Exploring the Effect of Device Aging on Static Power Analysis Attacks

Naghmeh Karimi\textsuperscript{1}, Thorben Moos\textsuperscript{2} and Amir Moradi\textsuperscript{2}

\textsuperscript{1} University of Maryland, Baltimore County, USA
naghmeh.karimi@umbc.edu
\textsuperscript{2} Ruhr University Bochum, Horst Görtz Institute for IT Security, Germany
firstname.lastname@rub.de

Abstract. Vulnerability of cryptographic devices to side-channel analysis attacks, and in particular power analysis attacks has been extensively studied in the recent years. Among them, static power analysis attacks have become relevant with moving towards smaller technology nodes for which the static power is comparable to the dynamic power of a chip, or even dominant in future technology generations. The magnitude of the static power of a chip depends on the physical characteristics of transistors (e.g., the dimensions) as well as operating conditions (e.g., the temperature) and the electrical specifications such as the threshold voltage. In fact, the electrical specifications of transistors deviate from their originally intended ones during device lifetime due to aging mechanisms. Although device aging has been extensively investigated from reliability point of view, the impact of aging on the security of devices, and in particular on the vulnerability of devices to power analysis attacks are yet to be considered.

This paper fills the gap and investigates how device aging can affect the susceptibility of a chip exposed to static power analysis attacks. To this end, we conduct both, simulation and practical experiments on real silicon. The experimental results are extracted from a realization of the PRESENT cipher fabricated using a 65 nm commercial standard cell library. The results show that the amount of exploitable leakage through the static power consumption as a side channel is reduced when the device is aged. This can be considered as a positive development which can (even slightly) harden such static power analysis attacks. Additionally, this result is of great interest to static power side-channel adversaries since state-of-the-art leakage current measurements are conducted over long time periods under increased working temperatures and supply voltages to amplify the exploitable information, which certainly fuels aging-related device degradation.

Keywords: Leakage Current · Static Leakage Analysis · Side-Channel Analysis · Device Aging

1 Introduction

Two decades after its introduction to the public domain \cite{KIJ99} and in view of numerous contributions of the scientific community towards a better understanding of its sources and mitigation mechanisms, side-channel analysis (SCA) attacks are known as a serious threat to devices which deal with cryptographic primitives. It can be considered as general knowledge that a cryptographic device, where a secret is stored and processed, is vulnerable to SCA attacks if it is not equipped with dedicated countermeasures. Amongst such attacks, power analysis attacks have attracted more attention due to their simplicity and high efficiency to recover the secrets.
Until recently, the dynamic power consumption has dominated the power behavior of CMOS integrated circuits and was, thus, the main source for their SCA leakage. Not surprisingly, almost the entire body of SCA research from the last 20 years, which consists of several 100s of publications, is based on the dynamic power consumption of the underlying circuit. However, through the continuous shrinkage of semiconductor technology, the static power consumption of CMOS circuits is becoming a major concern, while the dynamic power consumption (per logic unit) is decreasing due to smaller capacitances, shorter rising and falling times, and smaller supply voltages. Hence, it can be observed that in newer generations of CMOS devices the static power behavior contributes an increasingly large share to the overall power consumption [KAB+03, Hel09, WERR13, Sha12]. As a consequence, SCA attacks based on dynamic power will become increasingly more difficult, which can be viewed as a positive development. Therefore, in a couple of works it has been attempted to estimate the feasibility of SCA attacks through the static power consumption by means of transistor-level simulations [GSST07, LB08, AGST09, AGST10]. It is noteworthy, that such SCA attacks are sometimes referred to as \textit{leakage current} power analysis attacks. Afterwards, such experiments have been conducted in practice demonstrating successful key recoveries by means of the static power side channel [Mor14, PSKM15] under certain conditions. Further, the effectiveness of a masking countermeasure [NRS11] to mitigate dynamic power analysis attacks is practically compared to those based on static power [MMR17].

In contrast to traditional power analysis attacks, measuring the static power consumption requires a sophisticated setup and faces several engineering challenges. For example, the adversary should have control over the clock signal to force the target device into an idle state (i.e., no activity or change in the state of the circuit) to be able to measure the DC shift of its current consumption, i.e., its leakage current. This process also needs to be performed in a temperature-controlled environment, i.e., a heating chamber, since the leakage current is extremely sensitive to temperature variations. In addition, due to its very low amplitude, the leakage current signal needs to be amplified with a high gain and low-pass filtered to achieve measurements suitable for SCA attacks. Recently, a study on the effect of various elements of a measurement setup on the success of such SCA attacks has been published in [MMR18].

From another perspective, with the rapid scaling of process technology, aging-related degradation of integrated circuits has become one of the main challenges in nano technologies. Due to aging, electrical behavior of transistors deviates from the originally intended one, leading to performance degradation in the underlying device, and ultimate device failure [SKR+13, KHHC11]. Generally speaking, device aging leads to an increase of the threshold voltage of the transistors over time. Therefore, the gates exhibit longer propagation delays compared to their original state. Further, recent works show that the static power of a CMOS circuit reduces by aging. This has been examined by transistor-level simulation [RTY+17]. Although aging mechanisms and related mitigation schemes have received the lion’s share of attention from reliability perspective in recent years [LK11, EKD+03], their impact on the security of devices, in particular cryptographic devices, is yet to be investigated.

\textbf{Our Contribution.} In this work we examine the effect of aging on the exploitability of information leakage through the static power consumption of cryptographic devices as a side channel. We conduct both, simulations and a practical analysis, to study such an effect. Our practical investigations are based on analyses conducted on two samples of an ASIC chip which we have designed and fabricated in a 65 nm technology node, as well as on an available prototype in 150 nm technology. For transistor-level simulations of the 65 nm ASICs we model the post-layout netlist of the target core by means of SPICE models provided by the foundry that manufactured the ASIC prototype.
It is noteworthy, that in order to observe the effect of aging, we followed an aging-acceleration process by operating the device at a high temperature, a high supply voltage, and a dynamic workload [JED16]. Our analyses demonstrate that, although the concrete nature of the data dependency may shift due to aging, the amount of information leaked through the static power consumption is reduced when the device is aged. This highlights a common issue in static power analysis attacks. Since the amount of leakage current is increased in higher temperatures and for higher supply voltages, static power analysis attacks are usually conducted while the device is operated at a high temperature, e.g., 90°C, and an increased supply voltage, which leads to more easily exploitable leakages, i.e., lower number of traces for successful attacks [BCS+17, DBST17, MMR18]. However, during such a special measurement condition the device is aged faster, and as result of our research the exploitability of its leakage current is steadily decreased. This causes a mismatch between the samples collected at different measurement phases. Of course, this is only of concern if the measurement process takes a long time, e.g., at least a couple of weeks.

2 Preliminaries

In this section, we first present a background on device aging and its effect on circuit's characteristics, followed by a discussion on how aging affects the leakage currents in modern CMOS technology. Afterwards, we express the design architecture of the case study which we consider in our investigations.

2.1 Device Aging

Device aging results in performance degradation and eventual failure of digital circuits over time [Kim15]. Aging mechanisms include Negative Bias Temperature-Instability (NBTI), Positive Bias Temperature-Instability (PBTI), Hot Carrier Injection (HCI), Time Dependent Dielectric Breakdown (TDDDB), and Electro-Migration (EM).

In practice, BTI (including NBTI and PBTI effects) is one of the major causes of threshold voltage increase in transistors during their lifetime. NBTI and PBTI occur in PMOS and NMOS transistors, respectively. In practice, the impact of NBTI is more dominant than PBTI beyond the 45nm technology node. However, with the introduction of high-k gate dielectrics and metal gate transistors, PBTI effects have also received significant attention [ZKN+06, CPC+05]. NBTI occurs in a PMOS transistor when a negative voltage is applied to its gate. In this mechanism, positive interface traps are generated at the Si-SiO$_2$ interface. As a result, the threshold voltage increases and the PMOS transistor becomes slower and fails to meet timing constraints. In contrast to NBTI, PBTI occurs when a positive voltage is applied to the gate of an NMOS transistor. This results in generating traps at the interface of gate oxide and channel, and in turn increasing the transistor threshold voltage.

HCl occurs when hot carriers are injected into the gate dielectric during transistor switching and remain there. HCI mainly affects NMOS transistors, degrades the underlying circuit by shifting the threshold voltage and the drain current of transistors under stress [RFFT14a]. TDDDB relates to the creation of an electrical current conduction path through the gate oxide in the device-under-stress. It degrades the isolation properties of gate dielectric, increasing the tunneling current across the transistor gate terminal, and ultimately results in device breakdown [NBRR13]. On the other hand, high density currents result in EM aging. The currents create electron winds that cause metal atoms to migrate over time, gradually removing metal atoms from wires, thereby increasing interconnect resistance, and eventually resulting in an open circuit [Miz08].
Among all aging mechanisms, BTI and HCI are two leading factors in degradation of digital circuits [KDLG16]. Both mechanisms result in increasing switching and path delays in the circuit under stress [KGD18, KDG18]. What follows discusses these aging mechanisms in more detail.

### 2.1.1 NBTI Aging

NBTI affects a PMOS transistor when a negative voltage (i.e., \( V_{gs} < V_t \)) is applied to its gate. In fact, a PMOS transistor experiences two phases of NBTI depending on its operating condition. The first phase, so-called stress phase, occurs when the transistor is on, i.e., when a negative voltage \( (V_{gs} < V_t) \) is applied to its gate. In this case, positive interface traps are generated at the Si-SiO\(_2\) interface which lead to an increase of the threshold voltage of the transistor. The second phase, so-called recovery phase, occurs when a positive voltage \( (V_{gs} > V_t) \) is applied to the PMOS transistor's gate. As a result, the threshold voltage drift that occurred during the stress phase will partially recover.

![Figure 1: Threshold voltage shift of a PMOS transistor under NBTI effect.](image)

Threshold voltage drifts depend on the physical parameters of the transistor, supply voltage, temperature, and stress time [AKVM07, KCC+05]. The last three parameters (so-called external parameters) are generally used as acceleration factors of aging process. Figure 1 shows the threshold voltage drift of a PMOS transistor that is continuously under stress for 6 months and a PMOS transistor that alternates stress/recovery phases every other month. As shown, NBTI effect is high in the first couple of months but the threshold voltage tends to saturate for long stress times. The impact is exacerbated with thinner gate oxide and higher operating temperature [AKVM07, MSVK06].

Two prevalent theories, Reaction-Diffusion (R-D) and Trapping-Detrapping (T-D), have been proposed in literature to explain NBTI. The R-D model explains the NBTI phenomenon as the breaking and rebonding of hydrogen-silicon bonds at the silicon-gate dielectric interface of PMOS devices [Sch07, CCLM14]. The T-D model considers a number of defect states with different energy levels, and capture and emission time constants. In the T-D model, the threshold voltage increases when a trap captures a charge carrier from the channel of a PMOS device [SVRC15].

According to the R-D model proposed in [WYB+10], the NBTI-related increase in the threshold voltage of a PMOS transistor in the stress phase is evaluated as follows [TT08],

\[
\Delta V_{th_{stress}} = A_{\text{NBTI}} \cdot t_{ox} \cdot \sqrt{C_{ox}(V_{dd} - V_{th})} \cdot e^{\frac{V_{dd} - V_{th}}{E_0} - \frac{K}{E_0}} \cdot t_{ox}^{0.25},
\]

where \( t_{ox} \) is the oxide thickness, and \( C_{ox} \) the gate capacitance per unit area. The constants \( E_0 \) and \( E_u \) stand for device-dependent parameters, \( A_{\text{NBTI}} \) is a technology-dependent

---

3 The Y axis has not been shown to make the graph generic for different technologies.
constant, and $k$ the Boltzmann constant. $T$ represents the temperature, and $t_{st}$ the stress time.

As discussed, the threshold voltage drift of a PMOS transistor is partially recovered if the transistor is placed in the recovery phase. The following equation expresses the final change in the threshold voltage of a PMOS transistor [TT08].

$$
\Delta V_{th_{NBTI}} = \Delta V_{th_{0}} \times (1 - \sqrt{\frac{t_{rec}}{t_{rec} + t_{st}}}),
$$

where $\eta$ is equal to 0.35, and $t_{st}$ and $t_{rec}$ represent the stress and recovery time durations, respectively.

2.1.2 HCI Aging

Hot carriers refer to the electrons or holes in the substrate that attain energies above the average [RRFT16a]. These high energetic carriers, which are the result of high electric fields in the drain region of a transistor are injected into the gate oxide and form interface states and eventually result in performance degradation in the transistor under stress. HCI mainly affects NMOS transistors and has become more severe as the transistor features continue to shrink [CPS08].

HCI results in the change of the threshold voltage of the device under stress. Besides increasing the threshold voltage, HCI reduces the mobility of a device, which leads to a decrease in drain current. Unlike NBTI, there is no recovery for HCI.

HCI effect is due to the switching between ‘0’ and ‘1’ on an NMOS transistor. Thereby, HCI is highly sensitive to the number of transitions that occur in the gate input of the transistor under stress. In fact, the threshold voltage changes sublinearly with the number of transitions that occur in the input of an NMOS transistor. In practice, HCI has a sublinear dependency on the clock frequency, usage time, and the activity factor of the transistor under stress, where activity factor represents the ratio between the number of cycles the transistor is doing transitions and the total number of cycles the device is utilized. In addition, HCI effects depend on the operating temperature [OT12]. The equation given below evaluates the HCI-induced threshold voltage shift [WYB+10, TT08],

$$
\Delta V_{th_{HCI}} = A_{HCI} \cdot \alpha \cdot f \cdot \frac{V_{dd} - V_{th}}{t_{ox} \cdot E_{1} \cdot t_{ox}^{0.5}},
$$

where $t$ stands for time, $\alpha$ and $f$ for the activity factor and the frequency, respectively. In addition, $t_{ox}$ is the oxide thickness, and $E_{1}$ depends on the device specifications, the temperature, and $V_{dd}$. Further, $A_{HCI}$ is a technology-dependent constant.

As expressed with more details in Section 3, in order to extract the simulation results we deployed HS-MOSRA (MOS Reliability Analysis) [Syn16] to evaluate the impact of NBTI and HCI on a circuit under stress. MOSRA uses the Reaction-Diffusion (R-D) model discussed in [WYB+10].

2.2 Leakage Currents in MOSFETs

Metal-oxide-semiconductor field-effect transistors (MOSFETs) are by far the most common transistors in digital circuits and have experienced an aggressive scaling process during the last decades, mainly fueled by two forces. First, the increasing demand to put more computation capability onto a single chip, second the potential enablement of faster computation. Indeed, decreasing channel lengths not only reduces area, but also allows electrons (or holes) to pass through transistors faster, and decreasing distances between gates minimize the cell-to-cell travel time [Key05]. Also, due to the shrinkage of transistor and cell dimensions, capacitances decrease in size and, thus, can be charged faster. Besides improving the overall capability and speed of integrated circuits, a further implication
of the scaling process is a decrease in the dynamic power dissipation per logic unit due to smaller capacitances, shorter signal propagation delays and the allowance for lower supply voltages in general [Key03]. However, all these improvements come at the price of undesired side-effects like, for example, the static power consumption, a.k.a. leakage current.

Over the years it has become more and more apparent that nanometer-sized MOSFETs conduct a significant current between \( V_{dd} \) and ground even when being switched off. This current, which is often called leakage current or off-current, is not only limited to drain to source leakage, but can occur as leakage through the gate insulator and leakage from drain to bulk and source to bulk as well [Hel09]. Multiple physical effects responsible for these phenomena have been pointed out and investigated in the literature, e.g. subthreshold leakage, drain-induced barrier lowering (DIBL, a.k.a. punchthrough current), gate tunneling, hot carrier effects, diode currents and gate-induced drain leakage (GIDL) [Hel09]. For our investigations mainly the subthreshold conduction is of interest, since it constitutes the dominating source of drain to source leakage and exhibits an exponential dependency on the threshold voltage of transistors. An increase in the threshold voltage reduces undesired effects like the subthreshold leakage, but also degrades the performance of transistors in terms of signal propagation delays. A decrease in the threshold voltage on the other hand has the opposite effect, namely allowing transistors to switch faster while having a higher static power consumption. This trade-off between speed and average power consumption of transistors has fueled the introduction of multi-threshold voltage cell libraries [Hel09]. Indeed, logic cells in advanced technologies usually exist in multiple versions composed of transistors with either low, high or medium voltage thresholds. For signal delay optimizations in the critical path of a circuit cell with a low voltage threshold (LVT) can be selected, while in all paths of a design which are not timing critical, cells with a high voltage threshold (HVT) are chosen to minimize the overall leakage.

### 2.3 Data-Dependent Leakage Currents in CMOS Logic

CMOS logic gates are constructed in such a way that there exists at least one switched-off MOSFET (i.e., the transistor is not forming a conducting channel between source and drain) in any path between \( V_{dd} \) and ground during an idle state. This design decision was made in order to ensure that such logic gates only draw a significant current during switching behavior, enabling low power circuitry. However, due to the currents passing through individual inactive, i.e., switched-off, MOSFETs, there is also a perceivable current leaked by CMOS logic gates constructed from them. The magnitude of the leakage current exhibited by a CMOS logic cell depends on the type and formation of switched-off MOSFETs in the path between \( V_{dd} \) and ground and the different electric potentials across them. For example, considering a simple CMOS \texttt{NOR}, i.e., inverter, gate such as depicted in Figure 2, it can be observed that depending on the input signal either the NMOS or the PMOS transistor is inactive. When assuming that the active, i.e., switched on, transistor creates a perfect conduction channel (which basically means replacing them by ideal wires), one obtains the resulting schematics on the right side of Figure 2 for different inputs. Usually, PMOS and NMOS transistors have different leakage currents, leading to the fact that the static power consumption of this \texttt{NOR} gate depends on its input, i.e., on the processed data. This dependency becomes even more obvious in gates with multiple input lines, such as the two-input \texttt{NAND} gate in Figure 3. Here, four different formations of inactive transistors can be observed, depending on the input. Clearly, these four cases lead to different leakage currents exhibited by the \texttt{NAND} gate. Comparing the two cases (\( A=0, B=0 \) and \( A=0, B=1 \)), for example, it is obvious that the former, where two switched-off NMOS transistors are connected in series, has a significantly smaller leakage current than the latter. In particular, connecting inactive transistors in series causes a so-called stacking
Figure 2: CMOS NOT (or inverter) gate (left) and formation of inactive transistors across the power supply path for different inputs (right), assuming perfect conduction for active transistors.

Figure 3: CMOS AND gate (left) and formation of inactive transistors across the power supply path for different inputs (right), assuming perfect conduction for active transistors.

effect [RMMM03]. This effect reduces the current flowing through a stack of two inactive transistors by one order of magnitude compared to a single inactive one [RMMM03]. For this reason, transistor stacking is often used as a leakage current mitigation technique. The largest current is leaked by a CMOS AND gate, however, when the input combination \((A=1, B=1)\) is applied, since two inactive PMOS transistors, connected in parallel, are present between \(V_{dd}\) and ground, whose individual leakage currents cumulate.

Of course, assuming perfect conduction for the active transistors in the previous examples is meant as a simplification. In practice, even the two cases \((A=0, B=1)\) and \((A=1, B=0)\), which look identical in Figure 3, would lead to different leakage currents, caused by the difference in the electric potentials across them\(^1\). Thus, in reality the data dependency is even stronger than in the simplified scenario.

\(^1\)This depends for instance on whether a terminal is pulled up/down by another transistor or not.
2.4 Impact of Device Aging on Leakage Currents in CMOS Circuits

As described before, the magnitude of the threshold voltage of a transistor is increased when this transistor is subject to an aging process, due to effects such as BTI and HCl. Since any increase in the threshold voltage of transistors translates to a decrease of the subthreshold currents in the device under test, it has to be expected that aging affects the leakage currents in CMOS devices quite significantly. Indeed, it can safely be assumed that device aging reduces the overall cumulative leakage current of a digital CMOS circuit. Its impact on the exploitability of the static power consumption through side-channel attacks, however, is much harder to predict. In fact, it is not the case that the leakage currents of a circuit are reduced by a fixed factor for all possible inputs, which would maintain its input dependency. Caused by multiple factors, the data dependency of the static power consumption of a combinational circuit (such as a non-linear substitution box of a block cipher), may shift significantly. For example, two inputs to such a circuit, whose leakage currents are clearly distinguishable in the original state of the device may become less distinguishable due to the aging process. But the opposite situation might occur as well\(^2\). This depends on the concrete netlist of the circuit, the cells which are used, and whether HCI or NBTI are the predominant aging mechanism in the device under test, influencing NMOS or PMOS transistors respectively. Since HCI is caused by the active switching of NMOS transistors, but NBTI by stable values at the input of PMOS transistors (and partially recoverable), it also depends on the type of aging that is performed on the given circuit (i.e., low or high workload, random or biased input data, etc.). Considering Figures 2 and 3 it is clear that depending on whether the threshold voltage increase is more significant in some transistors than others, e.g. more dominant in NMOS than in PMOS transistors, the data dependency of the gate can change significantly in both directions (i.e., decreasing or increasing). Even two identical transistors (i.e., both NMOS or both PMOS) face a different leakage current reduction depending on the data they have processed in the past. This difference in the impact of aging on specific (types of) transistors is what makes it difficult to anticipate its influence on the vulnerability of devices to static power attacks. Thus, an accurate simulation of the impact of aging mechanisms on the data-dependency of the power consumption can be crucial to evaluate the long-term security of cryptographic devices before the tape-out of a design.

2.5 Targeted Device

The main target for our analysis is a (plain/unprotected) implementation of the PRESENT block cipher synthesized by a 65 nm CMOS standard cell library. PRESENT has been introduced in 2007 as an SPN-based ultra-lightweight block cipher for ubiquitous computing environments with extremely constrained resources [BKL+07]. It consists of 31 rounds and makes use of a 4-bit S-box. Its block size is 64 bits and two different key lengths, 80 and 128 bits, are supported. In this work we consider the 80-bit version, which is called PRESENT-80. In [RPLP08] a serialized hardware architecture of PRESENT-80 is proposed and synthesis results by a 350 nm standard cell library are claimed to require as few as 1000 gate equivalences (GE) for the full cipher. In order to achieve such an area-optimized implementation the data paths are serialized to 4-bit words (one nibble) and only one 4-bit S-box is physically implemented, which thus needs to be shared between the data path and the key schedule. Obviously, this area-driven optimization comes at the price of a greater delay in terms of clock cycles, i.e., less throughput. This particular implementation requires 563 clock cycles to encrypt one plaintext. A block diagram of the serialized architecture is shown in Figure 4.

\(^2\)Such a change in the data dependency is usually not to be expected when influencing the (global) operating conditions of a device.
We have implemented this PRESENT-80 architecture as one module of the 65 nm CMOS ASIC chip that we developed to serve as a state-of-the-art device under test (DUT) for advanced side-channel and fault injection evaluations. The fabricated ASIC features 27 different cipher cores (which are partially equipped with countermeasures against physical attacks) and combines them in a sophisticated configuration- and control-framework in order to operate them according to different application scenarios and to provide fast and easy communication with the outside. A picture of the layout of the chip and a microscopic photograph of the bottom right corner of the fabricated die can be seen in Figure 5. As it is apparent from the photo on the right side, no structural information (apart from the position of the bond pads) can be identified from microscope images of the chip due to the metal fill that is distributed over the whole top layer. However, to highlight where the targeted PRESENT-80 core is located on the chip we have circled and framed it in the layout schematic on the left side of Figure 5. This depiction of the layout was taken from the place-and-route software Synopsys IC Compiler (Version 2016.12) during the design process prior to fabrication.

For the tape-out a low-power, multi-threshold voltage, 65 nm CMOS library was used. We have applied the multi-threshold voltage power optimization technique in order to reduce the overall leakage current of the design. In particular, since we did not put any tight timing constraints on the PRESENT-80 core (indeed it was constrained to a clock frequency of 35 MHz) and as the combinatorial circuits (e.g., the S-box) are rather small
in this design, mostly standard cells with a high threshold voltage (HVT) and a low drive strength have been selected by the synthesis tool (Synopsys Design Compiler (Version 2016.12)). Cells with a low drive strength are smaller in terms of area, but cannot drive large output loads. Cells with a high voltage threshold have larger propagation delays, but exhibit a smaller off-current. Even though we have applied rather loose timing constraints, the architecture requires almost 90% more area (1873 GE) in our post-layout netlist than what is claimed in the original proposal [RPLP08]. Apart from the obviously different libraries that are used, this can on one hand be caused by the fact that in many synthesis scripts for area estimation no timing constraints at all are applied (i.e., the reported numbers always tend to be smaller compared to a manufacturable post-layout netlist). On the other hand, we have prevented the tool from performing specific optimization techniques that remove module boundaries or optimize paths across those boundaries.

3 Simulation Results

Our aim is to investigate the impact of aging on the success of static power analysis attacks through conducting both, simulations and silicon measurements. As detailed before, this impact on the exploitability of a device is hard to anticipate. A reduction of the overall leakage current of a device is expected, but the concrete influence on the data dependency and the available information in a side-channel attack needs to be investigated in terms of simulations and practical experiments. Since SCA attacks commonly target the leakage of combinational logic, which is also the case in static power analysis attacks [LB08, MMR17], our simulations only consider an S-box module of the target device described in Section 2.5. We deployed the same netlist used for the chip fabrication in the same 65 nm technology node. Note that as explained earlier, the design contains only one instance of the S-box module.

We used Synopsys HSpice for the transistor-level simulations and deployed the HSpice built-in MOSRA Level 3 model to assess the effect of NBTI and HCI aging [Syn16]. Static current values were extracted for the original circuit as well as the aged ones. The effect of aging was evaluated for 8 weeks of device operation in time steps of one week. We considered the operating temperature as 90°C during the aging process and 20°C during the measurement. The supply voltage \( (V_{dd}) \) was set to 1.416 V during the aging process and 1.2 V during the measurement. Please note that such operating conditions during aging (90°C and 1.416 V) accelerate the process of device degradation by a factor of about 80 (this is technology dependent) [ag17]. Thus, 8 weeks of aging under these conditions correspond to approximately 640 weeks of normal device operation, which is more than 12 years.

The netlist was fed with randomly generated stimuli during the aging simulation process. With respect to the data dependency of the static power consumption, the amount of leakage current of a combinational circuit solely depends on its input value, i.e., there is no dependency on the former input or on the transition between consecutive inputs. Thereby, in the measurement phase we simulated the target circuit (either the original or the aged one) for all possible input values, i.e., 16 cases due to the 4-bit input width of the PRESENT S-box.

Figure 6 depicts the simulation results representing the static current change in an original (0-week age) target circuit as well as the cases where it has been used between 1 and 8 weeks under the above explained aging conditions. In this figure, the static current has been shown when the circuit is fed with different input stimuli. As expected, the magnitude of the static current decreases during the device lifetime regardless of the input stimuli applied during the measurement. Yet, the factor by which the current is reduced

\[3\text{We excluded such kinds of optimization in order to not accidentally corrupt the (masking) countermeasures that have been applied to the other cores in the same ASIC.}\]
for each respective input differs significantly. In other words, the pattern of the data dependency changes.

As apparent from Equations (1) to (3) in Section 2, the influence of the stress time on the threshold voltage increase is strong in the beginning but declining quickly (asymptotically to the square root, respectively quartic root function). Increasing the threshold voltage in turn reduces subthreshold conduction exponentially [Hei09]. Thus, the decrease of the static current is dominant in the first week of aging ($\approx 24.6\%$). After that it continues to decrease, yet with a slower rate.

3.1 Metrics

In the following, we apply essentially two metrics to assess the side-channel leakage of the PRESENT-80 ASIC implementation in simulation and practice. The first one is the non-specific Welch’s $t$-test which has been proposed as a leakage assessment tool for cryptographic primitives in [GJR11, CDG+13] and further developed in [SM15, RGV17]. The second one is the Correlation Power Analysis (CPA) introduced in [BCO04]. The former one aims at being independent of any concrete attack scenarios, targeted intermediate values, and hypothetical leakage models. It can therefore be used to quantify the information leakage that is exhibited by a device through a particular side channel without the need to test many different parameter combinations and to perform an actual key-recovery attack. However, the drawback of this method is that, in case a leakage is detected, it does not provide any information about the hardness of an attack, the intermediate value that should be targeted, or the model that can be applied for a successful key recovery. It may even occur that the detected leakage is not related to any sensitive (key-dependent) intermediate value at all. In other words, a leakage that is detected by a $t$-test does not necessarily point to a vulnerability in the implementation. However, if no leakage can be detected by several non-specific $t$-tests, most likely the device does not exhibit any exploitable leakage. In this work, we do not only rely on the $t$-test as an evaluation metric, but additionally perform a CPA on the intermediate value that is processed by the single S-box instance of the design in order to recover (a part of) the key. In a CPA the adversary usually compiles a hypothetical power consumption for each recorded trace, by applying a hypothetical model to an intermediate value that partially depends on a known input and to the other part on a small chunk of the secret. By guessing the secret part, compiling the hypothetical power model under this guess and finally correlating the hypothetical leakage to the measured leakage traces a correctly guessed secret can often be recognized. This is true in case a large enough number of side-channel measurements has been collected, and
the chosen leakage model reflects the reality in a sufficiently accurate manner.

3.2 Analysis

In order to determine the influence of aging on the vulnerability of the PRESENT S-box circuit, we apply the previously introduced metrics. In this regard, we have sampled 500,000 values by adding Gaussian distributed noise (with standard deviation of $15 \times 10^{-10}$) to the simulated data in Figure 6, suitable for a non-specific Welch's t-test. To be more precise, half of these values have been sampled for randomly chosen inputs, the other half for one fixed input nibble$^4$. The t-test is then used to decide whether the two groups, i.e., fixed vs. random, can confidently be distinguished. It is a usual practice to set the threshold for a successful distinguishability to a value of ±4.5, since this corresponds to a confidence level of 99.999% to reject the null hypothesis [SM15]. We further made use of half of the collected measurements (those with random associated input) to conduct a CPA attack. To this end, we applied a Hamming weight (HW) model based on the S-box output targeting a key nibble. The corresponding results are depicted in Figure 7. Here, we considered three cases of the simulated static currents: (1) the original circuit, (2) the circuit aged for 4 weeks, and (3) the circuit aged for 8 weeks. As shown later in Section 4, we considered exactly these three cases in our practical investigations.

3.2.1 Discussion

The graphics show that the values for the t-statistic are reduced significantly by aging the circuit. In particular, the original circuit showed a $t$-value of about 16 after 500,000 measurements, while both, the 4-weeks and 8-weeks aged circuits show a value of around 9. This result confirms not only that aging reduces the data-dependency of the measured currents, but also that the most significant reduction takes place in the first weeks of aging.

$^4$Note that the order of giving fixed or random inputs is also randomized [CDG+13].
The correlation coefficient of the CPA, targeting the Hamming weight (HW) of the S-box output, reduces by 26.32% after 4 weeks of aging and further 20.22% after another 4 weeks of aging (8 weeks in total). Please note that the concrete values depend on the sampling of the Gaussian noise and thus may be subject to change when repeating the process. Yet, since we perform the same kind of analysis in the practical experiments, we chose to apply these metrics here.

Interestingly, in Figure 6, we can notice both of the following situations. On one hand, considering the leakage currents of the S-box prior to aging (week 0) for inputs 0x7 and 0x2 for example, it is obvious that these inputs can be distinguished easily by their leakage currents. However, already after the first week of aging, their corresponding leakage currents become very similar and continue to stay in the same range throughout the aging process. On the other hand, when taking a look at the leakage currents for inputs 0x1 and 0x3 for example, one may notice that prior to aging (week 0) they are very similar. But then, after aging the circuit the leakage currents become clearly distinguishable. Although the examples of the former kind (i.e., worse distinguishability after aging) are predominant, it is noteworthy that the latter kind (i.e., improved distinguishability after aging) exists as well. In general it can be concluded that the absolute leakage current difference between the individual input classes decreases by aging the circuit. This, in fact, corresponds to a decrease of the signal [MMR18] in the signal-to-noise ratio (SNR) [MOP07] frequently applied in side-channel analysis. However, due to the existence of the second kind of examples (i.e., improved distinguishability after aging) it can be assumed that it is possible to find combinatorial circuits that have a balanced static power consumption (i.e., showing no/small input dependency) before aging, which then in turn becomes imbalanced after aging the circuit. Thus, we presume that a (naive) leakage current balancing technique for combinatorial circuits is not an appropriate countermeasure against static power analysis attacks, since it can potentially be defeated by aging. Additionally this example demonstrates that template attacks will be difficult to carry out in such a scenario, since the input dependence, used to build the templates, changes over time.

4 Practical Analysis

The main objective of this work is the practical verification of the effects that device aging has on the exploitability of the data-dependent leakage currents in cryptographic hardware. To this end we first introduce the dedicated static power measurement setup that we used and the procedure that we follow to quantify the amount of information that is leaked through the static power side channel by our targeted ASIC implementations. As a next step we describe how the aging process of the device is artificially accelerated by controlling its operating environment. Afterwards, we compare the initial measurements to the ones recorded after the chips are aged for several weeks in order to determine the impact of the aging-related degradation on our ability to extract the secret key via the static power side channel.

4.1 Measurement Setup

As detailed in all previous publications which report results based on experimental static power side-channel analyses [Mor14, PSKM15, MMR17, BCS+17, MMR18], to extract sensitive information from the acquired leakage traces, a dedicated measurement setup is required. In addition to a digital sampling oscilloscope, the main components for such a setup include a precisely controllable climate chamber, a low-noise DC amplifier and a low-pass filter. We have used a similar setup as the one reported in [MMR18] consisting of a CTS climate chamber of type C-40/100, a custom amplifier with a $\times 1000$ gain and

\textsuperscript{5}Concrete values are compared in Section 4.
Figure 8: Custom measurement board carrying one of the 65 nm ASIC samples and the DC amplifier.

custom low-pass filter with a cutoff frequency of 100 Hz. For the sampling a LeCroy HRO 66 zi oscilloscope was used. As an interface for the communication between the PC and the ASIC we have used a Basys 3 board which features an Artix-7 FPGA [DIG]. In contrast to the setup in [MMR18], the FPGA board was not placed inside the climate chamber. We have developed a custom measurement board which holds the ASIC in a PLCC-44 socket and provides SMA connectors for the DC amplifier to measure the voltage drop over a 1 Ω shunt resistor. This board, which can be seen in Figure 8, is powered by the Basys 3 board and supplies the ASIC with two voltages via two linear voltage regulators, one for the core area (nom. 1.2 V) and one for the IO ring (nom. 3.3 V).

The custom board, together with the amplifier and the ASIC have been placed inside the climate controlled environment, while the remaining parts of the setup are placed outside of the chamber. For the static power measurements on the 65 nm ASIC, we have operated the core-region of the chip with nominal supply voltage of 1.2 V and a constant temperature of 20 °C.

4.2 Procedure

To measure the static power consumption of a device under test (DUT) the target implementation has to be kept in an idle state. More precisely, the attacker needs to control the clock signal in order to suspend the device at the desired cycle of the cryptographic operation and to measure its leakage current. This leakage current can be observed as a DC shift of the static signal, as soon as the effect of the dynamic power consumption is vanished. We follow the procedure introduced in [MMR18] to measure the DC shift that corresponds to the fixed state of the circuit in a particular targeted clock cycle. In our measurements, after the last edge of the clock we ignore the first 100 ms and average all values that are measured in the next 100 ms by the digital sampling oscilloscope to obtain a singular static power value. This procedure is depicted in Figure 9, where T1 denotes the first 100 ms which are ignored, and T2 denotes the next 100 ms which are averaged to obtain a quantitative value for the DC shift. The sampling rate was set to 10 MS/s, i.e., 1,000,000 samples were averaged over the 100 ms period. There was no explicit delay considered between the measurements. Yet, due to the communication between the PC and the Basys 3 as well as between the Basys 3 board and the ASIC a small implicit delay took place between the individual acquisitions.
Figure 9: Sample measurement illustrating the procedure. At the beginning of \( T_1 \) the clock is stopped, all values in \( T_1 \) are ignored. Then all values measured in \( T_2 \) are averaged to a singular value representative for amount of static power.

4.3 Results 65 nm ASIC

For these experiments we used two completely fresh and unused samples of the developed 65 nm ASIC. In particular, we took two of the dies that have been shipped by the foundry, packaged them using a semi-automatic die bonder into a JLCC-44 package, and started the experiments on those two fresh samples which never performed a single computation before and were power-upped for the first time. Prior to proceeding with the aging acceleration process on the ASIC samples, we first acquired the corresponding reference values for the PRESENT implementations. To this end, we collected 500,000 measurements using the previously described measurement procedure for randomly interleaved fixed and random plaintexts at a constant temperature of 20°C while the ASIC chips were powered by a 1.2 V supply voltage. The clock was stopped at the end of the first round, when the last state (plaintext $\oplus$ key) nibble is applied to the S-box circuit of the PRESENT-80 implementation. Acquiring those traces took roughly 3 days. After the first sample was finished we performed the same experiments on the second sample, which took another 3 days. Then we started the aging acceleration process on both chips in parallel.

In order to accelerate the device aging we have operated the ASIC chips for 4 consecutive weeks at a constant temperature of 90°C, a supply voltage that has been increased by 18% (i.e., 1.416 V instead of 1.2 V) and a continuously high workload (i.e., constantly giving random input to the targeted PRESENT-80 encryption core). Thus, due to the high activity factor, HCI aging will contribute a lot to the device degradation, while the NBTI aging of PMOS transistors is constantly changing between recovery and stress times, due to the randomized input data. In theory, such conditions lead to a much faster device degradation than a normal operation (factor of about 80) [JED16, agi17]. After this period, both chips were disconnected from the power supply and rested for two full days at room temperature without any operation in order to cool down. Afterwards we started another set of measurements at 20°C and 1.2 V on the first sample, while the second sample was still resting without being powered. Subsequently, we exchanged both chips so that the first ASIC was resting while the second one was measured. Please note that in order to avoid any influences in our measurements stemming from the aging of components in the measurement chain other than the targeted ASIC chips (e.g., the PCB and all electronic components, capacitors/resistors/voltage regulators/etc.), we have used different, but structurally identical, custom measurement boards for the aging acceleration process than the one which was used for the reference measurements. Furthermore, neither the shunt resistor nor the DC amplifier have been placed inside the climate chamber during the aging. In other words, none of the parts that have been present in the heating chamber during the aging acceleration process, with the exception of the ASIC chips themselves, are reused for the measurements, and equally none of the parts that are used in the measurements
Figure 10: Non-specific $t$-test and CPA results on static power measurements from the first sample of the 65 nm ASIC after 0, 4 and 8 weeks of aging (top to bottom). The CPA targets a key nibble in the first round using the HW of the S-box output.

have ever been exposed to increased temperatures, supply voltages or to an extremely large workload over a long time period. For the same reason we did not perform the measurements on the chips at increased operating conditions, but instead at 20°C and 1.2 V. Although, at higher temperatures and voltages the static power side-channel is more informative, we could not consider such conditions for our analysis. At higher temperatures or voltages the ASIC, the DC amplifier and the components on the measurement board would be subject to an accelerated aging process during the measurements, which certainly would influence the results and limit their comparability\textsuperscript{6}.

The described aging process and the subsequent measurements have been performed twice, which allows us to present results for 4 weeks and 8 weeks of intense aging respectively. These settings have been selected according to the simulation results shown in Section 3. For both, after 4-week and 8-week aging, we conducted the same evaluations as those performed on the simulated data, i.e., non-specific $t$-test and CPA based on the HW of an S-box output at the first round. The corresponding results, showing the development of the exploitability of the first sample over a period of 8 weeks of aging are shown in Figure 10. Similar to the simulation results it can be observed that the vulnerability of the implementation is reduced by aging the circuit. The $t$-test requires roughly twice as many measurements as on the unaged chip to overcome the 4.5 threshold when the device is aged, independent of whether 4 or 8 weeks of aging are considered. Furthermore, the $t$-test curves in the two aged cases behave similar up to approximately 400 000 measurements. After this limit the $t$-values start to decrease again in the 8-weeks aged case, due to some noise influence. Similar observations can be made for the correlation coefficient in the CPA attack on the traces measured for random inputs. Especially when comparing the results up to 400 000 traces (200 000 random ones) it is obvious that the most significant reduction of the exploitability occurs in the first aging period. However, this can still be observed after the whole set of traces in the $t$-test results, but with a smaller difference. In order to confirm these practical results we repeated the same analysis on the mea-\textsuperscript{6}Results for simultaneous aging and measuring of a chip are presented later in the section.
Figure 11: Non-specific t-test and CPA results on static power measurements from the second sample of the 65 nm ASIC after 0, 4 and 8 weeks of aging (top to bottom). The CPA targets a key nibble in the first round using the HW of the S-box output.

measurements recorded from the second (identical) ASIC sample as well. The corresponding results are depicted in Figure 11. Interestingly, the t-values are larger than on the other sample, while the CPA performs worse than before. This already shows the impact of process variations on the comparability of side-channel measurements taken from two structurally identical devices. The distinguishability of the fixed and random groups in the t-test is not only (positively) influenced by the leakage of the S-box circuit of the PRESENT implementation (although it certainly has a large contribution) but also by the leakage of the state register, multiplexers and other cells in the design. The CPA on the other hand, as it actively targets the leakage of the S-box, is affected negatively by those leakages as they contribute to the algorithmic noise (when associated to non-targeted state nibbles). Thus, it can be assumed that the cells not belonging to the S-box have a larger data-dependent leakage current on this sample of the 65 nm ASIC than on the other one. Apart from that observation, the results also confirm that aging reduces the available information. However, it can be noticed that the t-values on the measurements recorded after 8 weeks of aging the chip are slightly larger in comparison to the 4 weeks aged chip. We assume this to be a random occurrence. As predicted by the simulations there should be no large difference in the distinguishability between 4 weeks and 8 weeks anyway and both values are significantly smaller in comparison to the original state of the circuit. The CPA results confirm that the S-box circuit is more difficult to attack after 8 weeks than after 4. For a simple comparison of the simulations and the practical results on both ASIC samples we have listed the t-test and correlation values in Table 1.

4.4 Results 150 nm ASIC

In addition to the aging experiments on our self-made 65 nm chip, we have also performed measurements on another ASIC prototype chip which has been manufactured in a less recent technology node, namely 150 nm. In view of recently published results where static power measurements were performed under an increased working temperature and supply voltage
Table 1: Comparison of the simulation and practical measurements on both 65 nm samples.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation</td>
<td>Original device</td>
<td>15.941</td>
<td>0.02283</td>
<td>-</td>
</tr>
<tr>
<td>Simulation</td>
<td>4 weeks aged</td>
<td>8.818</td>
<td>0.01682</td>
<td>-</td>
</tr>
<tr>
<td>Simulation</td>
<td>8 weeks aged</td>
<td>8.590</td>
<td>0.01340</td>
<td>-</td>
</tr>
<tr>
<td>Measurements sample 1</td>
<td>Original device</td>
<td>12.514</td>
<td>0.02801</td>
<td>8.6 μA</td>
</tr>
<tr>
<td>Measurements sample 1</td>
<td>4 weeks aged</td>
<td>9.299</td>
<td>0.02410</td>
<td>8.0 μA</td>
</tr>
<tr>
<td>Measurements sample 1</td>
<td>8 weeks aged</td>
<td>6.359</td>
<td>0.01718</td>
<td>7.5 μA</td>
</tr>
<tr>
<td>Measurements sample 2</td>
<td>Original device</td>
<td>23.251</td>
<td>0.01472</td>
<td>7.5 μA</td>
</tr>
<tr>
<td>Measurements sample 2</td>
<td>4 weeks aged</td>
<td>13.647</td>
<td>0.01465</td>
<td>7.2 μA</td>
</tr>
<tr>
<td>Measurements sample 2</td>
<td>8 weeks aged</td>
<td>16.710</td>
<td>0.01447</td>
<td>6.9 μA</td>
</tr>
</tbody>
</table>

Figure 12: 150 nm ASIC chip, 90 °C, 10% over voltage, average number of measurements to disclosure (MTD) for a successful key recovery over 7 consecutive weeks under aging process.

over a long time period to amplify the exploitable signal in a side-channel attack [MMR18] we were eager to investigate whether a simultaneous aging and measurement process also leads to reduced exploitable information in the power traces over time. In this regard we took the identical test chip as was used in [MMR18] (but a new, unaged sample) and performed the same kind of measurements on the PRESENT core (i.e., 90 °C, 10% over-voltage and a 10 ms measurement interval) for a consecutive period of 7 weeks. In other words, during the measurement phase the device was continually aged. Afterwards, we have calculated average MTD (measurements to disclosure) values for each week respectively, corresponding to CPA results on disjoint subsets of the full measurement set per week. The results are presented in Figure 12. As shown, the average number of required measurements increases significantly in the first three weeks. Afterwards, it still increases, yet with a slower rate. A similar behavior can be observed regarding the correlation coefficients for the correct key candidate, as demonstrated in Figure 13. However, in this case the most significant reduction of the correlation coefficient can be observed between the first and the second week of simultaneous aging and measuring.

4.5 Discussion

Our practical analysis, based on real-silicon measurements, shows clearly that the information which is leaked through the static power side-channel is reduced when the circuit is aged. The transistor-level simulations have indicated that the predominant decrease of the exploitable data dependency occurs in the first week(s) of aging. The practical experiments on both 65 nm ASIC samples have confirmed this prediction, especially with respect to the t-test results.
The experiments on the 150 nm ASIC revealed that throughout a process of 7 weeks of simultaneous aging and measuring, the number of traces for a successful key recovery is increased by a factor of roughly 5 and the correlation coefficient is decreased approximately by a factor of 4. However, in order to argue that it generally still seems to pay off to conduct leakage measurements at high temperatures and supply voltages, we refer to [MMR18], where it is shown that on a chip with the same 150 nm technology node the number of measurements to disclosure can be reduced by a factor of approximately 25 compared to room temperature and nominal supply voltage.

5 Conclusion

Due to so-called device aging, the electrical specifications of transistors embedded in integrated circuits change over their device lifetime. This causes the power consumption and timing characteristics of the device to alter over time. Hence, there seems to be a thorough need to examine its effect on the physical security of cryptographic devices. It is noteworthy that such analyses have previously been performed on delay-based PUFs [MvdL14, MHZ16, KDSG17, KDLG16, RRFT16b, RFFT14b, Qu09] and template attacks [KGD18, KDG18].

Here we investigated the effect of device aging on the success of side-channel analysis attacks through the static power consumption (i.e., leakage current). The transistor-level simulations and practical investigations based on real-silicon experiments that we demonstrated indicate that the amount of exploitable information in the leakage current is reduced when the device is aged. Consequently, the corresponding attacks on aged devices require more measurements for a key recovery. Since static power measurements are usually performed when the target device is being operated at a high temperature (and sometimes with high supply voltage), the device is being aged at the same time. We have shown that in such conditions, when the measurement process takes a couple of weeks, the samples collected at different measurement phases do not correspond to each other. Thus, we do believe that all future publications which report analysis results based on static power side-channel attacks need to explicitly state whether and how long the corresponding measurements have been performed at aging-accelerating conditions. Additionally, it can be of interest to present information about the starting age of any device under test and the order of measurements in case they have been collected in multiple phases from the same chip.

The experiments we showed here were based on an unprotected implementation (i.e., no SCA-countermeasures have been applied). The effect of aging on static power analysis attacks is expected to be more destructive when higher-order leakages of an SCA-protected
implementation need to be exploited, since the estimation of higher-order statistical moments is highly sensitive to the noise level. We plan to practically investigate such cases in the future. A further scope for future work was already mentioned in Section 3. It can be of interest to verify whether a (leakage) power-balanced combinatorial circuit can be made vulnerable again by aging the device.

Acknowledgments

The work described in this paper has been supported in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy - EXC 2092 CASA - 390781972 and through the project 271752544 "NaSCA: Nano-Scale Side-Channel Analysis".

References


