficiency) for the experiments reported hereafter. According to this methodology, the explored injection parameters were the frequency of the AES and the timing of the applied EM disturbance with respect to the clock edges.

When it comes to discussing the overhead of detection sensors which can be expressed in terms of additional logic and time penalty, frequency is a key point. If a sensor requires the operating frequency to be set significantly below the target's maximum frequency, it involves a significant timing overhead. A low overhead sensor normally keeps its EMFI detection ability unimpaired when the operating frequency is close to its maximum, save for a minimal slack margin. In addition to that, testing an EMFI sensor for large injection settings also provides insights into the related mechanisms.

For each test series, the obtained results were analyzed using three metrics in alignment with our research objective:

- **Alarm:** if one of the 16 sensors was triggered, an alarm signal was raised.
- **Faulted Bits and Bytes (or FBB):** the number of faulted bits and bytes read from the AES ciphertext.
- **Alarm Failure (or AF):** expressing a failure in the sensor's detection ability (raised when an undetected fault is observed).

C. Detectors performance at several clock frequencies

In our first experiments, we set the pulse width to 4.5 ns, and its amplitude to 420 V (greater than the threshold amplitude of 380 V required to inject faults in every AES rounds and to trigger the sensors). Each campaign went through the whole AES rounds with a time step of 0.1 ns. In the following curves, the Alarm, FBB, and AF metrics are drawn as a function of the EM injection timing (expressed as the voltage pulser delay from a trigger signal). The results of 20 iterations are averaged at each time position.

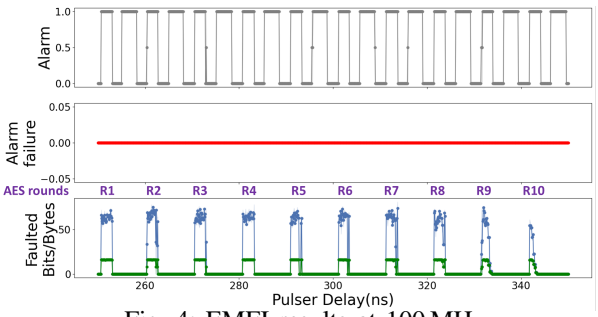


Fig. 4: EMFI results at 100 MHz.

1) Clock frequencies ≤ 150 MHz: Fig. 4 shows the results of the campaign launched at 100 MHz. Regarding sensor behavior, AF remained null throughout the campaign, as shown by the red curve, proving that all injected faults were detected. The gray curve, representing the alarm states, showed continuous Detection Windows (DW) with a width of 2-3 ns. The DWs are spaced with a half-clock period. Regarding the fault injection behavior in AES, we obtained 10 consecutive Injection Windows (IW) showing the FBB in blue and green curves respectively. These IWs indicated the position of the clock rising edge corresponding to the AES computation round. These windows were spaced with a period equal to the clock period, which was

set at 10 ns. Their width was 2-3 ns. As the number of faulted bytes remained 16 bytes for all round calculations except the last 2 rounds, we counted the number of faulted bits to improve the accuracy and interpretation of our results. Indeed, the last 2 IWs corresponding to rounds R9 and R10 revealed a decreasing trend in the number of FBB (as the injected faults had less opportunity to spread through the remaining AES operations). Key extraction techniques typically require a small number of FBB [1]. Therefore, it is feasible to configure the injection parameters to meet these specific criteria. The average number of faulted bits was 64 bits, oscillating between a range of 50 to 70 bits. The DWs aligned with the IWs mark the main clock rising edges. The other DWs mark the sensor's extended ability to detect EM disturbances around the main clock falling edges. Having two DWs for each clock period is a feature of the chosen sensor (see III-A). Hence, the faults injected during this test series are consistent with the sampling fault model (as expressed in section II-B2). This analysis is further supported by the fact that all the injected faults were detected by the sensors constructed on the same fault model.

Multiple campaigns were conducted, wherein the clock frequency was varied from 10 MHz to 140 MHz while maintaining the same parameters. All results from these campaigns showed consistent behavior with respect to the sampling fault model and the high detection rate. The widths of the Detection Windows (DW) and Injection Windows (IW) were recorded for later analysis to explore their relationship with the clock frequency.

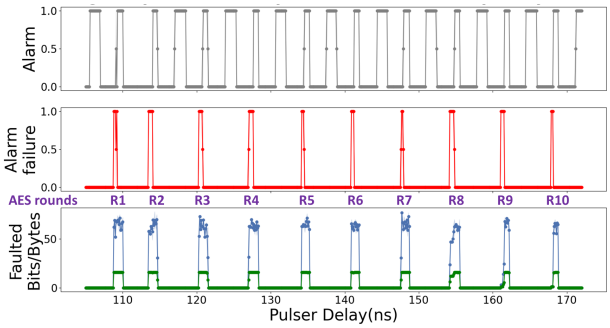


Fig. 5: EMFI results at 150 MHz.

2) Clock frequency = 150 MHz: At 150 MHz clock frequency, alarm failures started to emerge (i.e., EMFI that were not detected) as shown in Fig. 5. The appearing AF windows had a span of approximately 0.5 ns and were located around the clock rising edges. The width of the DWs simultaneously decreased to around 0.8 to 1.2 ns, which were narrower than at lower frequencies. For the AES injected faults, we obtained continuous IWs of 1.1 ns around the clock rising edges. Additionally, the number of faulted bits was stable when we got undetected fault windows. This phenomenon developed as the clock frequency was progressively increased from 150 MHz to 200 MHz close to the AES maximum frequency. By doing so, we observed a gradual increase in the width of the IWs that finally spanned entire clock periods. Therefore, while the sampling model faults were still partially held, another fault model clearly emerged that revealed performance limitations of these sensors.

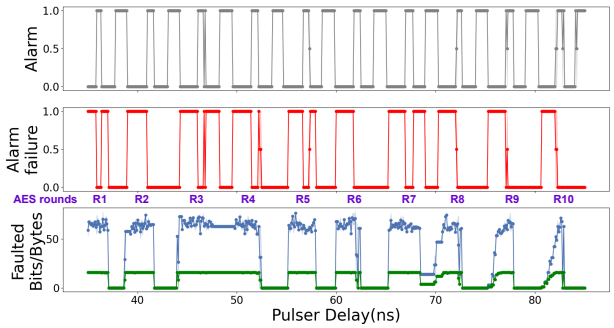


Fig. 6: EMFI results at 200 MHz (420 V voltage pulse).

3) Clock frequency = 200 MHz close to the DUT max. frequency: At 200 MHz, the AF windows developed significantly as shown in Fig. 6: most injected faults escaped the sampling fault-based sensors and the DWs were reduced to less than 1 ns. The faults injected into the AES computations proved repeatable and were consistent with the mechanism of timing faults violation as described in section II-B1c. This behavior is expected when running the target at its max. clock frequency as the slack approaches zero. Although most of the injected faults correspond to the timing fault model, the injection of faults according to the sampling fault model cannot be ruled out as the sensors were triggered for some timings (there are still effective DWs).

Moreover, it was observed that a significant EM stress was not necessary to inject faults during the target execution at 200 MHz. This was evident from the experiments conducted with a reduced voltage pulse amplitude of 340 V, as illustrated in Fig. 7.

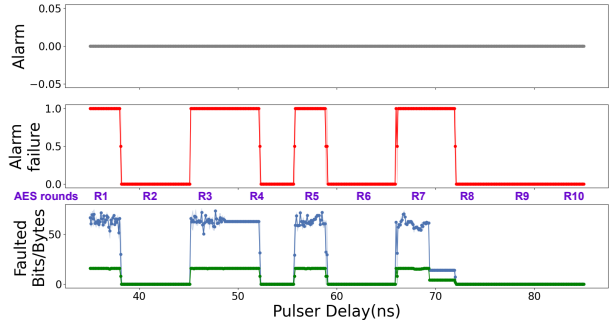


Fig. 7: EMFI results at 200 MHz (340 V voltage pulse).

For a 340 V pulse amplitude, the detection sensors failed to trigger the alarm (the gray alarm curve is flat) while faults following the timing fault model were injected. This voltage pulse amplitude is below the threshold of sampling faults. The obtained IWs spanned over a few rounds of the AES, signifying those with the higher propagation paths. When the input data (AES plaintext) was changed, the IWs moved to other rounds. This data dependency is a characteristic feature of timing faults (the propagation times of each AES round are different and data-dependent [13]). This clearly confirms that EMFI can follow a timing fault model at high clock frequency.

In addition, upon analyzing the IWs shapes of figures 6 and 7, it emerges that timing faults were also injected at clock rising edges for a voltage pulse of 340 V. Hence, for a higher-voltage pulse (420 V), both injection mechanisms interact which explains the faults injected around the clock rising edges and the fact that the IWs widths differed for clock frequencies in the range of 100 MHz to 200 MHz. It contradicts the sampling fault model hypothesis that the IW are constant against frequency variations [3], [9]. In summary, the fully digital detector proved effective at low clock frequencies up to 150 MHz, but was found powerless against faults injected at high frequency close to the maximum target frequency.

V. Analysis of Experimental Results

A. EMFI sensitivity variations with clock frequency

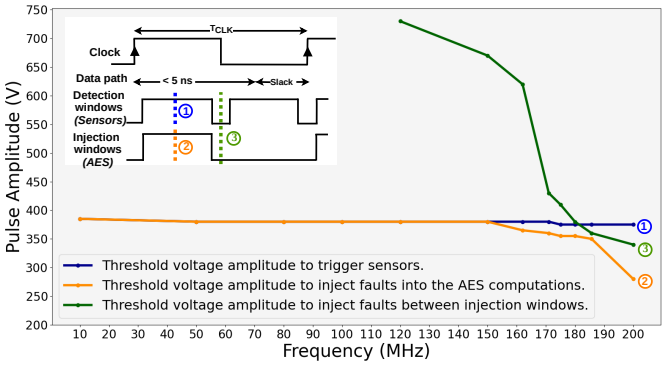


Fig. 8: Evolution of the threshold voltage amplitudes required to trigger sensors and to inject faults into the AES with respect to the clock frequency (measured at distinct EM injection time).

EMFI experiments were carried out for clock frequencies ranging from 10 MHz to 200 MHz. It made it possible to record the voltage pulse amplitude thresholds needed to inject faults into the AES computations and to trigger the sensors. These thresholds are drawn in Fig. 8 (respectively in orange and blue) for an injection timing set after the clock rising edge. The sensor's threshold, represented in blue, remained constant at 380 V for all frequencies, whereas the fault threshold was constant at the same 380 V value up to 150 MHz, before progressively decreasing to 280 V at 200 MHz. Beyond 150 MHz, undetected faults started to appear as represented by the orange line in Fig. 8 falling beneath the blue line.

We assumed that all faults injected at clock frequencies less than 150 MHz only correspond to the sampling fault model. Above 150 MHz, timing fault effects started to progressively increase with increasing clock frequencies. For an EM injection timing set between the sensor DWs, specifically around the clock falling edge, a different voltage pulse amplitude threshold resulted. This is represented by the green line in Fig. 8. It is consistent with the timing fault model: starting at 120 MHz it decreases from approximately 700 V to 450 V at 170 MHz. This is related to the decrease in timing slack of the AES with increasing the operating frequency. Around 180 MHz, it goes below the detection threshold (blue) to reach 340 V at 200 MHz. These findings conclusively indicate that there exist two distinct mechanisms of fault injection contributing to the occurrence of EMFI. They also blend as the shape of the fault threshold (orange) goes down after 150 MHz when the timing fault mechanism becomes more prevalent (decreasing green line) while the detection threshold (blue) related to sampling faults remains unchanged.

B. Limits of the sampling fault model

The previous experimental results provide clear evidence that the sampling fault model is not the only explanation for EMFI. Nonetheless, the observed results appear consistent with the research works describing it ([8]–[10]), namely that the corresponding IWs are periodically centered around the clock rising edges when the sampling occurs. However, some discrepancies between the model and our experimental results have arisen. These mechanisms are further analyzed in the subsequent sections:

1) Impact of the clock frequency: the widths of the DWs, which are associated with the sampling fault model, consistently reduced as the clock frequency increased. This observation pointedly contradicts the model description that was stated as constant with respect to frequency variations [3], [9].

2) Impact of the input data: variations of the IWs were recorded from one AES round to the other (each round computed different data), as well as when the AES plaintext was altered. This reveals a data dependency of the IWs width, which stands in contrast to the model [3], [9].

3) Impact of the critical path: To investigate the influence of the critical path on the width of the IW, the tested AES design has been modified in order to bypass the MixColumn transformation for the entire round, albeit this modification prevents it from producing a valid AES ciphertext. When bypassed, it has the effect of shortening the propagation paths of all AES rounds. Fig. 9 reports the injection results obtained

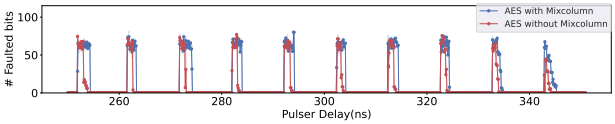


Fig. 9: Number of faulted bits in AES with/without Mixcolumns step (pulse width=4.5 ns, pulse amplitude=420 V).

at 100 MHz for the genuine AES seen in blue and the shortened version seen in red. As a result of the shortened propagation time of its rounds, the IWs of the shortened AES have a width that is reduced with respect to that of the genuine AES. This reveals that the IW width correlates with the propagation time of the targeted logic (contrarily to the theoretical sampling fault model [3], [9]).

4) Fault model at bit level: as every data of the performed AES encryptions were known, it makes it possible to analyze the injected faults at bit level (by reversing the encryption). The injected faults followed the bit-set and bit-reset fault models in similar proportions for every injection parameter. No correlation with the handled data was found. Hence, EMFI follows a bit-flip fault model.

This represents an additional deviation from the theoretical model [3], which states that a specific pulse polarity leads more often to either bit-set or bit-reset faults. For instance, it suggests that for positive voltage pulses more often result in bit-sets, and the reversing of the polarity inverts the type of injected faults (e.g. negative voltage pulses leading to a majority of bit-resets). Our tests revealed no effects of the pulse polarity on the fault type, which remained bit-flip in both cases.

At low, or moderate frequency, the injected faults should be investigated further to reconcile theory and practice.

VI. EMFI INDUCED GLITCHES IN THE CLOCK NETWORK

A. Study of the EMFI effects on the clock distribution network

In this section, we present the behavior of influential signals on an oscilloscope during EMFI attacks. All corresponding tests were conducted at 10MHz. Fig. 10 shows the signals used to control the EMFI process. The green signal illustrates the 10MHz (T=100 ns) clock period. The blue signal AES ON goes to '1' during the AES round calculations. It corresponds to 12 clock cycles required to complete the AES calculation. It is managed by a down counter. The width of the AES ON signal, when it reaches '1', is 1200 ns in normal operations as shown in Fig. 10. During the EMFI experiments, we observed that EM disturbances affect the width of AES ON signal. The behavior under attacks changes according to the polarity of the EM pulse.

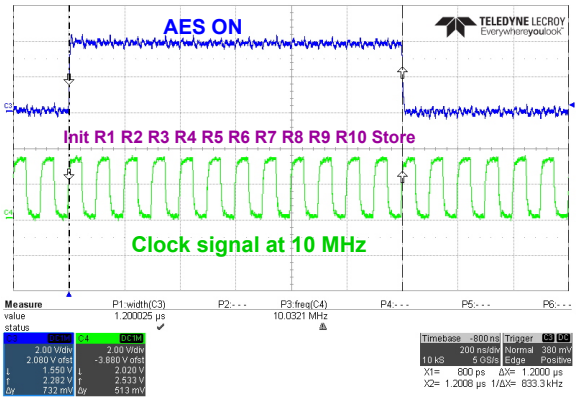


Fig. 10: Influential signals shown on an oscilloscope under normal operation.

1) Impact of the positive pulse in the clock network: Fig. 11 shows the impact of EMFI attacks, induced by a positive pulse on the AES ON signal. The left side of this figure displays the effects of EM disturbances on signals, while the right side contains a drawing that explains the observed impact. We noticed that when the clock signal is '0', there is no effect of EMFI on the width of the AES ON signal (upper left of Fig. 11). However, when the clock signal is '1', the width of the AES ON signal is reduced by one clock period (100 ns), as shown on the middle-left of Fig. 11. Additionally, the ciphertext is correctly received, indicating that all AES round calculations have been performed. We hypothesized that this reduction is due to an EM-induced negative voltage glitch on the clock signal when the EM disturbance is coupled with the target's

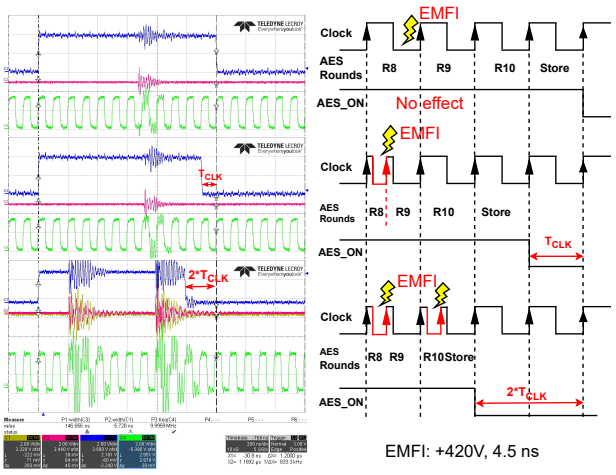


Fig. 11: EMFI effects on the AES ON signal (pulse width=4.5 ns, pulse amplitude=+420 V).

clock tree. A figure illustrating this assumption is presented in the middle section of the right-hand side. It shows how the negative voltage glitch turned a unique clock cycle (that of round R8) into two clock cycles (corresponding to R8 and R9) occurring in 100 ns. Because the resulting clock cycles have a duration longer than the AES critical time (around 4.5 ns), the AES calculation is completed without any fault.

In order to corroborate our theory, we triggered two EM disturbances, as shown in the lower section of the left-hand side, resulting in a reduction of the AES ON signal width by two clock periods.

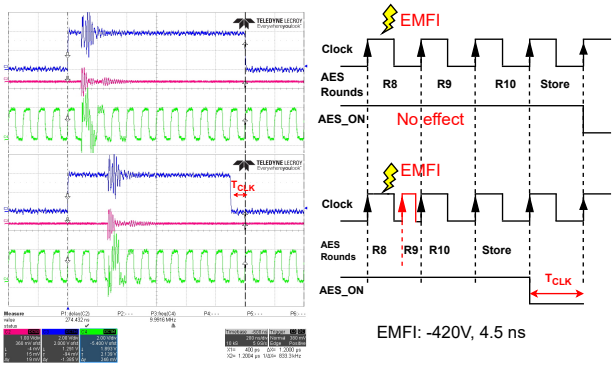


Fig. 12: EMFI effects on the AES ON signal (pulse width=4.5 ns, pulse amplitude= -420 V).

2) Impact of the negative pulse in the clock network: Fig. 12 illustrates the impact of EMFI attacks, induced by a negative pulse on the AES ON signal. Contrary to the behavior obtained in the previous section, it was noted that the AES ON signal is reduced by one clock period when the clock signal is '0', while no effect is observed when the clock signal is '1', as depicted in the left-hand side of Fig. 12. The drawing of this behavior is presented on the right side of Fig. 12. Our findings strongly indicate that EMFI induces voltage glitches when coupled with the CDN. Specifically, we noted a negative glitch in response to

a positive pulse and a positive glitch in response to a negative pulse.

B. Evidence that EMFI induced glitches in the clock network

Experimental difficulties have hindered the possibility of properly observing glitches in the clock signal on an oscilloscope during EMFI attacks. The clock signal transmitted from the FPGA output port is filtered and mixed with unwanted signals, such as disturbances and noise, which negatively affect the quality of the signal. Nonetheless, we have successfully demonstrated our approach through two different tests, which are detailed in this section.

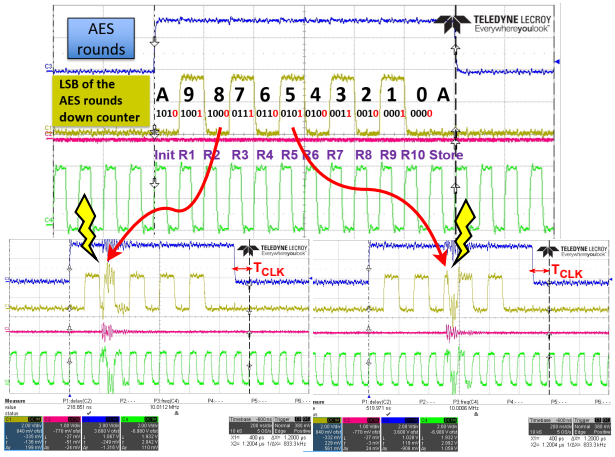


Fig. 13: LSB of the AES rounds down counter (Clock frequency= 10 MHz, pulse width=4.5 ns, pulse amplitude=420 V).

1) Test 1: LSB of the AES rounds down counter: To further investigate the effects of EMFI on the AES ON signal and the reason behind its reduced width, we transmitted the least significant bit (LSB) of the AES rounds down counter onto the oscilloscope, as shown by the yellow signal on the top of Fig. 13. The bottom of Fig. 13 illustrates the effects of EMFI attacks induced by a positive pulse on signals. We observed a reduction in the down counter value windows in both cases, i.e., at low level and high level as shown respectively in the lower left and lower right of Fig. 13. It accelerated the count and thus, reduced the AES ON width. The AES calculations were complete and the ciphertext was correctly received. This demonstrates the creation of negative glitches in the clock tree due to coupling with the CDN. Likewise, we demonstrated the creation of positive glitches under EMFI attacks induced by a negative pulse (-420V).

2) Test 2: Freezing the input clock of the AES calculation blocks: To definitively ascertain whether voltage glitches induced on the clock tree are effective in replacing a genuine clock rising edge, we conducted a test by freezing the AES clock signal. In other words, the clock has stopped. We used a BUFfer Global (BUFG) with a Clock-Enable (CE) to gate the clock signal. When the CE is asserted, the clock signal passes through the buffer (unfreeze case). However, the output of the buffer is held at logic '0' when CE is deasserted (freeze case). An FSM controlled the CE of the buffer. Once the AES down counter reached the predefined value, it entered into a closed state, deasserted the CE and the clock signal was held at a low level. This is the freeze case, as shown on the left side of Fig. 14. The clock signal is stopped, blocking the flow of the AES calculations. The FSM was designed in order to

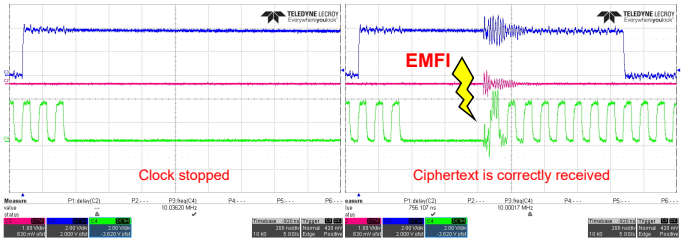


Fig. 14: Freezing the AES clock signal (Clock frequency= 10 MHz, pulse width=4.5 ns, pulse amplitude= -420 V).

automatically exit the freeze state on the next clock cycle, however as the clock signal was stopped in a low state, the FSM stays frozen in a dead lock: it needs a clock rising edge to exit the freeze state and resume the clock signal to normal. This allows us to prove that the positive glitches induced by a negative pulse (-420V) in the clock tree are effective in replacing a genuine clock rising edge. Indeed, it achieves its intended purposes, as shown on the right side of Fig. 14, under EMFI attacks. Minor adjustments to the EM probe were required to identify the optimal location where we could receive an accurate ciphertext. These tests attest to the fact that EMFI does induce clock glitches in the target's CDN. It can replace genuine clock rising edges.

Our discoveries reinforce the results provided in recent publications [6] which demonstrate the possibility of inducing clock glitches by EMFI attacks on a RISC microprocessor (LEON3-design, 180 nm TSMC).

VII. In-depth analysis of EMFI-induced clock glitches

A. EMFI-induced clock glitch principle

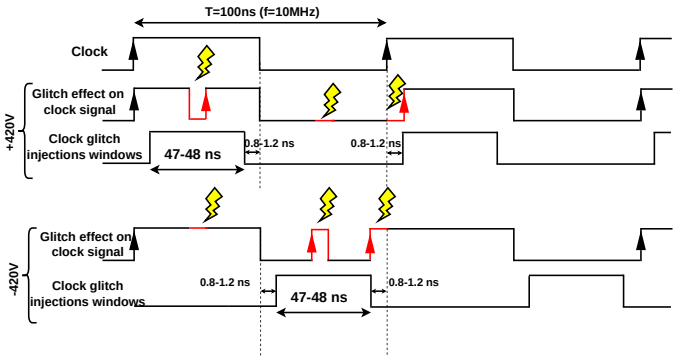


Fig. 15: EMFI-induced clock glitch (Clock frequency= 10 MHz, pulse width=4.5 ns, pulse amplitude=+/-420 V).

After presenting evidence that EMFI generated clock glitches in the clock tree, this section provides a detailed analysis of characteristics of EMFI-induced clock glitches. Specifically,

we show the conditions that resulted in positive or negative voltage glitches based on the pulse polarity, and the impact these glitches have on the affected signals. We examine the EMFI effects on the width of the AES ON signal as a function of the EM injection timing and its link to the clock signal. These effects vary with the pulse polarity. Fig. 15 shows these effects in both polarity pulse cases. It should be noted that the EM probe is wrapped in a counter-clockwise direction.

- Positive pulse induced negative glitches: three distinct behaviors are observed (highlighted in red from left to right in Fig. 15):
  1) When the clock signal is '1', negative glitches are induced on the clock signal and we observe a reduction in the width of the AES ON signal because an additional clock cycle is carved out into the positive half-period of the targeted clock cycle. This effect occurs during a 47-48 ns time window for a clock period equal to 100 ns.
  2) When the clock signal is '0', there is no effect on the AES ON signal width.
  3) When the EM injection timing is just after the rising clock edges, a shift in the clock signal edge is induced but with no effect on the AES ON signal width. We called k a constant margin for this case, which is measured between 0.8-1.2 ns. This is not an effective glitch during which the clock edges get a small shift. In fact, it extends a clock period and reduces the following one.
- Negative pulse induced positive glitches: three distinct behaviors are observed:
  1) When the clock signal is '1', there is no effect on the AES ON signal width.
  2) When the clock signal is '0', positive glitches are induced on the clock signal and we observe a reduction in the width of the AES ON signal because an additional clock cycle is created from the voltage glitch. This effect occurs during a 47-48 ns time window for a clock period equal to 100 ns.
  3) When the EM injection timing is just before the rising clock edge, a shift in the clock signal is noted but with no effect on the AES ON signal width. The constant margin k is measured between 0.8-1.2 ns. This is not an effective glitch, during which the clock edges get a small shift. In fact, it reduces a clock period and extends the following one.

Therefore, we established that the mechanisms of EMFI-induced clock glitches depend on the clock frequency, and the susceptibility window for injecting effective glitches (positive or negative) through EMFI attacks is related to T/2. The width of the susceptibility windows caused by clock glitches under EMFI attacks can be calculated in the following eqt. 2, where k is a constant margin during which clock edges get a small shift under EMFI attacks. We hypothesized that this specific time window k is related to the induced clock glitches width.

$$W_{EMFI\ susceptibility\ windows} = T/2 - 2k \qquad (2)$$

B. Evaluation and analysis of faults injected in AES computations due to EMFI-induced clock glitches

The insight of calculating the number of faulted bits and bytes present in erroneous ciphertext received during campaigns conducted at several clock frequencies is crucial for accurately interpreting experimental results. This knowledge can also help validate the corresponding characteristics of the EMFI-induced clock glitches and reconcile theoretical predictions with practical observations. In Fig. 16, the number of faulted bits and bytes in the last 3 injection windows corresponding to the final 3 rounds of the AES are shown as a function of the EM injection time. These IWs improved the accuracy and interpretation of our results. This figure illustrates how the number of faulted bits and bytes varies when changing the pulse amplitude polarity and when modifying the clock frequency to both 100 MHz and 170 MHz.

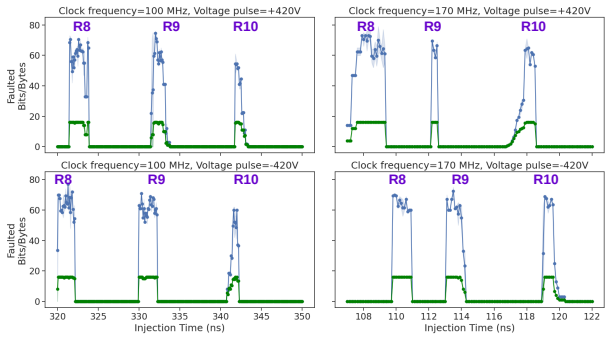
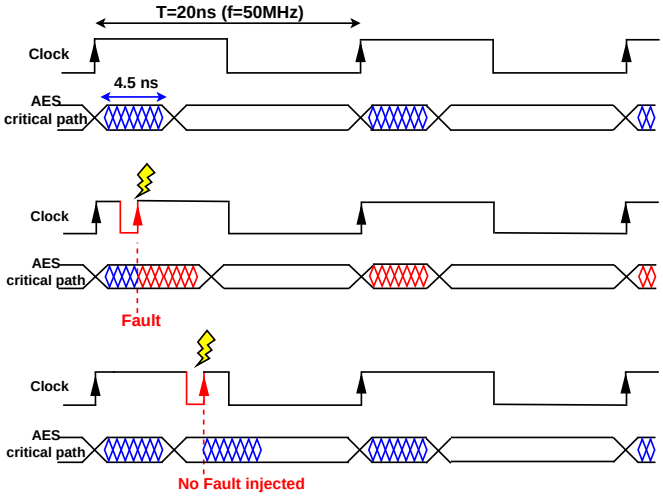


Fig. 16: A comparative analysis of faulted bits progression in function of EM injection time when varying the clock frequency and the pulse amplitude polarity.

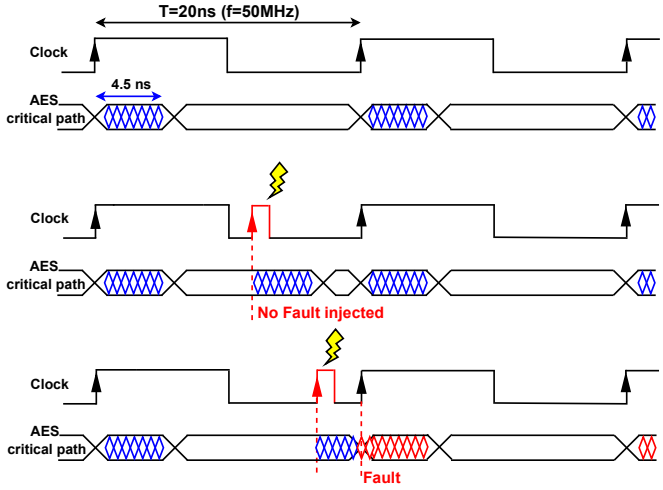
The EMFI-induced clock glitch can lead to timing violations in AES calculations, resulting in erroneous ciphertext. To investigate the behavior of faults in AES, it is crucial to study the relation between the critical path of AES's combinatorial logic ($\approx$4.5 ns) and the clock frequency as discussed in the section VII-A. We have divided the analysis of faults in AES into two cases:

1) Case 1: $t_{critical} < T/2$: Fig. 17 depicts the occurrence of faults in AES calculations due to the mechanism of EMFI-induced clock glitches at 50MHz when $t_{critical} < T/2$. In Fig. 17a, the impact of negative glitches is analyzed. When the EM injection time is close to the clock rising edge, there is no effect to generate a negative glitch and no fault was injected in AES calculations. However, when the EM injection time continues to increase after a constant margin k, a negative glitch occurs, halting a round of AES calculations and forcing the next round by the rising edge of the glitch. It leads to a massively faulted ciphertext with a significant number of faulted bits, which decrease progressively with the increased EM injection time as long as it does not exceed the critical path. Beyond this point, an effective glitch is obtained, but it does not result in a computational error within the AES operations.

This theoretical explanation was consistent with the experimental results obtained from a campaign launched at 100 MHz

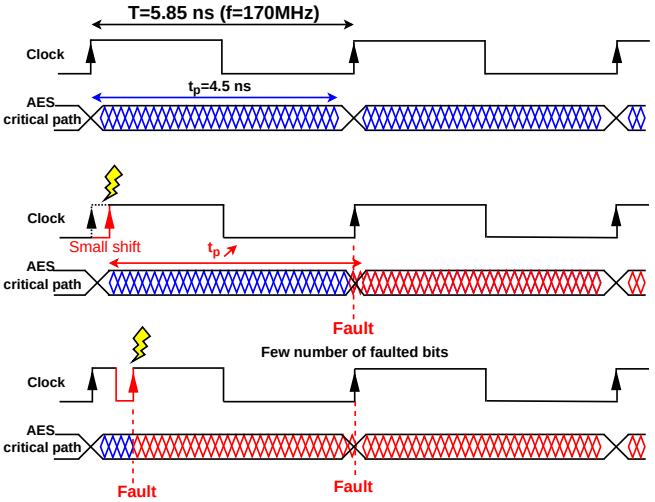


(a) Negative glitch, pulse amplitude=420 V.

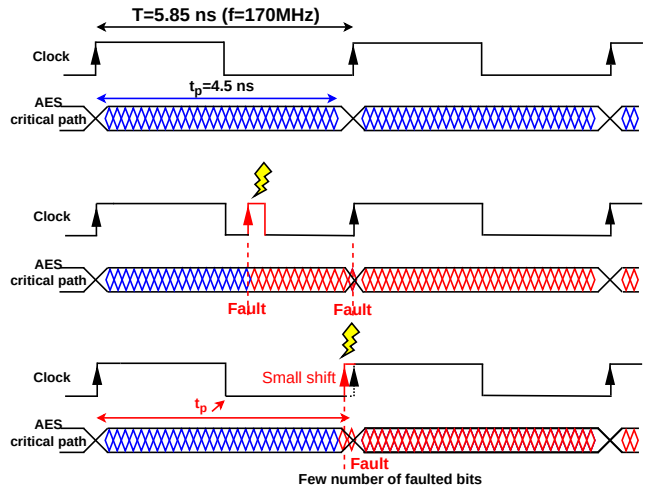


(b) Positive glitch, pulse amplitude= -420 V.

Fig. 17: EMFI-induced clock glitch at 50MHz ($t_{critical} < T/2$)



(a) Negative glitch, pulse amplitude=420 V.



(b) Positive glitch, pulse amplitude= -420 V.

Fig. 18: EMFI-induced clock glitch at 170MHz ($t_{critical} > T/2$)

($t_{critical} < T/2$) with a +420 V pulse amplitude as depicted in the upper section on the left side of Fig. 16. In Fig. 17b, we examine the positive glitch impact. An opposite behavior is observed compared to the negative glitch behavior. The clock rising edge of the positive glitch forces the early calculation of a round of AES calculations, and the next clock rising edge induces faults as long as it corresponds with AES computations of the round. The number of faulted bits increases progressively while increasing the EM injection time. This theoretical explanation was consistent with experimental results obtained from a campaign launched at 100 MHz ($t_{critical} < T/2$) with a -420 V pulse amplitude as depicted in the lower section on the left side of Fig. 16. Hence, the width of the injected fault windows in the AES computations depends on the critical path. It is consistent with the following equation:

$$W_{AES} = t_{critical} - k = constant \quad (3)$$

Upon comparing the width of IWs predicted by the eqt. 3 with experimental results, it has been established that eqt. 3 is both accurate and consistent with the IWs width saved during campaigns launched at clock frequencies lower than 120 MHz. This provides evidence that as long as $t_{critical} < T/2$, the width of IWs remains constant when varying the clock frequency.

2) Case 2: $t_{critical} > T/2$: Fig. 18 depicts the occurrence of faults in the AES calculations due to the mechanism of EMFI-induced clock glitches at 170 MHz when $t_{critical} > T/2$. As outlined in section V, EM-induced faults follow two different mechanisms beyond 150 MHz for 420 V voltage pulse: the timing fault model, which prolongs the logic propagation time, and the EMFI-induced clock glitches. Fig. 18a explores the case involving a negative glitch. When the EM injection time corresponds around the clock rising edge, an EM coupling with the target's CDN causes a small shift in the rising edge and shortens the clock period. This, in combination with EM coupling with the target's PDN which increases the logic propagation time, produces an erroneous ciphertext with a low number of faulted bits. It gradually increases during a constant

margin k. However, when increasing EM injection time, a negative glitch occurs, interrupting a round of AES calculations at the rising edge of the glitch. The next clock rising edge halts the next round during its calculations as depicted in Fig. 18a. This leads to a significant number of faulted bits in the ciphertext. The experimental results obtained from a campaign launched at 170 MHz ($t_{critical} > T/2$) with a +420 V pulse amplitude, as depicted in the upper section on the right side of Fig. 16, are consistent with the corresponding theoretical explanation proposed. The number of faulted bits and bytes progressively increases during a duration of approximately 1 ns, and then sharply increases, as shown by the last IWs.

In Fig. 18b, we examine the positive glitch impact which shows the opposite behavior from a negative glitch. Initially, an effective positive glitch causes a significant number of faulted bits that halts the current round of AES calculations on the glitch's rising edge, as well as the subsequent round on the clock rising edge, as depicted in Fig. 18b. When EM injection time approaches the clock rising edge, the number of faulted bits gradually decreases. A small shift in the clock edges combined with an increase in logic propagation interrupts the current round by the next clock rising edge and results in a low number of faulted bits in the ciphertext. The experimental results obtained from a campaign launched at 170 MHz ($t_{critical} > T/2$) with a -420 V pulse amplitude, as depicted in the lower section on the right side of Fig. 16, are consistent with the corresponding theoretical explanation proposed.

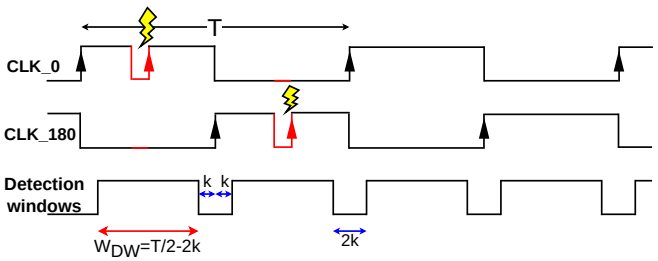
Hence, the width of the injected fault windows in the AES computations due to the EMFI-induced clock glitch depends on the clock frequency. It is consistent with the following equation:

$$W_{AES} = T/2 - 2k \tag{4}$$

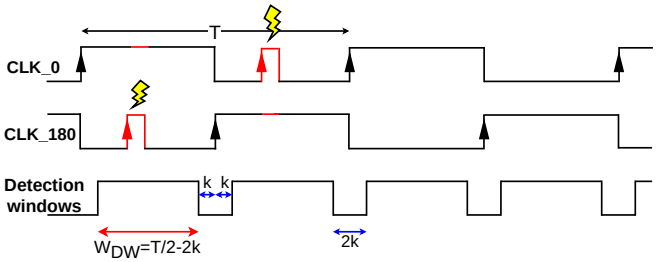
As the clock frequency increases, the clock period shortens and the width of the IWs windows in AES reduces. At high clock frequencies, this reduction becomes so significant that IWs due to the EMFI-induced clock glitches become negligible in size but the timing fault models stay the primary concern in inducing faults in the target when its frequency is close to its maximum.

C. How does the mechanism of EMFI-induced clock glitch explain the triggering of the sensors?

How does the digital detection sensor, which is designed to be sensitive to the sampling fault model in detecting EMFI attacks, prove to be effective at detecting the EMFI-induced clock glitches? Section III-A describes the detection sensor architecture designed by Elbaze et al. [10] as shown in Fig. 1 and its behavior under normal operation and attacks as shown in Fig. 2. The sensor's state transitions are related to the rising edges of two clock signals: the primary clock (CLK0) and the 180° phase-shifted clock signal (CLK180). If there are any deviations from its normal operation, a dedicated logic block raises an alarm signal. Our experimental results show that DW widths decrease with the increase of the clock frequency. This discrepancy challenges the sampling fault model's characteristics. Fig. 19 provides an explanation for these sensor behaviors in light of the clock glitches induced during EMFI attacks. Fig. 19a examines the impact of a negative glitch due to a



(a) Negative glitch, pulse amplitude=420 V.



(b) Positive glitch, pulse amplitude= -420 V.

Fig. 19: Sensors detection windows width

+420V voltage pulse. Under EMFI attacks, a negative glitch is induced when the clock is at a high logic level but has no effect when the clock signal is at a low logic level. Conversely, a positive glitch induced by a negative pulse amplitude has a contrasting effect, as illustrated in Fig. 19b. These two cases are sufficient to analyze that the DW width is related to the following equation:

$$W_{DW} = T/2 - 2k \tag{5}$$

When the clock frequency increases, the clock period decreases, as well as the DW width for triggering sensors. Upon comparing the width of DWs predicted by the eqt. 5 with experimental results, it has been established that it is both accurate and consistent with the observed DW width saved during campaigns launched at clock frequencies between 10 MHz to 200 MHz.

At low clock frequencies, in case of $t_{critical} < T/2$, the width of injected faults in AES is less than the detection window width when these sensors are triggered as shown in Fig. 20. This explains why all injected faults are detected at low frequencies.

Fig. 20 highlights three aspects regarding the sensor's detection capability in the case when $t_{critical} < T/2$. The green window represents the Detected Faults, indicating the faults injected into the AES computations that were successfully identified by the sensor. On the other hand, the orange window represents the False Alarm, indicating instances where the sensor reported faults that were not actually induced into the AES computations. The white window represents the No Alarm, indicating that no sensors have been triggered and no faults have been injected into the AES computations.

Through a comprehensive analysis of experimental results from multiple campaigns conducted across the complete frequency range of a target and variations in EMFI sensitivity with clock frequency detailed in sections IV and V respectively, we

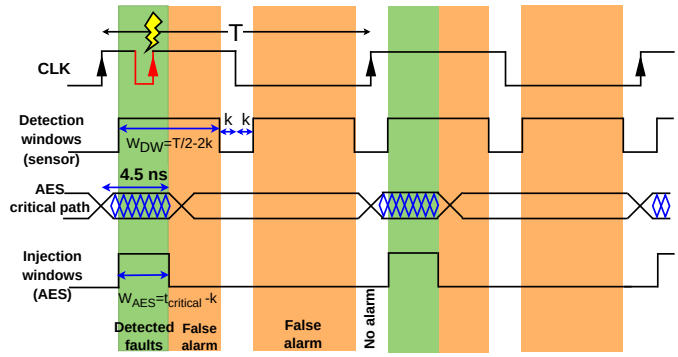


Fig. 20: Faults detected (green), False alarm (orange) and No alarm (white) windows in function of EM injection time at low clock frequencies ($t_{critical} < T/2$).

have examined the impact of changing clock frequencies and pulse amplitude on the sensor's fault detection performance. At low frequencies ($t_{critical} < T/2$), the sensor exhibits a high detection rate, as demonstrated in the green area of Fig. 21. However, as the clock frequency increases reaching the case when $t_{critical} > T/2$, this performance gradually diminishes from moderate (orange area) to low (pinkish red area) detection rate. Between 120 MHz and 150 MHz, the sensor's detection rate becomes moderate, as indicated in the orange area of Fig. 21. Beyond 150 MHz, the sensor fails to detect injected faults, aligning with the timing faults, conducted with a reduced voltage pulse amplitude less than 380 V, observed in the red area of Fig. 21.

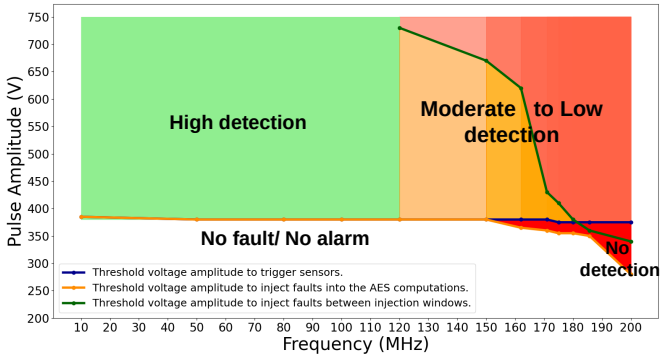


Fig. 21: Evolution of the impact of changing clock frequencies and pulse amplitude on the sensor's performance in detecting faults.

Therefore, this sensor has been proven to be effective against EMFI-induced clock glitches at low frequencies. Nevertheless, it was found ineffectual against faults injected following the timing fault model at high frequencies.

VIII. Conclusion

This paper explores the effectiveness of EMFI detection sensors based on the assumption that the sampling fault model can explain EMFI. The sensor efficiency began to falter for operating frequencies exceeding 150 MHz casting doubts upon the model's validity. At low, or moderate frequency, the injected faults generally adhere to the sampling fault model. However, certain discrepancies from the theory framework raise doubts about its legitimacy. This highlights the potential risks taken when relying on an incomplete fault model as the basis for placing a sensor. The reported experimental evidence has confirmed that EMFI can result from various mechanisms, challenging the notion that it can be attributed to a single fault model. Two mechanisms align with the timing violations fault model: At high frequency (i.e., for a low slack), timing faults are induced through a coupling of the EM disturbance with the target's PDN that results in an increase of the logic propagation times beyond the clock period. At low, or moderate frequency, EM-induced clock glitches occur due to coupling with the target's clock tree. This paper presents a comprehensive explanation of the timing violations fault model, involving the two mechanisms induced by EMFI across the target's full-frequency range.

This research paper improves the understanding of FIA mechanisms in the field of integrated circuits security and assists designers in developing effective detection sensors based on a complete EMFI fault model consolidated by rigorous experimentation results.

References

[1] C. Giraud, "Dfa on aes," in International Conference on Advanced Encryption Standard, 2004, pp. 27–41.

[2] A. Dehbaoui, J.-M. Dutertre, B. Robisson, and A. Tria, "Electromagnetic transient faults injection on a hardware and a software implementations of aes," in 2012 Workshop on Fault Diagnosis and Tolerance in Cryptography, Leuven, Belgium, September 9, 2012, 2012, pp. 7–15.

[3] M. Dumont, M. Lisart, and P. Maurine, "Modeling and simulating electromagnetic fault injection," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, pp. 680–693, 2020.

[4] L. Zussa, A. Dehbaoui, K. Tobich, J.-M. Dutertre, P. Maurine, L. Guillaume-Sage, J. Clediere, and A. Tria, "Efficiency of a glitch detector against electromagnetic fault injection," in 2014 Design, Automation & Test in Europe Conference & Exhibition (DATE). IEEE, 2014, pp. 1–6.

[5] M. Agoyan, J.-M. Dutertre, A.-P. Mirbaha, D. Naccache, A.-L. Ribotta, and A. Tria, "How to flip a bit?" in 2010 IEEE 16th International On-Line Testing Symposium. IEEE, 2010, pp. 235–239.

[6] M. Ghodrati, B. Yuce, S. Gujar, C. Deshpande, L. Nazhandali, and P. Schaumont, "Inducing local timing fault through em injection," in 2018 55th ACM/ESDA/IEEE Design Automation Conference (DAC). IEEE, 2018, pp. 1–6.

[7] S. Ordas, L. Guillaume-Sage, K. Tobich, J.-M. Dutertre, and P. Maurine, "Evidence of a larger em-induced fault model," in International Conference on Smart Card Research and Advanced Applications. Springer, 2014, pp. 245–259.

[8] S. Ordas, L. Guillaume-Sage, and P. Maurine, "Em injection: Fault model and locality," in 2015 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC), 2015, pp. 3–13.

[9] ——, "Electromagnetic fault injection: the curse of flip-flops," Journal of Cryptographic Engineering, vol. 7, no. 3, pp. 183–197, 2017.

[10] D. El-Baze, J.-B. Rigaud, and P. Maurine, "A fully-digital em pulse detector," in 2016 Design, Automation Test in Europe Conference Exhibition (DATE), 2016, pp. 439–444.

[11] L. Zussa, J.-M. Dutertre, J. Clédiere, B. Robisson, A. Tria et al., "Investigation of timing constraints violation as a fault injection means," in 27th Conference on Design of Circuits and Integrated Systems (DCIS), Avignon, France, 2012, pp. 1–6.

[12] Xilinx, "Digilent reference," in Nexys Video 7 FPGA board, 2020. [Online]. Available: https://digilent.com/reference/_media/reference/programmable-logic/nexys-video/nexys-video_rm.pdf

[13] L. Zussa, J.-M. Dutertre, J. Clédiere, and B. Robisson, "Analysis of the fault injection mechanism related to negative and positive power supply glitches using an on-chip voltmeter," in Hardware-Oriented Security and Trust (HOST), 2014 IEEE International Symposium on, 2014.