Gate-Level Side-Channel Leakage Assessment with Architecture Correlation Analysis

Pantea Kiaei, Student Member, IEEE, Yuan Yao, Zhenyuan Liu, Student Member, IEEE, Nicole Fern, Cees-Bart Breunesse, Jasper Van Woudenberg, Kate Gillis, Alex Dich, Peter Grossmann, and Patrick Schaumont, Senior Member, IEEE

Abstract—While side-channel leakage is traditionally evaluated from a fabricated chip, it is more time-efficient and cost-effective to do so during the design phase of the chip. We present a methodology to rank the gates of a design according to their contribution to the side-channel leakage of the chip. The methodology relies on logic synthesis, logic simulation, gate-level power estimation, and gate leakage assessment to compute a ranking. The ranking metric can be defined as a specific test by correlating gate-level activity with a leakage model, or else as a non-specific test by evaluating gate-level activity in response to distinct test vector groups. Our results show that only a minority of the gates in a design contribute most of the side-channel leakage. We demonstrate this property for several designs, including a hardware AES coprocessor and a cryptographic hardware/software interface in a five-stage pipelined RISC processor.

Index Terms—Pre-silicon, Side-channel leakage, Power estimation, Hardware security

1 INTRODUCTION

Power-based side-channel leakage occurs when a secure chip performs operations that depend on an internal secret value such as a secret key. An adversary who observes the chip power consumption can derive the internal secret value through differential analysis techniques that correlate a power model of the secret activity with the observed power consumption. In recent years, side-channel vulnerabilities have risen to prominence and successful side-channel attacks have been demonstrated on a wide range of devices from small IoT devices to large cloud computing systems. Therefore, the evaluation of the power-based side-channel leakage has become a critical component in the design flow of secure chips. It is particularly helpful to perform side-channel leakage assessment prior to manufacturing because it reduces the cost of post-manufacturing testing, and it reduces the probability of side-channel vulnerabilities in the chip tape-out.

Fig. 1 compares the Side-channel Leakage Assessment (SLA) process of a post-silicon assessment flow with a pre-silicon assessment flow. The starting point is identical in both cases and assumes that a Register-transfer Level (RTL) description of the design under consideration is available.

With a post-silicon SLA flow, the RTL design is first prototyped into a physical implementation. Power measurements are then collected from the design and statistically tested to confirm the presence of side-channel leakage or to estimate the quantity of side-channel leakage. In a pre-silicon strategy, the RTL design is synthesized into a gate-level netlist including optional parasitic effects from place-and-route. Next, power traces are simulated and then statistically tested to confirm the presence of side-channel leakage.

These two flows appear similar from a macro-level objective, but they have very different properties. A post-silicon flow is expensive because of the extra prototyping step, which slows down the verification cycle. The statistical tests are applied globally on the measured leakage of the overall design. In a large and complex design, it therefore remains difficult to pinpoint the leakage source.

In contrast, a pre-silicon flow makes use of simulated power traces, and it is able to perform side-channel leakage
discuss related work on each of these two aspects. The simulated traces of the pre-silicon flow are noiseless, and therefore they represent the attacker with the best possible observation. Due to the absence of noise, a pre-silicon flow can work with a fraction of the number of power traces compared to a post-silicon flow. On the downside, a pre-silicon flow must make a trade-off between accuracy and simulation speed. We use a gate-level power simulation methodology that is able to capture many technology-dependent effects (such as glitches [1] and static leakage [2]). Some side-channel leakage effects, including those based on coupling [3] or the long-wire effect [4], require a simulation accuracy beyond what gate-level power simulation can offer. Our proposed flow offers gate-level side-channel leakage assessment but makes no assertion of leakage below that abstraction level.

This paper presents the following contributions. We describe a methodology called Architecture Correlation Analysis (ACA) which determines the side-channel leakage of a design at the granularity of a single gate. The basic principle of ACA has been initially proposed in our previous work [5], [6]. This paper serves as an extension to our original publication. In this extended work, we propose two different side-channel leakage assessment techniques for use in ACA. The first one is based on a specific test and it demonstrates the presence of correlation between a specific power model and individual logic gates. The second is based on a non-specific test that ranks a gate’s power according to its ability to distinguish between two distinct groups of test vectors. The specific test is used to identify the gates that enable a specific side-channel attack, while the non-specific test is used to make a generic assessment on how much potentially harmful leakage can be produced by a gate. We apply our proposed ACA leakage assessment technique to two case studies: a cryptographic AES coprocessor, and the driver software for that coprocessor when running on a five-stage pipelined RISC processor. The side-channel leakage properties of AES are already well understood, and our experiments are specifically aimed at the ability of pre-silicon leakage assessment to identify the source of side-channel leakage.

The remainder of the paper is organized as follows. The next section describes related work. Section 3 provides the overall methodology of Architecture Correlation Analysis (ACA), highlighting both the specific and the non-specific testing strategy. We also discuss a prototype implementation of the flow. Section 4 applies our proposed methodology to a cryptographic coprocessor. Section 5 applies the methodology to cryptographic driver software running on a RISC processor. Section 6 evaluates the performance of the proposed methodology. We then conclude the paper.

## 2 RELATED WORK

Gate-level side-channel leakage assessment is built on two components of design automation: (a) power simulation under a set of selected test vectors, and (b) identification of a leakage source at the sub-module level or gate level. We discuss related work on each of these two aspects.

### 2.1 Power simulation for side-channel leakage analysis

To simulate a design’s power consumption, one needs a model of the design implementation details to estimate the power of the physical implementation under a set of test vectors. The model can be constructed at different abstraction levels, and there is a trade-off between modeling detail and simulation performance. Buhun et al. review many of the recent proposals to simulate side-channel leakage [7] and here we only describe the most representative ones.

The origin of side-channel leakage power simulation is found in simulators for smart cards, starting with PINPAS [8]. The objective of these instruction-level simulations is to generate a power trace corresponding to a software application running on an embedded micro-controller. These simulators are processor-specific, and require knowledge of the internal design of the processor. Recent research efforts have addressed power model construction techniques to handle the case when the internal design is unknown. This includes ELMO [9] and ROSITA [10] for power, and EMSIM [11] for Electromagnetic Radiation. In our approach, we build on the basic assumption that the hardware source code is available.

It is now commonly understood that instruction-level power modeling by itself is inadequate to accurately capture all aspects of power-based side-channel leakage, and that additional modeling detail is required to capture circuit-level effects. The CASCADE power simulation flow aims at a comprehensive simulation of power traces at the gate-level [12], while making the argument that gate-level power simulation hits a sweet spot for the known power-based side-channel leaks. A similar flow (and argument) is found in SCRIPT [13]. However, transistor-level power simulation has been investigated as well to address specific side-channel leakage assessments with a limited scope in time and in design size [14]. More recently, power simulation for side-channel analysis is starting to appear in commercial tooling.

### 2.2 Identification of the leakage source

By simulating power with a structural model, it becomes feasible to identify the *source component* of side-channel leakage. In traditional measurement-based side-channel leakage analysis, this type of analysis is not possible because the design under test remains a black box. We review several recent proposals aimed at identifying the structural source of side-channel leakage, starting at low abstraction levels.

Karna partitions the gates of a design according to a spatial grid over the circuit layout [15]. A gate-level simulation leads to a power trace per grid cell. A leakage metric then ranks different cells according to their contribution to the side-channel leakage. The resolution of the Karna method depends on the granularity of the grid cells on the layout, since all logic cells within the same grid cell receive the same leakage score. The Karna authors include around 150 logic cells in one grid cell. Another tool, RTL-PSC, works at the register-transfer level of abstraction [16]. RTL-PSC analyzes the power in terms of transitions on the state variables of a design, while sub-cycle effects such as glitches and the effects of physical routing are abstracted out. The leakage metric ranks the different modules of a design. A similar register-transfer level analysis tool is PARAM [17], where
the authors identify the sources of side-channel leakage in a processor’s micro-architecture.

Another view on the problem of leakage source identification is to formally prove that a design meets a predefined side-channel leakage criteria. This technique works well for designs based on masking, a countermeasure based on secret-sharing. One example is Coco, which combines event-driven simulation with a SAT-solver-based verification of the statistical distribution of the secret shares [18]. Strictly speaking, these tools do not simulate the power consumption, but they verify the statistical properties of design activity.

Compared to this related work, ACA creates a ranking of the cells in the design according to their contribution to side-channel leakage, with a user-defined leakage metric. ACA can handle both specific and non-specific leakage testing. ACA is processor-independent as well as technology-independent. ACA builds a flow on commercially available synthesis and power simulation tooling. To the best of our knowledge, no such tool has been presented by earlier work.

3 Architecture Correlation Analysis

In this section, we outline the strategy of Architecture Correlation Analysis. The objective of ACA is to identify a ranking among the cells of a design according to their contribution to side-channel leakage. The cell ranking does not have to reflect the absolute level of side-channel leakage generated by a cell. Knowing just a relative ranking already provides critical insight into the parts of a design that are most prone to side-channel attacks.

Traditional side-channel leakage assessment uses the overall power consumption of a design to make an assessment on the global design. In contrast, ACA uses not only the overall power consumption but also the internal design construction details to make an assessment of leakage on a part of a design and to rank these local assessments. ACA uses gate-level power simulation in order to capture power events with sub-cycle accuracy as well as structural effects such as wire-loading and static leakage. The challenge of ACA is to perform such a gate-level side-channel leakage assessment with reasonable accuracy but without exhaustively generating the power consumption trace for each individual cell in the design.

Side-channel leakage assessment aims to minimize assumptions regarding the specific strength and know-how of the attacker. This leads to the use of specific and non-specific tests. A specific test for side-channel leakage uses a high-level power model that the adversary would presumably use in a Differential Power Analysis. A non-specific test uses two groups of inputs and aims to demonstrate a statistically distinguishable power consumption difference between those two groups. The Test Vector Leakage Assessment (TVLA) methodology provides guidance on the selection of the two groups of input vectors [19]. Both specific and non-specific tests have their merits and limitations. Specific tests are limited by specific assumptions on the capabilities and activities of the attacker, but provide specific assertions on the

---

1. We use the term cell over gate as it reflects better the technology encountered in standard-cell based IC design. A single cell often corresponds to multiple primary gates.
test vector and internal design state needs to be simulated only once. The second factor is that each simulation run can be isolated from the next by re-initializing the design in between simulation runs. In a typical side-channel leakage assessment, we gather between a few hundred and a few thousand simulated traces and we aim for a turn-around of between simulation runs. In a practical assessment of hardware, we over-sample the clock cycle at least several times, in order to run the analysis on sub-cycle events. However, with long-running, complex simulations, we have already successfully down-sampled the frame size to as much as 80 clock cycles per frame, while still being able to demonstrate side-channel leaks [20]. Successfully finding side-channel leaks in an under-sampled power trace is due to the averaged sampling technique employed by power simulation tools [21].

In Phase 2 and Phase 3 of the ACA flow, we aim to identify the cells’ individual contribution to side-channel leakage. The challenge is to complete this task using only the global power traces. Indeed, the generation of per-cell power traces has quadratic complexity (namely design_size x frame_count) and is therefore not scalable to large applications. We solve this with a two-step approach. First, using the global power traces, we identify the leakage time interval, the time window within which a design leaks information. Next, within the leakage time interval, we use the activity traces to identify the contribution of an individual cell to a design-level leak. Therefore, Phase 2 and Phase 3 of the ACA flow are defined as Leakage Time Interval Selection and Leakage Impact Factor Computation respectively. Leakage Time Interval Selection uses design-global power traces, while Leakage Impact Factor computation uses per-cell event traces.

To test individual cells for side-channel leakage, we can
use two different testing scenarios. In the following sections, we clarify each testing scenario separately.

### 3.2 ACA for Specific Testing

Fig. 4 describes the two steps to apply ACA using a specific test. The specific test requires the definition of a leakage model, similar to the leakage model used in Differential Power Analysis or Correlation Power Analysis. The leakage model is correlated with the simulated power consumption to identify a window of high correlation as the Leakage Time Interval (LTI), the time window during which a design leaks. Next, each individual cell is ranked by computing its correlation to the leakage model within the LTI.

#### 3.2.1 Leakage Time Interval for Specific Testing

Given a cipher which computes an internal result $V_k = f(K, C_k)$ with a partial key $K$ and a controlled input $C_k$ from $k$-th test vector, a possible leakage model $L(V_k)$ is the Hamming Weight $HW(V_k)$. Alternately, for an internal tuple $(V_{1k}, V_{2k}) = f(K, C_k)$, the leakage model can be expressed as the Hamming Distance $HD(V_{1k}, V_{2k})$. We then compute the correlation between the leakage model $L(V_k)$ or $L(V_{1k}, V_{2k})$ and the simulated power $P_k[n]$ for each frame $n$ of the simulated power traces:

$$\rho[n] = \frac{cov(L(V_k), P_k[n])}{\sigma_L \sigma_P}$$

(1)

The LTI consists of the set of frames for which the absolute correlation is above a given threshold.

$$LTI = \{n\} : \text{abs}(\rho[n]) > \rho_{\text{threshold}}$$

(2)

The choice of the threshold is a sensitivity parameter that must be chosen such that the LTI covers design activity that contains a likely correlation peak. Table 2 shows several examples of threshold levels as a function of the number of test vectors $m$ and the confidence level. As expected, a requirement for higher confidence or the use of fewer traces will increase the confidence interval, which means that stronger correlation peaks must be identified before the frame is flagged as leaky and added to the LTI.

| Confidence Level | $m=600$ | $m=1000$ | $m=2000$ |
|------------------|---------|---------|---------|
| 99%              | $\pm 0.105$ | $\pm 0.081$ | $\pm 0.058$ |
| 95%              | $\pm 0.180$ | $\pm 0.062$ | $\pm 0.044$ |
| 90%              | $\pm 0.067$ | $\pm 0.052$ | $\pm 0.037$ |

#### 3.2.2 Leakage Impact Factor for Specific Testing

Within the LTI, we compute the contribution of each individual cell to the leakage. This contribution is quantified in the Leakage Impact Factor (LIF), which is a dimensionless number that expresses the relative amount of side-channel leakage from a cell.

The LIF of a cell is computed as the correlation of cell output activity and leakage model activity. The LTI is a set of frames that are considered leaky (Fig. 5, left). To investigate the cell leakage, the LTI is superimposed over the activity trace. For each net (or each net-driving cell), we then compute a toggle trace as follows. When a given net switches during the LTI, then that transition is counted as a +1 toggle. When a given net does not include such a transition in the LTI, then that is counted as a -1 toggle. Hence, the activity of each net $i$ under test vector $j$ is converted to a bi-valued signal $K_{ij}$ with values $\{-1, +1\}$. To compute the architecture correlation $C_i$ of net $i$, $K_{ij}$ is multiplied with the toggle trace $H_j$ of the leakage model $L(V)$ (Fig. 5, right).

$$C_i = \sum_{\text{stimuli}} K_{ij} H_j$$

(3)

This correlation can be computed for every frame within the LTI. A high value in $C_i$ indicates a strong correlation between the cell activity and the leakage model, and hence a strong indication that the cell contributes side-channel leakage. All cells in the design are ranked according to their $C_i$ from most leaky to least leaky. We also include an additional weighing factor for each $C_i$, defined as the ratio of the cells’ average power consumption $P_i$ during the LTI over the total average power consumption of all gates $P_T$.

This increases the weight of high-drive cells with a high correlation. This leads to the weighted per-cell LIF $F_i$:

$$F_i = C_i \frac{P_i}{P_T}$$

(4)
3.3 ACA for Non-specific Testing

When a specific test is hard to apply, or when the number of leakage models $L(V)$ becomes too numerous, it may be helpful to apply a more generic non-specific test for leakage. We can run ACA using a non-specific leakage model following a strategy as in Fig. 6. Similar to TVLA, non-specific ACA requires the definition of two groups of stimuli. These two groups are used to show some systematic difference in the design behavior. For example, Goodwill et al. suggest AES test vectors that are random for group 1, while introducing a specific bias within a middle-round state for group 2. Using these two vector groups, ACA then follows the same two-step strategy as for specific testing. First, the LTI is computed to bound the leakage in time, and next the LIF per gate is computed. The testing statistic is adjusted to a non-specific test.

3.3.1 Leakage Time Interval for Non-Specific Testing

The test-statistic compares the distribution of power values at a specific frame between two test vector groups. One solution is to use a Welch-t statistic, which tests the difference between the mean values of both groups. Another solution is to use the leakage model to measure the correlation of the power value to the group number. We use the leakage model $N = groupid$, with $groupid$ equal to -1 for vectors from group 1, and +1 for vectors from group 2. With this leakage model, we can compute a non-specific test statistic as a correlation value that can be compared against a threshold value $p_{threshold}$. With $P[n]$ the power consumed during frame $n$, the correlation is given by:

$$\rho[n] = \frac{\text{cov}(N, P[n])}{\sigma_N \sigma_P}$$  \hspace{1cm} (5)

The LTI is then defined using the same method as in Equation 2.

3.3.2 Leakage Impact Factor for Non-Specific Testing

Once the LTI is fixed, we proceed with computing the LIF using a similar strategy as for the specific test. The LIF is superimposed over the activity trace of the design. A net transition during the LTI counts as a +1 toggle, and an absence of transition counts as a -1 toggle. We compute the architecture correlation $C_i$ of net $i$ by correlating the toggle trace $K_{ij}$ with the groupid $N_j$.

$$C_i = \sum_{stimuli} K_{ij} N_j$$  \hspace{1cm} (6)

Again, a high value in $C_i$ indicates a strong correlation between the cell activity and the non-specific leakage model, and hence flags the cell $C_i$ as leaky. To find a cell’s non-specific leakage impact factor $F_i$, a weighing factor is introduced as in Equation 4.

The advantage of the non-specific test over the specific test is that no high-level leakage model is necessary. For example, we have used non-specific tests based on one or more state bytes at a middle round being 0 or else random. In the experimental results, we will demonstrate the selectivity of both the specific as well as the non-specific ACA method.

3.4 Implementation

Our flow is fully realized in commercial tooling along with customized scripting to implement the statistical post-processing. We use Cadence Genus 20.1 for logic synthesis from RTL, Cadence XCellium 20.09 for functional simulation, and Cadence Joules 10.1 for gate-level power simulation. Computation of ACA LIF and LTI is scripted on top of the Jlsc toolbox. We have used Skywater 130nm standard cells as technology targets during experiments.

4 ACA on a Cryptographic Coprocessor

In this section, we describe the application of ACA on an AES encryption/decryption coprocessor. The architecture selected for analysis is typical for a medium-throughput accelerator residing in an embedded SoC. The coprocessor...
handles encryption and decryption and uses an offline key schedule, which computes the roundkeys once upon loading of the key. A single round takes 5 clock cycles. In the first four cycles, the coprocessor computes 16 Sbox lookups in sets of 4, and in the fifth clock cycle, the remainder of the round is computed. The encryption/decryption core is encapsulated by a bus interface which contains software-accessible registers and a controller. The bus interface handles various modes of operation for the coprocessor. We synthesized this coprocessor for SkyWater 130nm standard cell technology and a 50MHz clock. Table 3 reports the type and number of cells used in the design, as well as their relative area.

### 4.1 Architecture Correlation Analysis

Stimuli selection plays an important role in ACA, as it enables a designer to choose which part of a design will be exercised. In this analysis, we will focus on Architecture Correlation Analysis for a single key byte in the AES coprocessor. The objective of the ACA is to determine which cells, among the 9,640 cells in the design, contribute to side-channel leakage of this key byte. We explicitly differentiate this objective (finding leaky gates) from a more traditional side-channel analysis of the hardware. There is no doubt that there is side-channel leakage in this design. However, the object of this experiment is to find what cells are most responsible for this leakage?

#### 4.1.1 Results

We performed ACA as follows. We selected a set of 1024 vectors under a random plaintext and a constant key. Next, we ran a gate-level simulation and a gate-level power simulation for ACA. Fig. 8 shows the average power trace of the first round at 64 frames per clock cycle, as well as the (sign-flipped) standard deviation. The clarity of this power trace

| Type     | Cell Count | Area (%) |
|----------|------------|----------|
| Sequential | 2,479      | 51.8     |
| Logic    | 7,161      | 48.2     |
| Total    | 9,640      | 100.0    |

Table 3: Cell type and area for AES coprocessor

We emphasize that this CPA experiment only verifies the selection of leaky cells. ACA is not a countermeasure but a detection tool. One cannot remove an arbitrary cell from a netlist without substituting it with an equivalent cell with identical functionality. However, ACA is useful in conjunction with countermeasure tools that protect individual cells or subsets of cells, such as Karna [15] or STELLAR [22]. A second observation is that a power simulation that collects individual power traces for each cell for all vectors and all frames. The per-cell power traces of the 412 selected leaky cells are then subtracted from the overall power traces to construct a modified set of power traces. Next, we apply a Correlation Power Analysis (CPA) on the original trace set as well as on the modified trace set, with the results summarized in Fig. 9(left: unmodified set, right: modified set). For the modified set, the number of measurements to disclosure increases with a factor of 3.

We next ran ACA using a specific leakage model on the Hamming Weight of the SBOX output of the first key byte. We selected a specific leakage model on the Hamming Weight of the SBOX output of the first state byte. We identify the LTI with a \( \rho_{\text{threshold}} \) of 0.2, which flags 230 frames out of the 512 frames as containing potentially leaky samples. By correlating the transitions by cells within LTI with the specific leakage models, 412 cells are then flagged as leaky. This group represents 4.3 % of the total number of 9,640 cells. We will analyze the relation of these 412 cells to the overall AES coprocessor in subsection 4.2.

#### 4.1.2 Result Verification

The assertion made by ACA is that the side-channel leakage under the selected leakage model is primarily caused by these 412 cells. To verify the correctness of the selection, we performed the following verification. We re-ran the simulation while collecting individual power traces for each cell for all vectors and all frames. The per-cell power traces of the 412 selected leaky cells are then subtracted from the overall power traces to construct a modified set of power traces. Next, we apply a Correlation Power Analysis (CPA) on the original trace set as well as on the modified trace set, with the results summarized in Fig. 9(left: unmodified set, right: modified set). For the modified set, the number of measurements to disclosure increases with a factor of 3.

We emphasize that this CPA experiment only verifies the selection of leaky cells. ACA is not a countermeasure but a detection tool. One cannot remove an arbitrary cell from a netlist without substituting it with an equivalent cell with identical functionality. However, ACA is useful in conjunction with countermeasure tools that protect individual cells or subsets of cells, such as Karna [15] or STELLAR [22]. A second observation is that a power simulation that collects an individual power trace for every gate is extremely complex both in disk space and in time. We found the overhead of single-gate power tracing (compared to standard ACA) to be around 4 orders of magnitude in disk storage and one order of magnitude in simulation time, and worsening with design size. The high cost of per-gate power tracing
4.1.3 Impact of Frame Size

We also performed an ACA analysis at a frame size of 2 frames per cycle rather than 64 frames per cycle. Fig. 10 shows the effect on the power trace. At wide frame size, the power converges to the average of the smaller frame size. In our experiments, we found that a wide frame size is less precise to pick out leaky gates. For the same trace set as the previous experiment, the 2-frame-per-cycle version flags only 122 cells (as opposed to 412 cells) as leaky.

However, the use of a wider frame size may still have advantages. The 122 cells that are found at a wider frame size are a subset of the 412 cells found at a smaller frame size, with an exception of a single cell. Furthermore, there is a significant performance gain in power simulation time at wider frame sizes (Section 6). We can thus think of power simulations at wide frame sizes as a quick assessment to determine the LTI and to scan the overall properties of side-channel leakage in the design.

TABLE 4

Leaky Gate Identification for AES Coprocessor

| Module      | File                        | # Cells | Sequential |
|-------------|-----------------------------|---------|------------|
| Top-level   | aes_comp_core.v             | 5       | 0          |
| Decryption  | aes_comp_decipher_block.v   | 29      | 0          |
| Encryption  | aes_comp_encipher_block.v   | 204     | 26         |
| Keymem      | aes_comp_key_mem.v          | 1       | 0          |
| SBOX        | comp_sbox.v                 | 130     | 0          |
| Bus Interface | picaes.v                  | 22      | 10         |

The list of cells in Table 4 is intriguing. ACA is able to identify non-trivial leakage, often occurring as a result of the integration of cryptographic functions. The following example illustrates this point. In the Bus interface, the IV register is flagged as a source of leakage, which is unexpected. However, upon inspection of the code, it can be shown that the IV register senses every output value of the encryption module. Furthermore, due to the sequential nature of the computation, the encryption module reflects intermediate round values, including each individual SBOX output. The results in SBOX-related leakage appear at the IV register. The identification of individual RTL files and line numbers as leaky, based on a gate-level simulation, is an important debugging tool in the hands of the designer.

4.3 Non-specific ACA

We illustrate how ACA identifies leaky cells with a non-specific leakage model. In the following example, we create a non-specific test on the state variable in round 6 of the encryption. We use two sets of test vectors, and both contain random plaintext and key values. However, the second group contains specially selected (plaintext, key) pairs that create an all-zero round-6 state variable. Such pairs are easy to create: select a random key, and decrypt an all-zero state starting at round 4 of the decryption. The resulting plaintext is the sought starting value.

Using non-specific ACA we can identify the gates that are most affected by this bias, and thus the gates that are responsible for side-channel leakage. Fig. 11 illustrates the LTI on round 6, where the bias occurs. The majority of the frames (251 out of 385 in the trace) are flagged as part of LTI. Furthermore, after ranking the cells, we identify 2,812 unique cells as correlated with the round-6 state bias. This is much more than the 412 cells selected using a specific leakage model on the first key byte. However, this result is not unexpected: zero-forcing an entire state word (128 bits) where the expected value would be random is a very significant bias, which has an impact throughout the datapath.

Among the flagged leaky cells 2,498 gates were traced back to their RTL design files. Fig. 12 illustrates the rank of flagged leaky gates from each design file with rank=1 belonging to the leakiest gate in the design.
We caution that a strongly biased test, such as this all-
zero round-6 non-specific test, always results in aggressive
leaky cell selection. However, a weaker form of the test is
easy to define, for example by biasing only a single state
byte of round 6. The non-specific ACA test lets a user
evaluate the impact of an arbitrarily chosen bias in the
cipher.

5 ACA ON RISC-V BASED SoC

To investigate the scalability of ACA we also applied the
methodology on the SoC shown in Fig. 13. A 5-stage
pipelined RISC-V core (fetch, decode, execute, memory,
write-back) integrates a collection of peripherals including
the memory-mapped AES coprocessor discussed in Section 4.
In a typical access sequence, the RISC-V software
uploads a key and a block of plaintext to the coprocessor,
and then uses the control/status register of the coprocessor
to start the encryption and monitor the completion flag. The
RISC-V software then retrieves the ciphertext.

These are caused by pipeline stall operations on the RISC-V
processor (Fig. 14 lines 3, 5, 7, 9, 12, 14, 16, 18). Next, the
coprocessor control register is configured to run a single block encryption. The software
then goes into a polling loop waiting for the coprocessor to
complete operation, about 50 clock cycles later.

5.1 Architecture Correlation Analysis

In this test, we are investigating the hardware/software
interface between the RISC-V software and the AES copro-
cessor. Therefore, we apply ACA with a specific leakage
model using the Hamming Weight on the output of the pre-
whitening round. This will enable the monitoring of any
interactions between the plaintext and the key on the path
from software to the hardware coprocessor. Fig. 14 shows a
portion of the driver software. In a 32-bit architecture, a 128-
bit block is loaded using 4 consecutive memory-mapped
writes. The driver first loads 4 plaintext words, followed
by 4 key words. Next, the coprocessor control register is
configured to run a single block encryption. The software
then goes into a polling loop waiting for the coprocessor to
to complete operation, about 50 clock cycles later.

5.1.1 Results

We selected a set of 1024 vectors under a random plaintext
and a constant key. We then ran a gate-level simulation
and a gate-level power simulation at 5 frames per clock
cycle. Fig. 15 shows a sample power trace from the sim-
ulation. The testbench covers software activity as well as
hardware activity. The hardware activity uses considerably
more power than software because of the higher parallelism
of the hardware coprocessor implementation. However,
because of the noiseless simulation, the overall operation is
visible with remarkable clarity. The trace starts with the
software transmitting a key value, followed by a plaintext
value. Fig. 15 shows a series of 8 notches in the power
trace which correspond to reduced power consumption.
These are caused by pipeline stall operations on the RISC-V
processor. Fig. 14 shows a series of 8 notches in the power
trace which correspond to reduced power consumption.

Next, the software triggers the hardware AES execution, which runs
the key schedule followed by 10 rounds. Finally, the RISC-V
software retrieves the ciphertext.

TABLE 5
Leaky Gate Identification using non-specific round-6 state bias

| Module          | File             | # Cells | Sequential |
|-----------------|------------------|---------|------------|
| Top-level       | aes_comp_core.v  | 79      | 0          |
| Decryption      | aes_comp_decipher_block.v | 351 | 0          |
| Encryption      | aes_comp_encipher_block.v | 1172 | 128        |
| Keymem          | aes_comp_key_mem.v | 0       | 0          |
| SBOX            | comp_sbox.v      | 870     | 0          |
| Bus Interface   | picoaes.v        | 277     | 0          |

Fig. 12. Leaky gate ranks identified by non-specific test in ACA on AES
coprocessor sectioned into RTL design files.

![Fig. 12. Leaky gate ranks identified by non-specific test in ACA on AES coprocessor sectioned into RTL design files.](image)

Table 5 shows synthesis results for SkyWater
130nm standard cells at 50MHz clock. The overall design is
time three larger than the AES coprocessor by itself.

![Table 6](image)

| Type       | Cell Count | Area (%) |
|------------|------------|----------|
| Sequential | 8,091      | 51.5     |
| Logic      | 21,484     | 48.5     |
| Total      | 29,575     | 100.0    |

5.2 Leaky Gate Analysis

Fig. 16 shows the distribution of leaky frames over the
testbench. The leakage model is $HW(key \oplus pt)$. Remarkably,
the bulk of the leaky frames occurs during the software driver

activity, while the key is being loaded into the coprocessor. In addition, there is also leakage during the first round of the encryption, which is expected.

ACA demonstrates that the RISC-V processor, the memory bus, and the memory-mapped AES coprocessor are all potential contributors to side-channel leakage. The origin of such leakage is caused by the interaction of values over shared storage and interconnect. The leakage occurs in the processor micro-architecture because the driver software writes the key after the plaintext. There is a minor hint of this issue in the software driver itself. The first key byte is loaded in register $a3$, which still contains a portion of plaintext. This leads to distance-based leakage conforming to the leakage model.

Table 7 shows the distribution of the 1,298 leaky cells over the design. There are indeed a large number of leaky gates located within the RISC-V processor. We analyze two examples below.

The highest-ranked leaky gate (with a correlation of 0.463) in the SoC is the pipeline register of the memory stage which, in its data-path components, transfers the contents of the second source register and the result of the ALU from the execute to the memory stage. Even for instructions that do not need ALU operation, the ALU result is written with the addition of the two source operands. Therefore the same transitional leakage discussed for $a3$ register can occur for the ALU result register.

The leakage from the peripherals is caused by the transmission of key and plaintext on the memory interface. At the connection point of each peripheral module to the memory interface there are multiplexers to decide whether the transmitted data should be admitted to the current peripheral. Such interconnect logic can manifest Hamming weight leakage of the plaintext and key.

These examples demonstrate that ACA is a powerful debugging tool, as it can highlight side-channel leakage of the gate-level implementation at the RTL level.
6 ACA PERFORMANCE CONSIDERATIONS

ACA adds a new design step to the overall design flow, and thus the cost of running ACA in comparison to other tools in the design flow must be considered. Table 8 summarizes the runtime performance of ACA analysis. There are three major components that consume the bulk of the execution time: logic synthesis, functional gate-level simulation, and gate-level power simulation. The ACA correlation component is minor and typically takes less than a minute to complete. Overall, we observe that gate-level power simulation is a dominant factor that is more complex than gate-level synthesis and gate-level simulation. The overall runtime is strongly affected by the design size and the total number of frames per trace. On the plus side, the power simulation step is embarrassingly parallel. Each test vector can be run independently from the other. In our experiments, we did not use any parallel execution.

7 CONCLUSIONS

Gate-level leakage assessment is a tool that supports a designer to identify leaky gates in a pre-silicon design context. Our methodology relies on industry-standard tools including logic synthesis, gate-level simulation, and gate-level power estimation, together with scripting on the intermediate results. Architecture Correlation Analysis, the underlying detection technique to support gate-level leakage assessment, can serve as a verification technique as well as a basis for countermeasure design. In particular, by moving the leaky gates flagged by ACA into a separate power domain, a low-cost countermeasure may be enabled that requires only selective replacement of cells in a design.

REFERENCES

[1] S. Nikova, C. Rechberger, and V. Rijmen, “Threshold implementations against side-channel attacks and glitches,” in International conference on information and communications security. Springer, 2006, pp. 529–545.

[2] A. Moradi, “Side-channel leakage through static power - should we care about in practice?" in Cryptographic Hardware and Embedded Systems - CHES 2014 - 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings, ser. Lecture Notes in Computer Science, L. Batina and M. Robshaw, Eds. vol. 8731. Springer, 2014, pp. 562–579. [Online]. Available: https://doi.org/10.1007/978-3-662-44709-3_31

[3] Z. Chen, S. Haider, and P. Schaumont, “Side-channel leakage in masked circuits caused by higher-order circuit effects,” in Advances in Information Security and Assurance, Third International Conference and Workshops, ISA 2009, Seoul, Korea, June 25-27, 2009. Proceedings, ser. Lecture Notes in Computer Science, J. H. Park, H. Chen, M. Atiquzzaman, C. Lee, T. Kim, and S. Yeo, Eds., vol. 5576. Springer, 2009, pp. 327–336. [Online]. Available: https://doi.org/10.1007/978-3-642-02617-1_34

[4] I. Giechaskiel, K. Eguro, and K. B. Rasmussen, “Leakier wires: Exploiting fpga long wires for covert- and side-channel attacks," ACM Trans. Reconfigurable Technol. Syst., vol. 12, no. 3, aug 2019. [Online]. Available: https://doi.org/10.1145/3224853

[5] Y. Yao, T. Kathuria, B. Ege, and P. Schaumont, “Architecture correlation analysis at register-transfer level," in 2020 IEEE International Symposium on Hardware Oriented Security and Trust (HOST). IEEE, 2020, pp. 188–196.

[6] Y. Yao, T. Tufan, T. Kathuria, B. Ege, U. Guler, and P. Schaumont, “Pre-silicon architecture correlation analysis (paca): Identifying and mitigating the source of side-channel leakage at gate-level," in CACR Cryptol. ePrint Arch., vol. 2021, p. 530, 2021.

[7] I. Buhari, L. Batina, Y. Yarom, and P. Schaumont, “SoK: Design tools for side-channel-aware implementations," Cryptology ePrint Archive, Report 2021/497, 2021, https://ia.cr/2021/497

[8] J. den Hartog, J. Verschuren, E. P. de Vink, J. de Vos, and W. Wiersma, “PINPAS: a tool for power analysis of smartcards," in SEC, 2003, pp. 453–457.

[9] D. McCann, E. Oswald, and C. Whitnall, “Towards practical tools for low-cost software engineering: grey box modelling for instruction leakages," in USENIX Security symposium, 2017, pp. 199–216.

[10] M. A. Shelton, N. Samwel, L. Batina, F. Regazzoni, M. Wagner, and Y. Yarom, “Rosita: Towards automatic elimination of power-analysis leakage in ciphers," in NDSS, 2021.

[11] N. Sehatbakhsh, B. B. Yilmaz, A. G. Zajic, and M. Prvulovic, “EMSim: A microarchitectural-level simulation of electromagnetic side-channel signals," in HPCA, 2020, pp. 71–85.

[12] D. Sijacic, J. Balasch, B. Yang, S. Ghosh, and I. Verbauwhede, “Towards efficient and automated side-channel evaluations at design time," J. Cryptogr. Eng., vol. 10, no. 4, pp. 305–319, 2020.

[13] A. Nahiyan, J. Park, M. He, Y. Iskander, F. Farahmandi, D. Forte, and M. Tehraniipoor, “Script: A cad framework for power side-channel vulnerability assessment using information flow tracking and pattern generation," ACM Trans. Des. Autom. Electron. Syst., vol. 25, no. 3, may 2020. [Online]. Available: https://doi.org/10.1145/3383445

[14] F. Regazzoni, S. Badel, T. Eisenbarth, J. Groschädl, A. Poschmann, Z. T. Deniz, M. Macchetti, L. Pozzi, C. Paar, Y. Leblebici, and P. Ienne, “A simulation-based methodology for evaluating the dpab-resistance of cryptographic functional units with application to CMOS and MCM technologies," in Proceedings of the 2007 International Conference on Embedded Computer Systems: Architectures, Modeling and Simulation (IC-SAMOS 2007), Samos, Greece, July 16-19, 2007, pp. 209–214. [Online]. Available: https://doi.org/10.1109/ICSMOS.2007.4285753

[15] P. SLPSK, P. K. Vairam, C. Rebeiro, and K. Veezhinathan, “Karna: A gate-sizing based security aware eda flow for improved power side-channel attack protection," in Proceedings of the International Conference on Computer-Aided Design, ICCAD 2019, Westminster, CO, USA, November 04-07, 2019.

[16] M. T. He, J. Park, A. Nahiyan, A. Vassilev, Y. Jin, and M. M. Tehraniipoor, “RTL-PSL: automated power side-channel leakage assessment at register-transfer level," in VTS, 2019, pp. 1–6.

[17] M. A. K. F. V. Ganesan, R. Boddu, and C. Rebeiro, “PARAM: A microprocessor hardened for power side-channel attack resistance," in 2020 IEEE International Symposium on Hardware Oriented Security and HOST, 2020, San Jose, CA, USA, December 7-11, 2020. IEEE, 2020, pp. 23–34. [Online]. Available: https://doi.org/10.1109/HOST45689.2020.9300263

[18] B. Gigerl, V. Hadzic, R. Primas, S. Mangard, and R. Bloem, “Coco: Co-design and co-verification of masked software implementations on cpus," in 30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021, M. Bailey and R. Greensstadt, Eds. USENIX Association, 2021, pp. 1469–1468. [Online]. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/gigerl

[19] G. Goodwill, B. Jun, J. Jaffe, P. Rohatgi et al., “A testing methodology for side-channel resistance validation," in NIST non-invasive attack testing workshop, 2013, pp. 315–323. [Online]. Available: https://doi.org/10.1109/HOST45689.2020.9300263

[20] P. Kiaei, Z. Liu, R. K. Eren, Y. Yao, and P. Schaumont, “Saidoyoki: Evaluating side-channel leakage in pre- and post-silicon setting," Cryptology ePrint Archive, Report 2021/1235, 2021, https://ia.cr/2021/1235

[21] P. Kiaei, Z. Liu, and P. Schaumont, “Leverage the average: Aver-
Pantea Kiaei (Student Member, IEEE) is a Ph.D. student in Electrical and Computer Engineering at Worcester Polytechnic Institute. She received her MS degree in Computer Engineering from Virginia Tech in 2019 and prior to that received her BS degree in Electrical Engineering from Sharif University of Technology, Iran, in 2017. She has reviewed papers for ACM TECS, ACM JETC, and IEEE TVLSI journals. Her research interests include secure hardware design, computer architecture, and trustworthy secure systems.

Yuan Yao received her bachelor’s degree in Electronic Engineering from Northwestern Polytechnical University, Xi’an, China in 2014. She got her Master’s Degree in Electrical and Computer Engineering from Cornell University, Ithaca, US in 2016. Currently she is a Ph.D. candidate at the Bradley Department of Electrical and Computer Engineering, Virginia Tech. She serves as reviewer for several IEEE and ACM journals. Her research area include pre-silicon side-channel analysis, side-channel attacks and countermeasures, fault attacks and countermeasures, secure hardware design.

Zhenyuan Liu (Student Member, IEEE) is a Ph.D. student in Electrical and Computer Engineering at Worcester Polytechnic Institute. She received her MS degree in Electrical and Computer Engineering from Worcester Polytechnic Institute in 2020 and prior to that received her BS degree in Engineering of Science from Trinity University, San Antonio, Texas, in 2019. Her research interests include side-channel attacks, leakage assessments and micro-architectural hardware security.

Nicole Fern is a Senior Security Analyst at Riscure. She received her undergraduate degree in Electrical Engineering from The Cooper Union for the Advancement of Science and Art (2011) and her PhD degree in Electrical & Computer Engineering from University of California, Santa Barbara (2016). She continued her research in hardware security as a post-doc before joining industry in 2018. She previously worked at Tortuga Logic, a hardware security startup, for 2.5 years before joining Riscure in June of 2021.

Cees-Bart Breunesse is a principal security analyst at Riscure North America. He received his Ph.D. in Computer Science from Radboud University, Netherlands, in 2005. Prior to that, he received his MSc degree in Computer Science in 2000 from Utrecht University, Netherlands. His research interests include security of embedded devices, side-channel, and fault injection.

Jasper van Woudenberg (@jzvw) currently is CTO for Riscure North America and half of the authors of the “Hardware Hacking Handbook: Breaking Embedded Security with Hardware Attacks”. He works with Riscure’s San Francisco based team to improve embedded device security through innovation.

In the past, Jasper worked for a penetration testing firm, where he performed source code review, binary reverse engineering and tested application and network security.

At Riscure, Jasper’s expertise has grown to include various aspects of hardware security; from design review and logical testing, to side-channel analysis and perturbation attacks.

Jasper has spoken at many security conferences including BlackHat briefings and trainings, Intel Security Conference, RWC, RSA, EDSC, BSides SF, Shakacon, IMC, Infiltrate, has presented scientific research at SAC, WISSEC, CT-RSA, FDT, ESC Design West-East, ARM Tech-Con, has reviewed papers for CHES and JCRYPTOJEN, and has given invited talks at Stanford, NPS, GMU and the University of Amsterdam.

Kate Gillis received a B.S. in Electrical and Computer Engineering from Worcester Polytechnic Institute (WPI) in 2016. She has been at Intrinsix Corp. in Marlborough, MA since the same year, where she is currently a Verification Engineer and has worked to design and implement block-level, SoC and mixed signal verification on a variety of ASIC designs. Her awards and honors include the Provost’s MQP (senior capstone) Award and the Salisbury Prize, both at WPI.

Alex Dich received a B.S. in electrical and computer engineering from Worcester Polytechnic Institute, Worcester, MA in 2014. He has been at Intrinsix Corp., Marlborough, MA since 2014, where he is a Senior Design Engineer focused on front-end digital ASIC design. His interests include digital signal processing, cryptography, hardware acceleration, and processor architecture.

Peter J. Grossmann received a B.S. in engineering from Harvey Mudd College, Claremont, CA in 2001, an M.S. in electrical engineering from the University of Washington, Seattle, WA in 2006, and a PhD in computer engineering from Northeastern University, Boston, MA in 2013. He has been at Intrinsix Corp., Marlborough, MA since 2019, where he is currently a Solutions Architect focused on advanced research. He worked as Associate Staff and Technical Staff at MIT Lincoln Laboratory from 2007 to 2019, and as an ASIC Design Engineer at Zilog, Inc. from 2001 to 2004. His research interests include architecture and CAD for field programmable gate arrays, low power digital circuit design, and enhancing electronic design automation flows for security. Dr. Grossmann’s awards and honors include the 2017 MIT Lincoln Laboratory Best Invention award and the University of Washington Top Scholar award.

Patrick Schaumont (Senior Member, IEEE) is a Professor in Computer Engineering at WPI. He received the Ph.D. degree in Electrical Engineering from UCLA in 2004 and the MS degree in Computer Science from Ghent University in 1990. He was a staff researcher at IMEC, Belgium from 1992 to 2000. He was a faculty member with Virginia Tech from 2005 to 2019. He joined WPI in 2020. He was a visiting researcher at the National Institute of Information and Telecommunications Technology (NICT), Japan in 2014. He was a visiting researcher at Laboratoire d’Informatique de Paris 6 in Paris, France in 2018. He is a Radboud Excellence Initiative Visiting Faculty with Radboud University, Netherlands from 2020. His research interests are in design and design methods of secure, efficient and real-time embedded computing systems.