GDS-II Trojan Detection using Multiple Supply Pad V_{DD} and GND I_{DDQ}s in ASIC Functional Units

I. Wilcox, F. Saqib* and J. Plusquellic
University of New Mexico, *Florida Institute of Technology

Abstract

We propose a parametric, side-channel-based method designed to detect malicious changes that have been made to the chip layout, i.e., the GDS-II representation, by an adversary. We measure steady-state leakage currents (I_{DDQ}) from multiple, topologically distributed power ports on the chip and propose a chip-averaging method for eliminating within-die variations and improving the Hardware Trojan (HT) signal-to-process-noise detection sensitivity of our statistical-based detection methods. The technique is validated for the first time by measuring I_{DDQ} from an ASIC with Advanced Encryption Standard (AES) and Floating Point Unit (FPU) macros, 16 V_{DD} and GND *ports and a set of special HT emulation circuits. I_{DDQ} data is measured from multiple copies of the IBM, 90 nm ASIC.*

1 Introduction

Globalization of the integrated circuit (IC) design, fabrication and test industries, as well as increased use of 3rd party IP, increases the ease at which adversaries can insert malicious circuits that are designed to leak confidential information or cause system failure [1][2]. Four general approaches are available as a means of detecting malicious circuits or Hardware Trojans (HTs), namely, parametric (side-channel-based) methods, logic testing methods, destructive IC inspection and watch-dog monitors.

The primary advantage of side-channel detection methods is their ability to detect 'partial' activations of the HT circuit, which are much more likely to occur than the 'full' activations attempted by logic-based detection strategies. Partial activations refer to switching activity that occurs within the logic gates defining the HT, e.g., gates that monitor circuit state and implement the trigger, but which have no effect on the current state of the chip. Moreover, parametric approaches are also able to detect HTs which cause no changes to the functional operation of the chip, but rather are designed to leak private information, e.g., through electromagnetic transmission mechanisms.

The biggest challenge of parametric approaches is developing methods with adequate *signal-to-noise*, where the signal is defined as the anomaly introduced by the HT and noise refers to uncompensated chip-to-chip and within-die process variations and measurement noise.

In this paper, we propose a parametric approach that is based on the analysis of a chip's I_{DDQ} (steady-state or quiescent current), which builds on work done by the authors of [3][4]. Our proposed technique measures I_{DDQ}s from multiple-supply ports (MSP) across the 2-D surface of the chip as a means of improving *signal-to-noise*, and is an important distinguishing characteristic of our proposed approach over others. MSP provides regional observability and directly addresses the adverse impact of increasing levels of process variations and leakage currents[1]. MSP scales to larger chips with smaller feature sizes that incorporate additional power ports to accommodate the increase in power consumption. The *n* supply ports available with MSP can improve sensitivity significantly over traditional single supply port (global) measurement methods, up to a factor proportional to *n*. Calibration methods such as those proposed in [3], further improve signal-to-noise. However, calibration does not account for within-die variations, which remains the biggest challenge for parametric HT techniques.

This paper proposes a 'chip-averaging' technique that calibrates for within-die variations and proposes an ellipse-based statistical technique to detect HT I_{DDQ} anomalies. The contributions of this paper are summarized as follows:

- Random, within-die variations in leakage current are nearly eliminated using a calibration method that averages I_{DDQ} data measured across multiple chips.
- Patterns in I_{DDQ} that occur in scatterplots constructed from currents measured from adjacent pairings of power ports are used in combination with outlier analysis to help distinguish between random defects and the leakage current anomalies introduced by HTs.
- A statistical ellipse-based detection method based on principal component analysis is proposed for detecting outlier data points produced by HTs.
- Experimental evaluation of the methods is carried out using data from a set of ASIC chips that incorporate large circuit macros and a set of 16 V_{DD} and 16 GND ports.

2 Background

The authors of [5] were the first to address the HT issue. They use transient power supply currents to identify HTs in chips. The authors of [6] propose a method that first determines a set of target 'hard-to-observe' sites for a HT with *q* inputs and then uses automatic test pattern generation (ATPG) to generate patterns to activate the HT. A HT detection method that measures the combinational delay of a large number of register-to-register paths internal to the functional portion of the IC is proposed in [7]. In [8], the authors propose a region-based stimulation strategy and analyze the global power consumption to detect HTs. In [9], the authors introduce special circuitry that enables the direct control of the least controllable nodes in the circuit

[1] A region is defined as a portion of the layout that receives the majority of its power from a set of surrounding power ports or C4 bumps.

978-1-4673-7421-7/15/$31.00 ©2015 IEEE

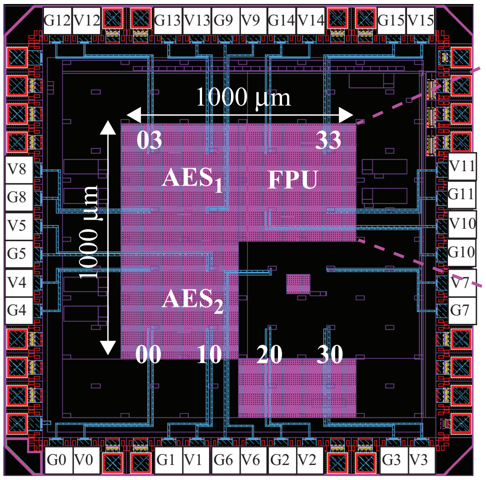

G12 V12 G13 V13 G9 V9 G14 V14 G15 V15
V8
G8
V5
G5
V4
G4
V11
G11
V10
G10
V7
G7
1000 μm
1000 μm
03
33
AES1
FPU
AES2
00
10
20
30
G0 V0 G1 V1 G6 V6 G2 V2 G3 V3

Fig. 1. Chip layout showing macros and 16 V_{DD} and GND ports.

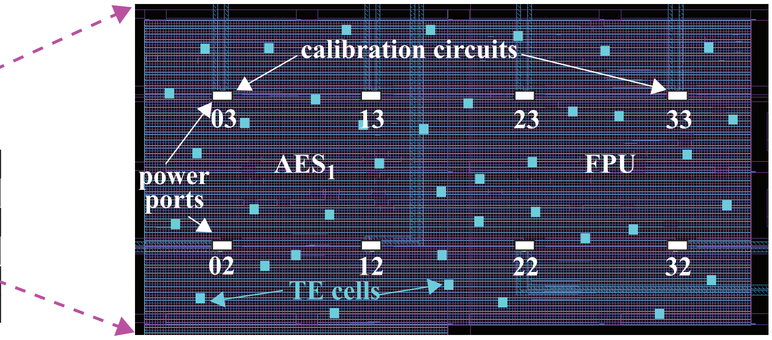
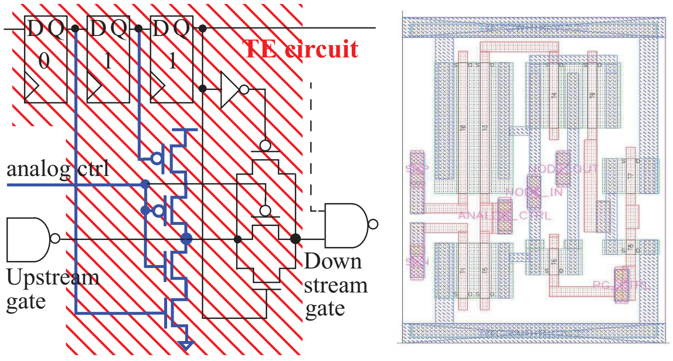


Fig. 2. AES_1 (left) and FPU (right) with Trojan Emulation (TE) cells and calibration circuits (not drawn to scale).



(a)

(b)

Fig. 3. Trojan Emulation Circuit (a) schematic and (b) layout.

as a means of triggering the activation of a HT. In [10], the authors build a path delay fingerprint of Trojan-free chips by running high coverage input patterns. In [11], the authors propose a methodology for reducing noise from circuit switching to improve detection of HTs designed to draw minimal power. A method that leverages the foundry process parameters to model and calibrate for delay variations introduced by process noise is proposed in [12]. [13] investigates the use of a side-channel signature for regions of the chip as a method to model systematic process variations to detect HTs if the measured results are outside of the estimation results computed from neighboring regions. A scalable circuit partition approach is proposed in [14] using gate-level delay measurements at all circuit locations to find HTs. [15] proposes to use a specific process of fault-injection to force a clock glitch that will decrease the clock period until the setup and hold time is violated while monitoring the output of an AES IP core. The authors of [16] propose an HT detection that encloses the three largest principal components within an ellipsoid of minimum volume. In [17], the authors propose an outlier HT detection method that compares the power signal analysis of a test chip with the training set derived from a genuine IC. The authors of [18] propose finding HTs using design dependent detection sensors to measure path delays on-chip without the need for a golden model. In [19], the authors propose to detect HTs by comparing the expected correlation of F_{max} and I_{DDT} with that of a golden chip.

3 Test Chip Design

A layout view of the test chip design is shown in Fig. 1. It consists of three macros, AES_1, AES_2 and FPU, each occupying an area of 500 μm^2. Large, low resistance M9 wires route from a set of peripheral V_{DD} and GND pads, called **power ports** or **PPs**, labeled V_0/G_0 through V_{15}/G_{15}, to a 4x4 grid of tap points distributed at 250 μm intervals across the macros. The M9 tap points connect down to the V_{DD} and GND grids, which are routed across the lower 8 metal layers. The M9 wires emulate an area I/O array configuration (also called a C4 array), which allows our MSP technique to fully leverage the regional observability available in commercial C4 implementations. The (x,y) coordinates of several of the PPs are given in the figure as, e.g., **00**. We use the (x,y) descriptors of the PPs as, e.g., PP_{00}, in the remainder of this paper.

Fig. 2 shows a blow-up of the upper portion of the layout, illustrating the V_{DD} and GND tap points, a set of 8 calibration circuits positioned underneath the tap points, and a set of 38 Trojan Emulation (TE) circuits (discussed below). Data from the calibration circuits is used in a process designed to eliminate chip-to-chip process variations. The calibration circuit design and process used in this paper are similar to those described by the authors of [3].

3.1 Trojan Emulation Circuit

The purpose of the Trojan emulation (TE) circuit is to enable a systematic approach to evaluating the sensitivity of our methods. We inserted 57 TE circuits in the layout (19 in each macro). The details of the TE circuit are shown in Fig. 3. The three scan chain FFs control the state of the TE circuit, which can be configured to enable one of several types of signal anomalies, including controlled impedance shorts and opens (opens are not investigated in this paper). The TE circuit is inserted in series between an Upstream and Downstream gate. The FF state shown as “011” disables the TE circuit and represents the Trojan-free state. We use the term **NT** for ‘no-Trojan’ to reference Trojan-free data collected under the Trojan-free state. The 2 shorting states investigated in this paper are “001” and “111” which enable the upper-most PFET and lower-most NFET in the 4 transistor stack, respectively. These two

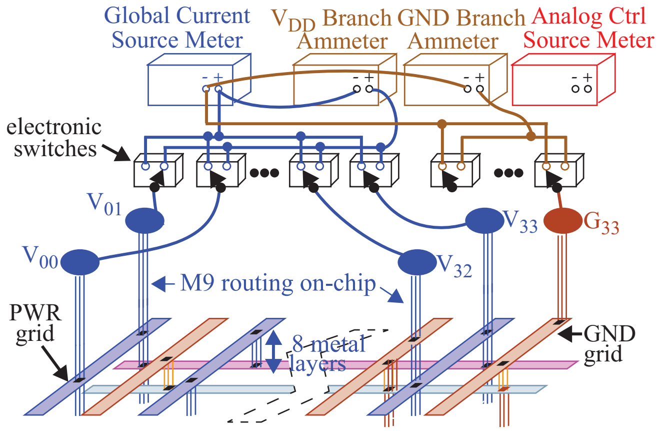


Fig. 4. Instrumentation Setup.

states in combination with the *analog ctrl* signal allow a controlled impedance short to be introduced between the V_{DD} or GND rails, resp., and the Upstream gate. Therefore, the TE circuit allows the controlled introduction of 'current anomalies' at various places on the V_{DD} and GND grid. We use the term **TR** to refer to the Trojan data.

The Cadence Encounter Tool is used to place and route the macros so the position of the Upstream gate is likely to be close to the corresponding TE circuit in the layout but will vary for each TE instance. Therefore, the (x,y) position in the layout where the current is sourced from the V_{DD} grid and where it is sinked into the GND will be different. Also note that only one of "001" or "111" will create a short, and this is determined by the output state of the Upstream gate. For example, if the output state is '0', then configuring the TE circuit with "001" will create a short in the uppermost PFET, with current proportional to the magnitude of the *analog ctrl* signal (with 1.2 V disabling the short and 0 V fully enabling it), and the state and geometry characteristics of the Upstream gate. The *analog ctrl* signal routes to all copies of the TE circuit and to an analog pin (not labeled) on the pad frame shown in Fig. 1. It can be controlled to any value between 0 and 1.2 V using an off-chip voltage source.

The external instrumentation setup for measuring the individual power port (PP) branch currents is shown in Fig. 4 for a subset of the PPs. We use a Keithley 2400 source meter as the Global Current Source Meter (GCSM) and two Agilent 34401A for the V_{DD} and GND Branch (current) Ammeters, each with resolutions of less than 1 µA. Any of the 16 branch currents can be measured through the Ammeters by configuring the electronic switches appropriately. We use an Agilent E3626A for the Analog Ctrl Source Meter to drive the *analog ctrl* signals at values of 0.0 V, 0.1V, ..., 1.2 V, in 100 mV steps, to emulate HT leakage sources of various magnitudes.

Fig. 5 plots the average TE circuit shorting currents as a function of the applied voltage on *analog ctrl*. The averages are computed using data collected from all chips. The upper portion of the x-axis is labeled with the **analog voltage** or **AV** that is applied to the NFET stack transistors, while the lower portion is labeled with the PFET voltages. The largest (smallest) analog voltage is restricted to 0.8 for NFET (0.4 for PFET) because actual Trojan leakage currents are likely to be small in practice, i.e., < 100 µA. The graph superimposes 57 curves, one for each of the TE circuits. The curves show that current monotonically increases to approx. 90 µA as voltage is increased for NFETs. At the smallest AVs, the average currents can be very small, i.e., in the range of 10 µAs or less.

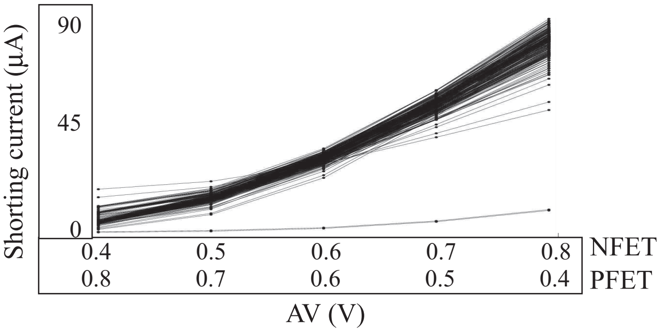


Fig. 5. Average TE circuit global current characteristics for different *analog ctrl* voltages.

We use the TE circuit to model the presence of extra gates, and/or a regional redistribution of gates in the layout, either or both of which would occur when an adversary tampers with the layout. The TE circuit injects current as a 'point source' and not over a region as would be true of extra gates or a redistribution of gates. However, from the perspective of the PP currents, the two different physical implementations are indistinguishable in cases where the modifications are constrained to a relatively small region, e.g., < 100 µm^2. Therefore, we believe the TE circuit is a good representation for HTs under these conditions, which in our opinion, covers the most likely scenarios.

4 Experiment Design

We collected data from 45 copies of the chip shown in Fig. 1. Similar to I_{DDQ} manufacturing test, we tested our chips using 4 different input vectors, each of which generates a unique leakage pattern in the PPs. Under each of these 4 leakage patterns, we tested 57 TEs at each of the 5 voltages shown in Fig. 5.

Our initial testing revealed that 3 of the 45 chips had some type of broad area leakage current defect, or a shorting defect, and therefore, we removed them from our analysis. We were able to identify these 'defective' chips quickly because each of them produced a unique pattern of leakage currents in the PPs that was distinct from the pattern produced by the majority of the chips (we will discuss this further in Section 5.1.1). The fact that we can easily and quickly identify and eliminate defective chips from the analysis is an important capability of our technique over traditional global I_{DDQ} methods. Being able to do so significantly improves the sensitivity of our method to very small leakage current anomalies introduced by Trojans. The analysis presented below therefore uses only 42 of the 45 chips.

4.1 Ellipse Statistical Method

We use an ellipse-based statistical technique for Trojan detection. The NT data is derived from the chips directly (with all TE circuits disabled), instead of using a simulation model as would be the case in practice. The golden model is derived from the chip data because deriving accu-

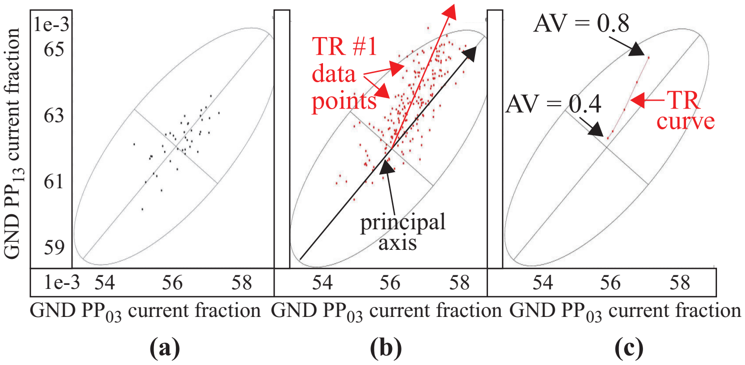


Fig. 6. (a) Ellipse Analysis on GND PP pairing 03-13, (b) TR #1 data points at all AVs and (c) average of TR #1 data points at each AV.

rate simulation models requires access to foundry data directly, which we do not have.

NT data is collected from each of the chips with all TE circuits disabled as a set of 16 V_{DD} (and GND) branch currents. Ellipse statistical limits are derived from the first 30 chips (of the 42) using a technique derived from principle component analysis [20]. The remaining 12 chips are used as **control** samples to help evaluate the false positive detection (**FPD**) rate of our method. The branch currents are 'calibrated' (see [3] for details) and then 'normalized' by dividing each of them by the global current (computed as the sum of the 16 branch currents). Therefore, the branch currents are converted into current fractions, with each PP value representing the fraction of total current sourced (or sinked for the case of GND PPs) by that PP.

Although the noise levels are less than a couple hundred nAs, the addition and division operations used to convert the branch currents to current fractions amplifies the noise levels, requiring an increase in the statistical limits to 4.5 σ (over the industry standard 3 σ) [21]. The TR data is collected by enabling one of the 57 TE circuits and then measuring the 16 V_{DD} and 16 GND branch currents. NT and TR data is collected in this fashion from each of the 42 chips, for each of the 57 Trojans and at each of the 5 AVs. Scatterplots of currents measured from pairs of adjacent power ports (PPs) are created for application of the ellipse statistical method. It is possible to construct a set of 42 PP pairings using adjacent sets of 16 GND (or V_{DD}) PPs, as illustrated by the dotted lines in Fig. 7.

The ellipse analysis process is illustrated in Fig. 6. In (a), we show the 42 NT data points, i.e., current fractions derived from the GND grid analysis, as a scatterplot for PP pairing 03-13. The GND current fractions for PP_{03} are plotted along the x-axis against the GND current fractions for PP_{13} on the y-axis. The ellipse, derived from the first 30 chips, encloses the data points from all 42 chips, including the 12 control samples. Therefore, there are no FPDs in this example, i.e., none of the NT data points fall outside of the ellipse.

The TR data points associated with the first TE circuit from the 42 chips and 5 voltages are plotted in the scatterplot of Fig. 6(b) (the NT data points are removed for clarity). Data points that fall outside the ellipse bounds are considered true positive detections.

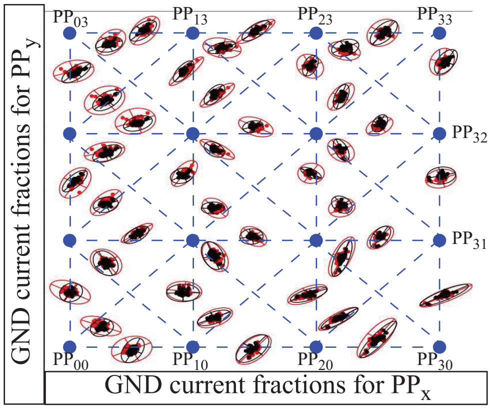


Fig. 7. Illustration of calibration process: red data and ellipses are uncalibrated, black is calibrated.

4.2 Directional Trending

The fixed position of the TE circuit #1 in all chips introduces a systematic component in the PP leakage currents. The arrow in Fig. 6(b) projects the overall trending of the TR data points. Fig. 6(c) depicts this directional trending much more clearly. The points on the TR curve in Fig. 6(c) are obtained by averaging the data points from the 42 chips at each of the 5 AV voltages. This **chip-averaging** process effectively eliminates random, within-die leakage current variations that occur in each of the chips. Therefore, **the chip-averaging technique is a very effective tool for identifying systematic current anomalies that occur across the entire set of chips**, which is precisely what we would expect if the layout has been manipulated by an adversary in all (or a large subset) of the chip-under-analysis. It is clear from the orderly progression of the TR data points beginning from a point close to the mean of the ellipse that the averaging process works well to capture the magnitude of the systematic current anomaly. Chip-averaging and directional trending are discussed further in Section 5.2.

4.3 Calibration

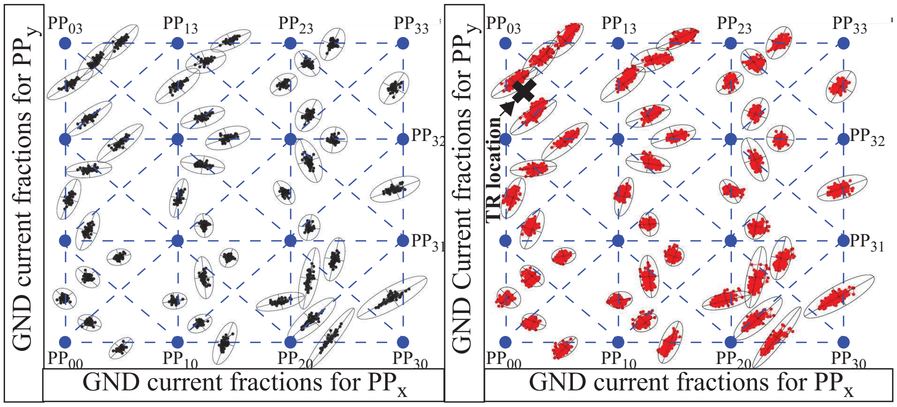
As mentioned above, Fig. 7 depicts the scatterplots and ellipses for the 42 pairings of PPs considered in our analysis. The scatterplots/ellipses are placed close to the dotted line that represents the PP pairings from which they are derived. We will use this technique as a means of illustrating the regional influence of the TE current anomaly. The superimposed red and black data points and ellipses represent the uncalibrated and calibrated data, resp. The black (calibrated) ellipses are smaller, illustrating the benefit of calibration, which increases the sensitivity of the method to smaller Trojan current anomalies. Space limitations prevent us from including a description of the calibration process, but the process is described in detail in [3].

5 Experimental Results

The ellipse-based statistical method is applied to both the raw data (Section 5.1) and chip-averaged data (Section 5.2) to demonstrate the benefits of chip-averaging on Trojan sensitivity and on reducing the number of FPDs.

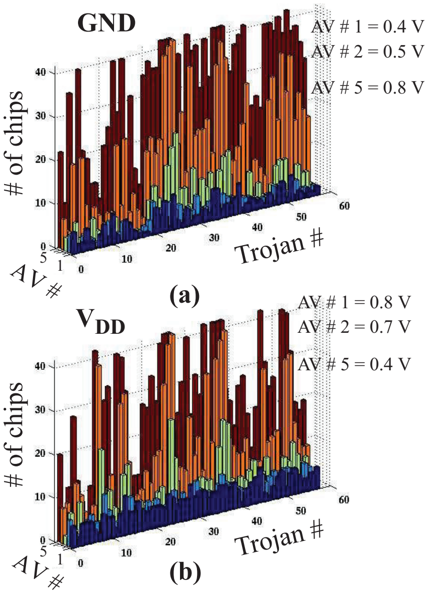
5.1 Outlier Detection using Individual Chips

The detection method described in this section is based on the analysis of data measured from individual chips.



(a) (b)

Fig. 8. Scatterplot analysis of TR #1: (a) Trojan-free (NT) data points and ellipses and (b) NT ellipses + TR data points.



(a)

(b)

Fig. 9. Histograms showing number of detections using the 1st leakage pattern for (a) GND and (b) V_{DD} analysis using individual chip data points.

Outlier analysis is carried out on the set of 42 scatterplots for each of the 57 Trojans, and at each of the 5 AVs separately. The ellipses are derived from the NT data of the first 30 chips while data from the remaining 12 chips is used as control samples for evaluating FPDs. A Trojan is counted as 'detected' if at least one chip (of 42) and at least one scatterplot (also of 42) has a data point that falls outside the ellipse, i.e., has an outlier.

The graphics in Fig. 8 illustrate our proposed detection approach. Following the presentation provided in Fig. 6, the NT data points and ellipses are shown in (a), while (b) shows the individual TR data points from the first Trojan experiment. The actual (x,y) position that TR #1 sinks current into the GND grid is marked with a 'X' in 6(b). Although TR #1 is detected at each of the 5 AVs by at least 3 chips, the small number of chips detecting it yields low confidence.

The detection results for all 57 Trojans using GND branch currents and under leakage pattern 1 are shown in Fig. 9(a) while those for the V_{DD} grid are shown in Fig. 9(b). The x-axis gives the Trojan # while the y-axis lists the AV number. The z-axis plots the number of chips that detect the Trojan/AV combination, with larger values representing higher confidence that a true Trojan I_{DDQ} anomaly exists. The trend in the histogram from smaller numbers of detections in row 1 (smallest AV) to larger numbers in row 5 is consistent with the expectation that the detectability of the I_{DDQ} anomaly is directly related to its magnitude.

All 57 Trojans at all 5 AV voltages (285 combinations) are detected in the V_{DD} analysis, while the GND analysis misses 7 Trojans, all at the lowest AVs. Table 1 shows the results for all 4 of the leakage patterns. In contrast to the results for leakage pattern 1, the GND analysis provides somewhat better results than the V_{DD} analysis. This is true because the noise levels are lower on the GND grid. There is 1 FPD for each of the GND and V_{DD} analysis for leakage pattern 2. Although these results show that Trojan current anomalies as small as 10 μAs can be detected, this approach misses a fair number of Trojans under some leakage patterns and produces FPDs under others. These two metrics can be traded-off by tuning the statistical threshold, i.e., increasing the statistical limit reduces FPDs while increasing the number of misses, and vise versa. However, this is no threshold that makes both of them 0.

Table 1: Trojan Results using Individual Chips.

	Leakage Pattern 1	Leakage Pattern 2	Leakage Pattern 3	Leakage Pattern 4
GND misses	7	13	0	0
GND FPD	0	1	0	0
V_{DD} misses	0	32	45	85
V_{DD} FPD	0	1	0	0

5.1.1 Distinguishing Defects from Trojans

Random, erratic behavior of data points in the individual scatterplots caused by defects can easily be distinguished from the systematic behavior of the data points caused by Trojans. Fig. 10 shows a scatterplot associated with PP pairing 00-01 using data from the 3rd leakage pattern. The larger ellipse labeled 'Original ellipse with outliers' is derived from a set of 45 chips, 42 defect-free chips and 3 defective chips. The data points from the 3 defective chips are labeled 'outlier' in the figure. Unlike the systematic anomalies created by Trojans, these 3 data points each appear as random, i.e., the direction of their displacement from the main cluster of points is unique. Therefore, it is easy to identify and eliminate the data points associated with these chips as potential Trojan candidates. In an actual application of our method, these data points would be produced by the chips-under-test (they would not be treated as Trojan-free as we do here) and therefore this type of identification process is extremely important. This is a good illustration of the importance of the following principle: **Methods that provide high detection sensitivity to Trojans must also be capable of dealing with a significant increase in false positives caused by subtle defects that have no effect on functionality**.

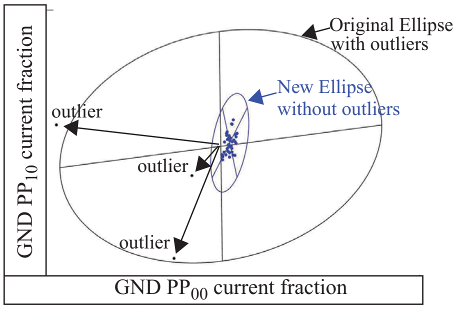


Fig. 10. NT outlier identification and elimination using data from leakage pattern 3.

5.2 Outlier Detection using Chip-Averaged Data Points

Deriving the statistical limits using data points from the individual chips as we did in the last section does not calibrate for within-die variations. Although the MSP technique is still robust enough to detect many types of Trojans under these conditions, eliminating within-die variations can both improve detection sensitivities and help with eliminating FPDs. Chip-averaging is very effective at accomplishing this goal as we show in this section.

The general idea of the chip-averaging technique is to compute an average using data from a subset of the chip population. In order to determine the noise levels and uncertainty associated with the NT data, this process is repeated using additional subsets of chips and the resulting collection of data points, each of which is a chip-averaged result, is then used to derive the ellipses for each power port pairing. Note that this process can be carried out in practice using data from the actual chips because it only defines the size of the ellipses and not their mean values. Therefore, whether the chips have a Trojan or not is irrelevant. The mean value of the ellipses, on the other hand, must be derived from a simulation of a golden model of the chip.

A significant benefit of this type of approach is that the Trojan-free leakage behavior for the chip is defined from a single nominal simulation for each leakage vector that is to be applied (reducing simulation time significantly) while the uncertainty defined by measurement noise and uncalibrated process variations is defined by the test data itself (significantly improving the ability of the method to accurately characterize measurement noise and process variations).

As mentioned earlier, lack of access to foundry data prevented us from developing an accurate simulation model in our experiments and instead, we derive the mean values of the ellipses from the chip data by disabling all of the TE circuits. We also needed to craft a method for defining the ellipse statistical limits. Ideally, each chip-averaged data point used to derive the ellipses would be obtained from a different set of chips in the population. Our small chip population required that we reuse chips in each of our subsets. The subset of chips used to define each data point is constructed by randomly partitioning the 42 chips into 2 subsets. The average values from each of the two subsets of 21 chips define two data points for each PP pairing. We repeat this process 21 times to produce 42 data points for each PP pairing. The ellipses for each PP pairings are derived using the first 30 data points while the remaining 12 data points are used as control samples. Chip-averaging allows us to reduce the statistical limit to 4 σ in these experiments, from 4.5 σ used in Section 5.1.

Fig. 11(a) shows the ellipse and all 57 chip-averaged TR curves for PP pairing 00-01 for GND analysis under leakage pattern 1. The 42 NT data points are enclosed within the ellipse, i.e., there are no NT outliers or FPDs, while most of the TR data points are detected as outliers. This graph clearly shows the benefits of chip-averaging. Interestingly, the normalization process which converts the currents into current fractions (see Section 4.1) makes it possible to detect Trojans that are not 'in the region' of this PP pairings (see curves labeled 'non-regional TR detections'). This occurs because the fraction of current in both PPs of the pairing reduces nearly equally for pairings that are not close to the actual Trojan, causing the data points to move toward the origin. Similar characteristics can be observed in the other PP pairing as shown by Fig. 11(b).

Directional trending, discussed earlier in Section 4.2, can be used to increase the confidence in the detection decision. The authors of [22] show that Trojan current anomalies present in the PPs provide information regarding the actual location of the Trojan. Location information is captured in the *angle* associated with the directional trending, as illustrated with the label 'DT angle' in Fig. 11(c), which shows the chip-averaged results for TR #1 only. Although directional trending can be identified as shown, collating the angle information from multiple scatterplots to determine if there is a common point of intersection in the layout would provide much stronger proof that the anomaly is systematic and potentially introduced by a Trojan as opposed to a random defect. We will investigate the usefulness of directional trending information in future work.

Similar to the format shown in Section 5.1, the histograms in Fig. 12 show the number of detections for each of the 57 Trojans and each of the 5 AVs. The z-axis in this analysis counts the number of PP pairings (of 42) that detect the Trojan, as opposed to the number of chips in Fig. 9 (chip-averaging eliminates chips as a dimension to our analysis). The improvement in sensitivity can be clearly observed by comparing the histograms in both figures. Table 2 summarizes the results, showing that only 6 Trojans are missed under leakage pattern 1 (all at the lowest AVs), and the number of FPDs is 0 for all leakage patterns.

Table 2: Trojan Results using Chip-Averaged Data Points.

	Leakage Pattern 1	Leakage Pattern 2	Leakage Pattern 3	Leakage Pattern 4
GND misses	6	0	0	0
GND FPD	0	0	0	0
V_{DD} misses	6	0	0	0
V_{DD} FPD	0	0	0	0

6 Conclusions

In this paper, we carried out hardware experiments in which Trojans are emulated in a set of 42 chips fabricated in a 90 nm technology. A MSP technique, in combination

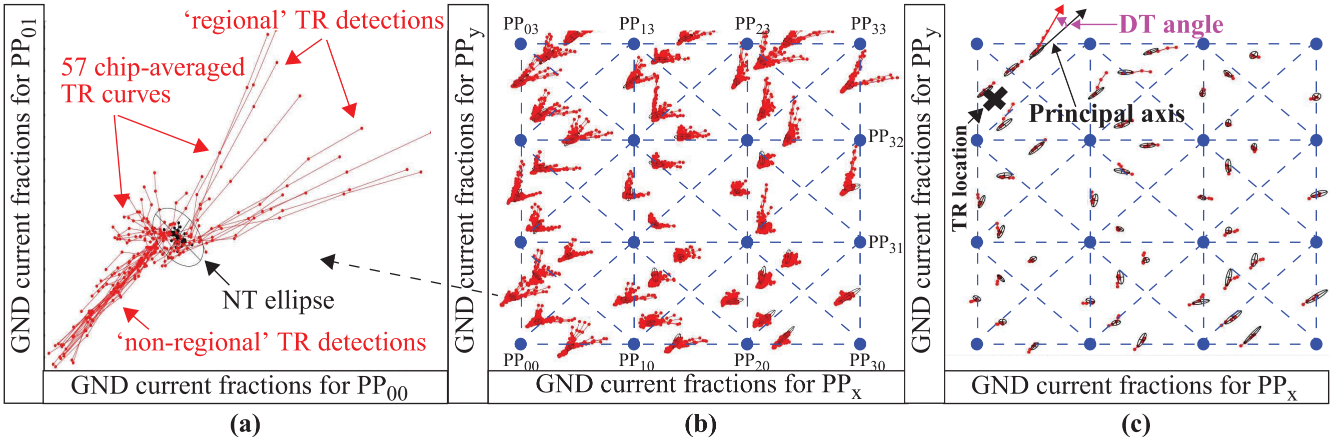


Fig. 11. Chip-averaged GND, leak pattern 1 data. (a) PP pairings 00-01 with NT ellipse and all 57 TR curves, (b) all PP pairings showing ellipse and all 57 TR curves and, (c) all PP pairings showing ellipse and TR #1 curves.

with a power signal calibration and a chip-averaging method, are shown to significantly reduce the adverse effects of chip-to-chip and within-die variations effects on Trojan detection sensitivities. We show that Trojans that introduce and/or redistribute currents as small as 10 µAs are detectable in a chip of size 2 mm x 2 mm.

References

[1] http://www.acq.osd.mil/dsb/reports/2005-02-HPMS_Report_Final.pdf

[2] http://www.darpa.mil/mto/solicitations/baa07-24/index.html

[3] J. Aarestad, D. Acharyya, R. Rad and J. Plusquellic, "Detecting Trojans Though Leakage Current Analysis Using Multiple Supply Pad IDDQs", *Trans. on Information Forensics and Security*, Volume: 5, Issue: 4, pp. 893-904, 2010.

[4] R. Rad, J. Plusquellic, M. Tehranipoor, "Sensitivity Analysis to Hardware Trojans using Power Supply Transient Signals", *HOST*, 2008, pp. 3-7.

[5] D. Agrawal, S. Baktir, D. Karakoyunlu, P. Rohatgi, B. Sunar, "Trojan Detection using IC Fingerprinting", Symposium on Security and Privacy, 2007, pp. 296 - 310.

[6] F. Wolff, C. Papachristou, S. Bhunia, and R. Chakraborty, "Towards Trojan-Free Trusted ICs: Problem Analysis and Detection Scheme", *DATE*, 2008, pp. 1362-1365.

[7] Jie Li and John Lach, "At-Speed Delay Characterization for IC Authentication and Trojan Horse Detection", *HOST*, 2008, pp. 8-14.

[8] M. Banga and M. S. Hsiao, "A Region Based Approach for the Identification of Hardware Trojans", *HOST*, 2008, pp. 40-47.

[9] R. S. Chakraborty, S. Paul and S. Bhunia, "On-Demand Transparency for Improving Hardware Trojan Detectability", *HOST*, 2008, pp. 48-50.

[10] Y. Jin and Y. Makris, "Hardware Trojan Detection Using Path Delay Fingerprints", *HOST*, 2008, pp. 51-57.

[11] H. Salmani, M. Tehranipoor, "Layout-Aware Switching Activity Localization to Enhance Hardware Trojan Detection", *Trans. on Information Forensics and Security*, Vol. 7, Issue: 1, Part: 1, 2012, pp. 76-87.

[12] C. Byeongju and S. K. Gupta, "Efficient Trojan Detection via Calibration of Process Variations", *Asian Test Symposium*, 2012, pp. 355-361.

[13] J. Zhang, Y. Haile and X. Qiang, "HTOutlier: Hardware Trojan Detection with Side-Channel Signature Outlier Identification", *HOST*, 2012, pp. 55-58.

[14] W. Sheng and M. Potkonjak, "Malicious Circuitry Detection using Fast Timing Characterization via Test Points", *HOST*, 2013, pp. 113-118.

[15] I. Exurville, J. Fournier, J.-M. Dutertre, B. Robisson and A. Tria, "Practical Measurements of Data Path Delays for IP Authentication & Integrity Verification", *International Workshop on Reconfigurable and Communication-Centric Systems-on-Chip*, 2013, pp. 1-6.

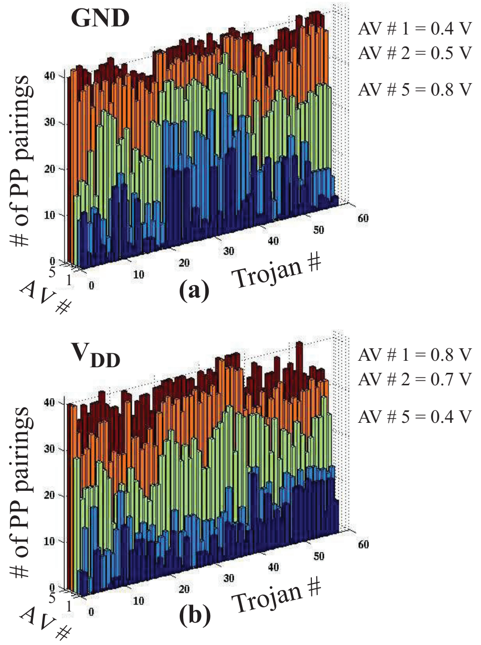


Fig. 12. Histograms showing number of detections using the 1st leakage pattern for (a) GND and (b) V_{DD} analysis using chip-averaged data.

[16] Y. Lu, Y. Jin and Y. Makris, "Hardware Trojans in Wireless Cryptographic ICs: Silicon Demonstration & Detection Method Evaluation", *ICCAD*, 2013, pp. 399-404.

[17] L. Wang, H. Xie and H. Luo, "A Novel Analysis Method of Power Signal for Integrated Circuits Trojan Detection", *International Symposium on Physical and Failure Analysis of Integrated Circuits*, 2013, pp. 637-640.

[18] A. Davoodi, L. Min and M. Tehranipoor, "A Sensor-Assisted Self-Authentication Framework for Hardware Trojan Detection", *IEEE Design & Test*, Vol. 30, Issue: 5, 2013, pp. 74-82.

[19] S. Narasimhan, D. Dongdong, R. S. Chakraborty, S. Paul, F. G. Wolff, C. A. Papachristou, K. Roy and S. Bhunia, "Hardware Trojan Detection by Multiple-Parameter Side-Channel Analysis", *Transactions on Computers*, Vol. 62 , Issue: 11, 2013, pp. 2183-2195.

[20] http://en.wikipedia.org/wiki/Principal_component_analysis

[21] http://en.wikipedia.org/wiki/Propagation_of_uncertainty

[22] J. Plusquellic and D. Acharyya, "Leveraging the Power Grid for Localizing Trojans and Defects", *International Symposium on Testing and Failure Analysis*, 2010.