Cheap and Cheerful: A Low-Cost Digital Sensor for Detecting Laser Fault Injection Attacks

Wei He, Jakub Breier and Shivam Bhasin

Physical Analysis and Cryptographic Engineering, Temasek Laboratories
Nanyang Technological University, Singapore
{he.wei,jbreier,sbhasin}@ntu.edu.sg

Abstract. Fault Injection Attacks (FIAs) have become a critical threat towards prevailing security embedded systems. FIA typically exploits the maliciously induced faults in security ICs for retrieving confidential internals. Since the faults are injected by disturbing circuit behaviors, FIA can possibly be detected in advance by integrating a sensitive sensor. In this paper, a full-digital detection logic against laser fault injection is proposed, which mainly consists of a high-frequency RO watchdog and a disturbance capture for sensing frequency ripples due to laser impact. Practical experiments on Virtex-5 FPGA show that the proposed sensor has fault detection rate of 100% for both regional and single CLB injection, protecting critical registers of PRESENT-80 cipher, with superior power/spatial security margin compared to a prior PLL-based sensor, while maintaining extremely low cost in hardware. The proposed logic is further applied to protect complete cipher over larger fabric, and the fine-grained fault injection using pulse laser shows a detection rate of 94.20%, and an alarm rate of 2.63 : 1 in this experiment. Owing to its simple digital architecture, this system can be easily applied into any security-critical ICs.

Keywords: Cryptography, Embedded System, Ring-Oscillator, Semi-Invasive Attack, FPGA

1 Introduction

Hostile implementation circumstances in security applications demand the security-critical circuits to be integrated with a strong protection against various attack threats. In modern cryptography, confidential data is protected by utilizing strong algorithms. However, the real-world implementation of these algorithms in devices inevitably draws numerous vulnerabilities in their applications. Various attack methodologies on the physical layer have been proposed for breaking crypto algorithms or other security-critical applications. The two commonly known methodologies are leakage-based side-channel attacks (SCA [12]), and abnormality-based fault injection attacks [7]. In SCA, the leaked physical information (like power consumption, timing, etc.) is exploited for extracting the secrets [12]. On the other hand, FIA retrieves confidential information by analyzing the faulty behavior or faulty outputs from the target when operated

under hostile environment. FIA can be widely used for serving different purposes, the most common one being the secret key retrieval [4]. Besides, it can also be used for reverse engineering purpose to deduce the internal architecture of the attacked chip by analyzing its faulty behavior [17]. Moreover, FIA is also a promising method to break the defense of the system, for assisting other hardware-level attacks [1]. Owing to its wide potential in various attacks, fault injection attacks have evolved to be a critical security threat against all kinds of security ICs. It is also commonly tested by certification bodies when evaluating security-critical devices.

The fault injection can be conducted at two levels. First, faults can be globally injected by imposing disturbances into global variables, such as the clock system or power supply of the device under test (DUT) [2]. In this approach, noticeable disturbances are induced in clock or power lines, which are distributed through the global network and affect the critical logic points that are vulnerable in exposure of disturbance. Typically, it is the critical logic path which can easily suffer from setup-time violation by a ripple in clock or power. Another approach to conduct fault injection is to affect the local chip fabric relying on high-precision injection methodologies, as laser (laser fault injection - LFI) or electromagnetic (EM fault injection - EMFI). The faults are injected by making an impact on the signal propagation by external means like EM or directly upsetting stored data bits in memory cells by using strong laser. Since the disturbance can be strictly constrained to a specific chip region, and it is easy to tune the injection time from the equipment, LFI and EMFI are superior to global injections in terms of both precision and controllability. The disadvantage compared to global methods is a high cost of injection equipment.

Protection against fault attacks can be done either at information level or circuit level. Error detection and correction codes find wide applications in information based fault protection [11]. Other kinds of information redundancy, like duplication, can also be used for fault protection. The circuit modification for information redundancy has a finite and non-negligible cost. Moreover, it is a reactive protection, which acts when the fault has already been injected and potentially exploited [1]. The other family of protection is proactive in nature and based on environmental sensors [22]. It monitors environmental parameters and raises an alarm in hostile conditions. Such protections are better for LFI or EMFI techniques which inject faults by controllable injection using high-energy electromagnetic or laser pulses. In this paper, a low-cost and fully digital sensor system is presented for detecting a semi-invasive laser fault injection on-the-fly. This sensor relies on a strict timing violation, for detecting the slight signal oscillation alteration (phase shift) in a watchdog ring oscillator (RO) from laser fault injection.

Contribution: The merit of the proposed technique resides in (a) its superior detection sensitivity and protection coverage against semi-invasive disturbance; (b) the capability of detecting bi-directional frequency ripple (i.e., either acceleration or deceleration of sensitive signal); and (c) fully digital and cost-

efficient architecture, which can be easily implemented into any digital/hybrid ICs and FPGAs, for high-security application.

Outline of this paper: The content of this paper is organized as follows: Sec. 2 recalls the technical backgrounds of laser-based fault attacks towards cryptographic primitives in hardware, and the prior countermeasures; In Sec. 3, the proposed low-cost digital sensor system against laser fault injection is elaborated; along with the FPGA implementation details. Sec. 4 describes a series of experimental evaluations in practical high-precision laser fault attacks, with a thorough comparison to a recently proposed PLL based sensor [13, 10]. Finally, the work conclusions are drawn in Sec. 5.

2 Background

2.1 Fault Attacks on Cryptographic Primitives

Integrated circuits (IC) can be easily affected by environmental conditions they operate in. One of the first phenomena observed in this direction was a higher number of failures in satellite systems caused by cosmic rays [5]. A new area, testing reliability of IC has emerged since then - *failure analysis*. More then 20 years later, Boneh, DeMillo, and Lipton [7] have shown that such failures can be used for attacking cryptographic primitives implemented in integrated circuits, naming this area *fault injection attacks*.

Currently, fault attacks are among the most popular physical attacks on cryptographic implementations, together with side-channel attacks. There are various techniques allowing attackers to influence electronic devices, ranging from low-cost solutions, such as voltage or clock glitches, to expensive ones, such as laser fault injection or focused ion beam [3]. If the attacker can control the device in order to make a precise errors during computations, some confidential internals, particularly as crypto keys, can be easily revealed. For example, it was shown that the full AES key can be recovered by injecting just one fault in the penultimate round [16].

For testing our countermeasure, we have chosen the laser fault injection technique, which provides very good spatial and timing resolution and therefore can accurately measure the effectivity of the fault protection.

2.2 Laser Fault Injection

Optical fault injection attacks were presented by Skorobogatov and Anderson in 2002 [19]. In the paper, authors used a flashlight for inducing faults in a microcontroller. However, such technique is not very precise, therefore, laser fault injection has quickly become the most used optical fault attack technique.

In this approach, a laser source attached to a microscope is placed over the chip, so the laser beam can lead to charge transmission in signal paths, or the ionization effect on transistors. In general, one can decide whether to approach the chip from the frontside or the backside. For the *frontside injection*,

either green (532 nm) or red (808 nm) lasers are used because there is no need to penetrate the silicon substrate. The advantage of this method is the direct visibility of components of the chip. However, metallic layers can completely nullify the effect of the beam. Especially modern ICs have several metal layers and therefore, it makes it infeasible to use this method. In the *backside injection*, one has to use at least near-infrared laser (1064 nm) because of the substrate. It is advisable to mill down and polish the substrate in order to make the components accessible and to achieve higher precision by avoiding the light refraction.

When it comes to effects on FPGAs, the resulted phenomena can cause a direct bit upset in memory cells, either in the flip-flops, in block memories, or in the configuration bits in reconfigurable circuit [18]. It can also affect the signal propagation, by either increasing or decreasing the signal transmission, causing timing violation in logic chain.

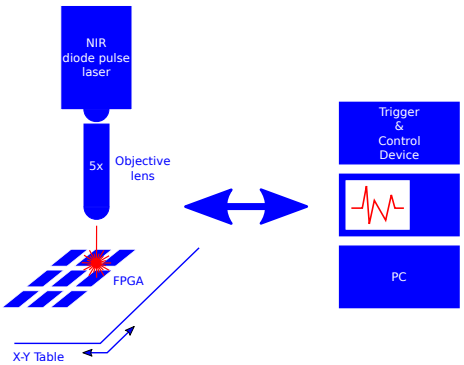


Fig. 1: Example of laser fault injection setup testbench.

An example of laser fault injection setup is depicted in Fig. 1. On the left side, we can see the laser source, with the power usually ranging in several Watts, attached to a magnifying objective lens (in our setup we use 5× magnification). There is also an X-Y table that is capable to precisely position the device under test (FPGA in the picture). On the right side, there are acquisition, communication and control devices. Normally, data is sent from the PC to the DUT, which sends a trigger signal before processing the data. Trigger & Control device captures this signal and sends a command to the laser source to perform the injection. To get precise timing and laser diode current, it is advisable to use a digital sampling oscilloscope.

2.3 Countermeasures

Numerous countermeasures against fault injection have been developed in prior literatures, which basically drop within two scopes: First, the cipher itself is fortified with capability of detecting data abnormality. In this approach, the cipher primitive needs to be merged with the detection logic, as the concurrent

error detection (CED) proposed by Karri et al in [11]. In this method, parity bits are computed in advance to predict and compare with the the parity of the output vector in each computation round. If they are equal, error check is passed, and otherwise, error/errors occurred in ciphering computation of this round. Another popular idea to detect the error is to simply duplicate the original cipher in parallel, and both are fed with the same plaintext. In the output side, the two outputs from the genuine and the duplicated rails are compared to see if any faults occurred in either rails. The pitfalls of these redundancy based detection can be summarized as follows:

1. **High-Cost**: The cost of these redundancy error check logics are resource consuming. This is because the detection needs to simulate the real data computation, or parity computation at each computation round, so as to be compared with real cipher outputs. Prior work reported roughly doubled hardware cost using these methods.
2. **Low-Detection Coverage**: Since these detection base on the data or parity comparison, a fatal problem arisen here is that not all the faults can be detected. For instance, parity comparison normally detect odd-number errors occurred inside the algorithm, and the duplication method cannot detect the faults that are simultaneously perturbated into the same logic points of the two rails.
3. **No Prediction Margin**: These detection logics can only detect the faults that have already been successfully injected into the cipher cores. In other words, the on-going injection campaign cannot be predicted in advance.

On contrary, sensor based countermeasures [13, 22] are alternatively used for detecting the fault injection on-the-fly. In this approach, an independent logic can be used as the injection sensor, being implemented together with the protected cipher. The sensor should have a higher sensitivity against the disturbance induced by the injection equipment, which should have logic (`alarm`) signal responding to injection turbulence earlier than the accomplished cipher faults. More precisely, the injection disturbance should have more significant impacts on the sensor, by inducing specific alarm signal. Moreover, the detection coverage of fault types should also be sufficiently high.

2.4 Previous Works on Sensor Based Countermeasures

As a summary, all the injections discussed above can cause change on signal propagation. Therefore, if a logic can be sufficiently sensitive in detecting abnormal frequency change, the malicious injections can be detected.

There are several techniques that can be employed in FPGA in order to detect disturbances by a laser. In the following, we will explain the works proposed so far in this area.

Glitch Detector Glitch detector is a timing-violation based sensor that was originally proposed for detecting any timing violation using power or clock global

fault injections [8]. Later research mentions its partial effectiveness against EM fault injections [22]. This logic consists in detecting the violation of a guarding delay prior to any timing violation. The clock signal is used as a reference to be able to draw comparisons between the guarding delay and the clock period to a flip-flop, as illustrated in Fig. 2(a). The output of flip-flop serves as the alarm signal which stays in low voltage level in absence of disturbance. In case the external disturbance increases the signal delay in `CK`, the setup-time will be violated which triggers a high voltage level in alarm signal, as illustrated by the timing diagram in Fig. 2(b). The pitfalls of this logic are twofold. Firstly, the detector is suited for global disturbances. However, using a network of detectors can also detect local injections to some extent [22]. Secondly, the detector is designed against injection method which increases propagation delay, while remaining insensitive to techniques which can accelerate the signal propagation as shown in Fig. 2(c).

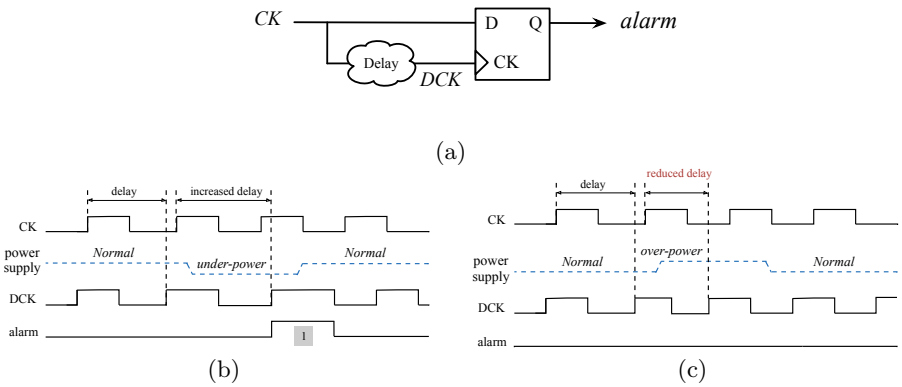


Fig. 2: (a) Topology of glitch detector; Timing diagram of disturbance detection by glitch detector under: (b) delayed signal propagation, and (c) accelerated signal propagation.

Ring-Oscillator with Frequency Counter As a low-cost oscillation generator, digital Ring-Oscillator (RO) has been widely used in security applications, such as the unclonable crypto key generation [21]. RO is a closed loop chained an by odd number of inverters, as sketched in Fig. 3(a). The oscillation frequency of a RO is determined by the summed-up signal propagation time in this loop. Any anomaly or disturbance would normally impact the RO resulting in change of oscillation frequency and phase. As aforementioned, many fault injections can cause timing change in signal path, hence RO can be potentially used to detect the on-going injection campaign.

Basically, the oscillation distortion in either phase or frequency can be captured by a digital counter [9], and the size (bit-width) of the counter can be determined by the used oscillation frequency and the time-window of the mea-

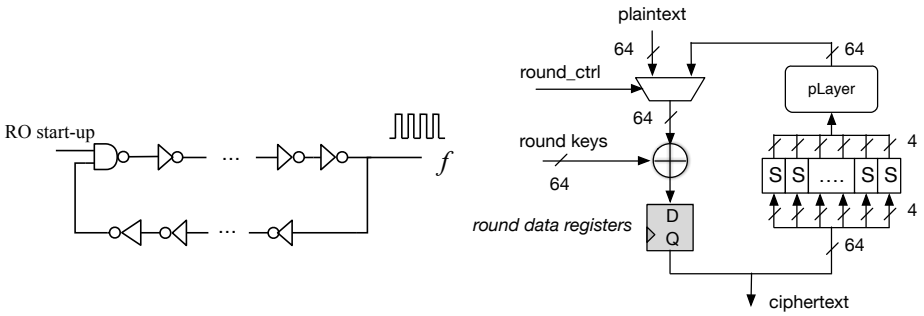


Fig. 3: (a) Inverter based digital ring-oscillator; (b) Round architecture of PRESENT-80 cipher.

surement. The drawbacks using frequency counter are clear. First, to enlarge the disturbance impact to a RO, the frequency of this RO should be high. Therefore, the required bit-width of the capture RO should be sufficiently big, in order to prevent any data overflow during the measurement in the time window. A smaller time window can reduce the size of the counter, however it risks the capture precision. In addition, the frequency measurement and comparison judgement by RO needs a significantly long time to be completed, hence the response to detected injection campaign cannot be immediate. And the large size of this logic is also vulnerable and easier to be affected by fault injections.

Ring-Oscillator with PLL Phase-Locked-Loop (PLL) was originally used by Miura et al. in [14] for detecting the phase shift disturbance in RO by EMFI. In this proposal, the frequency of a RO is fed into the frequency input of a PLL, hence any disturbance in phase shift comparison (must have two frequencies, one is reference to check the change of phase/frequency distortion).

A technique, using digital RO for detecting frequency disturbance caused by laser, and a PLL, allowing detection of frequency changes in RO, was published in [10]. By using this technique, authors were able to detect faults caused by the laser with the detection rate more than 92%.

Since the PLL is a scarce resource and not always available, we propose a fully-digital sensor which also allows us to achieve higher detection rates.

2.5 Lightweight PRESENT Cipher

To validate the effectiveness of the proposed countermeasure against LFIs, The *ISO/IEC* standardized `PRESENT-80` block cipher [6] is selected as the protection target. This cipher is a classic substitution permutation network (SPN), which consists of 64-bit AddRoundKey, 16 4-bit S-box and 64 bit pLayers, to en-/decrypt 64-bit plaintext/ciphertext using 80- or 128-bit key. In this work, we target

its 64-bit round data registers for injecting the cipher faults, as indicated in Fig. 3(b).

3 Low-Cost Digital LFI Sensor

As previously discussed, PLL-based LFI sensor [10] that senses laser injection through an underlying RO is an effective countermeasure. It is both low-cost and easy to integrate in a complex circuit. However, this countermeasure assumes availability of an existing PLL block. PLL is an analog block used for clock monitoring and generation which is often found in most, if not all, modern FPGAs. However, the need for PLL reduces the portability of the countermeasure to ASIC. Even if PLL are available in ASIC, being a scarce resource, it might not be viable to use it only for countermeasures due to area, power and cost consideration. To overcome this limitation, we propose a fully-digital low-cost LFI sensor. It precisely replaces the PLL with an all digital clock monitoring circuit while still keeping the watchdog RO. The fully digital nature of the sensor makes it versatile for different hardware platforms. The low-cost motivates the possibility of deploying several instances of the sensor if needed. As shown later, this all-digital sensor also shows a much higher detection rate than the original PLL-based solution. In the rest of the section, we discuss the design and features of the proposed sensor followed by its implementation details on FPGA platform. Being an all digital proposal, the cost in ASIC is also limited to only few gates.

3.1 Digital Fault Injection Detector

In this paper, we introduce a novel fault injection detector, as sketched in Fig. 4. This system consists of a multi-inverter RO serving as the frequency disturbance `Watchdog Sensor`, and a `Disturbance Capture` logic comprised of two flip-flops and a logic gate i.e. ($Q1\&\overline{Q2}$). The frequencies from two points (*f1.f2*) on this RO loop are fetched to be sampled by two flip-flops (*FF1.FF2*), being sampled by a derived frequency (*ck-delay*). The two-bit vectors from the two flip-flops manifest whether abnormality occurred in the RO. The function of the entire detection system is detailed in Fig. 5.

In this work, the outputs of three consecutive inverters in `Watchdog Sensor` RO are used as the inputs for the `Disturbance Capture` part, named as *f1*, *ck*, *f2* by signal propagation sequence. Given a stable electrical environment, the three signals will have the same frequency with fixed phase shift, and an opposite polarity to signal *ck*, w.r.t. *f1* and *f2*. FF1 and FF2 are both triggered by the `falling edge` of *ck*, as seen in Fig. 5(a). In absence of signal delay from RO to flip-flops, the sampled values for FF1 and FF2 are respectively `'1'` and `'0'`, as indicated by the blue dotted arrow lines in Fig. 5(a). **Noticeably**, the ripples in this RO will identically affect three frequencies, leading to no impact on the `Disturbance Capture` and thus giving false negatives.

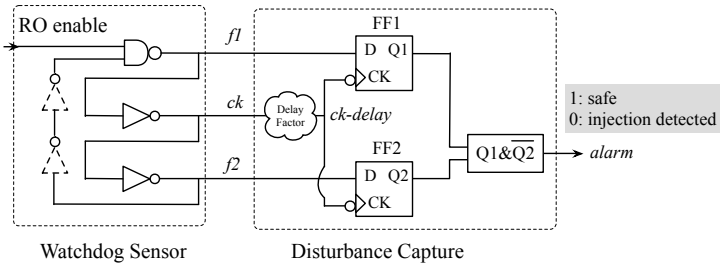


Fig. 4: Topology of the schemed fault injection sensor system.

In order to capture anomalies, a `delay factor` is intentionally inserted into the clock inputs of FF1 and FF2, which is used for introducing a propagation delay of signal *ck* by several clock cycles. In the sequel, each flip-flop is actually clocked by the falling-edge of a delayed *ck* cycle or *ck-delay*, as highlighted by the red dotted arrow lines in Fig. 5(a). The significant merit here is that the ripple in RO only affects the *f1* and *f2* at the injection moment, without immediately affecting the sampling frequency (*ck-delay*) on `Disturbance Capture`. In this way, this system is able to capture bi-directional abnormalities in RO frequency ripples, as explained in the following subsection. The area report is given in Tab. 1. The delay can also be configured by appropriate routing only.

Table 1: Area Report of the All-Digital LFI Sensor

Component	LUT	DFF
Watchdog Sensor	3	0
Disturbance Capture	1	2
Delay	1	0

3.2 Timing Violation Detection

In this part, we qualitatively analyse the proposed sensor against various timing impacts of laser injection to the RO.

Delayed Propagation In case the signal propagation is delayed by the LFI, the frequency of RO can be reduced shortly, as indicated by Fig. 5(b). In this situation, the duty cycles of *f1* and *f2* are temporarily extended. As discussed before, both FF1 and FF2 are clocked by the delayed clock signal *ck-delay*, hence the sampling time in flip-flops at the injection moment is not impacted by the RO disturbance, which is very likely to result in the set-up time violation at *f2*. As can be seen in Fig. 5(a), the sampled value vector from FF1 and FF2 is `'10'` under normal operation. Hence, the sampled vector in presence of timing violation from delayed signal propagation is `'11'`, as highlighted in Fig. 5(b).

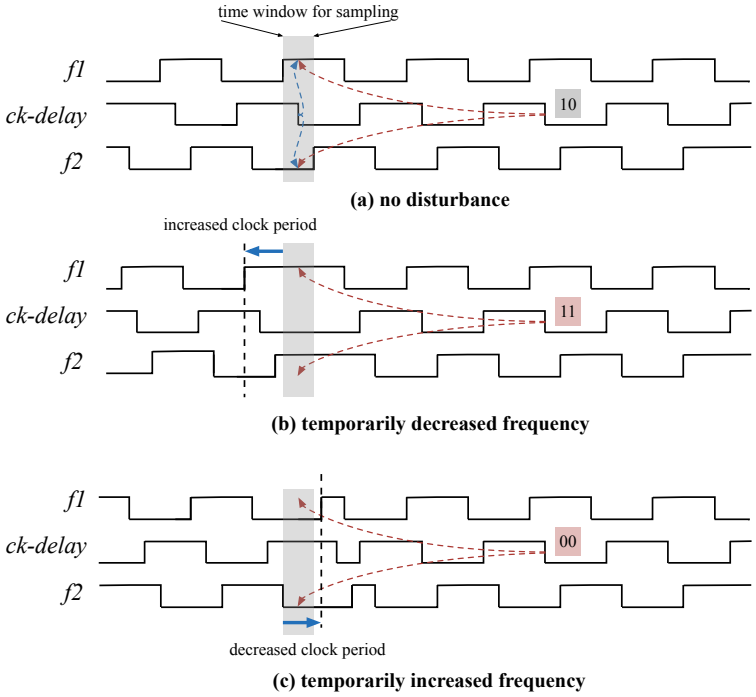


Fig. 5: Timing of low/high-frequency ripple detection.

Accelerated Propagation As aforementioned, the frequency can also be transiently increased by the LFI. In this way, the duty cycle of both *f1* and *f2* can be reduced when the injection affects the RO. Comparatively, the timing will be violated in FF1, rather than FF2, *cf.* preceding situation. As explained in Fig. 5(c), the sampled value vector from FF1 and FF2 becomes `'00'` from the normal `'10'`.

Complex Disturbances It should be emphasized that the timing analyses of disturbance in RO frequency above only considers a single frequency cycle. In a real scenario, the disturbance can be more complex and prevail for several clock cycles to produce a prolonged impact. Hence the extended or shortened duty cycle in *f1* and *f2* can be longer and more complicated than those single-cycle ripples illustrated in Fig. 5(a) and Fig. 5(b). Nevertheless, these complex event can be seen as a combination of several delayed and accelerated event. The timing violation will still be captured as the proposed countermeasure latches the first alarm glitch appearing in each disturbance-period. It allows to alert the main system and launch the fault recovery mechanism. This would also cover the less frequent sampled value of `'01'`. Hence, the complexity in alarm pattern dropping inside the disturbance time window does not impede the disturbance detection.

Since alarm signal is computed from $Q1\&\overline{Q2}$, both abnormalities can result in an alarm value change from `'1'` to `'0'` for alerting the cipher to respond the on-going injection campaign immediately. Here, '&' represents logical AND.

3.3 Target FPGA and Digital-Sensor Implementation

As one of the major FPGA vendors, Xilinx provides a wide range of commercial FPGAs with different technologies. In our work, we tested our circuit on Virtex-5 FPGA which is one of the most popular SRAM based FPGAs on market in recent decade. The basic architecture includes a massive Configurable Logic Block (CLB) array, and numerous peripheral functional logic modules, as Block RAM, Digital Signal Processor (DSP), Digital Clock Manager (DCM), Phase Locked Loop (PLL), as well as rich routing resource channels. In Xilinx terminology, each CLB is comprised by two slices for deploying the implemented logic. Four Look-up-tables (LUTs) in each slice are the main logic resource for implementing the synthesized logic gates, and 4 flip-flops can be configured as registers or latches. A switch-box is deployed besides each CLB for providing rich interconnected resources between the CLB logic to external routing channels. In this work, we mainly target the 64-bit round data registers of PRESENT-80 cipher (see Fig. 3(b)), which are implemented inside the 4 flip-flops in each slice.

The implemented circuit in FPGA-editor view is shown in Fig. 6. To evaluate the detection capability of the proposed sensor system against the previously proposed PLL-based LFI sensor [10], we have deployed both of them on the target Virtex-5 FPGA with similar implementation scheme. Since each slice in Virtex-5 FPGA has 4 flip-flops, we implemented the 64-bit `round data registers` of PRESENT-80 cipher into 16 slices (8 CLBs) as a rectangle. The RO routing path is forced to cross the 4 corners, so as to encompass the protected data registers, as shown in Fig. 6. As shown in Fig. 6, the all-digital `Disturbance Capture` using the 3 inverter outputs from the RO are deployed outside the RO routings. In the second implementation, the `Disturbance Capture` is simply replaced by PLL (not shown) to restore the reference implementation of [10].

4 Experimental Evaluation

4.1 Experimental Setup

The device-under-test (DUT) is a Xilinx Virtex-5 (VLX50T) FPGA, manufactured by 65 nm technology with a `flip-chip` package. The mother FPGA board (Digilent *Genesys*) is fixed on a motorized 2-dimensional (X-Y) stepper stage, with 0.05 μm minimum step size. As the chip substrate may significantly absorb the energy carried by laser photons, we have mechanically milled down the substrate of this FPGA to roughly 130 μm, in order to have sufficient energy penetrated into the active logic (i.e., transistor) layer. `Arduino Due` board is programmed to bridge the controller GUI in computer and the `cipher + countermeasure` system implemented on FPGA. This setting allows us to observe and record the real-time encryption outcome and the alarm signal, as well

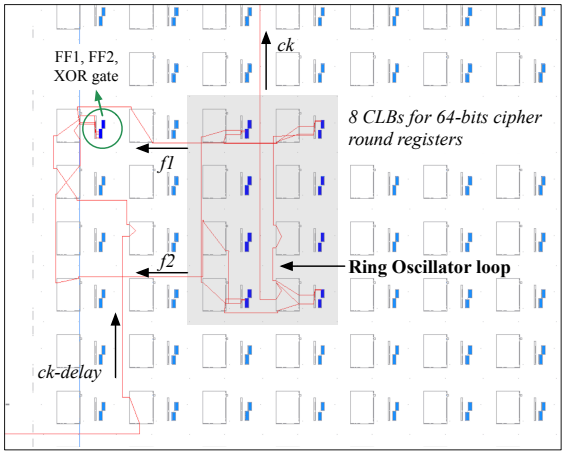


Fig. 6: FPGA implementation scheme of the proposed sensor system and the protected 64-bit round data registers of the PRESENT-80 cipher.

as the location coordinates for each injection of a LFI region scan. The setup is sketched in Fig. 7.

We used a diode pulse laser with 1064 nm wavelength. A 5× magnification reduced the spot size to 60×14 μm, but the effective size is roughly 10% of this size, allowing us to do a very precise laser injection. Injection time can be varied in nanoseconds.

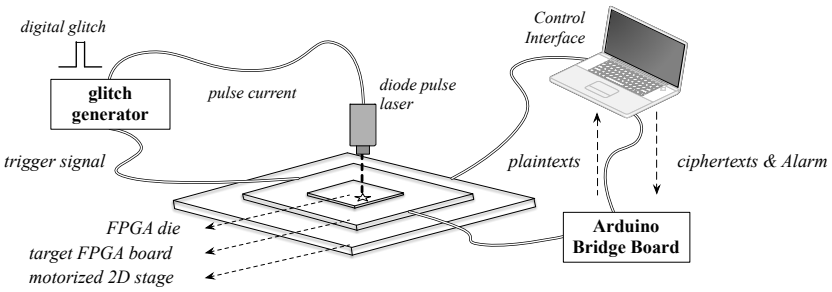


Fig. 7: Illustration of LFI experimental setup.

4.2 Timing Respnose

Fig. 8 shows the timing of the critical signals of this system. `Injection Trigger` is provided by the cipher which denotes the start of the target computation round

for a fault perturbation. `RO frequency` is a signal oscillation of the watchdog RO. In this figure, we captured the signal from a tiny RO with 357 MHz frequency. `Alarm` flags the occurrence of timing violation induced by laser injection. The trigger delay is comprised by (i) the fixed signal delay (from trigger signal on chip glitch generator), and (ii) the adjustable delay time from glitch generator to activation of the diode pulse laser. In our setup, the first fixed delay portion is roughly 100 ns and the second delay is properly set to ensure the injection occurs roughly at the next clock edge. The pulse length of each injection is set randomly between 100 and 200 ns to guarantee the laser is sufficiently powerful to cause bit upsets in registers. The time period of RO ripple is determined by the laser pulse length of each injection. The response time from the frequency ripple appearance to the rising edge of set of the alarm signal is affected by prolonged signal propagation from *ck* to *ck-delay* (see Fig. 4).

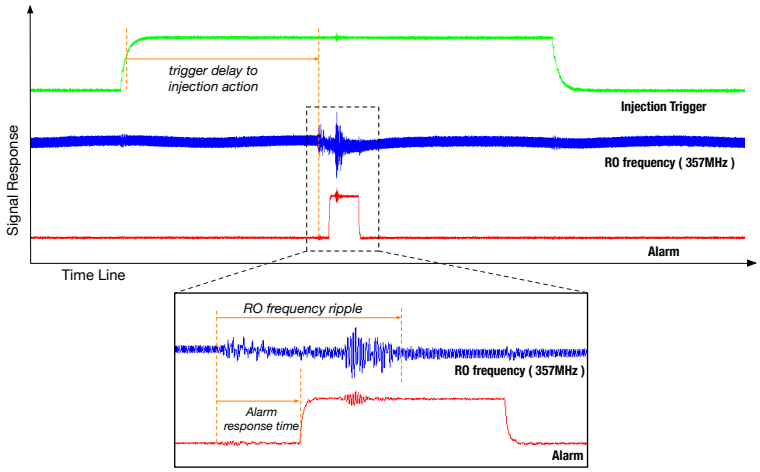


Fig. 8: Timing of signal response of a detected laser fault injection.

4.3 Scanning Results

We have performed the LFI on two implementations on the DUT. The first one was a laser scan of regional CLB array, and the second was a fine-grained single CLB scan. We categorized the faults into three types: (a) `Only Alarm` (Case_(1)) represent the detected injection without cipher faults; (b) `Fault + Alarm` (Case_(2)) refer to the detected injections that induced cipher faults, and (c) `Only Fault` Case_(3) denote the injections that induced cipher faults without triggering the alarm. Scanning results are stated in the following subsections.

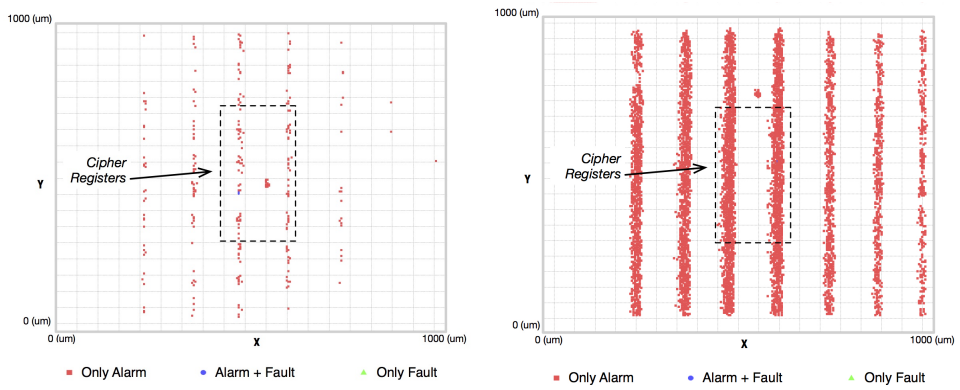
Regional Scan In the first scenario, the implementation details of the cipher and the device architecture are supposed to be unknown to adversaries. For

launching valid fault injection into the point-of-interest (POI), a coarse surface scan towards a big fabric region must be performed by adversaries for finding the POIs. In this experiment, the scan region is intentionally focused on a larger silicon region which does not just cover the RO circumvented cipher data registers, but also the neighbouring regions. The scan matrix is 300×300, which results in 90,000 scanned points with 1 injection per point. Fig. 9 shows the comparison of the LFI scan of the two implementations, and the dotted line rectangle indicates the 8 CLBs where the 64-bit PRESENT round data registers have been implemented.

As can be seen in Fig. 9(a), a PLL-based sensor detected the injection not just in the RO region, but also in the neighboring CLBs (`Only Alarm`=271). A few LFI injections incurred in cipher (Data) faults in the cipher registers, whilst all of them simultaneously triggered the alarm signal (`Fault+Alarm`=3), i.e., no cipher fault went undetected. The scan result for the cipher registers protected by the proposed digital sensor is given in by Fig. 9(b). Similarly, the alarm has been triggered from injections both inside and outside the watchdog RO (`Only Alarm`=5421), and all the induced cipher faults have been detected (`Fault+Alarm`=8). It can be clearly observed that the alarm density for this scan is much higher (5421 vs 271), which implies that this digital sensor system is more sensitive to laser injection *cf.* PLL sensor.

If we only consider the cipher faults, the `Detection Rate` of the sensor can be computed by $detection\ rate = \frac{Case\_(2)}{Case\_(2)+Case\_(3)}$. According to our experimental results, the `Detection Rates` for both regional LFI scans are 100%. Another metric that can be used for quantifying the countermeasure is the `Alarm Rate`, which gives the ratio between the triggered alarms and induced cipher faults. `Alarm Rate` is fair to be applied in a more realistic scenario, this is because the adversaries typically need to perform tedious scan over the chip for finding the exact location of POIs. Any triggered alarm (even without cipher faults) alerts the system to respond to the on-going LFIs, hence paralyzes the attackers. $Alarm\ rate = \frac{Case\_(1)+Case\_(2)}{Case\_(2)+Case\_(3)}$ is used to compute the ratio, which gives 91.33:1 for the PLL sensor, and 678.63:1 for the the digital sensor in this experiment. In addition, for the digital sensor, the lowest laser power to induce the cipher faults is 75% of its full strength, and the lowest power to trigger the alarm is 44%, which further certifies that the sensor is more sensitive to the LFIs, which offers a power security margin of 31%. The detailed comparison results are provided in the upper part of Tab. 2.

Single-CLB Scan A more rigid scenario was also evaluated, which assumes that the adversary knows the details of the implementation and device architecture, particularly the accurate location of the CLB on chip where the security-sensitive round data registers were situated. This way, the adversary is able to directly focus on the CLB to launch a fine-grain fault injection campaign. In this attack, we target a single CLB which has 4 cipher registers implemented inside. Since the effective region of the laser beam is smaller than the CLB size, the scan is still necessary, but the chance to induce cipher faults in registers is



(a) (b)

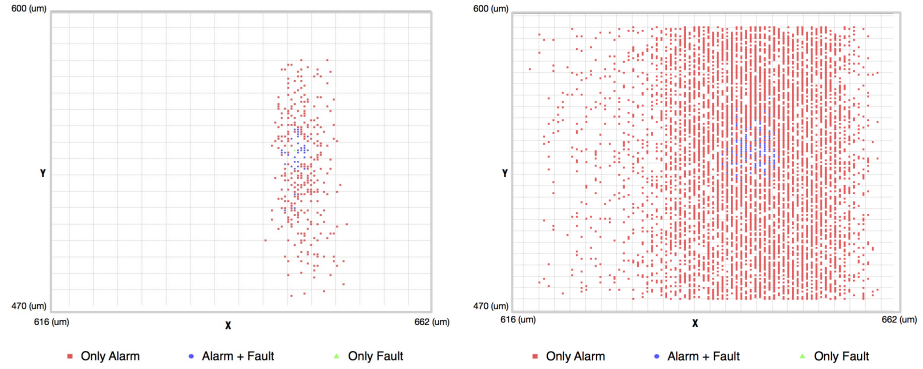
Fig. 9: Laser fault injection scan to regional silicon (a) PLL based LFI sensor; (b) the proposed digital LFI sensor.

much higher. Here, the scan matrix is reduced to 150 × 150, again with 1 injection per point. The experiment results are shown in Fig. 10. Similar to the region scan, injections to both implementation incurred cipher faults and alarm, as summarized in the lower part of Tab. 2. Results show that PLL sensor detected 284 injection without cipher faults, and 33 injections with cipher faults. Noticeably, 1 cipher fault went undetected. In comparison, 4461 injection without cipher faults have been detected using the proposed digital sensor, and all of the 99 cipher faults triggered the alarm. The result implies a higher sensitivity using this RO based digital sensor, *cf.* PLL sensor, under the assumption that the attack was performed by well-prepared adversaries. Similarly, for the digital sensor, the lowest power for triggering alarm (42%) is lower than the minimum power inducing cipher fault (63%), with a power security margin of 21%.

While one cipher fault was missed by the PLL-based sensor (97.06% detection rate), the digital sensor shows 100% detection rate. The general `Alarm Rate` is noticeably higher than the PLL counterpart (46.06:1 vs 9.32:1), as seen in Tab. 2. As explained before, any triggered alarm (detected injection either with or without induced cipher faults) would prevent the attack in a more realistic scenario, so it is safe to conclude that this digital sensor is superior in defending the LFI attacks. At the same time, it has much lower area cost than a scarce PLL block.

4.4 Full Cipher Protection

In total, 24 CLBs are covered by this watchdog RO. However, previous experiments have shown that the injections to neighboring CLBs are also able to trigger the alarm (see Fig. 9), so this RO can actually cover a larger fabric region. In this experiment, we deployed 2 PRESENT-80 ciphers in parallel for filling up the logic resources in a big area of a clock region, as indicated by `PRESENT 1` and `PRESENT 2` in Fig. 11. The higher logic density helps to yield more valid cipher faults.



(a) (b)

Fig. 10: Laser fault injection scan to a single CLB: (a) PLL based LFI sensor; (b) the proposed digital LFI sensor.

Table 2: Experimental results comparison between the PLL based sensor and the presented digital sensor using LFIs.

		Only Alarm Case_(1)	Fault+Alarm Case_(2)	Only Fault Case_(3)	Scan Matrix	RO freq. (MHz)
PLL LFI Sensor (Region Scan)	No.	271	3	0	300×300	≈ 220
	min.Power	54%	90%	n/a		
	Detection	Successful		Failed		
	Detection Rate	$\frac{Case\_(2)}{Case\_(2)+Case\_(3)} = 100\%$				
	Alarm Rate	$\frac{Case\_(1)+Calse\_(2)}{Case\_(2)+Case\_(3)} = 91.33:1$				
Digital LFI Sensor (Region Scan)	No.	5421	8	0	300×300	≈ 206
	min.Power	44%	75%	n/a		
	Detection	Successful		Failed		
	Detection Rate	$\frac{Case\_(2)}{Case\_(2)+Case\_(3)} = 100\%$				
	Alarm Rate	$\frac{Case\_(1)+Calse\_(2)}{Case\_(2)+Case\_(3)} = 678.63:1$				
PLL LFI Sensor (CLB Scan)	No.	284	33	1	150×150	≈ 220
	min.Power	60%	75%	n/a		
	Detection	Successful		Failed		
	Detection Rate	$\frac{Case\_(2)}{Case\_(2)+Case\_(3)} = 97.06\%$				
	Alarm Rate	$\frac{Case\_(1)+Calse\_(2)}{Case\_(2)+Case\_(3)} = 9.32:1$				
Digital LFI Sensor (CLB Scan)	No.	4461	99	0	150×150	≈ 206
	min.Power	42%	63%	n/a		
	Detection	Successful		Failed		
	Detection Rate	$\frac{Case\_(2)}{Case\_(2)+Case\_(3)} = 100\%$				
	Alarm Rate	$\frac{Case\_(1)+Calse\_(2)}{Case\_(2)+Case\_(3)} = 46.06:1$				

In this experiment, the LFI scan is launched towards the region of the two implemented ciphers, with the scan matrix of 300×300 with single injection per point. Similar to prior campaigns, the laser power level is set to random, between 40% to 100% of the full laser strength. Fig. 12 gives the distributions of different fault types. Due to the lower injection density, the number of observed faults is less than the preceding experiments, while information can still

be extracted. There were 69 injections resulting in cipher faults, and among those, 65 triggered the alarm (`Fault+Alarm`=65), leaving only 4 undetected (`Only Fault`=4). Besides, alarm has been triggered for 116 times without cipher faults (`Only Alarm`=116). Thus, the `Detection Rate`, computed using the equations from Tab. 2, for this experiment is 94.20%, and `Alarm Rate` is 2.63:1. This outcome demonstrates that the `Detection Rate` for protecting the whole cipher is still very high. Even with a reduced `Alarm Rate`, the chance to trigger the alarm is still 2.63 times of the chance to induce cipher faults for this densely implemented complete PRESENT-80 primitive. The faults marked as exceptional were faults observed on the I/O and power pads and not sensitive (unrelated to cipher) in nature.

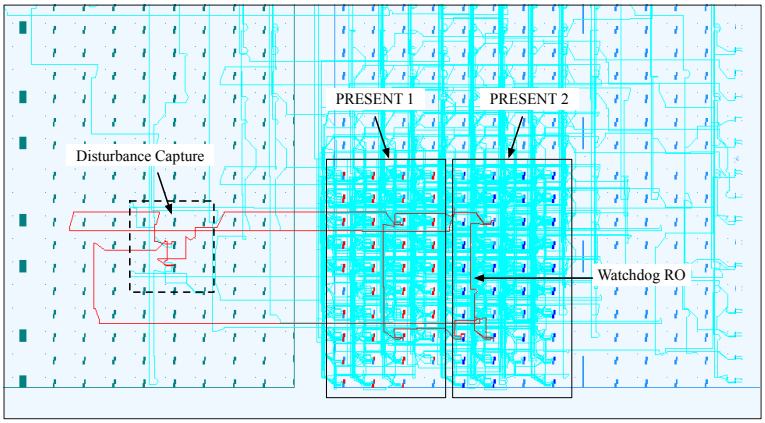


Fig. 11: Countermeasure configuration for protecting full PRESENT cipher.

4.5 Further Discussions

Timing Tuning of Delay Factor As discussed in Sec. 3, a prolonged delay from *ck* to *ck-delay* must be ensured, in order to enforce the falling-edge of *ck-delay* dropping between the rising-edges of *f1* and *f2*, in absence of laser disturbance. This proper timing can be easily achieved by adjusting the propagation time of the routing. Two methods can be applied for this purpose: First, the third-party toolkit can be relied on to control the routing delay for Xilinx FPGAs, such as RapidSmith and Torc [15] [20]. Another, easier method, is to insert a transparent LUT between *ck* and *ck-delay*, configured as `''Route-Thrus''` property, where the LUT has no logic function, only serving as a route point. By relocating the location of this LUT, the delay can be adjusted.

Detection Capability Against Other Fault Injection Methods In this paper, only laser based fault injection is discussed. However, the proposed logic is still promising to be used as a sensor against other fault perturbation techniques,

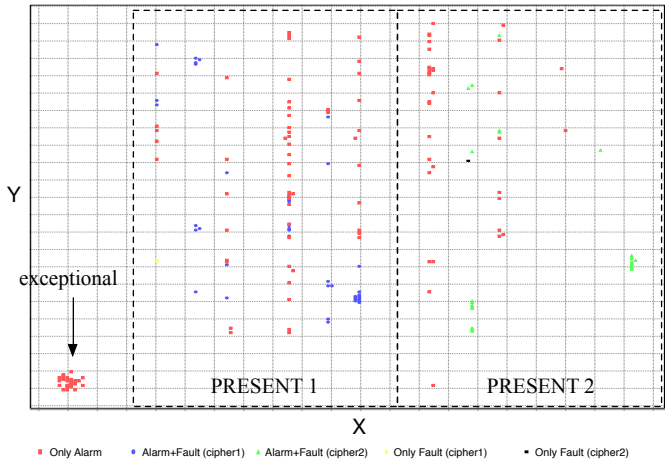


Fig. 12: LFI detection experiment on the two full PRESENT-80 ciphers.

such as EM based fault injection (EMFI). EMFI basically induces eddy current in circuit for causing signal errors, and the current direction relies on the direction of the pulse EM field, i.e., the position of the EM probe. If the current direction follows the signal propagation direction of the watchdog RO, RO frequency would be temporarily accelerated (high-frequency ripple), and otherwise, low-frequency ripple. Therefore, the bi-directional detection capability of the proposed digital sensor is specially useful to detect the EMFI. As well, glitches on power supply would change the RO frequency, hence it should be also effective against global fault injection on power supply on the chip.

False Positives One consideration for the proposed countermeasure is the unwanted false positives that may arise from neighbouring components or environmental variation. As shown in the results, the countermeasure can only be triggered when laser power is in medium to high ranges. Generating such high energy on board is not be obvious for a big range of devices. Moreover, environmental variations are gradual in nature and RO is inherently resistant to such changes. Thus the chances of false positives are quite low for the proposed countermeasure.

5 Conclusions

In this paper, a low-cost fully digital sensor for detecting the malicious laser fault injection in security-critical ICs is presented. This system consists of a multiple-inverter high-frequency RO for producing a stable frequency oscillation, and a disturbance capture logic for detecting the frequency ripple on this RO. In presence of any disturbance from an on-going laser injection, the frequency ripple on RO can be captured by timing violation in the two flip-flops, hence

alerting the system with an alarm signal. The effectiveness of this system is validated on Xilinx 65 *nm* Virtex-5 FPGA. Experimental results on both `round data registers` and full PRESENT-80 cipher show that the proposed digital sensor has a high fault `Detection Rate`, as compare to PLL-based sensor, and being significantly superior in terms of alarm sensitivity (`Alarm Rate`) against laser injections. Since the timing violation can be bi-directionally detected by the two flip-flops, both low-frequency and high-frequency disturbances can be captured, which exceeds the prior glitch-detector countermeasure. Owing to its pure digital and simple architecture, this system can be easily deployed into any digital/hybrid IC environments, particularly as Internet-of-Things (IoT) or embedded endpoints of Cyber-Physical System (CPS) with restricted power and hardware resources.

In the future work, we plan to validate its detection capability against EM and power/clock glitch injection. Moreover, it will be interesting to explore more precise laser setup and the physical limits of proposed countermeasure against laser spot size.

References

1. Amiel, F., Villegas, K., Feix, B., Marcel, L.: Passive and active combined attacks: Combining fault attacks and side channel analysis. In: Fault Diagnosis and Tolerance in Cryptography, 2007. FDTC 2007. Workshop on. pp. 92–102. IEEE (2007)
2. Bar-El, H., Choukri, H., Naccache, D., Tunstall, M., Whelan, C.: The sorcerer's apprentice guide to fault attacks. Proceedings of the IEEE 94(2), 370–382 (2006)
3. Barenghi, A., Breveglieri, L., Koren, I., Naccache, D.: Fault injection attacks on cryptographic devices: Theory, practice, and countermeasures. Proceedings of the IEEE 100(11), 3056–3076 (Nov 2012)
4. Biham, E., Shamir, A.: Differential fault analysis of secret key cryptosystems. In: Kaliski, BurtonS., J. (ed.) Advances in Cryptology CRYPTO '97, Lecture Notes in Computer Science, vol. 1294, pp. 513–525. Springer Berlin Heidelberg (1997), `http://dx.doi.org/10.1007/BFb0052259`
5. Binder, D., Smith, E.C., Holman, A.B.: Satellite Anomalies from Galactic Cosmic Rays. IEEE Transactions on Nuclear Science 22(6), 2675–2680 (Dec 1975)
6. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J., Seurin, Y., Vikkelsoe, C.: PRESENT: An ultra-lightweight block cipher. Springer (2007)
7. Boneh, D., DeMillo, R.A., Lipton, R.J.: On the importance of checking cryptographic protocols for faults. In: Proceedings of the 16th Annual International Conference on Theory and Application of Cryptographic Techniques. pp. 37–51. EUROCRYPT'97, Springer-Verlag, Berlin, Heidelberg (1997), `http://dl.acm.org/citation.cfm?id=1754542.1754548`
8. Endo, S., Li, Y., Homma, N., Sakiyama, K., Ohta, K., Aoki, T.: An efficient countermeasure against fault sensitivity analysis using configurable delay blocks. In: Fault Diagnosis and Tolerance in Cryptography (FDTC), 2012 Workshop on. pp. 95–102. IEEE (2012)
9. Hammouri, G., Akdemir, K., Sunar, B.: Novel puf-based error detection methods in finite state machines. In: International Conference on Information Security and Cryptology. pp. 235–252. Springer (2008)

10. He, W., Breier, J., Bhasin, S., Miura, N., Nagata, M.: Ring oscillator under laser: Potential of pll based countermeasure against laser fault injection. In: International Workshop on Fault Diagnosis and Tolerance in Cryptography 2016. pp. 1–12. IEEE (Aug 2016)
11. Karri, R., Kuznetsov, G., Goessel, M.: Parity-based concurrent error detection of substitution-permutation network block ciphers. In: International Workshop on Cryptographic Hardware and Embedded Systems. pp. 113–124. Springer (2003)
12. Kocher, P., Jaffe, J., Jun, B.: Differential power analysis. In: Annual International Cryptology Conference. pp. 388–397. Springer (1999)
13. Miura, N., Najm, Z., He, W., Bhasin, S., Ngo, X.T., Nagata, M., Danger, J.L.: Pll to the rescue: a novel em fault countermeasure. In: Proceedings of the 53rd Annual Design Automation Conference. p. 90. ACM (2016)
14. Miura, N., Najm, Z., He, W., Bhasin, S., Ngo, X.T., Nagata, M., Danger, J.L.: Pll to the rescue: A novel em fault countermeasure. In: To Appear in Proceedings of the 53rd ACM Design Automation Conference. Austin, TX, USA (2016)
15. Moradi, A., Immler, V.: Early propagation and imbalanced routing, how to diminish in fpgas. In: Cryptographic Hardware and Embedded Systems - CHES 2014 - 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings. pp. 598–615 (2014)
16. Saha, D., Mukhopadhyay, D., RoyChowdhury, D.: A diagonal fault attack on the advanced encryption standard. Cryptology ePrint Archive, Report 2009/581 (2009), http://eprint.iacr.org/2009/581
17. San Pedro, M., Soos, M., Guilley, S.: Fire: fault injection for reverse engineering. In: IFIP International Workshop on Information Security Theory and Practices. pp. 280–293. Springer (2011)
18. Selmke, B., Brummer, S., Heyszl, J., Sigl, G.: Precise laser fault injections into 90 nm and 45 nm sram-cells. In: International Conference on Smart Card Research and Advanced Applications. pp. 193–205. Springer (2015)
19. Skorobogatov, S., Anderson, R.: Optical Fault Induction Attacks. In: Kaliski, B., Ko, ., Paar, C. (eds.) Cryptographic Hardware and Embedded Systems - CHES 2002, Lecture Notes in Computer Science, vol. 2523, pp. 2–12. Springer Berlin Heidelberg (2003), http://dx.doi.org/10.1007/3-540-36400-5_2
20. Steiner, N., Wood, A., Shojaei, H., Couch, J., Athanas, P., French, M.: Torc: towards an open-source tool flow. In: Proceedings of the 19th ACM/SIGDA international symposium on Field programmable gate arrays. pp. 41–44. ACM (2011)
21. Suh, G.E., Devadas, S.: Physical unclonable functions for device authentication and secret key generation. In: Proceedings of the 44th annual Design Automation Conference. pp. 9–14. ACM (2007)
22. Zussa, L., Dehbaoui, A., Tobich, K., Dutertre, J.M., Maurine, P., Guillaume-Sage, L., Clediere, J., Tria, A.: Efficiency of a glitch detector against electromagnetic fault injection. In: Proceedings of the conference on Design, Automation & Test in Europe. p. 203. European Design and Automation Association (2014)