Exploiting the Difference of Side-Channel Leakages

Michael Hutter[1], Mario Kirschbaum[1], Thomas Plos[1], Jörn-Marc Schmidt[1], and Stefan Mangard[2]

[1] Institute for Applied Information Processing and Communications (IAIK), Graz University of Technology, Inffeldgasse 16a, 8010 Graz, Austria
{mhutter,mkirschbaum,tplos,jschmidt}@iaik.tugraz.at
[2] Infineon Technologies AG, Am Campeon 1-12, 85579 Neubiberg, Germany
stefan.mangard@infineon.com

Abstract. In this paper, we propose a setup that improves the performance of implementation attacks by exploiting the difference of side-channel leakages. The main idea of our setup is to use two cryptographic devices and to measure the difference of their physical leakages, e.g., their power consumption. This increases the signal-to-noise ratio of the measurement and reduces the number of needed power-consumption traces in order to succeed an attack. The setup can efficiently be applied (but is not limited) in scenarios where two synchronous devices are available for analysis. By applying template-based attacks, only a few power traces are required to successfully identify weak but data-dependent leakage differences. In order to quantify the efficiency of our proposed setup, we performed practical experiments by designing three evaluation boards that assemble different cryptographic implementations. The results of our investigations show that the needed number of traces can be reduced up to 90 %.

Keywords: Side-Channel Attacks, Power Analysis, Measurement Setup, DPA, SPA.

1 Introduction

Side-channel attacks are among the most powerful attacks performed on cryptographic implementations. They exploit secret information that physically leak out of a device. Typical side channels are the power consumption [11,12], the electromagnetic emanation [1], or the execution time of cryptographic algorithms [10]. The efficiency or even the success of an attack is largely determined by the used measurement equipment. The better the equipment, the less noise and the higher the side-channel leakage exploitation will be. Especially when countermeasure-enabled devices are analyzed, the setup is vital in order to limit the needed number of power-trace acquisitions to succeed an attack.

In this paper, we present a setup that improves the efficiency of side-channel attacks by measuring the difference of two side-channel leakages. Our setup is

W. Schindler and S.A. Huss (Eds.): COSADE 2012, LNCS 7275, pp. 1–16, 2012.
The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-642-29912-4_1

based on the idea to use two cryptographic devices (instead of one) and to measure the difference of their physical characteristics (e.g., the power consumption). If both modules perform the same cryptographic operation, their physical characteristics are the same so that the difference of both side-channel measurements becomes theoretically zero. However, if one module processes different data than the other module, a difference in both measurements can be observed at locations in time when data-dependent information is processed. The difference of both side channels therefore provides only data-dependent signals and eliminates static and (non data-dependent) dynamic signals (i.e., noise). Hence, the quality of the measurements can be significantly improved which results in the fact that less power traces have to be acquired in practice.

In order to perform side-channel analysis attacks using our setup, an attacker can choose from two possible attacking scenarios: (1) one device is fed with constant input data while the second device is fed with random data, or (2) one device is fed in a way such that the targeted intermediate value is complementary to the intermediate value of the second device. For both scenarios, we quantified the efficiency by performing practical experiments. We designed three evaluation boards where each board uses two devices (an AT89S8253 microcontroller, an ATmega128, and a custom 8051 ASIC design). In our experiments, we applied the Pearson Correlation coefficient and performed a classical Differential (or Correlation based) Power Analysis (DPA) attack [11,12] on the differential power trace. Our best results increased the correlation coefficient for the AT89S8253 from 0.64 to 0.99 (55 %), for the ATmega128 from 0.61 to 0.96 (57 %), and for the custom 8051 ASIC from 0.11 to 0.22 (100 %). Furthermore, we evaluated our method on countermeasure-enabled devices and performed attacks on an implementation that uses randomization techniques as well as a masked AES implementation. In this scenario, it shows that the setup reduces the number of needed traces up to 90 %.

The rest of this paper is organized as follows. In Section 2, we discuss related work. Section 3 gives a brief overview on side-channel measurements and describes how to improve the signal-to-noise ratio. After that, we present the new measurement setup and highlight the benefits. In Section 4, we describe the measurement process in detail and introduce two different measurement scenarios. The three evaluation boards are presented in Section 5. Section 6 describes the performed attacks. Results are given in Section 7 and Section 8. Conclusions are drawn in Section 9.

2 Related Work

There exist several side-channel analysis (SCA) measurement boards as well as SCA simulation tools and evaluation setups. SCA measurement boards aim at providing a common attack platform that eases the comparison of measurement results. Well-known attack platforms for SCA evaluation are the INSTAC boards from the Tamper-resistance Standardization Research Committee (TSRC) [13] and the SASEBO boards from the Research Center for Information Security

(RCIS) and Tohoku University [17]. The TSRC has released two boards, the INSTAC-8 with an 8-bit microcontroller and the INSTAC-32 with a 32-bit microcontroller and an FPGA. From the SASEBO boards there exist a variety of different evaluation platforms that contain Xilinx (SASEBO, SASEBO-G, SASEBO-GII) or Altera (SASEBO-B) FPGAs. The boards contain two FPGAs, one for the implementation of the cryptographic algorithm and one for handling control tasks. Since the FPGAs have processor cores integrated (powerPC processor cores), both hardware and software implementations can be evaluated with theses boards. An SCA simulation tool has been also presented by Eindhoven University of Technology. The tool is called PINPAS and allows analyzing the vulnerability of software algorithms against SCA attacks [7]. Commercial SCA evaluation setups are offered by companies like Cryptography Research (DPA Workstation [6]), Riscure (Inspector [16]), and Brightsight (Sideways [4]).

3 The Measurement of Side-Channel Leakages

A measurement of side-channel leakage involves various components. Besides components that are caused by the execution of an operation or due to data-dependent variations, there exist components that are caused due to different kinds of noise. Noise is produced by the equipment itself (e.g., quantization noise of the digital oscilloscope, an unstable clock generator, glitches and variations in the power supply, etc.), by the device (switching noise or noise due to leakage currents), or by the environment (radiated or conducted emissions, cosmic radiation, etc.). The higher the noise, the lower the measured side-channel leakage will be and the more traces have to be acquired to perform a successful side-channel attack. The signal-to-noise ratio is a measure to characterize the side-channel leakage of cryptographic devices. It is the ratio between the (data-dependent) signal and the noise component of a measurement [12].

In the following, we propose a new setup that can be used to increase the signal-to-noise ratio of side-channel measurements. Instead of exploiting the side-channel leakage of only one cryptographic device, we propose to use two devices to exploit the difference of their side-channel leakages. The setup therefore significantly reduces the number of needed power-consumption traces to succeed an attack.

3.1 The Proposed Measurement Setup

Figure 1 shows the schematic of the proposed setup. It consists of two cryptographic Integrated Circuits (ICs) (IC_1 on the left side and IC_2 on the right side of the figure). A resistor is placed in the ground line of each IC (GND_1 and GND_2) which allows to measure the voltage drop across the resistors.

In contrast to classical power-analysis setups, we propose to measure the voltage difference of both ICs, i.e., V_{Diff} in Figure 1. This can be simply done by using a differential probe which in fact implicitly subtracts the side-channel

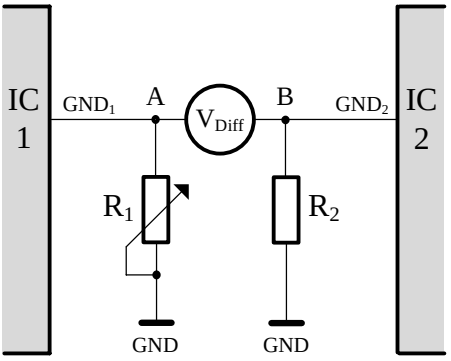

Fig. 1. Schematic of the proposed setup.

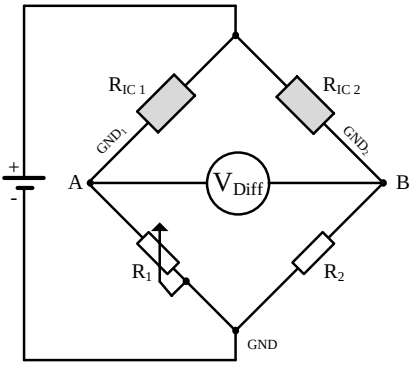

Fig. 2. Schematic of a Wheatstone bridge.

leakage of both devices and allows the efficient acquisition of their side-channel leakage difference.

In view of electrical metrology, the setup is actually equal to a bridge circuit which can be used to accurately measure very small variations of two circuit branches. Figure 2 shows the schematic of a Wheatstone bridge. The dynamic resistance R_{IC_1} of IC_1 and R_1 form one branch and the resistance R_{IC_2} of IC_2 and R_2 represent the other branch of the bridge circuit. The voltage difference of both branches is then measured between the points A and B.

The bridge can be manually balanced by varying the resistor R_1. It is balanced if a zero value is measured at V_{Diff} which means that the same amount of current flows through the branch $R_{IC_1} + R_1$ and through the second branch $R_{IC_2} + R_2$. Note that the voltage at point A is proportional to the ratio $\frac{R_{IC_1}}{R_1}$ and it is $\frac{R_{IC_2}}{R_2}$ at the point B.

If both ICs process different data, the measurement bridge becomes unbalanced. In this case, the measured voltage difference V_{Diff} is high and this causes a peak in the measured power traces. This voltage difference is in fact proportional to the processed data and can be therefore exploited in side-channel attacks.

3.2 What are the Advantages of the Proposed Setup?

The proposed setup provides three major advantages:

1. **Reduction of noise.** Constant and static power consumption (e.g., the clock signal or non data-dependent operations) are canceled out by the setup because the side-channel leakages of both devices are subtracted. Furthermore, noise from the environment is canceled out since both devices are exposed to the same noise level.
2. **Higher measurement sensitivity.** Since the power-consumption traces of both devices are implicitly subtracted by the setup, only their differences

are identified. This results in a much higher measurement sensitivity and acquisition resolution (we achieved a signal amplification by a factor of up to 7.3 in our experiments, cf. Section 7). Even very low power-consumption differences (that are caused by data-dependent operations for example) can be efficiently identified.

3. **Higher signal-to-noise ratio.** Since the noise level is reduced and the signal acquisition resolution is increased, the signal-to-noise ratio (SNR) is higher compared to conventional DPA attack setups. In fact, the higher the SNR, the less traces have to be acquired.

3.3 Applicability of the Setup

The setup can be applied in scenarios where both devices run synchronously, i.e., the devices process the same operation and data in the same instant of time. This is typically the case for devices that are fed by an external clock source or for devices that possess a very stable internal clock generator. In these cases, both devices can be easily synchronized by feeding the same clock source or by a simultaneous power up. In order to overcome the costly operation of the proposed setup due to synchronization issues, a simple yet effective synchronization circuit based on an FPGA could be used. The FPGA would just have to trigger a reset signal or to toggle the power supply of both devices if the first response (e.g., a power-up indicator) is asynchronous. Once implemented, such an automatic trial-and-error setup device would be universally usable and it would be able to provide a synchronous measurement setup in no time.

For many embedded systems like conventional smart cards, the setup may fail because both devices provide an asynchronous behavior which cannot be controlled by an attacker. This asynchronous behavior is caused by asynchronous designs, unstable clock sources, or by side-channel countermeasures such as clock jitters. However, in a white-box scenario, where the implementation is known and where the devices can be fully controlled, one can benefit from the higher signal-to-noise ratio of the setup to reduce the number of needed traces for a successful attack.

In this paper, we consider only contact-based power-analysis attacks even though the idea can be also extended to electromagnetic (EM) based attack settings. In such a scenario, the position of probes plays a major role in the efficient cancelation of uninteresting signals.

4 Measurement Methodology

In the following, we describe the measurement process to perform side-channel attacks using our proposed setup. First, the setup has to be calibrated in order to efficiently identify interesting side-channel leakages. In a second step, an attacker (or evaluator) has to choose from various attacking scenarios, e.g., keeping the key or input data of one device constant or choosing the inputs in such a way that the targeted intermediate value is complementary to the intermediate value of the second device.

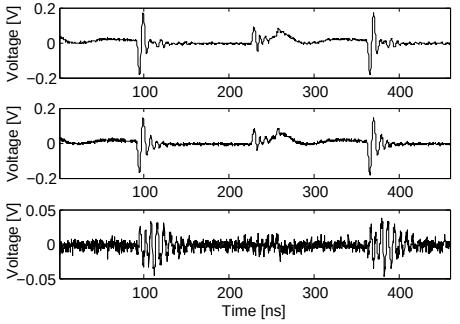

Fig. 3. Power-consumption traces of two devices that process *the same* data (first two rows from the top) are subtracted (difference trace at the bottom).

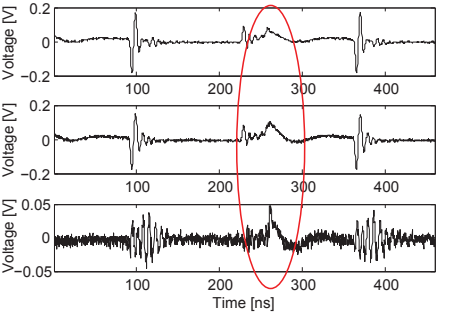

Fig. 4. Power-consumption traces of two devices that process *different* data (first two rows from the top) are subtracted (difference trace at the bottom).

4.1 Calibration of the Setup

In order to calibrate our setup, both ICs have to execute the same operations and the same data has to be processed (e.g., zero values). The resistor R_1 has to be adjusted such that a minimum voltage offset is measured at V_{Diff}. Figure 3 shows the result of the calibration step. In the upper two plots of the figure, the power-consumption traces of IC_1 and IC_2 are shown. Both ICs processed the same operation and the same data. The lower plot shows the result after subtracting both power traces. It shows that the signals are nearly canceled out (e.g., the clock signal or the signal between 200 and 300 ns is much weaker in the resulting power trace).

Figure 4 shows the subtraction of two power-consumption traces that are caused by devices which process different data. In this case, the setup becomes unbalanced and a significant voltage difference can be measured at V_{Diff}. A peak can be identified at locations in time when different data is processed.

After calibration, an attacker has to choose between the two possible attacking scenarios which are described in the following.

4.2 Scenario 1: Choosing a Constant Intermediate Value

In this scenario, one device is fed with constant input data such that the targeted intermediate value is also constant. The second device is fed with random input data. For both devices we assume a constant key.

This scenario is practicable for real-world attacks where the secret key of one device is not known. The second device can be fed with constant input data such that a difference in the power-consumption traces is caused that can be exploited in an attack.

The advantage compared to a classical DPA attack lies in a much higher signal-to-noise ratio of the performed measurement. Let P_{meas} be the measured power consumption of a single cryptographic device. Then, the power consumption can be separated into several components such as an operation-dependent part P_{op}, a data-dependent part P_{data}, noise from the environment $P_{env.noise}$, and noise caused by the device itself, i.e., $P_{dev.noise}$ (see [12] for a detailed description of power-trace characterization). P_{meas} can therefore be modeled as a sum of those components, i.e.,

$$P_{meas} = P_{op} + P_{data} + P_{env.noise} + P_{dev.noise}. \tag{1}$$

In view of our proposed setup, the measured power consumption can then be modeled as follows:

$$\begin{aligned} P_{meas} &= P_{op1} + P_{data1} + P_{env.noise1} + P_{dev.noise1} - \\ &\quad (P_{op2} + P_{data2} + P_{env.noise2} + P_{dev.noise2}) \\ &= (P_{data1} - P_{data2}) + (P_{dev.noise1} - P_{dev.noise2}). \end{aligned} \tag{2}$$

Since both devices process the same operation, P_{op1} and P_{op2} are equal and are therefore implicitly canceled out by the setup. The same holds true for the noise $P_{env.noise1}$ and $P_{env.noise2}$ that is caused by the proximity and that influences both devices with the same signal strength. Thus, the remaining power consumption consists only of the difference of their data-dependent components $P_{data1} - P_{data2}$ as well as the difference of their electronic noise, i.e., $P_{dev.noise1} - P_{dev.noise2}$.

4.3 Scenario 2: Choosing Complementary Intermediate Values

In this scenario, one device is fed in a way such that the targeted intermediate value is complementary to the intermediate value of the second device. Therefore, the power-consumption difference is maximized because both devices always process data that are complementary to each other.

This scenario is only practicable if the targeted intermediate value is known by the attacker because only then the complementary value can be generated. This is typically the case for design evaluators or compliance-testing bodies who are in possession of the entire implementation and the secret key. By knowing the targeted intermediate value, the complementary value can be easily calculated which is then processed by the second device.

Figure 5 shows an example where two ICs process different input data x and x'. The input values are chosen in a way such that the targeted intermediate value y' provides a maximum Hamming distance to y. This actually corresponds to flipping all bits of the intermediate value y or to perform an XOR operation of y with 255. For example, if the output byte y of IC_1 is 3 (0x03), the output byte y' of IC_2 is 252 (0xFC).

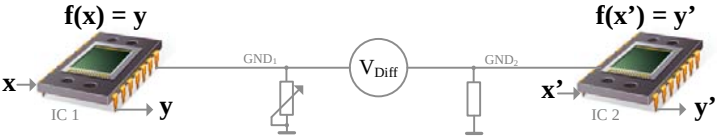

Fig. 5. The processing of different input data x and x' causes a voltage difference between both ICs which can be exploited in a side-channel attack.

4.4 Using Templates

Another big advantage of the proposed setup is the use of templates (cf. [5,2]). The setup can be effectively applied in scenarios where only one single acquisition trace can be measured and evaluated, e.g., in elliptic curve based implementations where the ephemeral key is changed in every execution. In this case, the setup efficiently reveals the power-consumption difference of the two devices in a single shot. This difference can then be compared with generated power-consumption templates in order to classify the leakage according to the processed data.

4.5 The ISO/IEC 10373-6/7 Test Apparatus

The proposed setup is similar to the test apparatus for compliance testing of identification cards specified in the ISO/IEC 10373-6 [8] (for proximity cards) or 10373-7 [9] (for vicinity cards) standard. Figure 6 shows a picture of the apparatus. It consists of a Radio-Frequency Identification (RFID) reader antenna in the middle of the setup and two so-called sense coils. The sense coils have the same distance to the reader antenna so that they measure the same signals emitted by the reader. Both sense coils are connected such that the signal from one coil is in phase opposition to the other coil. This theoretically cancels out the signal of the reader and allows the detection of load modulation signals of contactless identification cards (which are in fact much weaker than the RFID reader field).

5 Practical Evaluation of the Proposed Setup

In order to evaluate the efficiency of our proposed setup, we developed three prototyping boards. Each board assembles two equal ICs and allows the measurement of their power-consumption difference. We used the following processors: an 8051-compatible microcontroller (the AT89S8253 from Atmel), the ATmega128, and another 8051-compatible microcontroller that has been incorporated in an ASIC design fabricated as a prototype chip presented in [14,15].

Figure 7 shows a picture of the AT89S8253 board. It consists of two 8051 microcontrollers, a USB interface for communication, a BNC clock connector, a reset switch, and some I/O pins. The ATmega128 evaluation board (see Fig. 8) additionally features two JTAG interfaces, which allow the programming and debugging of both devices.

Fig. 6. The test apparatus according to the ISO/IEC 10373-6 standard [8].

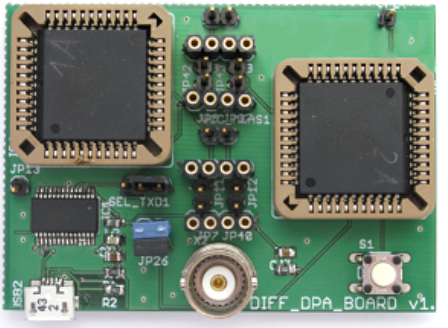

Fig. 7. The AT89S8253 evaluation board.

The ASIC prototype-chip evaluation board is shown in Figure 9. Each ASIC prototype chip contains an 8051-compatible microcontroller with an AES co-processor implemented in CMOS logic and in a masked logic style[3]. The ASIC evaluation board additionally contains voltage regulators and two ROMs for storing the programs executed in the microcontroller cores.

Both devices on the respective evaluation board are connected to the same clock source, whereby the clock wires have been routed in a way so that timing differences (i.e., clock skew) are minimized. All three evaluation boards provide the possibility to easily measure the core power consumption of each of the two devices over a measurement resistor either in the VDD or in the GND line, as well as to measure the power consumption difference of both devices.

6 Description of the Performed Attacks

We performed several attacks using the described evaluation boards. First, we evaluated the efficiency of our proposed setup by setting the intermediate value of one device to a constant value (further denoted as *Constant-Value Attack*). Second, we evaluated the performance of the setup by choosing complementary intermediate values (further denoted as *Complementary-Value Attack*). Third, we evaluated the efficiency of our setup regarding side-channel countermeasures and performed attacks on a randomized and a masked implementation using our custom 8051 ASIC chip.

In order to compare the results, we performed a reference attack for each setup, i.e., a classical Correlation Power Analysis (CPA) attack [3] on one IC of each setup. As a target of these attacks, we considered the output of a `MOV` operation (the input byte is moved from memory to a working register of the CPU). Note that this or similar memory operations are also performed in implementations

[3] As the type of the masked logic style implemented on our prototype chips is not important for this paper, we omit further details about it.

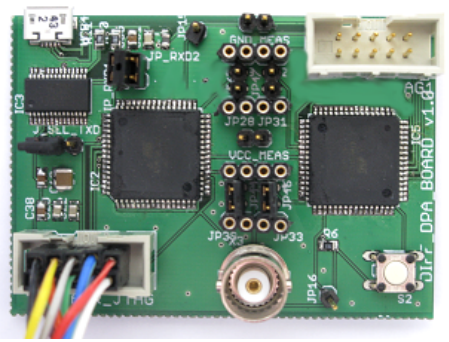

Fig. 8. The ATmega128 evaluation board.

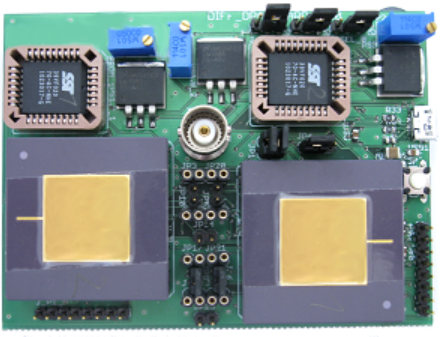

Fig. 9. The ASIC prototype-chip evaluation board.

of cryptographic algorithms such as DES or AES, e.g., moving the S-box output byte after the first round of AES from a register to the RAM.

All boards have been connected to a PC that runs Matlab [18] in order to control the entire measurement setup. The PC transmits three bytes over the serial connection to both ICs that are assembled on each board. IC_1 listens to the first byte, IC_2 listens to the second byte, and the last byte starts the operation on both ICs.

The power consumption of the ICs has been measured using the 2.5 GHz LeCroy WavePro 725Zi 8-bit digital-storage oscilloscope. For all experiments, we used a sampling rate of 5 GS/s. Each IC has been further programmed to pull a debug pin to high which triggers the oscilloscope and starts the measurement process. Furthermore, we used an active differential probe to measure the difference of both side channels. For this, we used the LeCroy D320 WaveLink Differential probe with 3.5 GHz bandwidth.

Processor Synchronization. It showed that the ICs of each setup are often not synchronized after startup and their trigger signals occur at different points in time. This is because both ICs are not powered up perfectly in parallel which causes one IC to get clocked earlier or later than the other IC. In addition, both ICs provide slightly different characteristics (power consumption, timing, etc.) which is due to variations in the fabrication process of the ICs. In order to minimize the differences, we recommend to use only ICs which provide at least the same revision number, production line, and year/month of fabrication.

In order to synchronize the two ICs, we needed to reset and power up the boards until they are synchronized (try and error). For example, for the 8051 microcontroller AT89S8253 the probability of synchronization is 1/24 since the processor requires 12 clock cycles (so-called T-states) to execute a single machine cycle.

Table 1. Result of the Constant-Value Attack using the Pearson Correlation coefficient.

	AT89S8253	**ATmega128**	**8051 CMOS ASIC**
Reference Attack	0.64	0.61	0.11
Constant-Value Attack	0.87	0.87	0.14
Improvement	**0.23**	**0.26**	**0.03**
Improvement [%]	**35.94**	**42.62**	**27.27**

7 Results of Attacks

This section presents the results of the performed attacks. All boards have been clocked at a frequency of 3.6864 MHz.

7.1 Choosing a Constant Intermediate Value

Table 1 shows the correlation coefficient for each measurement setup. For the AT89S8253 and the ATmega128, we measured 1 000 power traces. 10 000 traces have been measured for the 8051 CMOS core of the ASIC prototype chip.

It shows that our setup increased the correlation coefficient by 0.23 (about 36 %) compared to the result obtained from a classical CPA-attack setup. This means that the number of needed power traces is reduced by a factor of about 2.7 (from 50 to only 18). The y-coordinate resolution of the oscilloscope was increased from 81 mV/DIV (for the *Reference Attack*) to 11 mV/DIV (for the *Constant-Value Attack*) which is a factor of about 7.3. Similar results have been obtained for the ATmega128. The correlation coefficient increased by 0.26 (about 43 %), thus the needed number of traces is reduced by a factor of 3.2 (from 57 to 18). The acquisition resolution has been increased by a factor of about 3.8. About 27 % improvement has been obtained for the 8051 CMOS ASIC such that the needed number of traces is reduced by 1.6 (from about 2 300 to only 1 400). The acquisition resolution has been increased by the factor 3.3.

We also calculated the SNR in order to compare the signal level to the noise level. It shows that the SNR increased by a factor of 4.7 to 11.5 in our experiments (depending on the used device). An example for the SNR improvement on the ATmega128 is given in Appendix A.

Table 2. Result of the Complementary-Value Attack using the Pearson Correlation coefficient.

	AT89S8253	**ATmega128**	**8051 CMOS ASIC**
Reference Attack	0.64	0.61	0.11
Complementary-Value Attack	0.99	0.96	0.22
Improvement	**0.35**	**0.35**	**0.11**
Improvement [%]	**54.69**	**57.38**	**100.00**

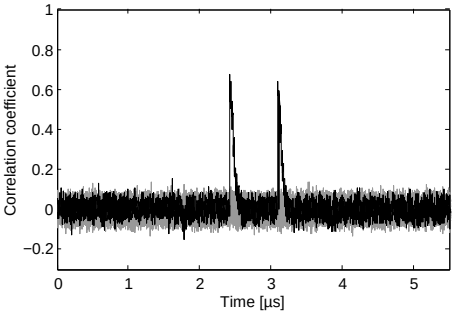

Fig. 10. Result of a classical CPA attack on one ATmega128 device (*Reference Attack*).

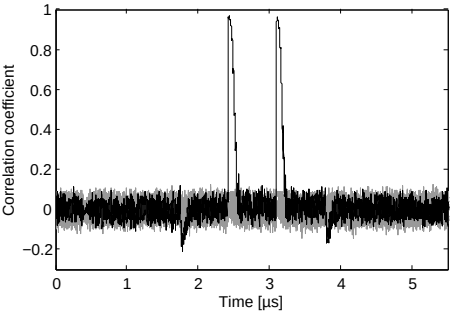

Fig. 11. Result of a CPA attack that exploits the difference of two side channels (*Complementary-Value Attack*).

7.2 Choosing Complementary Intermediate Values

Table 2 shows the result for the *Complementary-Value Attack*. The result shows a significant improvement of the correlation coefficient for every setup. The correlation coefficient has been increased by 0.35 for both the AT89S8253 and the ATmega128 setup, i.e., about 55-57 %. For the 8051 ASIC, a 100 % improvement has been obtained.

Thus, the needed number of traces has been reduced by a factor of 7.2 for the AT89S8253 (7 instead of 51 traces), a factor of 5.7 for the ATmega128 (10 instead of 57 traces), and a factor of 4.1 for the 8051 ASIC (about 550 instead of 2 300 traces).

Figure 10 presents the results of a CPA attack that has been performed on one ATmega128 microcontroller (*Reference Attack*). It shows two correlation peaks (two because the intermediate value has been moved two times in our implementation). The peaks occur between the second and fourth microsecond after the trigger signal. The maximum absolute correlation coefficient is 0.61 for the correct intermediate-value guess (trace plotted in black). All other incorrect guesses show no significant correlation (traces plotted in gray). Figure 11 shows the result of the CPA attack that exploits the difference of two side channels (*Complementary-Value Attack*). For the correct intermediate guess, a correlation of 0.96 has been obtained while no significant correlation can be discerned for incorrect guesses.

8 Results of Attacks on Countermeasure-enabled Devices

In order to evaluate the efficiency of our setup regarding side-channel countermeasures, we investigated two different types of countermeasures: randomization and masking. First, we present results of our ASIC prototype where the `MOV` operation is randomized in time. Second, we present the results of an attack on a masked implementation of a `MOV` operation as well as on the AES core.

8.1 Attacks on Randomization Countermeasures

We performed a *Constant-Value Attack* on a `MOV` operation using our ASIC prototype and compared the results with a *Reference Attack*. For the attack, we measured 10 000 power traces and applied a 50 % randomization in our measurement. This means that the `MOV` operation is randomized at two locations in time. The randomization experiment should indicate the performance of our proposed measurement setup in case of noisy environments (i.e., in case of a randomization countermeasure). Compared to the *Reference Attack* where we achieved a correlation coefficient of 0.11, corresponding to 2 300 traces, the randomization decreases the correlation coefficient to 0.07 (5700 traces). This results in a factor of approximately 2.5. Performing a *Constant-Value Attack* results in a correlation coefficient of 0.09 (3450 traces), i.e., the factor can be reduced from 2.5 to approximately 1.65. Most probably, a *Complementary-Value Attack* would decrease the factor even further.

8.2 Attacks on Masking Countermeasures

We also performed a *Constant-Value Attack* and a *Complementary-Value Attack* on our masked 8051 ASIC chip. First, we targeted a masked `MOV` operation. Second, we targeted the masked AES core. As a target for AES, we have chosen the first S-box output of the first round of AES.

As a result of the *Constant-Value Attack* on the masked `MOV` operation, it shows that the correlation coefficient increased from 0.05 to 0.10 in our experiments. This means that about 8 400 less power traces have be measured compared to a classical DPA attack, i.e., a factor of 4. For the *Complementary-Value Attack*, the correlation coefficient increased from 0.05 to 0.16. Thus, a factor 10 less power traces are needed, this corresponds to about 90 %.

We also performed an attack on the masked AES core that has been implemented on our ASIC prototype. As a reference, we measured the power consumption of a single chip (IC_1) during the execution of AES encryptions of known plaintexts. We performed a standard CPA attack on the AES coprocessor based on the Hamming distance (HD) of two consecutively processed S-box outputs in IC_1. Note that the device leaks the Hamming distance (HD) instead of the Hamming weight of the intermediate values.

After that, we performed a *Constant-Value Attack*. IC_1 performs the same operation as in the reference attack, i.e., AES encryptions of known random plaintexts. IC_2, in contrast, is fed with a constant plaintext. In our case, we set all bytes of the secret key stored in IC_2 to the value 82 (0x52). Moreover, the plaintext of IC_2 was chosen to be a zero value (0x00). This way, the output of the S-box transformation in the first round of AES was constantly 0. Also in this case, our CPA attack was based on the HD of two S-box outputs processed by IC_1.

Table 3 shows the results of the performed attacks. The table compares the results of the reference CPA attack on one single AES coprocessor (reference

Table 3. Summary of the CPA attacks on the AES coprocessor in the prototype chip implemented in CMOS logic; For the attacks, we applied the Hamming-distance power model.

	ASIC CHIP AES COPROCESSOR CMOS							
Byte transition	$2 \rightarrow 1$	$3 \rightarrow 2$	$4 \rightarrow 3$	$16 \rightarrow 4$	$1 \rightarrow 5$	$11 \rightarrow 6$	$3 \rightarrow 7$	$4 \rightarrow 8$
Reference attack	0.0174	0.0163	0.0164	0.0315	0.0133	0.0170	0.0155	0.0292
Constant-Value attack	0.0226	0.0239	0.0278	0.0436	0.0223	0.0293	0.0267	0.0466
Improvement	**0.0052**	**0.0076**	**0.0114**	**0.0121**	**0.009**	**0.0123**	**0.0112**	**0.0174**
Improvement [%]	**30**	**46**	**69**	**38**	**67**	**72**	**72**	**59**

attack) with the CPA results obtained from measuring the difference of the side-channel leakages in case the second chip always computes 0 (0x00) at the S-box output in the first round of the AES encryption. We targeted 8-byte transitions in the AES State and measured 200 000 power traces for the analyses.

The results show that our setup is able to improve the correlation coefficient between 30 % and 72 %. In five of the eight attacks, the correlation coefficient could be increased by more than 50 %. For the best attack, this means that 33 000 traces instead of about 97 000 traces have to be measured to succeed the attack which corresponds to a trace reduction of nearly 3.

9 Conclusion

In this paper, we presented a measurement setup that increases the efficiency of side-channel attacks. The idea of the setup is to use two cryptographic devices and to measure the difference of their side-channel leakages. If both devices perform the same operation synchronously and if they process different data, the static and the data-independent power consumption is canceled out and only data-dependent side-channel leakage can be effectively identified. This results in a much higher signal-to-noise ratio during the measurement where up to 90 % less power traces have to be acquired to succeed an attack as shown in practical experiments. Furthermore, the setup can be used to efficiently identify differences in the instruction flow of cryptographic implementations or to discover data-dependent variations which can be exploited in attacks. The setup further significantly increases the efficiency of template-based side-channel attacks that use only a single-acquisition power trace to reveal secret information.

Acknowledgements. The work has been supported by the European Commission through the ICT program under contract ICT-SEC-2009-5-258754 (Tamper Resistant Sensor Node - TAMPRES) and by Austrian Science Fund (FWF) under grant number P22241-N23 (Investigation of Implementation Attacks - IIA).

References

1. Agrawal, D., Archambeault, B., Rao, J.R., Rohatgi, P.: The EM Side-channel(s). In: Kaliski Jr., B.S., Koç, Ç.K., Paar, C. (eds.) Cryptographic Hardware and Embedded Systems – CHES, 4th International Workshop, Redwood Shores, CA, USA, August 13-15. LNCS, vol. 2523, pp. 29–45. Springer, Heidelberg (2003)
2. Agrawal, D., Rao, J.R., Rohatgi, P., Schramm, K.: Templates as Master Keys. In: Rao, J.R., Sunar, B. (eds.) Cryptographic Hardware and Embedded Systems – CHES, 7th International Workshop, Edinburgh, UK, August 29 - September 1. LNCS, vol. 3659, pp. 15–29. Springer, Heidelberg (2005)
3. Brier, E., Clavier, C., Olivier, F.: Correlation Power Analysis with a Leakage Model. In: Joye, M., Quisquater, J.J. (eds.) Cryptographic Hardware and Embedded Systems – CHES, 6th International Workshop, Cambridge, MA, USA, August 11-13. LNCS, vol. 3156, pp. 16–29. Springer, Heidelberg (2004)
4. Brightsight: Unique Tools from the Security Lab. http://www.brightsight.com/documents/marcom-materials/Brightsight_Tools.pdf
5. Chari, S., Rao, J.R., Rohatgi, P.: Template Attacks. In: Kaliski Jr., B.S., Koç, Ç.K., Paar, C. (eds.) Cryptographic Hardware and Embedded Systems – CHES, 4th International Workshop, Redwood Shores, CA, USA, August 13-15. LNCS, vol. 2523, pp. 13–28. Springer, Heidelberg (2003)
6. Cryptography Research: DPA Workstation. http://www.cryptography.com/technology/dpa-workstation.html
7. den Hartog, J., Verschuren, de Vink, E., de Vos, J., Wiersma, W.: PINPAS: A Tool for Power Analysis of Smartcards. In: Sec'03. pp. 453–457 (2003)
8. International Organisation for Standardization (ISO): ISO/IEC 10373-6: Identification cards - Test methods – Part 6: Proximity cards (2001)
9. International Organisation for Standardization (ISO): ISO/IEC 10373-7: Identification cards - Test methods – Part 7: Vicinity cards (2001)
10. Kocher, P.C.: Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems. In: Koblitz, N. (ed.) 16th Annual International Cryptology Conference – CRYPTO, Santa Barbara, CA, USA, August 18-22. pp. 104–113. No. 1109 in LNCS, Springer, Heidelberg (1996)
11. Kocher, P.C., Jaffe, J., Jun, B.: Differential Power Analysis. In: Wiener, M. (ed.) 19th Annual International Cryptology Conference – CRYPTO, Santa Barbara, CA, USA, August 15-19. LNCS, vol. 1666, pp. 388–397. Springer, Heidelberg (1999)
12. Mangard, S., Oswald, E., Popp, T.: Power Analysis Attacks – Revealing the Secrets of Smart Cards. Springer, Heidelberg (2007)
13. Matsumoto, T., Kawamura, S., Fujisaki, K., Torii, N., Ishida, S., Tsunoo, Y., Saeki, M., Yamagishi, A.: Tamper-Resistance Standardization Research Committee Report. The 2006 Symposium on Cryptography and Information Security
14. Popp, T., Kirschbaum, M., Mangard, S.: Practical Attacks on Masked Hardware. In: Fischlin, M. (ed.) Topics in Cryptology – CT-RSA, San Francisco, CA, USA, April 20-24. LNCS, vol. 5473, pp. 211–225. Springer, Heidelberg (April 2009)
15. Popp, T., Kirschbaum, M., Zefferer, T., Mangard, S.: Evaluation of the Masked Logic Style MDPL on a Prototype Chip. In: Paillier, P., Verbauwhede, I. (eds.) Cryptographic Hardware and Embedded Systems – CHES, 9th International Workshop, Vienna, Austria, September 10-13. LNCS, vol. 4727, pp. 81–94. Springer, Heidelberg (September 2007)
16. Riscure: Inspector – The Side-Channel Test Tool. http://www.riscure.com/fileadmin/images/Docs/Inspector_brochure.pdf

17. Side-Channel Attack Standard Evaluation Board: The SASEBO Website. http://staff.aist.go.jp/akashi.satoh/SASEBO/en/index.html
18. The Mathworks: MATLAB - The Language of Technical Computing. http://www.mathworks.com/products/matlab/

A Appendix: Example of SNR Improvement

We calculated the signal-to-noise ratio for the power measurements on the ATmega128 board (see Section 5 for a description of the board). Figure 12 shows three SNR plots according to three performed attacks: the *Reference Attack*, *Constant-Value attack*, and *Complementary-Value attack*. The SNR is defined as the ratio of signal power to the noise power. For the signal characterization, we calculated the variance of means for each of the 256 possible intermediate values (300 power traces for each value resulting in 76 800 power traces in total). The noise has been characterized by calculating the variance of constant value processing, cf. [12]. It shows that the SNR is improved by a factor of 21.6 (from 3 to about 65). For the *Constant-Value attack*, the SNR has been improved from 0.3 to about 14 (by a factor of 4.6).

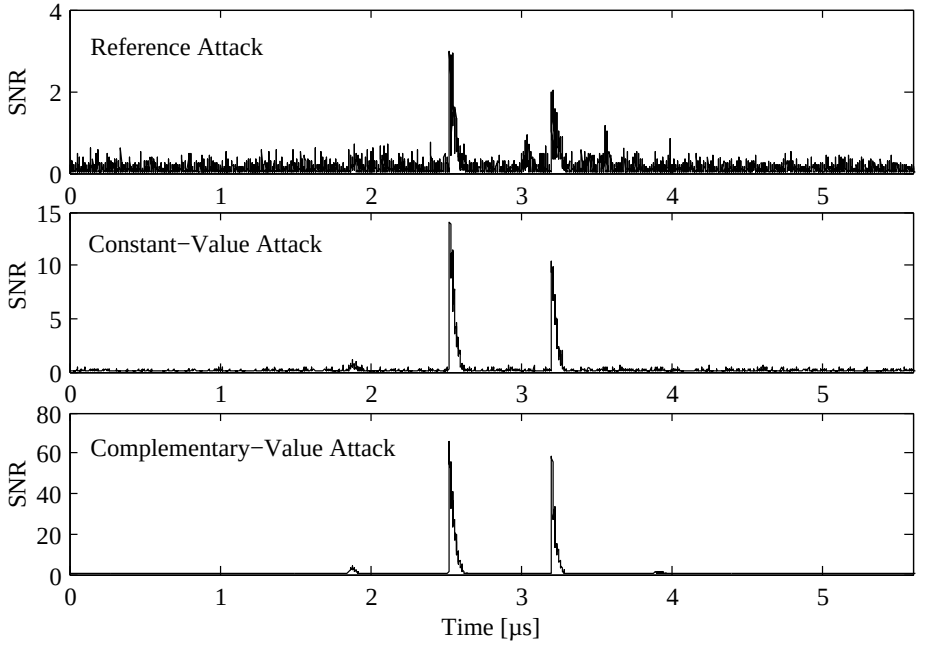

Fig. 12. Signal-to-noise ratio of the *Reference Attack*, *Constant-Value attack*, and *Complementary-Value attack* on the ATmega128.