Silicon Echoes: Non-Invasive Trojan and Tamper Detection using Frequency-Selective Impedance Analysis

Tahoura Mosavirik, Saleh Khalaj Monfared, Maryam Saadat Safa, and Shahin Tajik

Department of Electrical and Computer Engineering,
Worcester Polytechnic Institute, Worcester, MA, USA
{tmosavirik,skmonfared,msafa,stajik}@wpi.edu

Abstract. The threat of chip-level tampering and its detection has been widely researched. Hardware Trojan insertions are prominent examples of such tamper events. Altering the placement and routing of a design or removing a part of a circuit for side-channel leakage/fault sensitivity amplification are other instances of such attacks. While semi- and fully-invasive physical verification methods can confidently detect such stealthy tamper events, they are costly, time-consuming, and destructive. On the other hand, virtually all proposed non-invasive side-channel methods suffer from noise and, therefore, have low confidence. Moreover, they require activating the tampered part of the circuit (e.g., the Trojan trigger) to compare and detect the modifications. In this work, we introduce a non-invasive post-silicon tamper detection technique applicable to different classes of tamper events at the chip level without requiring the activation of the malicious circuit. Our method relies on the fact that physical modifications (regardless of their physical, activation, or action characteristics) alter the impedance of the chip. Hence, characterizing the impedance can lead to the detection of the tamper events. To sense the changes in the impedance, we deploy known RF tools, namely, scattering parameters, in which we inject sine wave signals with high frequencies to the power distribution network (PDN) of the system and measure the "echo" of the signal. The reflected signals in various frequency bands reveal different tamper events based on their impact size on the die. To validate our claims, we performed measurements on several proof-of-concept tampered hardware implementations realized on FPGAs manufactured with a 28 nm technology. We further show that deploying the Dynamic Time Warping (DTW) distance can distinguish between tamper events and noise resulting from manufacturing process variation of different chips/boards. Based on the acquired results, we demonstrate that stealthy hardware Trojans, as well as sophisticated modifications of P&R, can be detected.

Keywords: Tamper Detection · Hardware Trojans · Backscattered Side-channel · Physical Layer Security · Scattering Parameters · Impedance Characterization

1 Introduction

Malicious modifications to the designs of application-specific integrated circuits (ASICs) and field programmable gate arrays (FPGAs) endanger the security of several applications. Tampering with the design can be carried out at different phases of the Intellectual Property (IP) design, IP integration, or fabrication, depending on the target platform. FPGAs are vulnerable to tampering even after fabrication and testing due to their reconfigurable natures. These tamper events are often referred to as hardware Trojans (HTs). However,

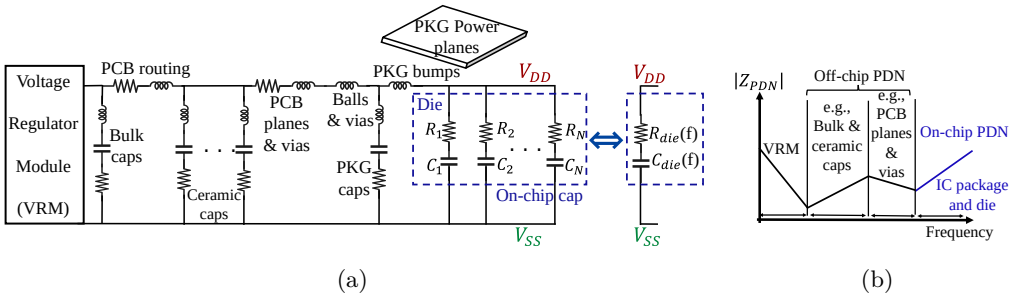
there are other tamper events, which cannot necessarily be classified as hardware Trojans, but still can compromise the security of the system. For instance, an attacker might remove side-channel/fault countermeasures (e.g., reducing the order of masking or redundancy) or manipulate the placement and routing (P&R) of a design without affecting the functionality of the design. Such modifications could make a provably secure design vulnerable to physical attacks. In these cases, no conventional Trojan triggers or payloads can be discovered using functional verification. The impact of such tampering might be observable only under certain physical conditions, e.g., specific temperature, supply voltage, or frequency range.

Several side-channel methods, as well as imaging techniques, have been proposed in the literature to detect such HTs and tamper events. Virtually all passive and non-invasive side-channel techniques [ABK+07, HMLZ20, SKMH14, LL08] suffer from resolution and are incapable of detecting all types of tamper events. Moreover, there is no guarantee to detect dormant Trojans using such passive measurements. In some detection methods, additional measurement circuitry is added to the design to facilitate post-silicon testing for Trojans [CG13, LFM17]. However, this additional circuitry increases the circuit size, manufacturing cost, and system power consumption and makes the detection technique incompatible with legacy systems. Semi- and fully-invasive techniques, on the other hand, are more powerful for detecting stealthy tampering. For instance, laser-assisted side channels (e.g., LLSI [KST21] and TLS [KKTS21]) or imaging techniques (e.g., scanning electron microscopy [VLS+18]) can be deployed to confidently detect very sophisticated tamper events on the die. However, such methods are slow, require package preparations, and are destructive in some cases, and hence, they might not be scalable.

Recently, a non-invasive Trojan detection method, namely EM backscattering, has been introduced, in which EM waves at a *certain frequency* are injected into the chip by an antenna, and its modulated reflection due to the switching activity of transistors is captured by another antenna [NCPZ19, AJN+20, JKPZ22, NYPZ20]. This method assumes that the Trojan trigger causes subtle changes to the impedance of the die, and therefore, such a change should have an influence on the activity and current consumption of the neighboring circuits. Consequently, changes in activity in the time domain should modulate the backscattered signal differently. However, to capture a Trojan through the backscattered EM signal, one needs to apply advanced signal processing and machine learning techniques on several measurements as well as find the best carrier and modulation frequencies as they differ among various technologies and circuit requirements. Moreover, the setup requires a complex and customized EM measurement setup to isolate the signal from RF noise, temperature variations, and other wireless activities in the room.

Driven by the limitations mentioned above, the following question arises: *Does a general non-invasive side-channel technique exist that is able to detect various classes and sizes of tamper events confidently without requiring the triggering or activation of any parts of the circuit under test?*

Our contribution: In this work, we present a non-invasive generic chip-level tamper detection method, which is applicable to all types of tamper events without the need to activate the Trojan trigger or any other part of a malicious circuit. Our approach is predicated on the observation that the impedance of the chip is altered by all classes of physical alterations, regardless of their physical, activation, or action features. We employ scattering parameters (known from the RF field) to characterize the impedance of the chip in the frequency domain, and thus, detect the smallest impedance variations caused by tampering. In this case, we inject sine wave signals into the power distribution network (PDN) of the system and listen to their "echoes," which reveal the changes in the impedance when compared to a reference obtained from a golden sample. We demonstrate that the impedance of the system's PDN over frequency is impacted by various components of the system, from PCB to chip level. Hence, by finding the right frequency bands, *we can actively probe the impedance of the die in a non-invasive manner.* Moreover, we show that



(a) (b)

Figure 1: (a) Equivalent RLC circuit model of the power distribution network of the PCB and chip. (b) Contribution of different parts of the PDN to the impedance over frequency.

various Trojans and tamper events can be detected in different frequency bands depending on their area overhead. In other words, for smaller Trojans, one should move to very high frequencies (GHz bands) to detect them.

To demonstrate the effectiveness of our approach, we performed extensive measurements on several proof-of-concept hardware Trojan benchmarks and tamper events realized on AMD/Xilinx Artix-7 FPGAs manufactured with 28 nm technology. We show that using the Dynamic Time Warping (DTW) distance, we can differentiate between tamper events and noise resulting from manufacturing process variation of different chips/boards. Finally, based on the obtained results, we show one can set different detection thresholds for different frequency bands to detect the stealthy hardware Trojans as well as sophisticated small tamper events confidently.

2 Technical Background

2.1 Power distribution network (PDN)

The PDN consists of electronic components and interconnects from the voltage regulator module (VRM) to the power rails on the chip. Each component plays a role in delivering low noise and constant voltage supply to the power rails on the die. Figure 1a shows the equivalent circuit model of the system's PDN. The PDN covers not only the off-chip components (e.g., bulk capacitors, PCB routing, ceramic capacitors, PCB planes, and vias) but also the on-chip components such as package bumps, on-chip power planes, and transistors' capacitance. The impedance contribution of these components to the overall PDN's impedance is different at various frequency bands. While the equivalent impedance of the PDN at lower frequencies is dominated by the voltage regulator's and off-chip components' impedance, the on-chip components contribute mostly to the impedance at higher frequencies [ZAB+18], see Figure 1b. The parasitic inductance that already exists on each capacitor is the primary cause of this impedance behavior. An ideal capacitor can be modeled as a short circuit at high frequencies. However, the existing parasitic inductance on the capacitor's metals causes a resonance at a particular frequency depending on its capacitance and inductance values. In this case, the capacitor becomes an open circuit at very high frequencies due to its impedance increase at frequencies greater than its resonance frequency. Due to their reduced physical size and lower parasitic inductance, smaller capacitors resonate at higher frequencies. As a result, as the frequency increases, all sets of capacitors, from the large to the small ones, become open circuits and have less of an impact on the PDN impedance. As depicted in Figure 1b, the PDN impedance in higher frequencies are dominated by the on-chip structures due to their smaller dimensions.

The dashed blue region in Figure 1a shows the equivalent RC model of the on-chip capacitance. The wideband on-chip behavior of the circuit can be approximated as N narrowband parallel RC circuits, which are connected to the V_{DD} and V_{ss}. In [ZSS+23,

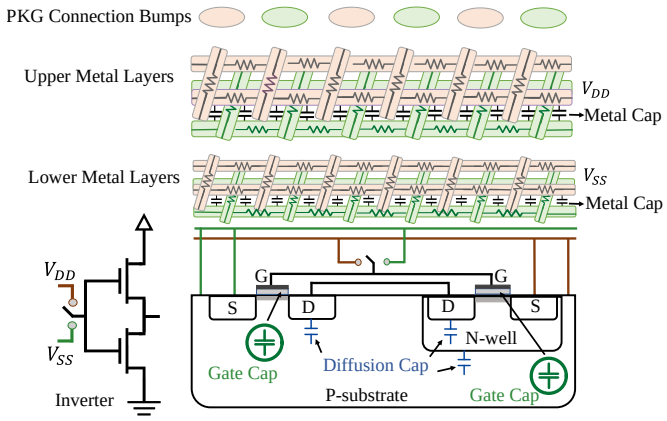


Figure 2: The physical representation of a CMOS inverter cross section and the locations of different types of on-die capacitors. The black capacitors show the capacitance of metal lines, the blue ones show the p-n diode junction diffusion capacitance, and the capacitance shown in green color corresponds to non-switching gate capacitance.

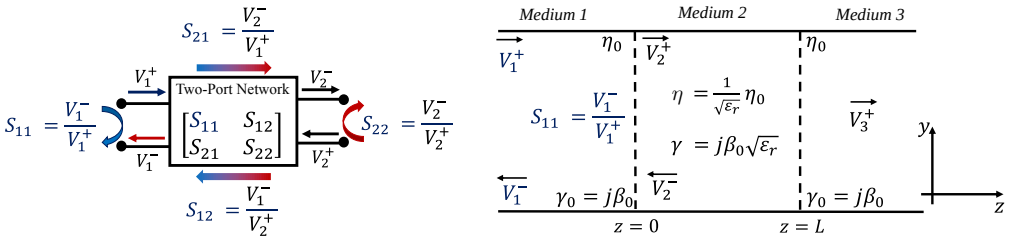
MST23, MGST22], it was shown that the characterization of PDN's impedance in the frequency domain enables the detection of PCB-level tamper events. Naturally, it is conceivable that any tamper event inside the IC should also have an impact on the PDN's impedance. Tampering with the logic gates of the circuit, placement, and routing would change the on-chip capacitance in specific frequency bands depending on the size, location, and nature of the tampering on the chip. In the next subsection, we elaborate on the sources of the on-chip PDN's impedance.

2.2 Sources of On-die Impedance

Several regions of an integrated circuit (IC) contribute to the on-die impedance. As discussed in the previous subsection, the impacts of the impedance of an IC package and its die on the PDN are revealed in high frequencies (see Figure 1b). The ranges of such frequency bands are determined based on the chip's technology and size. On-die capacitance C_{die} and resistance R_{die} are the dominant features of the on-chip impedance in these frequency bands [SSS+11, HKF+18]. Here, we explain the sources of on-die capacitance using the physical structure of a CMOS inverter.

Figure 2 shows the cross-sectional view of an inverter, metal power grid layers, and the locations of the corresponding on-die capacitors. An inverter consists of an NMOS and a PMOS transistor whose drain contacts are connected, as shown in Figure 2. A transistor is nothing more than a switch with an infinite off-resistance and a finite on-resistance. A PMOS transistor consists of an n-well, which is the positively doped source, the drain, and the gate regions. The metal layers grid network, the non-switching gate, and p-n diode junction diffusion are the fundamental contributors to C_{die} [SSS+11]. There also exist sources of R_{die}, which include the power net resistance, transistor channel resistance, transistor gate resistance, and resistance of contacts of n-well and P-substrate [MWK17].

In Figure 2, the colors show the location of each contributing capacitance. The metal capacitance (black color), C_m, is the capacitance associated with the on-die power/ground metallization grid network. The size of these capacitors depends on the density of the grid network, the distance and width of the metal layers, and the permittivity of the materials. To be more specific, in the upper metal layers, C_m is usually larger due to the density of the power and ground meshes. In the lower metal layers, C_m tends to be slightly smaller because the power traces are sparser and thinner. The diffusion capacitance (blue color), C_d, is associated with the p-n diode junctions. It should be noted that C_d and C_m contribute to a small portion of the total C_{die}, and the main contributor is the



(a) (b)

Figure 3: (a) Scattering parameters definition in a two-port network model. (b) The simplified (ideal) transmission line model for normal uniform plane wave incidence on different media (the characteristic impedance of medium 2 is different from medium 1 and 3).

non-switching gate capacitance, C_g.

All non-switching and powered-on circuits contribute to C_g in the chip's PDN. This is because a powered-on transistor has a channel under the gate and contributes to C_{die}, while a powered-off transistor has an inactive channel and does not significantly contribute to the on-die capacitance. When the device is not powered-on, the gates' decoupling capacitance effect is unrecognizable, but as the device gets turned on, the channels start forming, and consequently, C_g dominates C_{die}. When the chip's design is modified, different parts of C_{die} (especially, C_g) would change based on the size, location, and nature of the tamper event, and this changes the equivalent circuit of the on-chip PDN, thus impacting the measured signatures from the chip.

2.3 Impedance Characterization using Scattering Parameters

To characterize the impedance of the PDN in different frequencies, S (Scattering) or Z (Impedance) parameters are deployed [Bog10, Pup20]. Every circuit/electronic component can be represented as a one/two-port network, as depicted in Fig. 3a. S parameters are spectrally measured over the frequency domain and are typically used in RF/microwave engineering to obtain the reflection/transmission properties of the circuit to the applied electromagnetic field [Poz11]. In frequency domain analysis, waveforms are represented by sine waves. Frequency, amplitude, and phase are the three terms that can fully characterize a sine wave. Therefore, we leverage both the amplitude and phase response in the frequency domain to accurately characterize the chip at each frequency point. A Vector Network Analyzer (VNA) is an instrument that can measure the transmitted and/or reflected power of a signal going into and coming back from a component. We use a VNA to inject sine waves into the chip at every frequency sample and record the signal's reflection response from the chip's PDN. The impedance profile can be easily derived from the reflection coefficient. Equation 1 expresses the relationship between the input impedance of the device under test (DUT) and the reflection coefficient:

$$Z_{DUT} = Z_0 \frac{1 + S_{11}}{1 - S_{11}}, \tag{1}$$

where S_{11} is the reflection coefficient, Z_0 represents the reference impedance of the VNA which is 50 Ω, and Z_{DUT} corresponds to the impedance obtained from S_{11}. Depending on the measurement conditions, it might be more convenient to use one of the Z or S parameters and then convert it to the other one. We only deploy S_{11} in our proposed method as the VNA can directly measure it from the chip. However, based on Equation 1, it is observable that the reflection coefficient is another representation of the impedance.

3 Methodology

3.1 Tamper Detection using Scattering Parameters

We explain the changes that occur to the injected voltage wave by the VNA into the chip by analyzing the ideal transmission line model. This model is the backbone of more complex circuits, and understanding its theoretical foundation clarifies our methodology's mechanism. Figure 3b illustrates an ideal transmission line model where there is a change in the characteristic impedance and propagation constant of medium 2 that are represented by η and γ, respectively. For simplicity, we assume that medium 1 and medium 3 are lossless, thus giving a characteristic impedance of η_0 and a corresponding propagation constant of $\gamma_0 = j\beta_0$. We consider medium 2 a non-magnetic ($\mu_r = 1$) medium with a relative permittivity of ε_r. Considering $\beta_0 = \omega\sqrt{\varepsilon_0\mu_0}$, we can rewrite the second medium's propagation constant as $\gamma = j\beta_0\sqrt{\varepsilon_r}$. Considering $\eta_0 = \sqrt{\mu_0/\varepsilon_0}$, we can rewrite the characteristic impedance of medium 2 as $\eta = \sqrt{1/\varepsilon_r}\eta_0$. ε_0 and μ_0 are the permittivity and permeability of the free space, respectively, and β_0 denotes the free space wave number. The VNA injects a voltage wave with the known amplitude of V_1^+ in medium 1, and the reflected voltage wave has an amplitude of V_1^-. After V_1^+ is injected, multiple reflections and transmissions occur in the lines. Based on the model in Figure 3b, the lines' voltages can be written as [AHM+13, CON+04, Poz11]:

$$V_1(z) = V_1^+ e^{-j\beta_0 z} + V_1^- e^{+j\beta_0 z}, V_2(z) = V_2^+ e^{-\gamma z} + V_2^- e^{+\gamma z}, V_3(z) = V_3^+ e^{-j\beta_0 z}, \quad (2)$$

where V_i^+ and V_i^- $(i = 1, 2, 3)$ are forward and backward voltage waves through/from the medium i; however, we assume that there exists no backward voltage wave in medium 3, for simplicity. V_1^+ is a known parameter (injected by VNA), whereas V_1^-, V_2^+, V_2^-, and V_3^+ are unknown values. By enforcing the boundary conditions on the voltage wave components at the interfaces of the media, all these four unknowns can be found. We are interested in obtaining the S_{11} in medium 1 which can be derived as

$$S_{11}(f, \varepsilon_r, L) = \frac{V_1^-}{V_1^+} = \frac{(\eta^2 - \eta_0^2)(1 - e^{+2j\gamma L})}{(\eta_0 + \eta)^2 - (\eta - \eta_0)^2 e^{+2j\gamma L}} \quad (3)$$

where L is the length of the path that the injected wave voltage travels. From Equation 3, it can be concluded that the reflection coefficient depends on three parameters: the frequency band of interest, the relative permittivity of the sample, and the length of the wave's traveling path. On the other hand, the dependence of S_{11} on the frequency has another aspect: frequency and wavelength are inversely proportional to each other. This explains why we can detect smaller size changes in the chip's configuration at higher frequencies. When different chip configurations with different sizes and P&R are exposed to the incident wave injected from the VNA, the changes occurring in Equation 3 parameters will result in a change in the S_{11} profile at distinct frequencies. For example, when the P&R of the circuit is altered, L is changed, and this would cause the chip's reflection response to be different for different P&Rs. If a part of the chip's package or its heat sink is removed, there would be a change in the dielectric properties, and consequently, this variation in the ε_r would result in the reflection coefficient change (see section 6.3).

Tamper events would change the on-chip impedance, which can also be explained by Figure 4, from the chip's equivalent circuit model perspective. This figure demonstrates how the on-chip impedance changes when a tamper event, e.g., the change in the routing of the design, occurs. When the injected signal from the VNA travels through a different routing path, the binary values saved in the SRAM cells are changed (from the left-side to the right, in Figure 4). This would alter the die's equivalent circuit. The change is shown by changing from R_2 and C_2 to R_3 and C_3, as an example for better clarification.

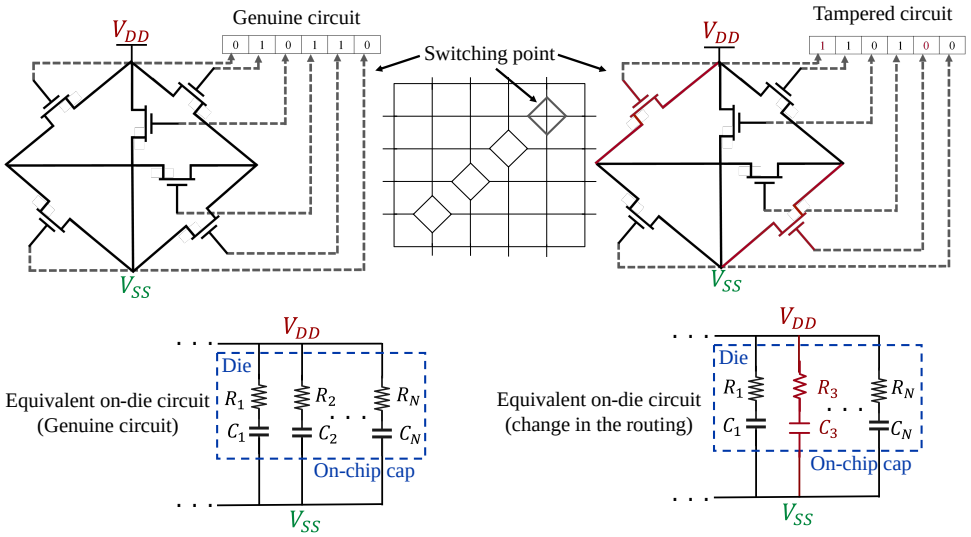


Figure 4: Illustration of a tamper event and how it affects the on-die impedance [Yu15, MWK17]. The figure shows an example of a switch box in an FPGA and how the change in the signal routing leads to a change in the equivalent on-chip impedance.

3.2 Statistical Analysis

Temperature and manufacturing process variations could affect the measurements, leading to noisy traces. While the adverse influence of temperature variation can be mitigated to a certain degree by measurement repetitions and averaging, the noise resulting from manufacturing process variations between different samples requires more sophisticated techniques. We first deploy a straightforward statistical metric, namely, difference in means, to explore the resulting impedance discrepancy from genuine and tampered designs on the same chip soldered to the same PCB. Second, we employ Dynamic Time Warping (DTW) distance for Trojan/tamper detection in a more realistic scenario, where a verifier should compare different systems with each other to reach a conclusion. Note that difference in means is a special case of DTW when two signatures are not experiencing any shifts in frequency (e.g., due to manufacturing process variations).

3.2.1 Difference in Means

In this paper, the number of measurement repetitions for frequency f_i is represented by N_m. We can define ${\mathfrak{S}_{11}}_i^{Gen}$ and ${\mathfrak{S}_{11}}_i^{Tamp}$ as random variables corresponding to the reflection coefficient of the chip at the frequency f_i for the genuine (untouched) and tampered cases, respectively. We use the mean difference (MD) of N_m measurements for each f_i as our statistical measure to differentiate between genuine and tampered chips. The difference in means is a standard statistical metric that measures the absolute difference between the mean values for ${\mathfrak{S}_{11}}_i^{Gen}$ and ${\mathfrak{S}_{11}}_i^{Tamp}$ (for both amplitude ($|\mathfrak{S}_{11_i}|$) and phase of ($\angle\mathfrak{S}_{11_i}$)) can be calculated as follows

$$MD_{Mag}(f_i) = \left|\mu(|{\mathfrak{S}_{11}}_i^{Gen}|) - \mu(|{\mathfrak{S}_{11}}_i^{Tamp}|)\right|, \tag{4}$$

$$MD_{Phase}(f_i) = \left|\mu(\angle({\mathfrak{S}_{11}}_i^{Gen})) - \mu(\angle({\mathfrak{S}_{11}}_i^{Tamp}))\right|, \tag{5}$$

where, $\mu(.)$ is the mean function. Note that the phase responses ($\angle\mathfrak{S}_{11_i}$ values at each frequency point) are constrained between $-\pi$ and π and show a periodic behavior over frequency. To calculate the MD of the phase values, we should perform the unwrapping process for each phase point. Phase unwrapping is used to reconstruct the signal's original

phase values by adding multiples of 2π to each phase input. For each frequency band of interest, distinct experimentally-tuned thresholds of TH_{Mag} and TH_{Phase} are assigned to the corresponding MD profile over frequency. These threshold values enable the verifier to compare the results.

3.2.2 Dynamic Time Warping

The existing manufacturing process variation between different PCBs and chips reveals itself as resistance and dielectric variations. Such variations cause constant shifts in the resonance frequencies and amplitude of the overall scattering parameter profile. Naturally, such shifts in the signature profiles could be interpreted as tampering and raise false alarms. Based on our observations in section 5.4, while these shifts exist, the overall pattern of the signatures over frequency remains similar. To measure such similarities and reduce the impact of such consistent shifts, we deploy Dynamic Time Warping (DTW), which is a similarity measure between time series [Vin68, SC78]. The DTW distance for ${\mathfrak{S}_{11}}_i^{Gen}$ and ${\mathfrak{S}_{11}}_i^{Tamp}$ (for both amplitude ($|\mathfrak{S}_{11i}|$) and phase of ($\angle\mathfrak{S}_{11i}$)) can be expressed as follows,

$$DTW_q(\mathfrak{S}_{11}^{Gen}, \mathfrak{S}_{11}^{Tamp}) = \min_{\delta \in A(\mathfrak{S}_{11}^{Gen}, \mathfrak{S}_{11}^{Tamp})} \left(\sum_{i,j \in \delta} d(\mathfrak{S}_{11}^{Gen}, \mathfrak{S}_{11}^{Tamp})^q\right)^{\frac{1}{q}} \tag{6}$$

where, an alignment path δ is a sequence of index pairs and $A(\mathfrak{S}_{11}^{Gen}, \mathfrak{S}_{11}^{Tamp})$ is the set of all admissible paths.

3.3 Threat Model

In our threat model, we make the following assumptions. The adversary can tamper with the internal design of an ASIC or FPGA prior to the verification. Tampering includes adding/removing logic gates to/from the design, modifying the P&R of the design without any logic addition/removal, or both. The goal of tampering could be to disrupt a system's specific or entire functionality, weaken cryptographic implementation's security or directly leak confidential assets. We further assume that the verifier possesses a golden sample to compare its scattering signature with the suspicious chip's signature. The requirement of golden sample could be eliminated if the verifier can accurately simulate the impedance of a design and use it as the golden signature. However, this would then necessitate having access to the IC/PCB design files, which may be challenging in some scenarios. The verifier needs to have neither control over the design nor specific internal support test circuitry for verification. However, we assume the verifier can halt the clock signal and freeze the chip in a specific state for frequency response measurements. Furthermore, the verifier needs to have access to the chip's power rails to perform the measurements. If the design contains power gating, we assume the verifier can control and disable power gating for verification purposes.

4 Experimental Setup

4.1 Device Under Test

For our experiments, we chose NewAE CW305 boards (NAE-CW305-04-7A35-0.10-X) [CW3] containing an AMD/Xilinx Artix-7 FPGA (XC7A100T), built with a 28 nm technology, see Figure 5. The direct access to the FPGA's PDN on CW305 boards was the primary reason for the selection of these kits. These boards have multiple power domains, namely, a 1 V domain supplying the core (V_{CCINT}) of FPGA, a 3.3 V domain supplying the FPGA I/O banks (V_{CCO}), and a 1.8 V domain as the auxiliary supply voltage (V_{CCAUX}). In this paper, we perform our measurements on V_{CCINT} power domain.

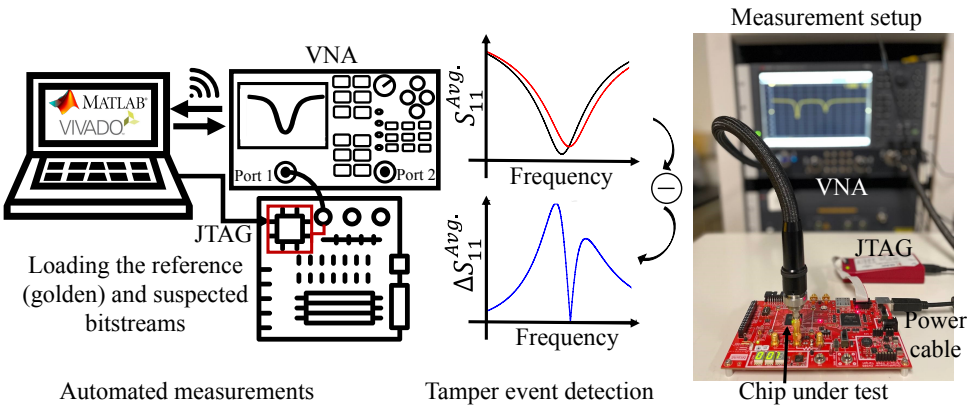


Figure 5: Measurement setup for detecting chip-level tamper events.

CW305 has three SMA connectors providing access to both the high and low sides of a shunt resistor, as well as a 20 dB low noise amplified low side signal. We chose the SMA port on the low side of the shunt resistor, which gives us direct access to the PDN of the FPGA. We deployed the CW305 version, which does not have any decoupling capacitors on V_{CCINT} power rail.

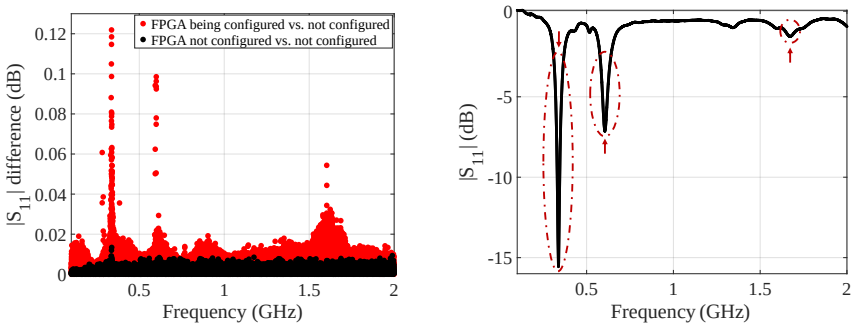
4.2 Measurement Setup

We utilized a Keysight N5227B PNA, which is a microwave VNA capable of operating within 10 MHz - 67 GHz bandwidth. We used N4697J characterization cables [Ope], [N52], which are shielded cables for the same frequency range, see Figure 5. The used VNA has an internal capacitor to filter out the DC voltage on the V_{CCINT}, and therefore, no Bias Tee is needed. We conducted systematic experiments using various implementations on the AMD/Xilinx Artix-7 FPGA to assess the performance of the proposed tamper detection technique.

The experiments can be described as follows: first, the bitstream of the reference (golden) configuration is loaded into the FPGA. The measurements are carried out automatically using a MATLAB script which sends commands to the VNA and reads back the reflection responses. We use this script to load the desired bitstream into the FPGA using JTAG and save the reflection response data on a computer for analysis. We performed the measurements in different frequency bands (the results and details of different bands will be explained in Sect. 5). Within each frequency band, we configured the VNA to measure equally-spaced 100,000 frequency points to ensure the maximum spectral resolution. Before performing the measurements, we precisely calibrated the setup using Keysight's N4694A electronic calibration module. We carried out the calibration until the end of the measurement cable for every frequency band. Then, VNA connects to the computer via WiFi through a TCP/IP link and starts to capture the 1-port reflection response traces of the loaded configuration. Thereafter, the same procedure is performed for the suspicious (possibly modified) configuration, and the reflection responses are saved by the computer in the desired frequency band for further analysis. Please note that all the experiments are performed when the circuit is in its idle state (i.e., no clock signal was provided), and hence, no switching activity exists on the chip.

5 Results

We performed S_{11} measurements for various classes of hardware Trojans and tamper events. Unlimited malicious tamper events occur to a design; naturally, we cannot cover



(a) (b)

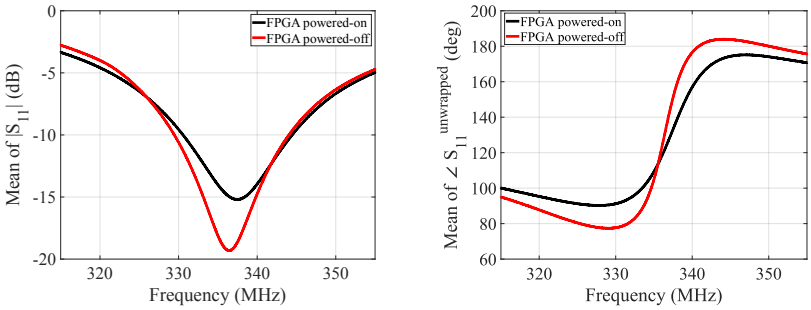
Figure 6: (a) The difference in the amplitude of the reflection responses between the case where the chip is powered on but not configured and the case the chip is being configured. (b) The amplitude of the reflection response of the chip and the highlighted frequency ranges of interest (this response is shown here to show the significant frequency bands, and it corresponds to the chip being powered on but not configured).

all of them. However, we select a few tamper events, which can represent various tamper categories to show our frequency-selective method capability in covering different threats. We prioritize our experiments based on the change in the size of the modified circuit, from the maximum to the minimum change in size. Furthermore, we investigated the effect of different placement and routing of the same implementation on the chip's S_{11} response. For Trojan and tamper experiments, we perform the measurements after the completion of the FPGA configuration. To compare two signatures, both golden and tampered circuits should be in identical states. Hence, we provide no clock signals to the implementation after the completion of the configuration and perform the measurements in state 0.

One of the important considerations in the process of developing our tamper detection technique is to achieve a high SNR. Typically, the SNR of a VNA, which operates in the frequency domain, is constant over its entire frequency range. In our method, to increase the SNR, and consequently, reduce the effect of thermal noise and environmental changes, we took three actions: first, we set the VNA power to 10 dBm for all of the experiments to ensure a strong reflection signal. This way, the reachability of the signal was improved. Second, we performed an external integration during the measurements using the measurement setup, shown in Figure 5, by measuring each reflection response 100 times at each frequency sample over the desired bandwidth and taking an average of responses at each frequency point. Based on the intra-genuine mean difference results (the difference in means between two genuine chip measurements at different trials), we defined an experimentally-tuned threshold for the amplitude and phase response at each sensitive band of interest. Finally, we used the difference in means (see section3.2) of these 100 measurements to compare the reference and tampered chip signatures using the corresponding threshold in every frequency band.

5.1 Finding the Frequency Bands of Interest

Localizing the frequency bands of interest is of great significance since performing measurements in narrower bandwidths increases the SNR and mitigates the effect of environmental changes. The frequency bands of interest are those in which the impedance is sensitive to any logic addition/removal and P&R changes. Hence, we designed a set of experiments and performed several wideband $|S_{11}|$ measurements within 100 MHz to 2 GHz to discover the frequency bands of interest. For this end, we first performed two separate measurements for the case where the FPGA is powered on but not configured. The $|S_{11}|$ difference of chip signatures for these two experiments are shown in Figure 6a in black color. It should



(a) Amplitude profile. (b) Phase profile.

Figure 7: The mean of reflection response of the chip in case of powering on and off the FPGA over 315 MHz - 355 MHz.

be noted that as there is no configuration loaded into the FPGA in these experiments, the FPGA is in its idle state, meaning there is no switching activity on the chip. Then, we performed another $|S_{11}|$ measurement for the case that the FPGA is being configured by a bitstream containing an AES-128 IP. The $|S_{11}|$ difference profile of the experiment in which the FPGA is during configuration and the one with no configuration is calculated and given in Figure 6a in red color.

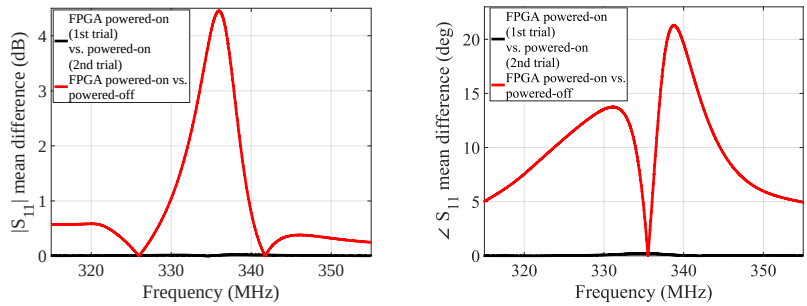
The comparison of $|S_{11}|$ difference profiles in Figure 6a guides us to where we should conduct narrowband experiments to detect the chip-level tamper events with high confidence. The baseline of this approach is the fact that transistors can be modeled as switches, and when the chip is being configured, the input of transistors would shift to a high state (from 0 to 1), connecting a number of switches to prepare the chip for establishing the configuration. These newly created connections would modulate the injected incident voltage wave by VNA, and the comparison between these two experiments' reflection responses would reveal the sensitive ranges of the spectrum. Based on this analysis, we found three frequency bands where we can perform our systematic experiments. The first band is found to be within 315 MHz - 355 MHz, the second band is between 560 MHz - 650 MHz, and the third band falls within 1.65 GHz - 1.7 GHz. These are the spectral ranges where we have the highest sensitivity and SNR for the reflection response experiments, as we can observe a high peak in the $|S_{11}|$ difference profiles in the aforementioned bands in Figure 6a. Figure 6b demonstrates the original $|S_{11}|$ profile for the case study where the FPGA is powered on but not configured. Please note that the response shown in Figure 6b is intended to illustrate the significant frequency bands.

5.2 Case Studies

We designed three categories of experiments to demonstrate the effectiveness of our approach. In the first category, we start with simple experiments where the entire FPGA die is involved, e.g., comparing the powered-off and powered-on FPGA, as well as configured vs. unconfigured FPGA. In the second group, we continue with tampering experiments, in which we compare the scattering signatures of masked AES DOM implementations with different security orders. Moreover, we change the P&R of the same AES design to observe whether it affects the impedance of the system. In the last category of experiments, we conduct experiments on three HT benchmarks from Trust-Hub, namely AES-T1100, AES-T1600, and AES-T1800.

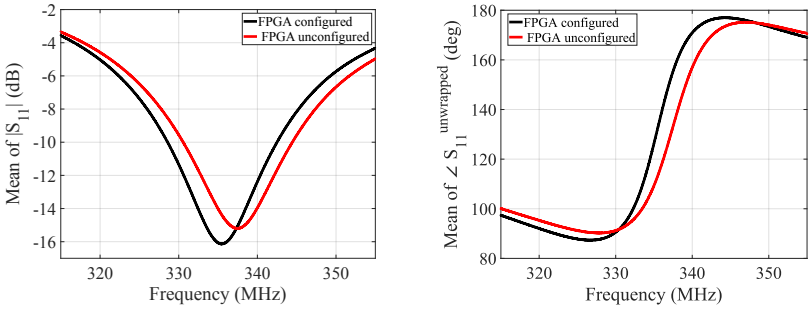
5.2.1 Changes in the Overall FPGA State

Our first set of experiments involved the state of the entire FPGA. We started with a case study to compare the impedance of the FPGA in the powered-on and powered-off



(a) Amplitude profile. (b) Phase profile.

Figure 8: The mean difference of reflection response of the chip in case of powering on and off the FPGA over the band of 315 MHz - 355 MHz.



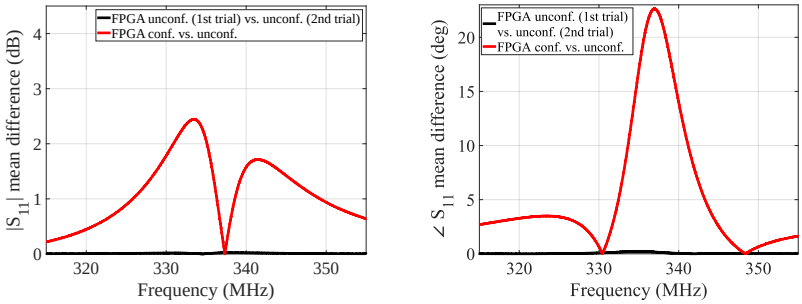
(a) Amplitude profile. (b) Phase profile.

Figure 9: The mean of reflection response of the chip for the configuration mode of the FPGA over 315 MHz - 355 MHz.

states. In the powered-on case, the FPGA is powered on using V_{CCINT} without any bitstream configurations. The first portion of the spectrum where we observe a change in the scattering signature is the 315 MHz - 355 MHz band. This band is where the global resonance frequency of the circuit takes place around 338 MHz, see Figure 6b. The mean of the reflection response of the chip (for 100 measurements) in case of powering on and off the FPGA is shown in Figure 7. The mean difference profiles of 100 measured S_{11} signatures for this case are given in Figure 8.

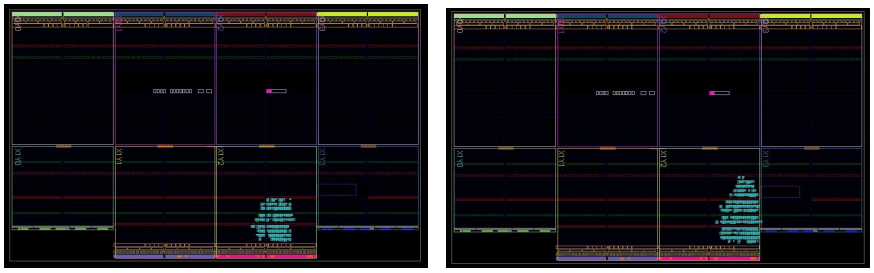
Second, we compare the scattering signature of the case where the FPGA is in the unconfigured powered-on state with the case where the FPGA is configured with a bitstream containing an AES-128 circuit. The mean difference profiles of 100 measured S_{11} signatures for the configuration mode experiments are shown in Figure 10. The reference (golden) configuration is the case that chip was powered on (with no configuration), and experiments using this implementation are performed in different trials to obtain the intra-distance between the golden configuration signatures (the graphs shown in black color in Figures. 8 and 10). The mean of the reflection response of the chip for the configuration mode of the FPGA is shown in Figure 9. The mean of S_{11} signatures are given in Figures 7 and 9 for the first two experiments to show the S_{11} signatures' magnitude and unwrapped phase response before the subtraction is performed in Figures 8 and 10.

When the FPGA is powered off, the portion of the FPGA's equivalent RLC model responsible for creating connections to power the chip is not connected to the PDN. When the chip is powered on, this portion of the circuit is present, and this is the main reason for the detected changes shown in Figures. 8 and 10. The paralleled on-die capacitance and package inductance are dominant features within the PDN within the 50 MHz to 500 MHz frequency band [MWK17, SSS+11]. This can be clearly seen in Figures 7 and 9 with



(a) Amplitude profile. (b) Phase profile.

Figure 10: The mean difference of reflection response of the chip for the configuration mode of the FPGA over the band of 315 MHz - 355 MHz.



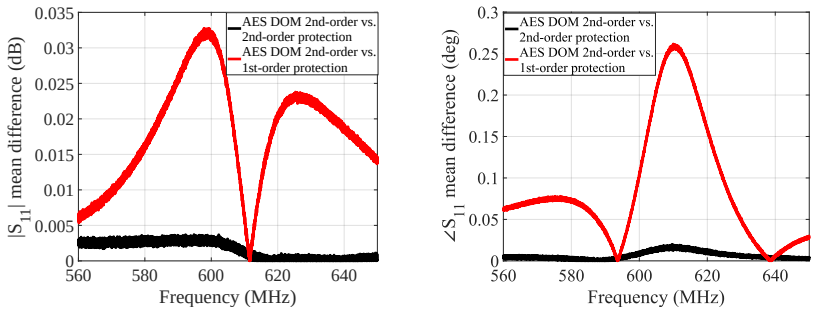
(a) AES DOM first-order protection. (b) AES DOM second-order protection.

Figure 11: AES DOM layout on the FPGA.

the resonances created around 338 MHz. It can also be seen that the reflection response profiles are less noisy in 315 MHz - 355 MHz band than the ones in higher frequencies. As mentioned in section 3.2, we unwrapped the phase responses in both experiments to make the distribution of phase continuous and be able to compare the corresponding points to each other. We can set an experimentally-tuned threshold of 0.05 dB and 0.5 deg for the amplitude and phase responses based on the mean differences for this lower portion of the spectrum in Figures 8 and 10.

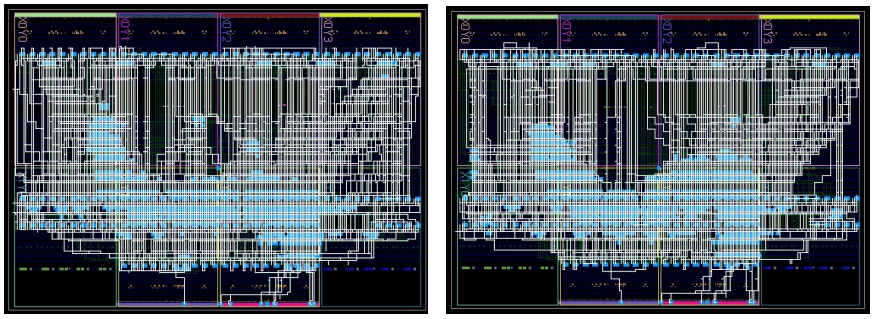
5.2.2 Tampering with Security Order and P&R of a Cipher

In the next round of experiments, we intended to verify the efficacy of our proposed tamper detection method for cases a) where the impact size of tampering is significantly smaller than our previous case studies, and b) the situation where the entire FPGA die



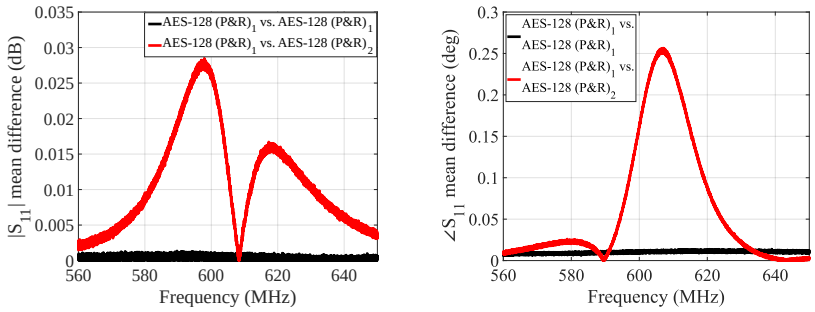
(a) Amplitude profile. (b) Phase profile.

Figure 12: The reflection response of the chip in case of different protection orders of AES DOM over the band of 560 MHz - 650 MHz.



(a) $(P\&R)_1$ (b) $(P\&R)_2$

Figure 13: Change of P&R in a genuine AES layout on the FPGA.

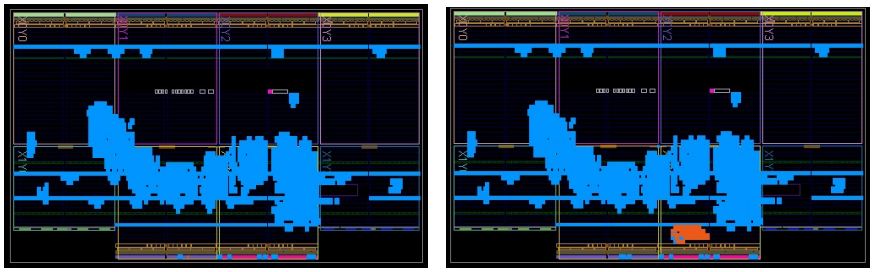


(a) Amplitude profile. (b) Phase profile.

Figure 14: The reflection response of the chip in case of different routing of the AES implementation over the band of 560 MHz - 650 MHz.

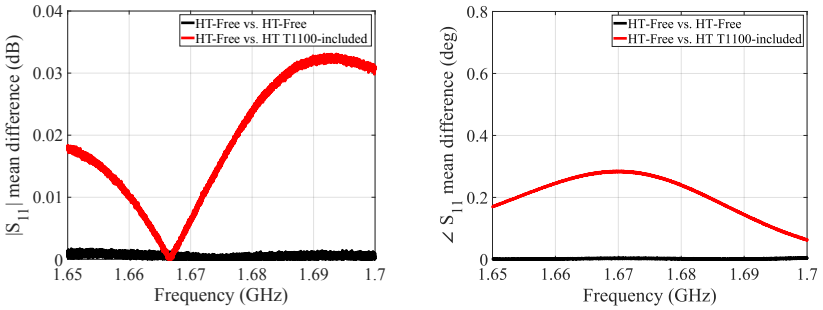
was involved (see Sect. 5.2.1). First, we implemented domain-oriented masked (DOM) AES with first and second-order protection. DOM is a generic masking scheme, introduced in [GMK16], that enables hardware designs to have arbitrary protection orders. The primary reason for this experiment is to see what would happen if an attacker tampers with a side-channel protected implementation and reduces its protection order. We performed our experiments to see in which frequency band we could distinguish between different orders of protection with more confidence. The designs of AES DOM circuits with different protection orders are given in Figure 11. Moving from first to second-order protection would increase the size of the circuit 55.31% compared to the first-order AES DOM. It should be noted that for these two implementations, we let the EDA tool perform the placement and routing and we did not fix the design location, as one of the goals of this experiment was to assess our method's effectiveness in detecting the change in the logic elements and placement and routing of the circuit. The S_{11} mean difference profiles of the chip for these two experiments are shown in Figure 12.

For the next case study, we considered a tamper event, which alters the P&R of the design. For example, in [EGMP17], it has been shown that a Trojan could be realized by only changing the P&R of the configuration. Therefore, in the next set of experiments, we keep the size of the circuit and logic elements unchanged and intend to alter the circuit's placement and routing to assess our method's efficacy in detecting these types of modifications. We implemented an AES-128 circuit and let Vivado compile, place, and route the design in two different trials. The implementations of the AES circuit with different placement and routing are given in Figure 13. We performed the experiments for these two circuits, and the results are reported in Figure 14. Based on the mean differences for this portion of the spectrum, we set an experimentally-tuned threshold of 0.005 and 0.05 for the amplitude and phase profiles, respectively.



(a) HT-free AES circuit. (b) HT-included AES circuit.

Figure 15: Genuine (HT-free) and HT T1100-included AES circuits.



(a) Amplitude profile. (b) Phase profile.

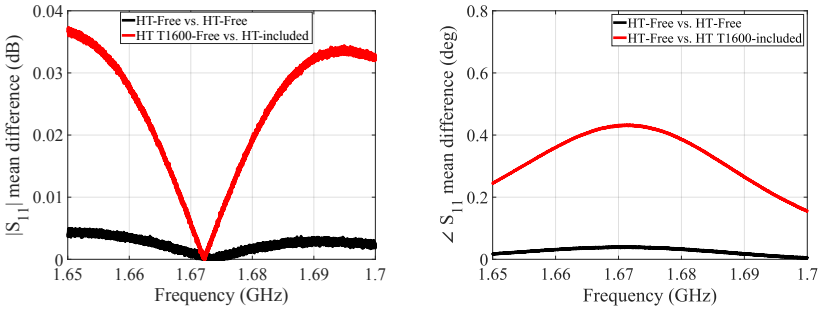
Figure 16: S_{11} of the chip for adding HT T1100 within 1.65 GHz - 1.7 GHz.

In FPGAs, signals are routed using switch points whose structure is shown in Figure 4. Different configurations result in different signal routings when the bitstream values are set to 0 or 1. This would lead to a change in the on-chip capacitance, and subsequently, the characteristic impedance and reflection response of the chip in this bandwidth. We also can consider the chip interconnects as transmission lines. When the stimulus signal is injected by the VNA into these two designs, it will experience different lengths through the transmission line when the voltage wave is traveling through different routes (see section 3.1).

5.2.3 Hardware Trojans

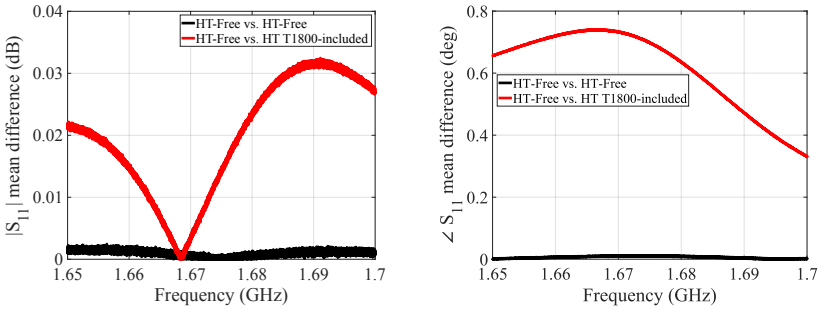
We performed another set of experiments to assess the effectiveness of the proposed detection method for detecting small and inactive alterations, and the best examples of such alterations are dormant HTs. In principle, as discussed in section 3.1, we should rely on higher frequencies for detecting smaller circuit alterations since the increase in frequency would decrease the detection wavelength. We implemented three HT circuits to differentiate between HT-free and HT-infected circuits by measuring the reflection response from the chip. It should be noted that the HT is not activated, and the circuit is in its idle state in all experiments to generalize the method's applicability to dormant HTs, which are highly challenging to detect due to their stealthy nature. We utilized the AES-T1100, AES-T1600, and AES-T1800 benchmarks (register-transfer level (RTL) level HTs) from Trust-Hub [Tru]. In these HT implementations, the original HT-free design is an AES-128 cryptographic IP, which uses an 11-stage pipeline to perform the ten stages of AES encryption on the 128-bit block. We chose these three HTs as independent case studies because they show different trigger and payload mechanisms. However, the applicability of our method extends to other HT benchmarks as well. It is worth mentioning that these three HTs are physically realized through addition of transistors or gates.

In implementing the HT-included circuits, if the modified bitstream is subjected to



(a) Amplitude profile. (b) Phase profile.

Figure 17: S_{11} of the chip for adding HT T1600 within 1.65 GHz - 1.7 GHz.



(a) Amplitude profile. (b) Phase profile.

Figure 18: S_{11} of the chip for adding HT T1800 within 1.65 GHz - 1.7 GHz.

the default compilation, placement, and routing, the addition of the HT causes the EDA tool to change the placement and routing of most logic elements in the overall circuit, and this extensive change makes the alteration easier to detect regardless of the Trojan's size and activity. To make the HT more stealthy, we let Vivado compile, place, and route the HT-included circuit, and we fixed all the cells and logical elements in their location. We then directly remove the HT circuit logic elements and their connections. This way, the HT-free circuit is created while leaving the placement of logic elements unchanged. To illustrate this process, the HT-free and the HT-included circuit implementations are shown in Figure 15 for one of the HT case studies (AES-T1100). In Figure 15b, the logic elements shown in blue color demonstrate the genuine (HT-free) AES circuit. The area in orange color is the HT circuit consisting of the trigger and payload, whose size is 2.9% of the AES circuit. The payload of AES-T1100 modulates its activity using a spread-spectrum technique to create a power consumption pattern that leaks the AES key. The trigger is a sequential circuit that looks for a predefined sequence of values at the input of the AES circuit to activate the payload.

The other two HT circuits are implemented similarly, but their layouts are not shown for the sake of brevity. AES-T1600 is structured in such a way that its payload modulates an unused pin to generate an RF signal on the chip. This signal can be used to transmit the AES key bits. The HT's triggering circuit consists of sequential logic elements for activating the payload when a predefined sequence of values is detected at the input of the AES circuit. The measured reflection response mean differences for AES-T1100 and AES-T1600 are given in Figures 16 and 17, respectively. It should be mentioned that the size of the HT circuit is 2% of the entire AES circuit in AES-T1600.

The last design is HT T1800, whose payload is a shift register that continuously rotates after the Trojan activation phase is complete, thus resulting in an increase in the on-chip power consumption and a decrease in its expected lifetime. The measured reflection

Table 1: The ratio of the maximum difference in means of tampered $|S_{11}|$ to the maximum difference in means of genuine $|S_{11}|$ for different experiments (one experiment from each tampering group from subsections 5.2.1, 5.2.2, and 5.2.3 is selected).

Frequency	FPGA powered-on vs. off to powered-on vs. on	AES-DOM 2nd-order vs. 1st-order protection to 2nd-order vs. 2nd-order protection	AES-T1100-included vs. HT-free to HT-free vs. HT-free
around 338 MHz	**82.23**	7.48	2.82
around 596 MHz	5.09	**10.36**	6.98
around 1.66 GHz	29	6.98	**26.17**

response mean differences for AES-T1800 are given in Figure 18. The size of AES-T1800 is 1.8% of the AES circuit. Based on the mean differences for this portion of the spectrum, we set an experimentally-tuned threshold of 0.005 and 0.05 for the amplitude and phase profiles, respectively. From the results, it is observable that all inactive HTs are successfully detected with high confidence.

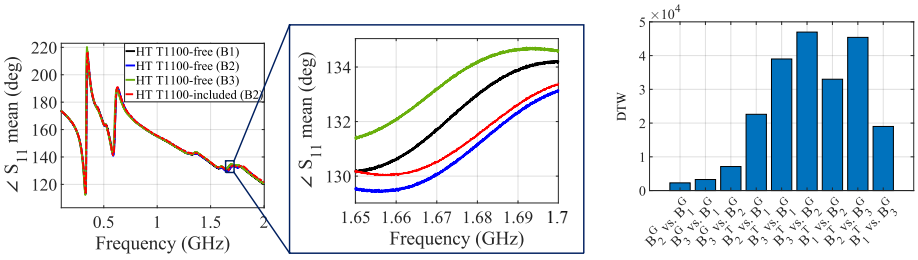
5.3 Tamper Area Overhead vs. Frequency Bands

Overall, the analysis of the scattering signatures at different frequency bands shows that the changes in amplitude of the reflection response are lower than the changes in the phase values, and in higher frequencies, the phase signature suffers from less noise than the amplitude of S_{11}. Please note that we chose the thresholds for each frequency band differently because the size of the circuit and the nature of tamper events are different in the three spectral ranges of interest. We did not use the wideband measurements for all tamper events discussed in this work because, in that case, the SNR would be greatly impacted by noise and environmental variation. We offered the frequency selection approach to increase the resolution (the number of sampled points) in each band. Random errors can arise from the uncertainty of the VNA measurements. Since the uncertainty of measurement is inherent, the impact of such noise can be reduced by repeating the measurements and averaging the extracted signatures.

We find each experiment's most sensitive bandwidth so that we can carry out high-resolution measurements in those bands. We calculated the ratio of the maximum difference in means of tampered $|S_{11}|$ to the maximum difference in means of genuine $|S_{11}|$ for different experiments and reported the results in Table 1. One experiment from each tampering group is selected to perform the analysis. We can see that for the tamper events with larger circuit sizes, the ratio is higher in lower frequencies, and as we go to higher frequencies, we can detect tamper events with smaller sizes with more confidence. For the first case study, where we power off and power on the FPGA, the size of the circuit, which is added to the circuit after powering on the chip, is greater (ratio = 82.23) than in other case studies. Therefore, it highly impacts the first frequency band and resonates around 338 MHz. For the next experiment with an AES circuit with different protection orders, we see its effect at around 596 MHz as the size of the added circuit is 35.6% of the entire circuit. Finally, in case of the addition of an inactive HT, the size of the added circuit is 2.9% of the entire circuit; hence, Trojan can be detected at GHz band at around 1.66 GHz.

5.4 Impact of Manufacturing Process Variation

In a real-world scenario, the verifier receives a chip for verification, which is different from the golden sample. Therefore, we must ensure that the effect of Trojan insertion on the impedance is greater than the manufacturing process variation between different chips. Therefore, we performed S_{11} experiments on three CW305 boards (B_1, B_2, and B_3). For our experiments, first, we configured the FPGAs with the same AES implementation,



(a) (b)

Figure 19: (a) Examples of wideband and narrowband (zoomed-in view) unwrapped $\angle S_{11}$ signatures in case of HT-free designs and HT T1100 added to B2. (b) Cross DTW distances from $\angle S_{11}$ in case of HT-free designs and adding HT T1100 for different chips within 1.65 GHz - 1.7 GHz.

including the HT T1100-free design, and performed the characterization (B_1^G, B_2^G, and B_3^G). Afterward, we configured the same FPGAs with the HT T1100-included design and then performed the measurements (B_1^T, B_2^T, and B_3^T). Figure 19a shows examples of the phase response signatures of genuine designs compared to one of the Trojan-inserted designs, B_2^T. We selected the Trojan-infected design for these experiments as the size of the Trojan circuits is a small fraction of the total AES implementation (in the order of 2%). Hence, if the Trojan circuit can be detected from chip-to-chip measurements, the technique's applicability can be extended to larger circuit sizes as well. As described in section 3.2.2, we used DTW to define a quantitative metric for our measurements to reduce the impact of manufacturing process variations. Here, we only deploy the mean of unwrapped phase responses to calculate DTW as the environmental noises and process variations demonstrated a lower impact on the phase than the amplitude response. Due to having access to a limited number of FPGA samples, we cross-checked the genuine design on different FPGAs with genuine and tampered designs implemented on the other FPGA samples. We used MATLAB's *dtw* function to calculate this distance. Figure 19b shows DTW calculated from the phase responses within 1.65 GHz - 1.7 GHz bandwidth. The first three bars in Figure 19b show the DTW distance of the genuine design on different FPGAs. The other bars indicate the distance between all combinations of genuine and tampered designs on different FPGAs (B_n^T vs. B_m^G, where m,n=1,2,3). It can clearly be observed that in HT-included experiments, the DTW distance has higher values compared to the HT-free experiments for chip-to-chip measurements.

Based on these observations, the verifier can compute the DTW distances between all pairs of genuine signatures in the profiling phase. The verifier should choose the maximum DTW and set a threshold for detecting Trojans above or equal to this value. In the verification phase, the verifier calculates the distances between the signature of the suspicious sample and all genuine samples. Afterward, she compares these obtained distances with the threshold set in the profiling phase. The suspicious sample is considered tampered if all distances are above the threshold. On the other hand, if all these distances are less than the threshold, the sample is considered genuine. If some of the distances are above the threshold and some below it, the verifier should apply majority voting to conclude. Naturally, if there is a tie, no decision can be made.

6 Discussion

6.1 Comparison with Related Works

Table 2 compares the proposed tamper detection method in this work with other sensing methods in the literature [ABK+07, SKMH14, LL08, CG13, SSF+14, NCPZ19, KST21,

Table 2: Qualitative comparison between chip-level tamper detection methods.

Method	Legacy Syst. Compat.	Invasiveness	Complexity/ Cost	Activation Req.	Meas. Time
Power SCA [ABK+07]	Yes	No	Low	Yes	Low
EM SCA [SKMH14]	Yes	No	Low	Yes	Low
Timing SCA [LL08]	Yes	No	Low	Yes	Low
Delay-based [CG13]	No	No	Low	Yes	Low
FIB Imaging [SSF+14]	No	Yes	High	No	High
EM Backscattering [NCPZ19]	Yes	No	High	Yes	Low
Laser Probing [KST21]	No	Yes	High	No	High
SEM [VLS+18]	No	No	High	No	High
Optical Imaging [ZAV+21]	No	Yes	High	No	High
This work	**Yes**	**No**	**Low**	**No**	**Low**

VLS+18, ZAV+21] in terms of system compatibility, invasiveness, complexity, measurement time, and Trojan activation requirement. The proposed method in this work is compatible with legacy systems as it only needs a connection to the PDN of the system. Moreover, it is non-invasive, and the required measurement setup is a VNA, which is available in many test and characterization labs. Besides, as it directly characterizes the impedance, it does not require any HT triggering or active circuits to detect tampering. Finally, the proposed method can capture S-parameters in the order of a few milliseconds to seconds, depending on the configured frequency resolution on the VNA.

6.2 Applicability of Method in Various Cases

Applicability to ASICs: Since every Trojan insertion and tamper event would change the overall impedance of the die (see section 2.2), the proposed method in this work could, in principle, be applied to ASICs as well. One of the main structural differences between FPGAs and ASICs is the number of interconnects and routing resources. Thus, a specific tamper event might involve more routing resources on FPGAs compared to ASICs. Nevertheless, the same tamper event involves transistors, logic gates, layout placement, and routing of an ASIC and, consequently, impacts the die impedance. The impact of individual CMOS logic gates on the impedance of ASICs has been previously verified in [KKH+20]. Moreover, the changes in the impedance of ON/OFF states of the chips other than FPGAs also were shown to be detectable in [SMT23].

Impact of Unique IDs on the Impedance Profile: Several chips, including FPGAs, have unique IDs [AMD23]. Such IDs are stored on specific non-volatile memories, such as eFuses. In this case, a question regarding their adverse impact on the impedance signature and, consequently, on the effectiveness of the proposed method in this work arises. While such IDs affect the impedance profile, their impact will probably be observed at a different frequency band due to their distinct technology and size compared to the CMOS logic. Note that in our experiments, each FPGA had its own Device DNA value; still, we were able to detect the Trojans. However, the impact of various identifiers on very small tamper events still needs to be investigated more thoroughly, and it deserves its own study.

Applicability of other Impedance Characterization Methods: Other impedance characterizations, such as time-domain reflectometry (TDR) [FNH+18], might also be useful as an alternative method for detecting chip-level tamper events. The main requirement for an instrument to detect highly sophisticated and small-size tamper events is the high-frequency resolution, supporting frequencies higher than 1 GHz, and high output powers. Thus, inexpensive and portable VNAs [Edy] or on-FPGA VNAs [MST23] may not be able

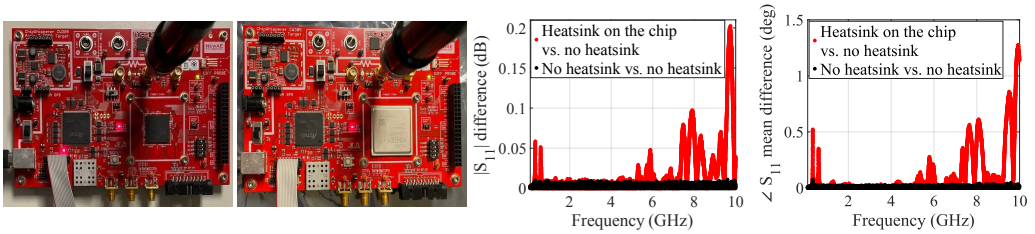


Figure 20: Placement of a heat sink on the FPGA and its effect on the chip's reflection response over 100 MHz - 10 GHz. (a) DUT without the heat sink. (b) DUT with the heat sink. (c) amplitude profile. (d) phase profile.

to detect such stealthy Trojans.

6.3 Sensing External Changes to the Chip Environment

Out of curiosity and to understand whether the proposed approach is also applicable to the IC package, we designed another experiment. In this case, we investigated the effect of changes to the chip's environment beyond its die by placing an object on the chip's package surface and recording the chip's scattering response. Detecting such tampering is of great importance since adversary might have tampered with the chip's package to prepare it for SCA or FI attacks, but the modification to the package might not be visible to the verifier during the verification.

We first performed two measurements for the normal case where no object is placed on the chip (see Figure 20a). The $|S_{11}|$ and $\angle S_{11}$ differences between these two experiments are shown in Figure 20c and 20d, respectively (black graphs). Then, we performed another experiment where we placed a heat sink on the chip during the measurement (see Figure 20b). The S_{11} difference of chip signatures, with and without heat sink, is shown in Figure 20c and 20d in red color. In these experiments, the chip is powered on, but no operation is being performed on the chip. Please note that the heat sink is taken from another Xilinx FPGA family, and therefore, the marking on the heat sink shown in Figure 20b does not show the model of our DUT. It is clearly observable that the heat sink's placement on the chip's surface is detectable at high frequencies with more confidence. The observable effect can be explained by the dependency of S_{11} parameter to the dielectric property (permittivity). When the heat sink is directly placed on the chip, there would be a change in the dielectric properties of the overall signal which is being reflected, and consequently, this variation in the ε_r would result in the reflection coefficient change (see section 3). On the other hand, the permittivity is itself a function of frequency ($\varepsilon_r(f) = \varepsilon_r'(f)$ - $j\varepsilon_r''(f)$); hence, the addition of the heat sink would show its effect at higher frequencies due to the higher dielectric losses (dissipation factor (f) = $\varepsilon_r''(f)/\varepsilon_r'(f)$) at the GHz regime[Poz11].

6.4 Evading the Detection or Reversing Tamper Events

A question might be raised about the feasibility of evading the detection using sophisticated tamper events (e.g., including a single transistor Trojan) or undoing the tampering effect on the impedance by other gates/FFs. While we were able to detect modifications in order of 2% of the total circuit, a single transistor tampering might not be detectable. Still, most Trojans consist of more than one transistor and change the routing of the design. Consequently, the change to the impedance would be significant, leading to a successful tamper event detection. On the other hand, each chip-level tamper event impacts the die's PDN impedance over the entire spectrum. Hence, if the adversary intends to evade

the detection, the two-dimensional (S_{11}, f) signature of the chip needs to be equalized using other gates/FFs, which is a challenging (if not impossible) task due to the unique physical locations of each gate/FF and their parasitics impacting the wave reflection. However, reversing chip-level tampering might still be feasible if the tampering impact on the spectrum is less than the detection threshold of the detection method.

7 Conclusion

In this work, we presented a frequency-selective chip-level tamper detection method based on the reflected scattering parameter analysis of the chip. We demonstrated that the impedance of the chip's PDN can be used as a reliable feature to detect configuration modifications, including the change in the logic elements, placement, and routing. We deployed the scattering signatures in high frequencies to directly probe the chip's die and obtain its frequency response. The reflection response of the chip's die in various frequency bands reveals different tamper events based on their impact size on the die. We performed extensive experiments on several tamper events on various FPGA implementations. We demonstrated that the effect of even small and dormant hardware Trojans and modifications of P&R on the impedance could be observed. By employing a statistical metric, namely difference in means (MD), we showed that these tamper events could be detected with high confidence. We also addressed the manufacturing process variation for different chips using DTW distance. We further showed that we could detect even external tamper events to the package, such as placing a heat sink on the surface of the chip. However, such package tampering deserves a separate thorough study and is beyond the scope of this work.

Acknowledgment

This work was sponsored by Electric Power Research Institute (EPRI). In addition, we gratefully thank James Eakin, Jacob Bouchard, and Douglas Petkie from the Laboratory for Education and Application Prototypes (LEAP) at Worcester Polytechnic Institute (WPI) for providing us with access to the N5227B PNA and shielded characterization cables for our experiments. We also thank Mohammad Hashemi from WPI for his help with FPGA implementations.

References

[ABK+07] Dakshi Agrawal, Selcuk Baktir, Deniz Karakoyunlu, Pankaj Rohatgi, and Berk Sunar. Trojan detection using IC fingerprinting. In *2007 IEEE Symposium on Security and Privacy (SP '07)*, pages 296–310, 2007.

[AHM+13] S Arslanagić, Troels Vejle Hansen, N Asger Mortensen, Anders Heidemann Gregersen, Ole Sigmund, Richard W Ziolkowski, and Olav Breinbjerg. A review of the scattering-parameter extraction method with clarification of ambiguity issues in relation to metamaterial homogenization. *IEEE Antennas and Propagation Magazine*, 55(2):91–106, 2013.

[AJN+20] Sinan Adibelli, Prateek Juyal, Luong N. Nguyen, Milos Prvulovic, and Alenka Zajic. Near-field backscattering-based sensing for hardware Trojan detection. *IEEE Transactions on Antennas and Propagation*, 68(12):8082–8090, 2020.

[AMD23] AMD/Xilinx. 7 series fpgas configuration user guide ug470. 2023.

[Bog10] Eric Bogatin. *Signal and Power Integrity–Simplified*. Pearson Education, 2010.

[CG13] Byeongju Cha and Sandeep K Gupta. Trojan detection via delay measurements: A new approach to select paths and vectors to maximize effectiveness and minimize cost. In *2013 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pages 1265–1270. IEEE, 2013.

[CON+04] Lin-Feng Chen, Chong Kim Ong, CP Neo, Vasundara V Varadan, and Vijay K Varadan. Microwave electronics: measurement and materials characterization. *John Wiley & Sons*, 2004.

[CW3] CW305 Artix FPGA Target. NewAE. https://rtfm.newae.com/Targets/CW305%20Artix%20FPGA/.

[Edy] Edy555. NanoVNA. URL: https://nanovna.com/.

[EGMP17] Maik Ender, Samaneh Ghandali, Amir Moradi, and Christof Paar. The first thorough side-channel hardware Trojan. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 755–780. Springer, 2017.

[FNH+18] Daisuke Fujimoto, Shota Nin, Yu-Ichi Hayashi, Noriyuki Miura, Makoto Nagata, and Tsutomu Matsumoto. A demonstration of a HT-Detection method based on impedance measurements of the wiring around ICs. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 65(10):1320–1324, 2018.

[GMK16] Hannes Groß, Stefan Mangard, and Thomas Korak. Domain-oriented masking: Compact masked hardware implementations with arbitrary protection order. In *ACM Workshop on Theory of Implementation Security*, 2016.

[HKF+18] Chulsoon Hwang, Jingook Kim, Jun Fan, Joungho Kim, and James L Drewniak. Modeling of on-chip power distribution network. In *Noise Coupling in System-on-Chip*, pages 93–138. CRC Press, 2018.

[HMLZ20] Jiaji He, Haocheng Ma, Yanjiang Liu, and Yiqiang Zhao. Golden chip-free Trojan detection leveraging Trojan trigger's side-channel fingerprinting. *ACM Transactions on Embedded Computing Systems (TECS)*, 20(1):1–18, 2020.

[JKPZ22] Erik J Jorgensen, Andrew Kacmarcik, Milos Prvulovic, and Alenka Zajić. Hyperspectral image recovery via reliability-weighted compressed sensing for hardware Trojan detection. *IEEE Access*, 10:96568–96580, 2022.

[KKH+20] Toshiki Kanamoto, Koki Kasai, Kan Hatakeyama, Atsushi Kurokawa, Tomoyuki Nagase, and Masashi Imai. A simple yet precise capacitance estimation method for on-chip power delivery network towards emc analysis. *IEICE Electronics Express*, 17(14):20200198–20200198, 2020.

[KKTS21] Thilo Krachenfels, Tuba Kiyan, Shahin Tajik, and Jean-Pierre Seifert. Automatic extraction of secrets from the transistor jungle using laser-assisted side-channel attacks. In *USENIX Security Symposium*, pages 627–644, 2021.

[KST21] Thilo Krachenfels, Jean-Pierre Seifert, and Shahin Tajik. Trojan awakener: Detecting dormant malicious hardware using laser logic state imaging. In *Proceedings of the 5th Workshop on Attacks and Solutions in Hardware Security*, page 17–27, 2021.

[LFM17] Maxime Lecomte, Jacques Fournier, and Philippe Maurine. An on-chip technique to detect hardware Trojans and assist counterfeit identification. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 25(12):3317–3330, 2017.

[LL08] Jie Li and John Lach. At-speed delay characterization for IC authentication and Trojan horse detection. In *2008 IEEE International Workshop on Hardware-Oriented Security and Trust*, pages 8–14. IEEE, 2008.

[MGST22] Tahoura Mosavirik, Fatemeh Ganji, Patrick Schaumont, and Shahin Tajik. Scatterverif: Verification of electronic boards using reflection response of power distribution network. *ACM Journal on Emerging Technologies in Computing Systems*, 18(4):1–24, 2022.

[MST23] Tahoura Mosavirik, Patrick Schaumont, and Shahin Tajik. Impedanceverif: On-chip impedance sensing for system-level tampering detection. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 1:301–325, 2023.

[MWK17] Leonard MacEachem, Xin Jie Wang, and Tad Kwasniewski. On-die power grid broadband model determination using a priori narrowband measurements. In *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pages 1493–1496, 2017.

[N52] N5227B PNA Network Analyzer. Keysight. https://www.keysight.com/us/en/assets/9018-04327/technicalspecifications/9018-04327.pdf.

[NCPZ19] Luong N Nguyen, Chia-Lin Cheng, Milos Prvulovic, and Alenka Zajić. Creating a backscattering side channel to enable detection of dormant hardware Trojans. *IEEE transactions on very large scale integration (VLSI) systems*, 27(7):1561–1574, 2019.

[NYPZ20] Luong N. Nguyen, Baki Berkay Yilmaz, Milos Prvulovic, and Alenka Zajic. A novel golden-chip-free clustering technique using backscattering side channel for hardware Trojan detection. In *2020 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, pages 1–12, 2020.

[Ope] Operation & Service Manual - N4697J/K 1.85 NMD-1.85 mm -f- to 1.85 mm Flexible Test Port Cables. Keysight. https://www.keysight.com/us/en/assets/9018-04766/service-manuals/9018-04766.pdf.

[Poz11] David M Pozar. *Microwave engineering*. John wiley & sons, ISBN: 9781118213636, 2011.

[Pup20] Peter J Pupalaikis. *S-parameters for Signal Integrity*. Cambridge University Press, ISBN: 9781108784863, 2020.

[SC78] Hiroaki Sakoe and Seibi Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1):43–49, 1978.

[SKMH14] Oliver Söll, Thomas Korak, Michael Muehlberghuber, and Michael Hutter. EM-based detection of hardware Trojans on FPGAs. In *2014 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*, pages 84–87, 2014.

[SMT23] Maryam Saadat Safa, Tahoura Mosavirik, and Shahin Tajik. Counterfeit chip detection using scattering parameter analysis. In *2023 26th International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS)*, pages 99–104. IEEE, 2023.

[SSF+14] Takeshi Sugawara, Daisuke Suzuki, Ryoichi Fujii, Shigeaki Tawa, Ryohei Hori, Mitsuru Shiozaki, and Takeshi Fujino. Reversing stealthy dopant-level circuits. In *Cryptographic Hardware and Embedded Systems–CHES 2014: 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings 16*, pages 112–126. Springer, 2014.

[SSS+11] Larry D. Smith, Shishuang Sun, Mayra Sarmiento, Li Zhe, and Karthik Chandrasekar. On-die capacitance measurements in the frequency and time domains. *DesignCon, Santa Clara, CA*, 2011.

[Tru] Trust-Hub. Hardware Trojan Benchmarks. https://trust-hub.org/#/home.

[Vin68] Taras K Vintsyuk. Speech discrimination by dynamic programming. *Cybernetics*, 4(1):52–57, 1968.

[VLS+18] Nidish Vashistha, Hangwei Lu, Qihang Shi, M Tanjidur Rahman, Haoting Shen, Damon L Woodard, Navid Asadizanjani, and Mark Tehranipoor. Trojan scanner: Detecting hardware Trojans with rapid SEM imaging combined with image processing and machine learning. In *ISTFA 2018: Proceedings from the 44th International Symposium for Testing and Failure Analysis*, page 256. ASM International, 2018.

[Yu15] Huang Yu. Circuit design for FPGAs in sub-threshold ultra-low power systems. *Master thesis, School of Engineering and Applied Science, University of Virginia*, 2015.

[ZAB+18] Shuze Zhao, Ibrahim Ahmed, Vaughn Betz, Ashraf Lotfi, and Olivier Trescases. Frequency-domain power delivery network self-characterization in FPGAs for improved system reliability. *IEEE Transactions on Industrial Electronics*, 65(11), 2018.

[ZAV+21] Boyou Zhou, Aydan Aksoylar, Kyle Vigil, Ronen Adato, Jian Tan, Bennett Goldberg, M. Selim Ünlü, and Ajay Joshi. Hardware Trojan detection using backside optical imaging. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 40(1):24–37, 2021.

[ZSS+23] Huifeng Zhu, Haoqi Shan, Dean Sullivan, Xiaolong Guo, Yier Jin, and Xuan Zhang. Pdnpulse: sensing pcb anomaly with the intrinsic power delivery network. *IEEE Transactions on Information Forensics and Security*, 2023.