Thermal Scans for Detecting Hardware Trojans

Maxime Cozzi, Philippe Maurine, Jean-Marc J.-M. Galliere

Thermal Scans for Detecting Hardware Trojans

Maxime Cozzi, Jean-Marc Galliere, and Philippe Maurine

LIRMM, 161 rue ada, Montpellier, France
maxime.cozzi@lirmm.fr

Abstract. It is well known that companies have been outsourcing their IC production to countries where it is simply not possible to guarantee the integrity of final products. This relocation trend creates a need for methodologies and embedded design solutions to identify counterfeits but also to detect potential Hardware Trojans (HT). Hardware Trojans are tiny pieces of hardware that can be maliciously inserted in designs for several purposes ranging from denial of service, programmed obsolescence etc. They are usually stealthy and characterized by small area and power overheads. Their detection is thus a challenging task.
Various solutions have been investigated to detect Hardware Trojans. We focus in this paper on the use of thermal near field scans to that aim. Therefore we first introduce and characterize a low cost, large bandwidth ($20\,kHz$) thermal scanning system with the high detectivity required to detect small Hardware Trojans. Then, we experimentally demonstrate its efficiency on different test cases.

Keywords: Trojan detection, lock-in thermography, thermal mapping, thermal modeling

1 Introduction

Hardware security recently emerged as an important research problem. Attacks such as Side Channel proved that it was possible to break trusted ciphering algorithms such as Rijndael and therefore raised the problem of securing electronic devices [1]. This concern is even greater in the economic context where the quest for better performances pushes CMOS technology close to its limits and to an exponential growth of Integrated Circuits (IC) complexity and cost. Consequently, more and more companies are fabless and are outsourcing their production to foreign countries. As a result, ensuring the integrity of integrated products has become a critical issue because most of electronic systems, even critical ones, rely on ICs. We obtained these last years more and more evidence that counterfeits, cloning and Trojan insertion have become a credible vector of attack against electronic systems [2], [3], [4].

The increasing complexity of ICs and the scaling of technology have made Trojan detection a particularly challenging task as both their size and power overhead have become infinitesimal in their respective applicative context, thus creating the need for high performance methodologies (inspired from Side Channel Attacks for most of them) [2] [3], [4] and embedded design solutions [5], [6].

Infra-Red (IR) thermography has proven to be efficient in detecting small defects in ICs [7]. It has also demonstrated to be efficient, *by simulation only,* for Trojan identification in [8]. One drawback is their reliance on IR camera set-ups which have a very limited frame rate, a limited number of pixels and are costly. Within this context we propose in this paper a low cost and high detectivity thermal platform characterized by a bandwidth of $20\,kHz$ as well as Side Channel Attacks (SCA) inspired techniques to exploit IR data collected.

The organization of this paper is as follows. Section 2 provides a theoretical background on thermal emissions. It also gives a state of the Art relative to our application domain and illustrates that DC silicon thermal response can be modeled by a first order system to deduce a usage policy of tunable IR platforms to manage their spatial resolution and detectivity. Section 3 details the proposed low cost and high detectivity experimental IR set-up. Then, section 4 gives experimental results demonstrating the efficiency of the proposed IR set-up. Finally, in section 5, performance of the proposed platform regarding Trojan detection is given, as well as the SCA inspired techniques defined for this purpose.

2 State of the Art

This section aims at introducing ICs thermal mapping. Many other thermal investigation methods have been proposed, such as thermoreflectance presented in [9]. Here, we only present techniques that are relevant to our measurement system.

2.1 Light Emission From Above 0 °K Bodies

It is well known that every body above absolute $0\,^{\circ}K$ emits light [10]. This principle is described by Plank's law which shows that the wavelength of the light emitted by a black body is linked to its temperature by the following formula:

$$I_{\lambda,b} = 2.h.c_0^2.\lambda^{-5}.e^{\frac{-h.c_0}{k.\lambda.T}} \tag{1}$$

where c_0 is the electromagnetic radiation propagation speed in a vacuum, h and k respectively are the Plank and Boltzmann constants, and λ is the wavelength of the emitted light. Considering a classical environment for the Device Under Test (DUT), e.g a room temperature of $25\,^{\circ}C$, we get from (1) that light emission should be observed in the (IR) spectrum. Silicon is transparent to wavelengths above $1100\,nm$. It is therefore possible to detect hot spots using IR sensors, through the backside of DUTs [7].

One of the challenges in IR thermography is compensating for natural emissivity of materials. Indeed, if every body does emit light depending on its temperature, it does not radiate the same intensity depending on its constitution. For that matter, we define emissivity as the ratio between the intensity of the radiation emitted by the studied material and the intensity of the radiation emitted by a black body at the same temperature. A precise thermal map of a DUT

composed of different materials with high contrast in emissivity can be difficult to obtain as weak thermal sources can be concealed by surrounding hot spots emissions. This is particularly true for mordern ICs because of the high emissivity contrast between metals and silicon. To overcome this phenomenon, we use lock-in thermography techniques as proposed in [11] and detailed in Sect. 2.3.

2.2 DC Measurements

The simplest method in order to detect circuit activity is to directly acquire all thermal emissions from the chip using an IR camera. Work in [12] presents a methodology for post silicon power characterization. Based on temperature measurements obtained using a -196°C cooled SC5600 FLIR IR camera with a resolution of 640×512 pixels, the authors managed to retrieve a power density cartography of a die for different workloads. For that, they show that the heat diffusion equation can be approximated by the following linear matrix formulation

$$Rp + e = t \tag{2}$$

where R is the matrix of the thermal resistivities between different locations, p is the desired power map, e is the error in temperature measurement and t is the temperature matrix. Previous methodology used least squares estimation to find the p value that gives temperatures as close as possible to measured temperatures t. According to the authors, this technique poses several problems because of the inherent thermal spatial low-pass filter effect of silicon dies that leads to critical loss of information, especially in high frequencies. Hence, many power patterns can lead to the same thermal image, thus rendering the problem of temperature to power conversion ill-posed. To replace this method S. Reda et al. proposed instead minimizing the total squared error between temperature computed using (2) and measured ones combined with techniques from regularization theory [14].

In [13] and [15] authors managed to obtain a high resolution thermal map of a dual-core AMD Athlon II 240 running at $2.1\,GHz$, using the same IR acquisition platform. The circuit has a power consumption of $65\,W$ and measured temperature gradients were up to 16 °C. In these papers, the authors demonstrated how different workloads can lead to variations in hot spot location. Several configurations, assigning the workload only to one core or both of them, were used by S. Reda et al., highlighting the possibility of active area tracking by IR thermography because hot spots were found on top of active areas while sectors of lower activity such as memory remained cooler. They then applied their method of temperature to power inversion in order to recover the power density map of the chip.

We find that this method of IR image acquisition is flawed, as a DC offset is generated by the static power consumption of the chip (including constant power consumption of the IC such as the clock tree), and the diffusion of the heat generated by this phenomenon can lead to weak spot concealing. Moreover, this method requires steady environmental conditions as both the detector and the

thermal emissions are sensible to room temperature variations. This is especially true because our area of investigation mainly includes ICs such as FPGAs or microcontrollers which have a significantly lower power consumption than ICs considered in [13] and [15]. These devices consume few hundreds of milliampere, whereas microprocessor can draw up to several dozens of ampere. So it is obvious that this methodology (DC measurements) is highly unsuitable for weak thermal spot detection because of the high contrast of material emissivity, static thermal emissions, and heat diffusion.

However, if DC measurements are not suitable for hot spot mapping, they are of a great utility to learn about the thermal behavior of the DUT and thus for guiding dynamic measurements, i.e. to apply the lock-in thermography approach described in the next section. Indeed, a few DC measurements of the DUT step response enables us to quickly set up a first order model of its thermal behavior; such a model is of great help in deciding which (f_{lockin}, gain) couple should be used to obtain lock-in maps of high quality.

By way of illustration Fig. 1 gives, for several current steps, the responses of the DUT considered in the rest of the paper as well as the responses deduced from the identified first order model. In the present case, the IC thermal behavior of the DUT is characterized by a cutting frequency of $5\,mHz$. This is extremely low and implies the use of an amplification chain of at least $60\,dB$ to obtain lock in thermography maps at $10\,Hz$. Implementation and protocol used to produce Fig 1 is described in Sect 4.

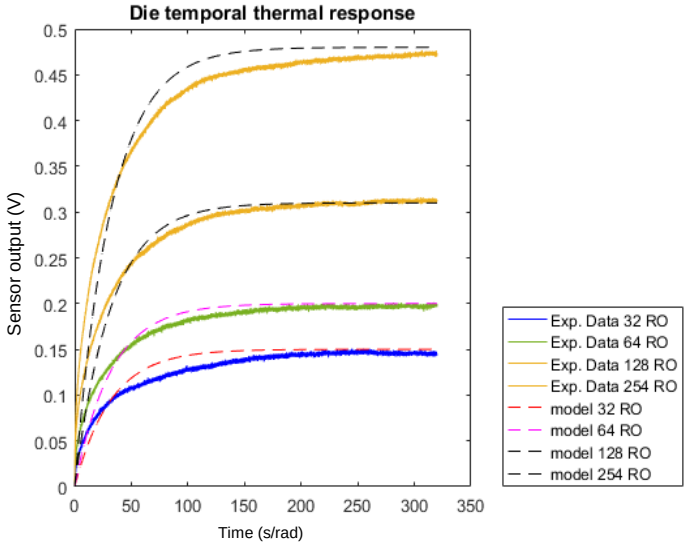

Fig. 1. Experimental and calculated thermal responses of the DUT to several step inputs.

2.3 Lock-in Thermography

Lock-in thermography is a correlation technique that allows retrieving periodic signals deeply drowned in noise. The principle, which is illustrated in Fig. 2, is very close to lock-in detection. It mainly consists of imposing a periodic thermal modulation to the DUT. With only the knowledge of the modulation frequency, it is then possible to retrieve the amplitude A and phase ϕ of the thermal signal and thus to fully rebuild the thermal behavior of the DUT [11].

Considering two processing channels, lock-in correlation consists of integrating the multiplication of sensor output by the correlation signal on the first channel and by the 90 ° phase shifted correlation signal on the second channel. Results are respectively named S_0 and S_{90} and are given by eq. (3) - (5) where K_j is the correlation signal, $F_{i,j}$ is the incoming signal, n is the number of samples and N is the number of lock-in periods the measurement is averaged over.

$$S = \frac{1}{n \cdot N} \sum_{i=1}^{n} \sum_{j=1}^{N} K_j F_{i,j} \tag{3}$$

$$S_0 = A \cdot cos(\phi) \tag{4}$$

$$S_{90} = A \cdot sin(\phi) \tag{5}$$

From (4) and (5) we easily infer eq. (6) and (7) that provide amplitude and phase of the thermal wave.

$$A = \sqrt{S_0^2 + S_{90}^2} \tag{6}$$

$$\phi = Arctan(\frac{S_0}{S_{90}}) \tag{7}$$

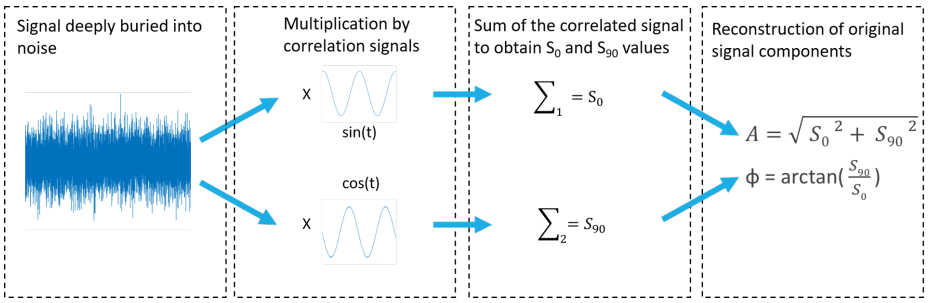

Fig. 2. Discrete lock in process

This methodology was first implemented by G. Busse et al. in [16] and used to implement the first lock-in camera system in 1992. This was then deepened by O. Breintenstein et al. in [17] and [11] to investigate small resistive defects in solar panels. In this work, authors show that lock-in thermography can be used to highlight small hot spots created by resistive paths in silicon dies, dynamic operations of a particular circuit element turned on and off at $54\,Hz$, and gate oxide integrity defects in Cu-grown silicon MOS structures.

Following on [12], [13] and [15], the authors demonstrated in [18] that increasing the lock-in frequency reduces considerably the heat diffusion distance comparing results from DC to $8\,Hz$. According to O. Breitenstein in [11] the lock-in frequency must be chosen respectively to a trade off between the Signal to Noise Ratio (SNR) and the spatial resolution. Indeed, if raising the lock-in frequency certainly reduces the heat diffusion distance, it also impacts the thermal load's duration on the die, thus the amount of signal that can be measured. In [19], it was demonstrated both theoretically and experimentally, that for their specific IR acquisition platform, SNR rises at first with the lock-in frequency but starts decreasing after a corner frequency around $3\,Hz$. The corner frequency obviously relies on detectivity, bandwidth and the amplifiers of the measurement chain.

So, using lock-in correlation to create a thermal map of the DUT brings forward several advantages. First, as mentioned earlier, it allows detection of signals deeply buried in noise which is critical in low power IC characterization. Secondly, the use of lock-in thermography discards any thermal emission that is not modulated at f_{lockin}. This means that, not only the final thermal map is free from any DC offset, but the user is able to target a specific area of the chip by adapting the modulation (induced through power supply modulation, software modulation, data or address modulation, etc) used to create the thermal wave. In addition, heat diffusion is controllable by modifying the lock-in frequency. Finally, retrieving the signal phase is a tremendous advantage that allows us to completely discard the emissivity contrast, which is a critical problem while facing complex ICs as aforementioned in 2.1.

3 Experimental Set-Up

From the state of the Art, we get that IR cameras have been widely preferred to single pixel sensors as they provide faster image acquisition and easier calibration. On the other hand, the latter advantages are at the expanse of cost (around $70\,k$ USD), bulk and bandwidth as their frame-rate rarely exceed $100\,Hz$ using full resolution [18], [11].

In this paper, we propose a low cost compact measurement set-up, based on a mono pixel IR sensor providing a large acquisition bandwidth and a higher detectivity at equivalent temperatures. Our testbench is composed of a InAs IR sensor working in the $1 - 3.8\,\mu$m spectrum at $-60\,°$C, a trans-impedance amplifier providing a 2.10^8 V.A gain and a remote controlled oscilloscope, for a

total cost of $3.5\,k$ USD, not counting the oscilloscope which is basic measurement equipment. This set-up is able to detect signals from DC up to $20\,kHz$.

In order to draw a thermal map, we use the lock-in correlation algorithm to compute amplitude and phase values at every position on the die. One drawback of our system is the acquisition time of a full map, which is around 12 hours for a 160×160 pixels thermal map (acquiring 10 measurements at $f_{lockin} = 10\,Hz$ at each position). However, the mapping time is customizable by modifying the number of acquired traces, the cartography spatial step, the trace length, and f_{lockin}.

4 IC Thermal Characterization

As explained in Sect. 2.3, the f_{lockin} value has an influence on several parameters. The higher the frequency the shorter is the heat diffusion distance [18]. Therefore, to increase the spatial resolution of thermal maps, it is necessary to increase f_{lockin}. On the other hand, increasing the heat modulation frequency leads to shorter periods of heating and thus to weaker IR amplitude and SNR. In this section we demonstrate that this trade off can be managed rationally through a first order modeling of the thermal behavior of ICs.

4.1 Experimental Protocol

For many designs or research objectives, FPGAs are suitable integration targets as they are nearly 100 % customizable. This is the case for our work. We have thus implemented our several designs on a Xilinx Virtex 5 FPGA after having removed the metallic package to get a direct access to the backside. This FPGA has a die area equal to $16 \times 16\ mm^2$ and is designed with a $65\,nm$ CMOS technology.

The aim of our first experiment was to estimate the detectivity of our platform and to observe the thermal behavior of the FPGA. We integrated 255 Ring Oscillators (RO) to use them as micro-heaters [18]. Each RO was composed of two inverters and one Nand2 gate allowing us to enable/disable it. All ROs were placed as homogeneously as possible in a constrained area. The main idea was to integrate a local and controllable source of heat by driving the number of active ROs. Indeed, ROs are constant micro-heater thanks to their constant power consumption. By modifying the number of active ROs we were able to linearly control the local power density.

The lock-in toggling frequency of ROs was fixed at $10\,Hz$. This toggling imposed with an external signal generator creates a current variation and thus a heat wave. The amplitude of this current variation was measured after removal of the on-board voltage regulator. The toggling of a single RO generated a current variation equal to approximately $3.23\,mA$, while the core was biased by a $316\,mA$ current under a voltage of $1\,V$.

4.2 Electrical Activity Detection by Heat Detection

Thermal maps with $n = 1, 8, 16, 32$ and 255 active ROs were collected and drawn. In order to diminish experimental measurement time, only the top right quarter of the die was mapped. The results for $n = 1, 8, 16$ and 32 ROs are presented Fig. 3.

On the amplitude map, heat generation is clearly and visually distinguishable when at least 16 ROs are activated. Even if a few heat sources can be spotted on the 8 ROs map, they can not be directly separated from heat diffusion of the surrounding hot spots or measurement noise without application of statistical or signal processing techniques.

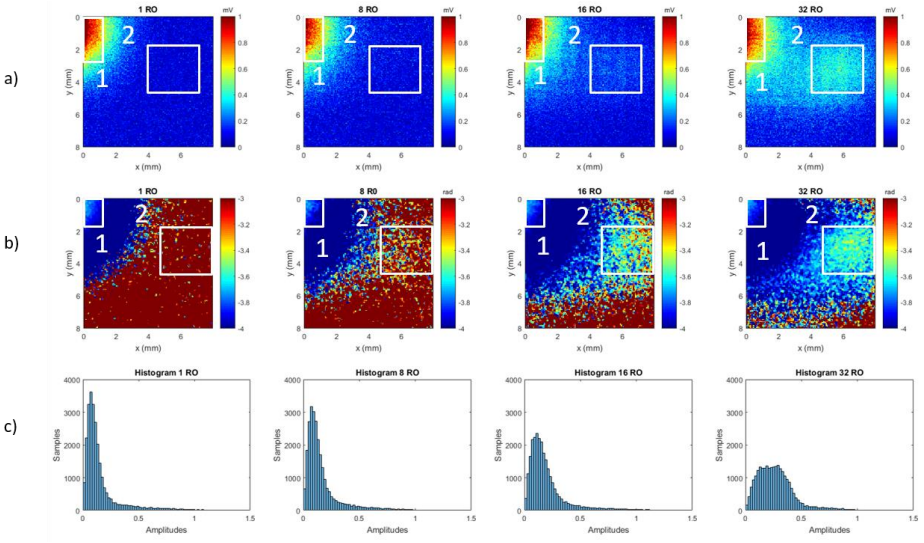

Fig. 3. a) Amplitude map, b) Phase map, c) Amplitude histograms; 1: Heat source generated by the control logic. 2: Localization of implemented ROs used as thermal heaters.

On the other hand, looking at phase maps in Fig. 3 b), one can observe that the presence of the 8 ROs is more visible than on the amplitude maps reported in a). In b), dark blue areas correspond to heat diffusion zones whereas dark red zones represent areas free of heat diffusion and hot spots. This confirms that the study of phase signal is a key element in separating diffusion heat from heat sources. The higher information of the reported phase maps is a direct illustration of former comments related to emission contrast. It proves that the phase image can provide much more accurate information in several situations.

However, the main point here is that the distributions in Fig. 3 c) of the lock-in amplitudes over the IC surface are also highly interpretable. Indeed, the effect of 8, 16 and 32 ROs on the distribution shapes is clearly visible. From

these observations, we believe it is possible to extrapolate whether a circuit is infected or not by a stealthy hardware Trojans (HT) using statistical means. This point will be further discussed in Sect. 5.

Using the same IR measurement platform, we were able to acquire thermal maps for $n = 255$ ROs with a f_{lockin} up to $210\,Hz$. In comparison, works reported in [18] and [19] used a maximal f_{lockin} of $8\,Hz$. This considerably increases our detection capability (as shown Fig. 4) as we are able to detect weaker hot spots diffusing on a very limited area at higher f_{lockin} frequencies. This is presented in the top left map of Fig. 4 that shows a regular pattern in the target FPGA heat diffusion, which is less visible in the 10 Hz bottom left thermal map. Without any access to the layout of this FPGA we are not able to explain the origin of this pattern at this time.

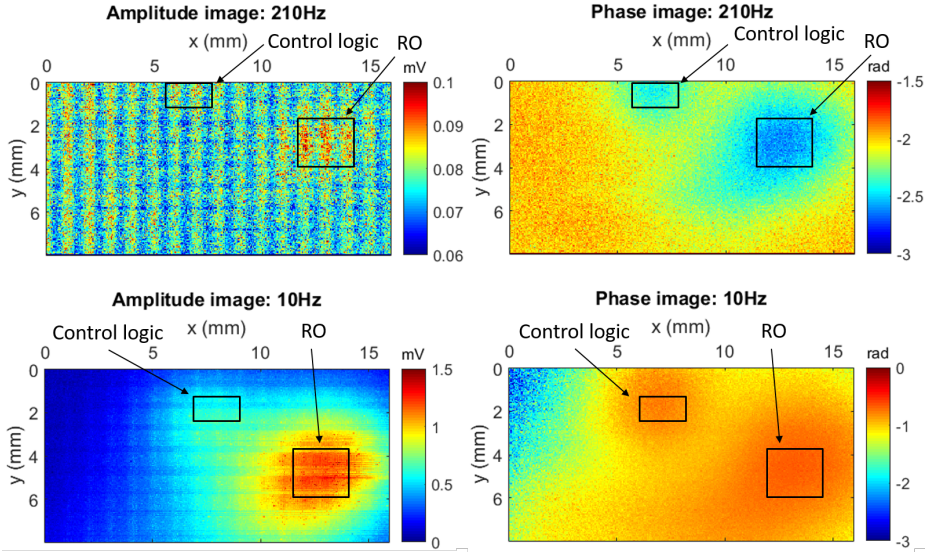

Fig. 4. Thermal maps at 10 Hz and 210 Hz

5 Trojan Detection

This section describes the methodology we proposed to detect rough and stealthy hardware Trojans. It also gives and discusses experimental results obtained on the considered DUT.

5.1 Hardware Trojan Characteristics

A Trojan is a tiny integrated circuit maliciously added to an existing design without knowledge of the company owing its intellectual propriety. This circuit's

purpose can be variable, including denial of service, programmed obsolescence etc. They usually aim to avoid being detected, both from a power overhead and surface point of view. As a consequence, detecting them is a challenging task and requires a comprehensive knowledge on the subject.

A Trojan is made up of two components: the trigger and the payload [20]. The trigger is the part of the circuit waiting for the occurrence of an event to activate the malicious function of the Trojan, i.e. its payload. This trigger could be 'always-on' or active when a part or a functional block of the IC is under use. This is, a priori, the only part of the HT we can detect since the payload remains quiet before its triggering. The trigger could be sequential or combinatorial. In the rest of the paper we consider a sequential trigger, i.e. a trigger waiting for a sequence of states.

The payload is the hostile part of the HT which is activated by the trigger when the firing condition is met. A HT can be spread inside the circuit as well as restrained to a particular area. It could be hidden in the functional block where it is waiting the triggering condition (e.g. a particular sequence of values of different registers in the block). In that case we say it is 'stealthy'. It can also be placed far from the functional block(s) from which it is waiting for the triggering condition. In that case we say it is 'externalized'. Most former papers on HT detection using SCA focused on externalized HTs. In the following paragraphs, we consider both externalized HTs and stealthy HTs, placed in a restraint area of the device, and with a sequential trigger.

5.2 Testchips and Emulation of HTs

To demonstrate the efficiency of thermal maps in detecting HTs we chose to emulate the infection by an HT of a simple design mapped into a Xilinx virtex 5 FPGA. This simple design is made of a hardware 128-bit AES block and its associated control logic. This AES is clock gated. This means that its electrical activity, and thus its heating effect, can be fully stopped by disabling its clock signal.

Two different implementations of this simple design were done. The resulting floorplans are shown in Fig. 5. For both implementations, the HT is a 16-bit Linear Feedback Serial Register (LFSR) clocked with the same clock signal as the AES as explained in [6]. It is therefore only active when the AES is operating. It occupies 4 slices of the FPGA among the 17280 available slices. This represents less than 0.023% of the total resources (surface). For a convincing demonstration that the proposed lock-in thermography platform is able to detect stealthy HTs, we implemented the HT with an enable signal to be sure that the golden design and the infected designs are exactly the same from a hardware routing point of view and thermally differs only when the HT is enabled.

The two implementations of the design differ by the placement of the HT with respect to the AES block. In the first implementation, the HT is placed far from the AES block (Fig 5 left). This case corresponds to an HT externalized in a block which is inactive when the AES is operating (or externalized in an empty place of the circuit, an improbable situation in real ICs). In the second implementation

(Fig 5 right), the place and route constraints were set so that the HT is merged in the middle of the AES. This situation corresponds to an adversary trying to hide the HT activity within the activity of another functional block, i.e. trying to render its HT to be as stealthy as possible. This case corresponds to a more challenging situation regarding HT detection.

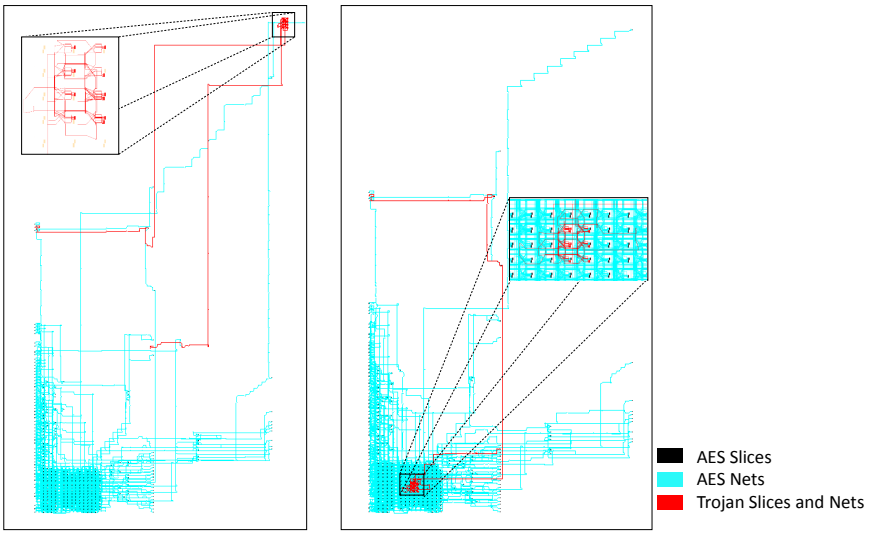

Fig. 5. Left: testchip with a externalized HT. Right: testchip with a stealthy HT.

5.3 Detection Methodology

The principle of the detection methodology consists of comparing golden thermal maps drawn from measurements done on a IC from a trusted production lot (i.e. a golden chip) with the corresponding thermal maps obtained above the DUT, i.e. above an IC coming from a potentially infected lot.

These thermal maps correspond to several computational activities of the design. The latter should be chosen so that it activates all the functional blocks of the design or has high coverage of its surface.

According to the size of the infection which is expected or researched, i.e. according the stealthiness of the HT, the comparison between the golden IC and the DUT can be done with different levels of accuracy. It can be done visually to detect externalized un-stealthy HTs or by simple difference of means between corresponding positions of the maps for quite stealthy HTs. Alternatively, it can be done using statistical tests such as the Welch's t-test for stealthy HTs. The next sections detail the application of this methodology to our two testcases and give the obtained experimental results.

5.4 Experimental Results

Case 1: externalized HT This first case considers a HT externalized in an inactive area when the HT is active. This is thus the case of a rough HT. To detect it, a basic use of our IR measurement platform could be sufficient. Such a use, which can also be applied to detect rough counterfeit products (but not clones), consists of (a) visually comparing a golden thermal map with the ones obtained for a potentially infected IC or (b) computing a basic difference of means between the two thermal maps.

Thermal maps were therefore acquired on the same FPGA with HT activated or not. The acquisition of one map consists of collecting $n = 10$ lock-in traces (vectors length of $N = 1000$). The lock-in frequency was fixed at $10\,Hz$. This means that the AES (or the AES and the HT in case of an infection) is disabled during $0.05\,s$ (cold phase of the lock-in process) and active during $0.05\,s$ (hot phase of the lock-in process).

Fig. 6 gives the thermal maps of the whole die. They are made of 160×160 pixels. The first left column of this figure gives the amplitude of the thermal wave at each coordinate with a color scale corresponding to the dynamic of all measurements. The second column gives the same results but with a color scale allowing us to detect the HT which is in the rectangle labeled (3) in the maps. The rectangles labeled (2) point out areas where large buffers are used to drive IO pads allowing to get out output values of the AES. The third column gives the phase of the thermal wave with respect to the lock-in signal. One can observe the significant impact of the HT on the phase map, impact which is much more visible than on the amplitude map. These results demonstrate that our IR platform is efficient in detecting rough or externalized HTs (and probably rough counterfeits) by simple visual inspection of thermal maps and this especially by considering the phase map.

Case 2: stealthy HT Figure 7 gives the thermal maps (amplitude only) of the quarter of the IC surface containing the AES and the stealthy HT which is hidden in the AES. For this experiment, the output signals of the AES were gated using a Nand gate to suppress the electrical activity of the IO pad's buffers. As shown, there are no visual difference between these two maps. The heat generated by the HT is masked within the AES's heat. Thus an enhanced comparison technique must be used to detect the thermal impact of the HT.

The main idea to compare these thermal maps is to apply a Welch's t-test between corresponding positions of the maps in order to detect small heat differences due to the HT. However this cannot be done in a straight forward manner. Indeed, thermal maps were not done the same day and in a controlled environment. Thus temperature changes significantly during their acquisitions and one must take these changes into account prior to applying the Welch's t-test application.

Because we observed that this global shift of the room temperature during the cartography process acts as a multiplicative coefficient on the lock-in amplitudes, the applied procedure to conceal the effects of temperature changes

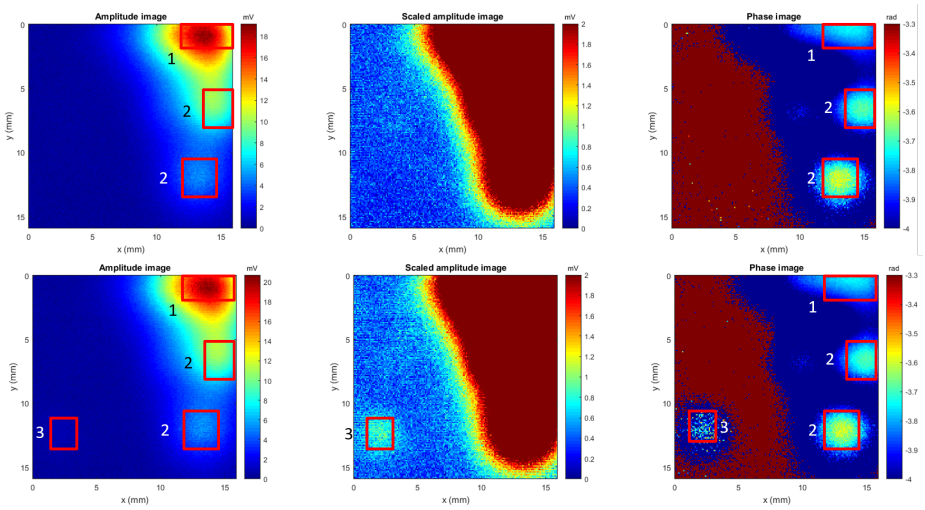

Fig. 6. First row : thermal maps (amplitude and phase) obtained with the golden circuit. Second row: thermal maps obtained with the infected circuit. Label (1) shows the position of the AES, label (2) shows the position of large output buffers and label (3) shows the position of the externalized HT.

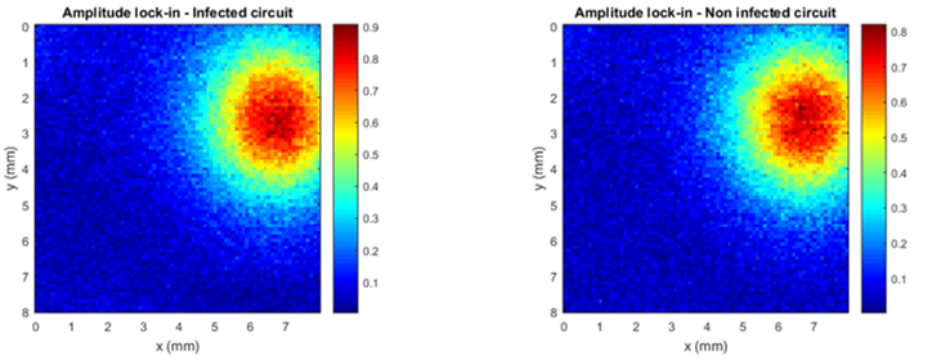

Fig. 7. Left : amplitudes of the thermal waves collected above the golden IC. Right: amplitudes of the thermal waves collected above the infected IC.

is quite simple. It consists of considering the amplitudes of thermal maps as statistical distributions and standardizing them in order to get the best match between the two distributions; one distribution being associated with the golden IC and the other with the DUT. Concealing the room's temperature shift allows us to minimize, as best as possible, the differences between the thermal maps (distributions) prior to applying the Welch's t-test.

The standardization of all the lock-in values, $A_l(x, y)$, obtained at coordinate (x, y) is done using the following formula:

$$A_l^S(x, y) = \frac{A_l(x, y) - < \bar{A}_l >}{\sigma(< \bar{A}_l >)} \tag{8}$$

where $< \bar{A}_l >$ is the empirical mean of mean amplitudes obtained over the whole map; $\sigma(< \bar{A}_l >)$ is the standard deviation of the mean amplitudes obtained over the whole map and $A_l^S(x, y)$ is the standardized lock-in value of $A_l(x, y)$ at coordinate (x, y). By way of illustration Fig. 8 gives the cumulative density functions (cdf) associated with two set of measures above the same circuit before and after concealing effects of the temperature variation.

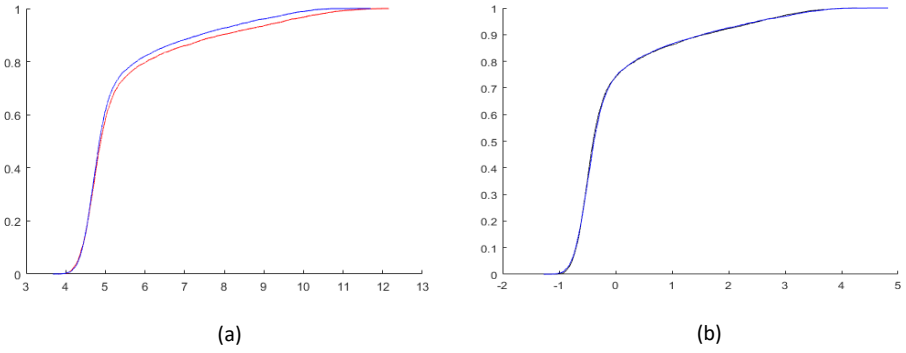

Fig. 8. (a) cdfs of two thermal maps of the same IC before temperature effect concealing (b) after concealing of temperature effect

After correction, by standardization of the effect of room temperature change, the Welch's t-test can be applied to detect the remaining changes due to the presence of an HT. Applying this test means herein computing, for each (x, y) coordinate of the maps, the statistic of the the Welch's t-test, $T_{(x,y)}$ between two samples of A_l^G and A_l^D.

Then the obtained $T_{(x,y)}$ value is compared to a critical value T_{crit} defined according to the chosen confidence level fixed by α that sets the critical p-value for the test. Typically, α is set to 0.05 or 0.01. This means that we accept 5% (or 1%) of chance that the detected difference is a false positive. If $\mid T_{(x,y)} \mid > \mid T_{crit} \mid$,

the samples do not have the same mean and one can conclude that at this coordinate there is an extra source of heat, i.e. in our application case an HT.

Figure 9 (a) gives the $T_{(x,y)}$ map obtained by comparing two thermal maps performed with the same golden IC. From this map, it clearly appears that the means of all corresponding samples of the two maps are the same. This result indicates that the IC are the same.

Figure 9 (b) gives the $T_{(x,y)}$ map obtained by comparing with the Welch's t-test the thermal maps associated to the stealthy HT of Fig. 5 with the one associated to the golden IC. From this map it is clear that there is an HT close to $(x, y) = (70, 30)$ which is close to the effective HT position. This demonstrates the correctness of the proposed HT detection technique and the interest of lock-in thermography for detecting HT and for locating small electrical activities in IC in general.

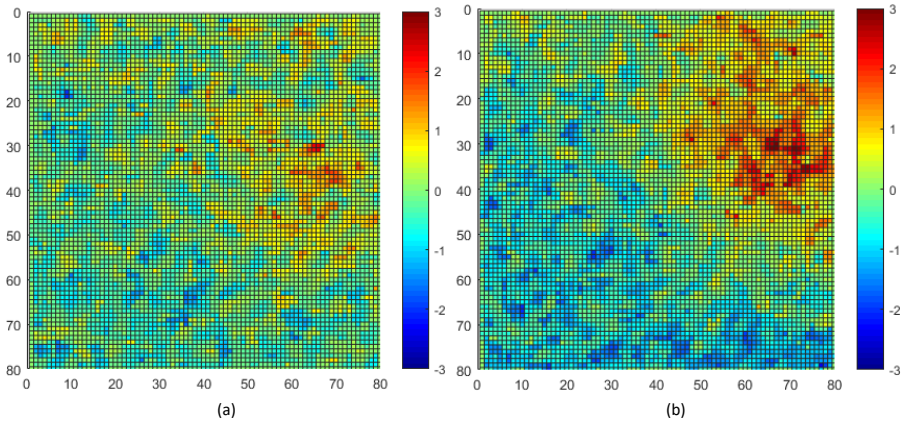

Fig. 9. (a) Welch t-test between two golden chips, (a) Welch t-test between a golden IC and a IC infected by a stealthy HT.

6 Conclusion

In this paper we have introduced a cost effective IR measurement platform characterized by a large bandwidth and a high detectivity. It has been designed to be able to locate small electrical activities within ICs using lock-in correlation. An application has been shown to the detect a stealthy hardware trojan hiden in a functional block. Results obtained are very encouraging and demonstrate the usefulness of lock-in thermography in the field of secure device characterization.

References

1. A. Loai, H. Houssain, et T. F. Al-Somani, "Review of Side Channel Attacks and Countermeasures on ECC, RSA, and AES Cryptosystems".
2. A. N. Nowroz, K. Hu, F. Koushanfar, and S. Reda, "Novel Techniques for High-Sensitivity Hardware Trojan Detection Using Thermal and Power Maps". In: "IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems", vol. 33, n 12, pp. 1792-1805, December 2014.
3. X. T. Ngo, Z. Najm, S. Bhasin, S.Guilley and J. L. Danger. "Method taking into account process dispersion to detect hardware Trojan Horse by side-channel analysis". J. Cryptographic Engineering, vol. 6, pp. 239-247, 2016
4. J. Balasch, B. Gierlichs and I. Verbauwhede, "Electromagnetic circuit fingerprints for Hardware Trojan detection, "IEEE International Symposium on Electromagnetic Compatibility (EMC), Dresden, 2015, pp. 246-251.
5. X. Zhang and M. Tehranipoor, "RON: An on-chip ring oscillator network for hardware Trojan detection," Design, Automation and Test in Europe, Grenoble, 2011, pp. 1-6.
6. M. Lecomte, J. Fournier and P. Maurine, "An On-Chip Technique to Detect Hardware Trojans and Assist Counterfeit Identification," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 25, no. 12, pp. 3317-3330, Dec. 2017.
7. M.C. Tan, M. Y. Tay, W. Qiu, and S. L. Phoa, "Fault localization using infra-red lock-in thermography for SOI-based advanced microprocessors". In: "Physical and Failure Analysis of Integrated Circuits" (IPFA), pp. 15), 2011.
8. K. Hu, A. N. Nowroz, S. Reda, and F. Koushanfar, "High-sensitivity hardware Trojan detection using multimodal characterization". In: Design, Automation and Test in Europe Conference and Exhibition (DATE), pp. 12711276, 2013.
9. G. Tessier, M. Bardoux, C. Bou, C. Filloy, and D. Fournier, "Back side thermal imaging of integrated circuits at high spatial resolution". In: Applied Physics Letters, vol. 90, no. 17, pp. 171-172, April 2007.
10. Frank P. Incropera, David P. Dewitt, "Fundamentals of Heat and Mass Transfer", 5^{th} edition, pp. 700-746, 2001.
11. O. Breitenstein, W. Warta, and M. Langenkamp, "Lock-in Thermography", vol. 10. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.
12. R. Cochran, A. N. Nowroz, and S. Reda, "Post-silicon power characterization using thermal infrared emissions". In: "Proceedings of the 16th ACM/IEEE international symposium on Low power electronics and design", pp. 331-336, 2010.
13. S. Reda, R. Cochran, and A. N. Nowroz, "Improved Thermal Tracking for Processors Using Hard and Soft Sensor Allocation Techniques". In: "IEEE Transactions on Computers", vol. 60, n 6, pp. 841-851, June 2011.
14. M. Bertero and P. Boccacci. "Introduction to Inverse Problems". In: Imaging. Institute of Physics Publishing, 1998.
15. S. Reda, Thermal and Power Characterization of Real Computing Devices , IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 1, n 2, pp. 76-87, June 2011.
16. G. Busse, D. Wu, and W. Karpen, J. Appl. Phys. 71, 3962, 1992.
17. S. Huth, O. Breitenstein, A. Huber, D. Dantz, U. Lambert, et F. Altmann, Lock-in IR-thermography-A novel tool for material and device characterization , in Diffusion And Defect Data Part B Solid State Phenomena, pp. 741-746, 2002.
18. A. Nowroz, G. Woods, and S. Reda, "Improved post-silicon power modeling using AC lock-in techniques". In: "Design Automation Conference" (DAC), 48th ACM/EDAC/IEEE, pp. 101-107, 2011.

19. A. N. Nowroz, G. Woods, and S. Reda, "Power Mapping of Integrated Circuits Using AC-Based Thermography". In: "IEEE Transactions on Very Large Scale Integration (VLSI) Systems", vol. 21, n 8, pp. 1398-1409, August 2013.
20. R. S. Chakraborty, S. Narasimhan, and S. Bhunia. "Hardware Trojan: Threats and emerging solutions". In: High Level Design Validation and Test Workshop, 2009. HLDVT 2009. IEEE International, pp. 166-171, 2009.