Post-Layout Estimation of Side-Channel Power Supply Signatures

Sushmita Kadiyala Rao, Deepak Krishnankutty, Ryan Robucci, Nilanjan Banerjee and Chintan Patel
CSEE, University of Maryland, Baltimore County

***Abstract*—Two major security challenges for integrated circuits (IC) that involve encryption cores are side-channel based attacks and malicious hardware insertions (trojans). Side-channel attacks predominantly use power supply measurements to exploit the correlation of power consumption with the underlying logic operations on an IC. Practical attacks have been demonstrated using power supply traces and either plaintext or cipher-text collected during encryption operations. Also several techniques that detect trojans rely on detecting anomalies in the power supply in combination with other circuit parameters. Counter-measures against these side-channel attacks as well as detection schemes for hardware trojans are required and rely on accurate pre-fabrication power consumption predictions. However, available state-of-the-art techniques would require prohibitive full-chip SPICE simulations. In this work, we present an optimized technique to accurately estimate the power supply signatures that require significantly less computational resources, thus enabling integration of Design-for-Security (DfS) based paradigms. To demonstrate the effectiveness of our technique, we present data for a DES crypto-system that proves that our framework can identify vulnerabilities to Differential Power Analysis (DPA) attacks. Our framework can be generically applied to other crypto-systems and can handle larger IC designs without loss of accuracy.**

***Index Terms*—Hardware Security, trojan detection, Side-channel attacks, Power Supply analysis**

I. Introduction

THE high cost of custom IC fabrication has given rise to a large number of fabless IC firms that outsource their design for manufacturing at third party foundries. This entails design transfer to untrusted, off-site foundries thus making fabricated ICs vulnerable to security compromise, malicious hardware modifications, and proprietary information leakage.

These vulnerabilities manifest themselves as hardware trojans or side channel leaks post-fabrication. In military systems, financial infrastructure, transportation and automotive devices, as well as household appliances these vulnerabilities can have deleterious effects. There is a need, therefore, for efficiently assessing these security vulnerabilities during design phase and aiding post-fabrication device testing to verify IC authenticity [1]. Conventional techniques for detecting trojans or side channel leakage that rely on processing power supply information require generation of *golden signatures* [1], [2], [3], [4]. These signatures quantify the power consumption of security conscious designs and untampered ICs. The golden signature is predominantly determined by performing exhaustive tests on a select set of ICs and measuring the power consumption. During device testing, the measured power consumption from the Chip-Under-Test (CUT) is compared against the golden signature to identify either malicious hardware insertions or side channel leaks. Such techniques suffer from two drawbacks: 1) they assume that this select set of ICs do not have malicious circuits and 2) exhaustive post-fabrication testing can be prohibitively expensive, even for a small subset of ICs.

Researchers have proposed on-chip power monitoring systems e.g. using ring oscillators [3] that measure the dynamic power in combination with off-chip measurement equipment to derive the golden signatures. These methods, though computationally cheaper, fundamentally suffer from the previously identified drawbacks. Additionally, these techniques cannot localize circuitry on the chip that is being attacked. This information can be used to identify and address the security vulnerability.

In this paper we present a pre-fabrication CAD framework to derive the golden signatures for chip power consumption that is computationally efficient and highly accurate. Our technique augments the IC design flow to integrate Design for Security (DfS). The design and evaluation of our technique presents the following research contributions.

- **An Efficient CAD Framework for Deriving the Golden Signatures that Addresses Security Vulnerabilities:** We describe a technique that can be used for deriving the golden signatures for detecting IC-level security vulnerabilities such as side-channel attacks and hardware trojans. The framework can also be leveraged for applications in delay and transition based testing as well as power supply noise (PSN) estimation [11]. The technique optimizes power consumption estimation using a partitioning technique that characterizes the power grid independent of the chip logic. It precomputes the gate-level current transients using path simulation which are then convolved with the Current-to-Current impulse responses to estimate the power consumption at each power pad in the IC.
- **Localize IC Circuit Elements Under Attack:** Our framework can estimate power consumption at each power pad on an IC accurately, and we leverage differences in the power signatures at multiple supply pads to localize the circuit element under attack. We demonstrate this feature in §IV by localizing the S-box that processes the bit under a Differential Power Attack (DPA).

II. Background

Side Channel Attacks allow an attacker to extract secret keys from a target device by monitoring the power supply, electromagnetic radiation or timing information.

Simple Power Analysis (SPA) is the most basic among all side channel attacks. This technique involves direct interpretation of the power supply traces from the operation of interest. Differential Power Analysis (DPA) attack was introduced by Kocher et al. [5] to identify secret keys from the CUT's power traces. They have since been used widely in the cryptographic community [8]. Single-bit and multi-bit DPA attacks have been demonstrated. The difference is in the requirement for a larger number of guessed keys due to the additional bits. These extra bits generate more intermediate values, unlike the two possibilities in the single-bit case, implying more than two groups to sort power traces. The groups are simplified and combined to create a single DPA trace for the guessed key. Correlation Power Analysis, first introduced in [6] and implemented in [7] utilizes a statistical approach to compute the correlation between power traces observed from the Device Under Test (DUT) and a power model based on Hamming distance or Hamming weight.

DPA attacks against the Data Encryption Standard (DES) algorithm are primarily targeted at the first or final round of encryption. A detailed description of the DES algorithm is available from [9]. For the purpose of analysis, this paper focuses on the initial round of encryption.

III. Power Supply Prediction Technique

A convolution-based framework is presented that accurately predicts power consumption at each supply pad in the IC.

A. System Partitioning

Our estimation techniques are based on a modular framework proposed in [10]. They show that a digital chip can be partitioned into two independent subsystems, namely the linear Power Grid Circuit (PGC) and the non-linear Core Logic Circuit (CLC) as shown in Figure 1.

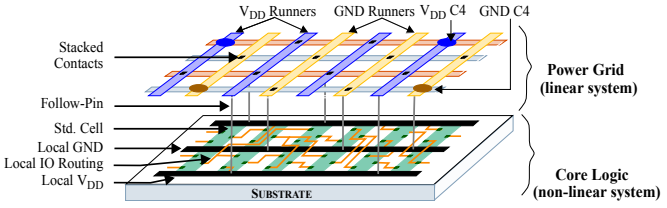

Fig. 1: Linear and non-linear components of a chip [10]

The power grid network is a multi-input, multi-output linear system. Since any linear time invariant (LTI) system can be characterized by its impulse response (IR), it is possible to characterize the response of the power grid to any arbitrary input signal (in this case, switching logic in the crypto-system).

B. Power Estimation Framework

Grid simulations are carried out to compute IR responses between input and output locations of the power grid. The grid characterization process is shown in Figure 2. The current response to an input is computed using Current-to-Current impulse responses (C2C) that are computed by applying a step input and differentiating the response at an output. C2C IR provides the relationship between the current source applied at the input and the corresponding currents measured at any output location on the power grid. Switching gates are identified using a Verilog logic simulation and are extracted from a detailed RLC netlist. These isolated paths are simulated and provide current inputs to the power grid. Figure 3 illustrates how power transients are estimated by convolving grid IR responses with individual gate transients.

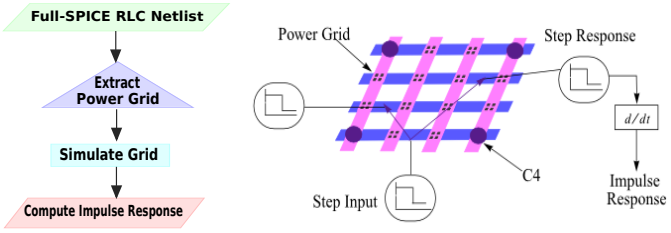

Fig. 2: Power Grid Characterization Flow

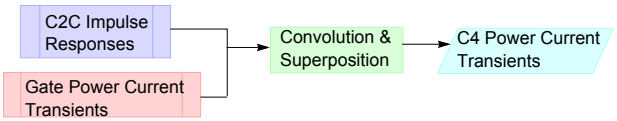

Fig. 3: Computation of Power Pad Transient Current

The layout of DES core is implemented in 180-nm technology using a commercial place-and-route tool. The chip has four V_{DD} and four GND power pads. In order to cater to much larger circuits, a partitioning strategy as shown in [11], where both the partitioning and signal redistribution methods have been applied to two test chips based on the c6288 ISCAS'85 benchmark circuit to estimate the transient response at the power pads of the test circuits for different test stimuli.

IV. Evaluation

In this section, we analyze the effectiveness of our power consumption estimation technique. The ground-truth is generated from full-chip SPICE simulations. Additionally, we evaluate the accuracy of our framework for application to DfS using a security metric described below.

A. Computational Performance

Our framework that encompasses a partitioned system as well as full-chip SPICE simulations were performed using the Cadence Spectre Accelerated Parallel Simulator^{TM} (APS)[12]. Table I compares the time required to simulate a full-chip SPICE netlist of the DES core, an individual path and the power grid.

The grid simulation for each input takes about 27 seconds, however for larger circuits the number of inputs will be high. Since this characterization is performed *only once* for a design, the run-time will be amortized over all the predictions. Thus, the computational advantage of our framework versus the full-chip SPICE will be substantial for large designs.

TABLE I: Simulation Time Comparison

Component Simulated	Time
Full Chip SPICE	7hr 50m
Path SPICE	4hr 11m
Grid simulation for each input	27s

B. Correlation-Based Analysis of Side-Channel Attacks

The purpose of our simulation model is to predict power supply variations in response to both input data and design variations. The ability to predict design dependent perturbations allows for an iterative design cycle which involves an initial design, evaluation of security vulnerabilities and mitigating these problems. We evaluate our framework using a security metric based on correlations and one that compares absolute similarity described below.

A system perfectly invulnerable to side channel exploitation would either offer an immutable supply responseor at least that the side-channel responses are uncorrelated to the input (perhaps randomized). Such an ideal system is likely impossible to design—though getting closer at a reasonable cost requires a tool for estimating leakage. Therefore, an effective comparison metric for the similarity/difference in waveforms is the correlation factor: $\rho_{x,y} = \frac{\sum_i (x[i]-\mu_x)(y[i]-\mu_y)}{\sqrt{\sum_i (x[i]-\mu_x)^2 \sum_i (y[i]-\mu_y)^2}}$

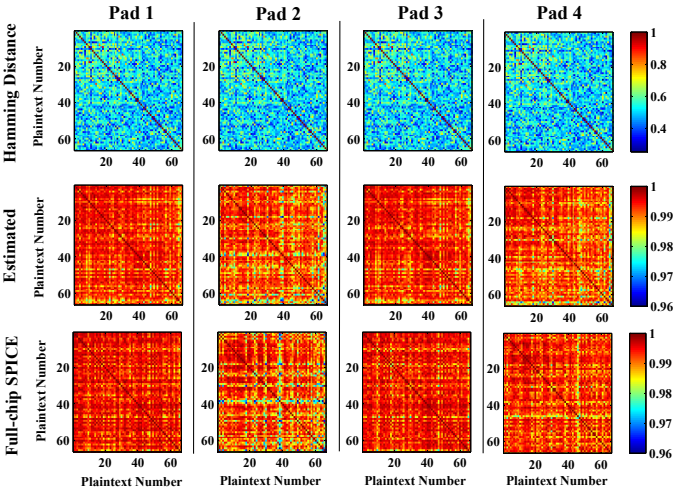

Fig. 4: Correlation analysis for 66 plain-texts Row 1) Hamming Distance, Row 2) our framework and Row 3) Full-chip SPICE

In Figure 4 the last two rows show the cross-correlation between current transients (with respect to each of the four power pads) for 66 different plain-texts for our framework and full-chip SPICE, respectively, while the first row estimates the same using Hamming Distance (which are independent of which power pad we are examining since it does not encode any information about position). Such matrices convey the mutual information between the power supply and the inputs.

The values from the full-Chip SPICE present the challenge to the attacker. The correlation values are all high, meaning there is a low variation between waveforms as compared to the magnitudes of the transients. This means it is difficult to infer an input from a measurement of the power supply, particularly in the presence of noise in a physical system.

Even so, previous work has shown that once enough transients are captured, hidden values may be uncovered such as private keys [5]. Therefore, evaluating this weak yet critical dependency is important. The values produced using our framework in Figure 4 show the predicted data-dependent correlations.

As Table II shows, the values in rows two and three of Figure 4 are similar. In particular, we draw attention to the similarities between row 1 and 3 which predicts the effectiveness of a particular Hamming Distance-based DPA attack. Therefore, *our system will effectively predict data-dependent perturbations as well as predict vulnerabilities to power supply-based side-channel attacks.* We next present data that evaluates the accuracy of our framework.

TABLE II: SPICE-Estimation power transient correlations

Data Dependent Correlations Comparison	P1	P2	P3	P4
Hamming Distance vs SPICE	.1289	.1161	.1067	.1351
Estimated vs SPICE	.6321	.6883	.6234	.5556
Hamming Distance vs Estimated	.1433	.1316	.1491	.1687

C. Power Supply Estimation

To demonstrate the temporal similarities between full-chip SPICE and by our power supply estimation framework absolute, we present data for 66 plain-text input for a fixed secret key. For a particular combination of encryption key and input plain-text, the current waveforms estimated (represented by dotted lines) at the four power pads are compared to full-chip SPICE waveforms (represented by solid lines). Shown in Figure 5 are waveforms (magnified) for plain-text 1 during the initial rounds of encryption. As seen in the figure, the estimated waveforms match full-chip SPICE waveforms closely.

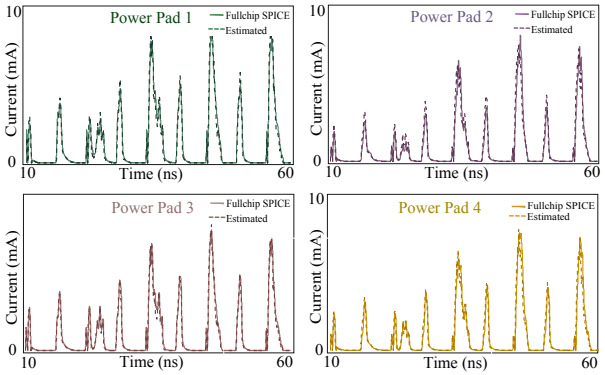

Fig. 5: Current Transients of plain-text 6 for all Power Pads

The metric used to evaluate the effectiveness of our prediction techniques is Mean Square Error (MSE). The MSE (%) is computed to compare the difference between the full-chip and estimated waveforms as defined in Equation 1.

$$\frac{\sum_{n=0}^{N-1}(I_{FullChip}[n] - I_{Estimated}[n])^2}{\sum_{n=0}^{N-1}(I_{FullChip}[n])^2} \times 100\ \% \quad (1)$$

In Eq.1, $I_{FullChip}$ is the full-chip SPICE transient current, $I_{Estimated}$ is the estimated transient current and N is the number of data points in the waveforms. The maximum MSE % observed across all four power pads over 66 paths is 11 %. Figure 6 shows the MSE % for all 66 plain-text inputs per power pad and the average is indicated by a horizontal line corresponding to each power pad.

To demonstrate applicability, this technique can be utilized in vulnerability assessments of encryption systems, the DES current transients are subjected to DPA attacks described in the background section.

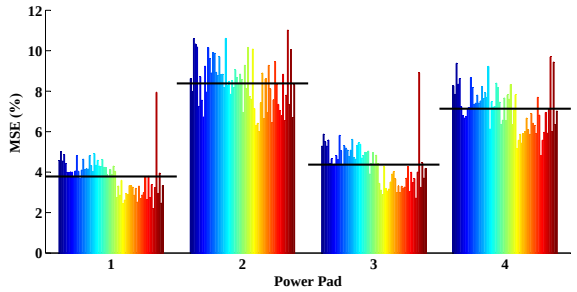

Fig. 6: Mean Square Error (%) for 66 paths for all Power Pads

The current peaks after the first round of DES for a correct key guess versus the incorrect guessed key at power pad 1 can be seen in Figure 7. Also noticeable are secondary peaks indicating the state of the intermediate values before moving over to the next round. Similar observations can be made for power pads 2, 3 and 4. Figure 8 shows results for DPA attack based on 1000 guesses of a bit-sequence for the left input register to a given DES encryption clock cycle. actually be found to match one of the sequence of bits from the 32 left-registers.

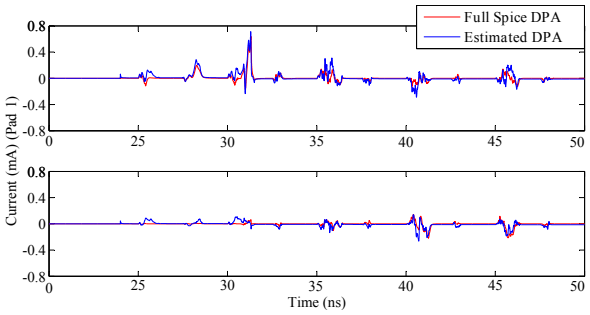

Fig. 7: DPA on a correct key guess (top) Vs incorrect key guess (bottom) for Power Pad 1

Proven analysis of DPA attacks, show that if a bit sequence hypothesis matches a sequence found in operation, a correlation between the hypothesis and the power supply transients can be found. In the analysis, 32 of the hypothesized sequences can

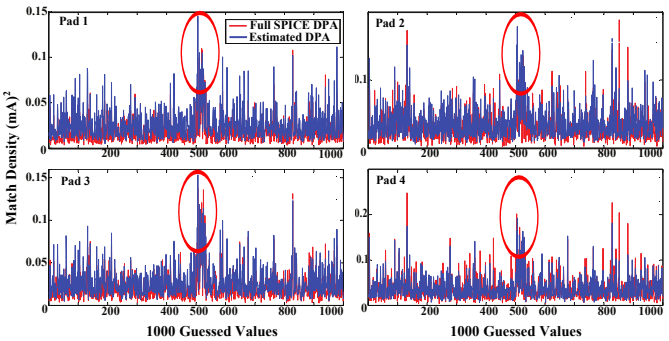

Fig. 8: DPA signal distribution over 1000 guessed values for all power pads

As expected, these show a higher correlation to the power supply, as compared to the total population of guesses, according to both the full-spice and our estimation method. The peaks (marked by red ovals) clearly indicate the matching sequences of bits occurring between the 500^{th} and 531^{st} guessed sequences. The DPA attack for Full-chip SPICE simulation traces are indicated in red and the Estimated Simulation traces are indicated in blue. Peaks appearing prior and after the aforementioned interval are also indicative of partially matching sequences which occur with some probability at other points in the simulation.

Inferences can be drawn, from the observation that some of the peaks in the interval appear more prominently in comparison with other peaks. Registers in the layout with a closer proximity to a specific power pad tend to have higher current peaks in comparison to other registers placed at a distance from the power pad. As an example, guessed bit number 506 (which corresponds to the 7th s-box's bit after the first round of DES) at power pad 4 has a higher peak in relation with the other matched bits. This gives hints on localizing the position of core elements under attack in a Chip-Under-Test.

V. Conclusion

In this work we show results using our framework to accurately estimate power supply signatures of a DES system without running full-chip SPICE simulations. The estimated waveforms compared well against full-chip SPICE results. The predicted power supply signatures can be used to predict side-channel leakage, help in designing and evaluating counter-measures against attacks and provide golden signatures for trojan detection schemes. Furthermore, the current transients observed on each power pad also gives some hints on the proximity of core-elements being observed or attacked. The technique can be generically applied to other crypto-systems as well as industrial scale ICs by leveraging our partitioning scheme.

References

[1] Tehranipoor, M; Koushanfar, F; , *A Survey of Hardware Trojan Taxonomy and Detection*, Design & Test of Computers, IEEE , no.99, pp.1, 2010

[2] Agrawal, D.; Baktir, S.; Karakoyunlu, D.; Rohatgi, P.; Sunar, B.; , *Trojan Detection using IC Finger-printing*, Security and Privacy, 2007. SP 07, pp. 296-310, 20-23 May 2007.

[3] Xuehui Zhang; Tehranipoor, M.; ,*RON: An on-chip ring oscillator network for hardware Trojan detection*, Design, Automation & Test in Europe Conference & Exhibition (DATE), 2011 , pp.1-6, 14-18 March 2011.

[4] Davoodi, A.; Min Li; Tehranipoor, M. *A Sensor Assisted Self Authentication Framework for Hardware Trojan Detection, Design & Test*, IEEE, On page(s): 74-82 Volume: 30, Issue: 5, Oct. 2013.

[5] P. Kocher, J. Jaffe, B. Jun. *Differential power analysis*, CRYPTO,LNCS 1666, pp.388-397,1999.

[6] J.-S. Coron, D. Naccache and P. Kocher,. *Statistics and secret leakage, ACM Transactions on Embedded Computing Systems*, Vol. 3, No. 3, pp. 492-508, August 2004,

[7] E. Brier, C. Clavier, F. Olivier, *Correlation Power Analysis with a Leakage Model*, proceedings of CHES 2004, LNCS, 3156 , pp. 16-29, 2004

[8] E. Peeters, Advanced DPA Theory and Practice - Towards the Security Limits of Secure Embedded Circuits, Chapter 2, 2013.

[9] National Bureau of Standards, "Data Encryption Standard", Federal Information Processing Standards Publication 46, January 1977.

[10] Abhishek Singh, Jim Plusquellic, Dhananjay Phatak and Chintan Patel, *Defect Simulation Methodology for iddt Testing*, J.Electron Test.,22:255-272, June 2006.

[11] Rao, S.K.; Robucci, R.; Patel, C.; , *Framework for Estimation of Dynamic Power-Supply Noise and Path Delay*, Defect and Fault Tolerance in Symposium, Oct.2-4, 2013.

[12] http://www.cadence.com/products/cic/accelerated_parallel/pages/default.aspx, (Accessed October 2014)