A Formal Study of Power Variability Issues and Side-Channel Attacks for Nanoscale Devices

Mathieu Renauld, François-Xavier Standaert,
Nicolas Veyrat-Charvillon, Dina Kamel, Denis Flandre.

UCL Crypto Group, Université catholique de Louvain.
Place du Levant 3, B-1348, Louvain-la-Neuve, Belgium.

Abstract. Variability is a central issue in deep submicron technologies, in which it becomes increasingly difficult to produce two chips with the same behavior. While the impact of variability is well understood from the microelectronic point of view, very few works investigated its significance for cryptographic implementations. This is an important concern as 65-nanometer and smaller technologies are soon going to equip an increasing number of security-enabled devices. Based on measurements performed on 20 prototype chips of an AES S-box, this paper provides the first comprehensive treatment of variability issues for side-channel attacks. We show that technology scaling implies important changes in terms of physical security. First, common leakage models (e.g. based on the Hamming weight of the manipulated data) are no longer valid as the size of transistors shrinks, even for standard CMOS circuits. This impacts both the evaluation of hardware countermeasures and formal works assuming that independent computations lead to independent leakage. Second, we discuss the consequences of variability for profiled side-channel attacks. We study the extend to which a leakage model that is carefully profiled for one device can lead to successful attacks against another device. We also define the perceived information to quantify this context, which generalizes the notion of mutual information with possibly degraded leakage models. Our results exhibit that existing side-channel attacks are not perfectly suited to this new context. They constitute an important step in better understanding the challenges raised by future technologies for the theory and practice of leakage resilient cryptography.

Introduction

Side-channel attacks are one of the most important threats against modern cryptographic implementations. Since the apparition of power [11] and electromagnetic analysis [6, 21], the design and evaluation of countermeasures allowing to withstand such physical attacks has become an increasingly important research topic. The security assessment of commercial products (such as smart cards) has also implied major developments in the industry of secure hardware devices. Various solutions purposed to increase the security against side-channel attacks have been proposed, at different abstraction levels. They range from the modification of the hardware [31] to generic techniques using the formalism of

modern cryptography [20]. Significant progresses have also been made in better understanding the statistical aspects of power analysis and its connection with countermeasures such as masking and hiding, as detailed in the DPA book [14].

By contrast to classical cryptanalyis, that targets abstract mathematical objects, side-channel cryptanalysis is implementation-specific. The gain of such a specialization is a significantly increased power. Cryptographic algorithms that are assumed (or proven) secure against classical adversaries, even with intensive time and memory complexities, often turn out to be completely insecure against physical attacks, if implemented in an unprotected device. As a consequence, technological dependencies are at the core of both the theory and practice of side-channel analysis. On the one hand, solutions to attack cryptographic implementations are most efficient if they can exploit a good understanding of the underlying physics. On the other hand, solutions to (provably) ensure the security of leaking devices need to rely on assumptions that correctly capture the peculiarities of actual hardware. In this paper, we tackle this issue of technological dependency and show that some of the common assumptions used in power analysis attacks are not going to hold anymore in future cryptographic hardware.

In particular, the scaling of the CMOS technology, that is the basis of most present microelectronic devices, is a permanent trend since the apparition of integrated circuits in the late 1950s. Shrinking transistors is generally motivated by the need of increased performances and reduced energy per operation. But when reaching the nanometer scale, two major detrimental side effects also arise. First, the relative importance of so-called static currents increases (i.e. energy is consumed, even if no computation is performed) [25]. Second, device variability becomes important (i.e. it becomes increasingly difficult to engineer identical chips) [1, 17]. As a consequence, the goal of this paper is to investigate the impact of these effects, with a focus on power variability, from the point of view of side-channel attacks. More precisely, our contributions are as follows.

1. A classical tool in DPA is to use Pearson's correlation coefficient in order to compare key-dependent leakage predictions with actual measurements performed on a chip [2]. These (so-called) correlation attacks are most efficient if a good leakage model is available for the predictions. And a very common solution is to use the Hamming weight (or distance) of the manipulated data for this purpose. We show that such models are not accurate anymore for 65-nanometer and smaller technologies. Hence, their use may lead to overestimate the security of a (protected or unprotected) implementation.
2. Recent works in the area of leakage resilient cryptography frequently assume that independent computations lead to independent leakage. We put forward that this assumption is not fulfilled anymore for 65-nanometer technologies. In particular, we show that linear leakage models that only depend on the input/output bits of an S-box are not able to capture parasitical effects occurring during the computations. We then discuss the consequences of this observation and highlight that they are different for works such as [5], which assume independence at the gate level, and works such as [4], which assume independence at a larger scale, e.g. between functional blocks.

3. Profiled attacks, e.g. using templates [3] or stochastic models [27], are an important class of side-channel attacks in which an adversary first characterizes a target device (in order to obtain a precise knowledge of the leakage probability distributions), and then uses this knowledge in a very powerful online phase. In this context, it is important to know whether a profile obtained from one device can be used against other similar devices. We discuss this question in light of the increased variability of recent technologies. For this purpose, we define the perceived information, which is a generalization of the mutual information that allows quantifying degraded leakage models.
4. Finally, we provide a careful empirical evaluation of both the information leakage and the success rates of various implementations and attacks. Our results are based on a set of 20 implementations of the same AES S-box in a 65-nanometer low-power CMOS technology. We use these experiments to discuss the impact of the power supply on the information leakage, and the selection of meaningful time samples in the traces. We also take advantage of this case study to compare real measurement traces with simulated ones.

Summarizing, while an important literature covers the impact of nanoscale technologies from a microelectronic point of view, e.g. [7], only a few works consider its consequences in terms of security. To the best of the authors' knowledge, the simulated experiments in [13] are the only available reference. In this paper, we extend these preliminary investigations, and show that technology scaling implies new challenges for the theory and practice of side-channel attacks, that are not completely solved by present statistical tools, proof techniques and assumptions.

1 Preliminaries

1.1 Target implementation

Our analysis is based on simulated and actual power traces obtained from the execution of an AES Rijndael S-box, full-custom designed in a low power 65-nanometer CMOS technology, and measured under two different supply voltages: 1.2V and 0.5V. We used an area-optimized S-box architecture based on composite field arithmetic, described in [16], of which the design is detailed in [9].

Measurements were performed on 20 prototype chips implementing this S-box, each of them made of 1,530 transistors in static CMOS logic style, with a maximum logic depth of 22. The S-box delay is 3 ns at 1.2V supply voltage, meaning a maximum operating frequency of 200 MHz (taking a security margin of 2 ns). This maximum clock frequency drops down below 10 MHz when decreasing the supply to 0.5V. In our experiments, we monitored the voltage drop on a resistor introduced in the supply circuit of the chips, using a high sampling rate oscilloscope (1 Gsample/second), while running the chip at 2 MHz (motivated by interface constraints of our prototype board). Post-layout simulations were performed using Spice models provided by the same industrial foundry as for actual measurements, for the chosen technology node.

1.2 Notations

Let a power trace l be the output of a leakage function L. In our experiments, the leakage function will essentially depend on three input arguments: X, C and N. The (discrete) random variable X denotes the input value of the S-box under investigation, the (discrete) random variable C denotes the index of the chip under investigation, the (continuous) random variable N denotes the noise in the measurements. As a result, we denote the random variable representing the leakage traces as $L(.,.,.)$, where the arguments are written as capital letters if they are variable, and as small caps if they are fixed. For example, $l(x, c, n)$ is a single measurement trace, corresponding to input x and chip c; $L(x, c, N)$ is a random variable representing the noisy traces corresponding to input x and chip c. We also denote the t^{th} time sample in a leakage trace as $L_t(x, c, n)$. Finally, it is sometimes convenient to consider noise-free mean traces, that are defined as:

$$\overline{L}(X, C) = \mathop{\mathbf{E}}_{n} L(X, C, n),$$

where $\mathbf{E}$ denotes the mean operator, which is to be replaced by a sample mean operator (denoted as $\hat{\mathbf{E}}$) when applied to actual measurement traces. Simulation environments such as Spice do not directly allow parametrizing the noise level in the power traces. Therefore, they provide noise-free traces by default. In this case, and in order to analyze the impact of noise on the security of our AES S-box, our evaluations considered an additive Gaussian noise (which is a reasonable starting point for the simulated analysis of side-channel attacks). We denote with $\mathcal{N}(l|\mu, \sigma^2)$ the probability density function (pdf) of a normal random variable L with mean μ, variance σ_n^2 and evaluated on input x. It yields:

$$L_t(X, C, N) = L_t^{sim}(X, C) + N,$$

where N has mean 0 and variance σ_n^2. When considering multiple time samples in the traces (i.e. $L_{t_1:t_d}(X, C, N)$), the mean and variance are replaced by a mean vector and a covariance matrix. Our simulated evaluations assume the same noise distribution for all inputs, chips and time samples. By contrast, when considering actual power traces, the noise is directly present in the measurements obtained from the oscilloscope. In this case, our evaluations characterized its distribution, in order to take possible correlation between different time samples into account.

As an illustration, Figure 9 in Appendix A shows noise-free power traces corresponding to 4 different inputs, measured for 10 different chips, under 1.2V and 0.5V supply voltages, obtained from simulations and actual measurements.

1.3 Noise distribution

The preliminary analysis of a set of leakage traces usually starts with the characterization of the noise. For this purpose, we first applied the filtering described in Appendix B, in order to remove some parasitic frequencies from the traces. Then, we tested the distribution of the residual noise. In recent works on side-channel attacks, this distribution is usually assumed to be normal, with mean

zero and variance σ_n^2 [14]. Using a normality test like the Pearson's chi-square test telled us that, formally, the residual noise does not exactly follow a normal distribution. However, the ratio between the entropy of the estimated normal distribution and its Kullback-Leibler divergence with the actual distribution is smaller than 0.5%, meaning that the residual noise distribution is very close to Gaussian. As will be seen in the following section, this assumption is also validated from a side-channel point of view, when comparing the information leakage computed with the actual noise distribution and with a Gaussian estimate.

1.4 Physical variability

The power consumption traces of an electronic device can be divided into a static part and a dynamic part. These parts can be informally identified by visual inspection: the static power corresponds to the constant parts of the traces, the dynamic power corresponds to their variable parts. Dynamic power is usually the most useful in side-channel attacks, because its strong input-dependency can be used to accumulate information about a secret value manipulated by a device. As discussed in [10], physical variability of the dynamic energy in nanoscale devices can be explained by capacitance fluctuations that are magnified when the computation delays increase, because of the random glitches that are generated by variability-induced unbalanced logic paths. In the following, we will mainly be interested in two parameters that influence the physical variability.

First, the supply voltage can be scaled down, resulting in a reduced dynamic power at the cost of an increased delay, hence implying a higher variability. Second, different time samples can be selected in the traces. Because of the impact of the computation delays on the random glitches in these traces, the samples corresponding to the beginning of the computations have less variability than the ones corresponding to the end of the computations. These parameters can be illustrated by looking at the variance of the power traces, over the input plaintexts and chips, in Figure 1. One can see that the variance over the chips (caused by physical variability) increases when moving from 1.2V to 0.5V supply voltage. In addition, for the 0.5 supply, i.e. when variability becomes significant, this variance is quite localized in the late time samples. Note that this effect is particularly visible when considering the actual measurements.

1.5 Dimensionality reduction

One difficult task when performing a side-channel attack is to select the samples of interest in the traces. Many heuristics have been proposed for this purpose. A straightforward solution is to apply the attacks to all the samples in the traces and to select the samples where they perform best. This is possible, e.g. when applying Kocher's DPA [11], correlation attacks [2] or template attacks [3] (as long as the templates are only built for a reduced number of samples). Alternatively, it is also possible to use dimensionality reduction techniques such as Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA) [29]. These

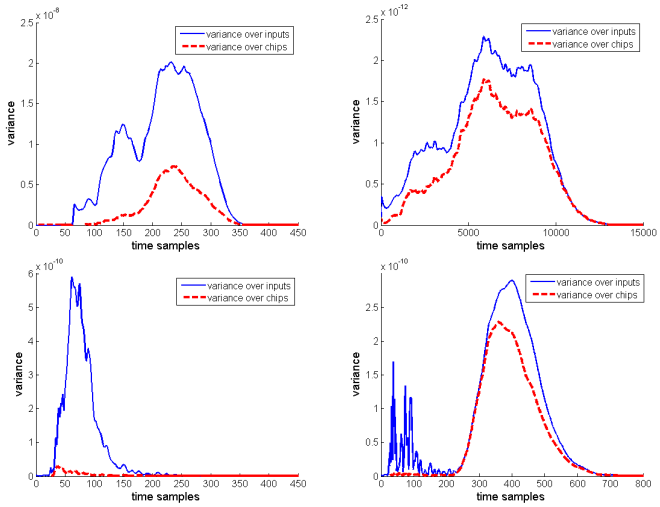

Fig. 1. Variances of the power traces over the input plaintexts and chips. Left: 1.2V power supply, Right: 0.5V power supply / Up: simulations, Down: actual measurements.

are linear transforms that can be used to project the traces in a subspace of small dimensionality, with the goal of "summarizing" the useful information in a few samples. PCA uses the inter-class variance as optimization criteria, while LDA uses the ratio between inter- and intra-class variance. Figure 11 in Appendix C plots the eigenvectors corresponding to the principal component produced by PCA and LDA, for simulated traces. It shows that physical variability makes the application of PCA irrelevant, as it cannot distinguish between inter-plaintext and inter-chip variances. By contrast, LDA does a good job in this case, and only selects early time samples in the traces, where inter-plaintext variance is large and inter-chip variance is small. In order to simplify the interpretation of the results, our analyzes in the following sections will reduce the dimensionality by selecting one to three meaningful time samples, with small, medium and large variability (examples are give in the upper left part of Figure 9 in Appendix A).

2 Information theoretic analysis

The goal of this paper is to investigate how inter-chip variability affects the application of side-channel attacks. For this purpose, we start with an information theoretic analysis. As detailed in [28], it allows to quantify the security of an implementation against an adversary who can perfectly profile the leakage pdf. In our context, we will consider the information between a secret S-box input X and the corresponding leakage L. We analyzed three types of leakage. First, we used simulations $L_t^1 = L_t^{sim}(X, C) + N$. Second, we used actual measurements $L_t^2 = L_t(X, C, N)$. Third, we considered a hybrid situation combining the average traces obtained from the oscilloscope with simulated noise: $L_t^3 = \overline{L}_t(X, C) + N$.

Interestingly, the presence of inter-chip variability implies new concerns regarding the profiling of a leakage pdf. In our 65-nanometer technology, two pieces of silicon implementing the same functionality can give rise to different power consumptions. And this variability can even occur intra-chip, e.g. two S-boxes within the same implementation of the AES can have different leakage models. As a consequence, this section will focus on two main scenarios. In the first one, the profiling and attack are performed on the same chip. This scenario reflects the classical assumption that two chips produced from the same design leak in a similar way. It corresponds to a worst case situation in which all the information leaked by an implementation can be exploited by the adversary. In the second (more realistic) one, different chips are used for profiling and attacking. We study the possibility of building templates from a set of n chips and to attack a $n{+}1^{\text{th}}$ chip, in order to infer the effect of process variability. Doing this, we introduce a new notion of "perceived information", which allows capturing the information loss that is due to the degradation of an adversary's templates.

This section will also consider two additional questions. First, we evaluate the assumption of "independent leakage" that is frequently required by formal security analyzes in physically observable cryptography, e.g. [4, 5]. Then, we discuss the notion of model soundness and its relation with the scenarios of standard DPA attacks [2, 3, 11, 15] and algebraic side-channel attacks [23, 24, 26].

2.1 Worst case scenario: profiling and attacking the same chip

Analyzing the information leakage of a cryptographic implementation first requires to choose a profiling technique, in order to estimate the leakage pdf. In this section, we use the template attacks introduced in [3], which are the most generic solution for this purpose[1]. Template attacks essentially work in two steps. In a first profiling phase, the adversary builds 256 Gaussian templates, denoted as $\hat{\Pr}_{\mathtt{model}}[L|x] = \mathcal{N}(l|\hat{\mu}_{x,c,N}, \hat{\sigma}^2_{x,c,N})$, corresponding to the 256 maximum likelihood estimates of the conditional density functions $\Pr_{\mathtt{chip}}[L|x]$. Then, in a second online phase, he uses these templates to recover information from a leaking chip, for which he will select the maximum likelihood input candidate:

$$\tilde{x} = \underset{x^*}{\operatorname{argmax}}\ \hat{\Pr}_{\mathtt{model}}[x^*|l]. \tag{1}$$

The information theoretic analysis introduced in [28] consists in evaluating the posterior probability of different inputs and computing the mutual information:

$$\mathrm{MI}(X;L) = \mathrm{H}[X] - \sum_{x\in\mathcal{X}} \Pr[x] \sum_{l\in\mathcal{L}} \Pr_{\mathtt{chip}}[l|x].\log_2 \Pr_{\mathtt{chip}}[x|l], \tag{2}$$

where $\Pr_{\mathtt{chip}}[x|l]$ is derived from $\Pr_{\mathtt{chip}}[l|x]$ using Bayes' formula and $\mathcal{X}, \mathcal{L}$ are the sets of all possible input values and leakage. In practice, the real leakage distribution is a priori unknown, both for adversaries and evaluators. Hence, the

[1] An alternative is to use stochastic models [27] and is discussed later in the paper.

probability of the leakage l conditioned on input x is replaced by a sample estimate $\hat{\Pr}_{\mathtt{chip}}[l|x]$ (i.e. typically, one divided by the number of measured traces). And the probability of the input x conditioned on leakage l is replaced by the adversary's model estimate $\hat{\Pr}_{\mathtt{model}}[l|x]$. In general, one assumes that the adversary's model is reasonably close to the actual chip leakages, which allows to formally compute the mutual information. As demonstrated in [30] in the context of the masking countermeasure, the mutual information provides an excellent indicator of the template adversary's success rate. However, if the adversary's model degrades for some reason, and differs from the actual chip leakage distribution, the mutual information cannot be computed anymore. In order to capture such situations, we introduce the following definition of perceived information:

$$\hat{\mathrm{PI}}(X;L) = \mathrm{H}[X] - \sum_{x\in\mathcal{X}} \Pr[x] \sum_{l\in\mathcal{L}} \hat{\Pr}_{\mathtt{chip}}[l|x].\log_2 \hat{\Pr}_{\mathtt{model}}[x|l]. \tag{3}$$

When profiling and attacking the same chip with sufficiently accurate templates, $\hat{\Pr}_{\mathtt{chip}}[l|x]$ and $\hat{\Pr}_{\mathtt{model}}[l|x]$ are the same, and the perceived information reverts to the mutual information. From a side-channel point of view, the intuitive meaning of the perceived information is close to the one of mutual information: it captures the information about a latent variable X obtained when observing leakages L, generated according to a density $\Pr_{\mathtt{chip}}[L|x]$, and interpreted with the model $\hat{\Pr}_{\mathtt{model}}[L|x]$. This implies that the perceived information is lower or equal to the mutual information, and may have a negative value, meaning that the leakage is misinterpreted by the adversary's model. In this case, the side-channel attacks do not converge towards their correct result (i.e. they don't output the correct key). The perceived information can also decrease with measurement noise. Such a counterintuitive behavior will be observed in the next sections: less measurement noise may increase the misinterpretations of the model, as the probability of the correct event $\hat{\Pr}_{\mathtt{model}}[x|l]$ will be closer to zero in this case. Note finally that, in the case of simulations, the sum over the leakages l, in Equations (2) and (3), becomes an integral, as an analytical description of the pdf is available.

The results of our analysis for the worst case scenario where we profile and attack the same chip are displayed in Figure 2 (averaged over 20 chips), with models using 1, 2 or 3 samples (denoted as 1D, 2D and 3D in the plots). They do not exhibit deviations from previous information theoretic analyzes (e.g. the perceived information is always positive). They also confirm the intuition that reducing the power supply reduces the information leakage, and that higher dimension leakage provides more information to the adversary [22]. In fact, the most interesting observations in these experiments relate to simulations:

1. *Simulated noise.* As witnessed by the right part of the figure, average measurements plus simulated noise (i.e. L_t^3) provide an excellent approximation of actual measurements with real noise (i.e. L_t^2), from an information leakage point of view. This is in line with our observation of Section 1.3.
2. *Simulated traces.* As witnessed by the differences between the left and right parts of the figure, the information leakage of simulated traces reasonably

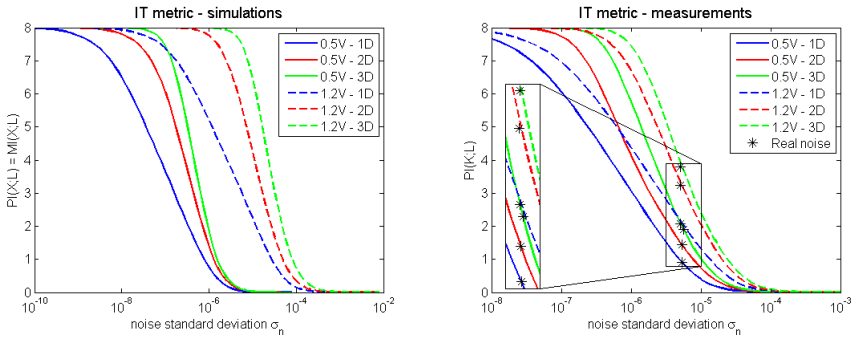

Fig. 2. Mutual information between an input X and corresponding leakage L in function of the noise std. deviation, for 1D, 2D and 3D leakages. Left: simulations. Right curves: measurements + simulated noise. Right stars: measurements.

corresponds to the one of actual traces at 1.2V, and exhibit more deviations at 0.5V (i.e. when variability increases). This can be explained by the difficulty to capture all physical effects in simulation models (including the ones related to our measurement setup). While the intuitions given by simulations are sound (e.g. decreasing the supply voltage reduces the information leakage) the numerical values they provide need to be considered with care.

In the rest of the paper, we will systematically consider averaged measurements plus simulated noise in our evaluations, since this behaves very similarly to the actual measurements while allowing modifications in noise levels[2].

2.2 A note about the "independent leakage" assumption.

An important assumption found in several formal works in the area of leakage resilient cryptography is that independent computations give rise to independent leakage. Our experiments suggest that such an assumption may not hold in practice. One first reason for this, discussed in [18], is cross-talk: the current flowing in one wire of a bus may significantly influence the one of adjacent wires, both in terms of delays and power consumption. More generally, the coupling between any locally connected parts of an integrated circuit, like our S-box implementation, has an important impact in this respect. For example, the leakage traces of different chips in Figure 9 are significantly different. The main cause of these different shapes are glitches, i.e. random transitions at the gates inputs/outputs that are caused by signals arriving at different times. As these arrival times depend on all the paths of the signals before they reach a gate, glitches are a clear expression of leakage dependencies between different parts of a circuit.

In general, it is difficult to quantify the exact impact of each type of coupling that can occur within an integrated circuit. This is because only the combination of all these effects can be observed in a measurement trace. However, it is possible

[2] For each experiment, we additionally checked that simulated noise did not introduce any significant deviation from the real measurement noise, as in Figure 2.

to show that simple models that are linear combinations of the S-box input or output bits are not able to capture the full complexity of the leakage samples in our traces. The stochastic models introduced by Schindler et al. [27] are a very useful tool to quantify this claim. The principle of stochastic models is to perform a regression in order to find the function $\hat{\mathsf{L}}_t = \sum \alpha_i \cdot g_i(x)$ that will best approach the actual leakage function, with $[g_1(x), g_2(x), \ldots, g_N(x)]$ representing the basis used in the regression. If one uses the S-box output bits as base vectors, the approximated function will be linear. And by adding quadratic, cubic, ... terms in the basis, it is possible to refine the approximation. Eventually, a stochastic model using all possible terms of degree equal to or smaller than 8 has enough degrees of freedom to assign an independent value to the leakage of each S-box input, i.e. it is strictly equivalent to the exhaustive construction of 256 templates.

Note that, when evaluating the information leakage with a stochastic model using small bases, the model used by the adversary $\hat{\Pr}_{\mathtt{model}}[L|x]$ may not anymore correspond to the actual leakage pdf $\hat{\Pr}_{\mathtt{chip}}[L|x]$. This happens, e.g. if the stochastic model is not able to capture all the leakage dependencies in the traces.

The left part of Figure 3 plots the information leakage corresponding to different stochastic models. For low noise standard deviations and low degree bases, it shows that the stochastic models are not accurate, as the perceived information decreases below zero. The figure also exhibits that the impact of adding terms in the basis varies with the time samples. Again, the intuitive meaning of the non linear terms in the basis is not easy to give, as they relate to various physical effects. As illustrated in the right part of Figure 3, a combinatorial circuit connects several gates and any intermediate value may be used as an additional base vector. But our experiments at least show that, for any time sample in the traces, even late ones that are mainly influenced by the S-box output bits, various features cannot be predicted by a linear combination of those bits.

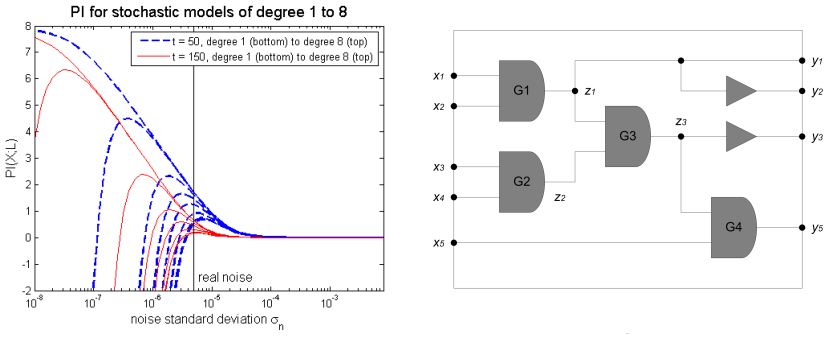

Fig. 3. Left: mutual information between an input X and corresponding leakage L, in function of the noise, obtained using stochastic models with bases containing functions of various degrees of the S-box output bits (1.2V supply). Right: gates combination.

These observations have important consequences for formal works in the area of physically observable cryptography. First, they contradict the assumption in [5], where the security proof requires that different gates generate independent leakage. In general, it is unlikely that this condition can hold for locally connected parts of a circuit. The integration of coupling effects (e.g. leakage functions with quadratic, cubic, ... terms) in such analyzes is an interesting scope for further research. More theoretically, our experiments suggest that the assumption in [4, 19] may not always hold either. These works assume independence at a higher abstraction level, e.g. by requiring that two PRGs lead to independent leakage. In view of the importance of coupling in deep submicron technologies, fulfilling this requirement would at least require to ensure a sufficient (time or space) distance between their executions, so that local dependencies become negligible.

2.3 Realistic scenario: profiling and attacking different chips

As inter-chip variability increases in recent CMOS technologies, the worst-case analysis in the previous section no longer corresponds to an actual attack scenario. This is because the templates profiled for one implementation may not be optimal anymore for attacking other implementations. As a consequence, we now concentrate on a more realistic situation, where one profiles and attacks different chips. The success rate of the side-channel key recovery will then essentially depend on the extent to which the templates built during profiling are sufficiently close to the actual power consumption of the target implementation. Again, this can be measured with the perceived information. But in order to build sound leakage models $\hat{\Pr}_{\mathtt{model}}[L|x]$, we first need to improve the profiling, in order to take the process variability into account. For this purpose, a natural approach is to extend the template profiling as illustrated in Figure 4. That is, classical template attacks estimate the conditional leakage distributions with $\hat{\Pr}_{\mathtt{model}}[L|x] = \mathcal{N}(l|\hat{\mu}_{x,c,N}, \hat{\sigma}^2_{x,c,N})$, where $\hat{\mu}_{x,c,N}$ (resp. $\hat{\sigma}^2_{x,c,N}$) denotes the sample mean (resp. variance) of the leakage variable, for a fixed plaintext x, chip c and a random noise N. In order to take the inter-chip variability into account, one

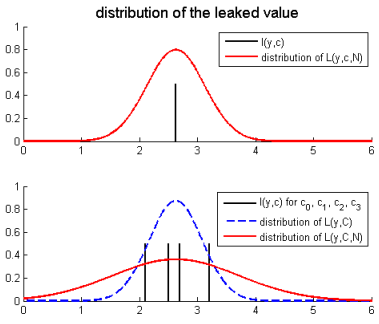

Fig. 4. Example of a template built from 1 chip (top) and 4 chips (bottom).

can simply accumulate these sample means and variances, considering multiple chips rather than a single one. This means using the following estimates:

$$\hat{\Pr}_{\texttt{model}}[L|x] = \mathcal{N}(l|\hat{\mu}_{x,C,N}, \hat{\sigma}^2_{x,C,N}),$$

where $\hat{\sigma}^2_{x,C,N} = \hat{\sigma}^2_{x,c,N} + \hat{\sigma}^2_{x,C,0}$ when (additive) simulated noise is used for N. The left part of figure 5 shows the information leakage of different templates, obtained using different sets of profiling chips. It shows that variability has important consequences for the application of side-channel attacks. First, we see that profiling a large set of chips allows avoiding situations in which the perceived entropy is negative (see, e.g. the "10 vs. 1" curve). But this is at the cost of a reduced information leakage: by inferring the inter-chip variability directly into the templates, one also obtains models that are less accurate for any single chip. This can be observed by the significantly higher information curve of the worst case scenario (denoted as "1 vs. 1 (same chip)"). Second, the right part of the figure illustrates the effect of the selected time sample on the attack: the information decreases both when the input variability decreases (time sample 75 to 110) and when the chip variability increases (time sample 75 to 300).

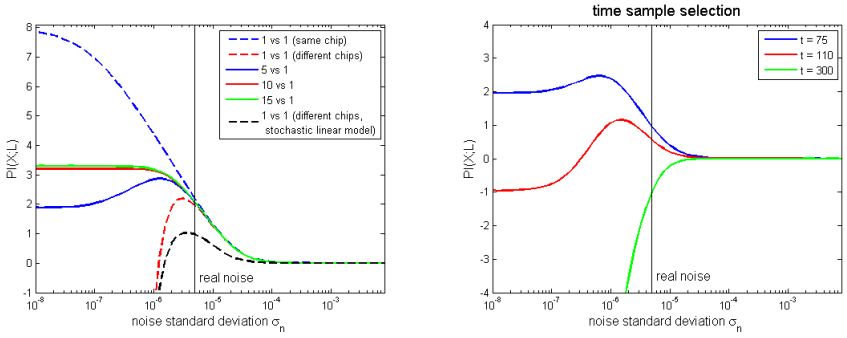

Fig. 5. Left: information theoretic analysis for various sets of learning chips (1.2V supply). Right: information theoretic analysis for various time samples (0.5V supply).

It is essential to properly understand the meaning of these different curves. What they essentially show is that modeling inter-chip variability with a straightforward extension of template attacks (as we did in this section) leads to significant information losses. Hence, it underlines the need to develop new side-channel distinguishers, that can better cope with such situations. One possible solution, discussed in the next section, is to use non-profiled distinguishers and to perform model estimation "on-the-fly", while performing the attacks.

Another important remark is that these experiments do not reduce the relevance of the worst case curve in security evaluations: perfectly profiling one chip and evaluating the perceived information in this context (i.e. the mutual information), remains useful to determine the security limits of an implementation. From a cryptographic designer point of view, the good news is that technology scaling will generally make this limit harder to reach for actual adversaries.

2.4 Model soundness versus DPA soundness

The previous section suggests that inter-chip variability makes the building of accurate templates a challenging task. There exist cases in which the adversary's leakage model is not even sound, in the sense that it leads to negative perceived information values. Following [28], a leakage model is sound if the asymptotic success rate of a Bayesian adversary exploiting it in order to recover a secret target value equals one. In our present case study, the model is sound if all inputs x can be recovered thanks to their corresponding leakage.

This definition of soundness is very strict: a single inversion in the templates (*e.g.* $\hat{\mu}_{x_i,C,N} < \hat{\mu}_{x_j,C,N}$ when $\mu_{x_i,C,N} > \mu_{x_j,C,N}$) is enough for a model not to be sound. However, it is of particular interest for the application of algebraic side-channel attacks [23, 24, 26], in which errors in the leakage information usually makes the solving of the system of equations containing the secret key impossible. As a consequence, we now discuss solutions to obtain sound leakage models.

For this purpose, we use the notion of key class. That is, in the previous section, we always considered leakage models built for all the 256 possible S-box inputs x. However, it is also possible to build less informative models, by considering a lower number of templates. Formally, we define a function $\delta : \mathcal{X} \rightarrow \mathcal{S}$ that maps each input value x to a key class $s = \delta(x)$. The number of key classes can be equal (in the case of an identity mapping $\delta(x) = x$) or lower than the number of possible inputs: $|\mathcal{S}| \leq |\mathcal{X}|$. The mutual information between a mapping variable $S = \delta(X)$ and an input variable X, is defined as:

$$\mathrm{I}(X; S) = \mathrm{H}[S] - \mathrm{H}[S|X] = \mathrm{H}[S].$$

Given a key class variable S, it is then possible to check the soundness of the corresponding leakage model with the conditional entropy matrix defined in [28]:

$$\begin{aligned} \hat{\mathbf{H}}_{s,s^*} &= -\sum_{l \in \mathcal{L}} \hat{\Pr}_{\mathtt{chip}}[l|s] \log_2 \hat{\Pr}_{\mathtt{model}}[s^*|l], \\ &= \begin{pmatrix} \hat{h}_{1,1} & \hat{h}_{1,2} & \dots & \hat{h}_{1,|\mathcal{S}|} \\ \hat{h}_{2,2} & \hat{h}_{2,2} & \dots & \hat{h}_{2,|\mathcal{S}|} \\ \dots & \dots & \dots & \dots \\ \hat{h}_{|\mathcal{S}|,1} & \hat{h}_{|\mathcal{S}|,2} & \dots & \hat{h}_{|\mathcal{S}|,|\mathcal{S}|} \end{pmatrix}, \end{aligned}$$

where s and s^* respectively denote the correct key class and a key class candidate. The model is sound if and only if the minimum value for each line of the matrix is the diagonal value $\hat{h}_{i,i}$. Having defined these tools, we can study the tradeoff between the informativeness of a key class $\mathrm{I}(X; S)$ and the soundness of the corresponding leakage model $\hat{\Pr}_{\mathtt{model}}[L|s]$. For this purpose, we considered consecutive key classes with $|\mathcal{S}| = 256, 255, \dots 1$, with the mapping function δ grouping close leakages, and the templates built from a set of 5 chips, as illustrated in the left part of Figure 6. The right part of the figure then shows how

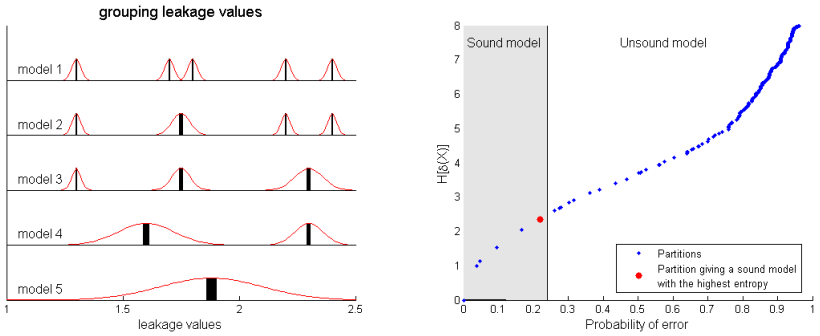

Fig. 6. Left: building more robust / less informative models. Right: maximum information provided by a model in function of the classification error rate (1.2V supply).

the informativeness and soundness of these successive key classes evolves with the error probability of the template attack that we define as:

$$\Pr_{\mathtt{error}} = \Pr[\underset{s^*}{\mathrm{argmax}}\ \hat{\Pr}_{\mathtt{model}}[s^*|l] \neq s]. \tag{4}$$

It illustrates that it is possible to build a key class with 6 possible values for which the model $\hat{\Pr}_{\mathtt{model}}[L|s]$ is sound. Such a key class could be directly used in an algebraic side-channel attack. We note, however, that the same comment as in the previous section applies. Namely, classical template attacks are probably not the best solution to build sound and informative leakage models.

To conclude this section, we finally mention that model soundness is a necessary condition for successful algebraic side-channel attacks. But for standard DPA types of attacks [15], it is only a sufficient condition. That is, standard DPA attacks will generally exploit the leakage corresponding to the execution of the S-box for several plaintexts, i.e. $\mathsf{S}(x \oplus k)$, in order to recover the secret key k. Hence, from an information theoretic point of view, the relevant metric is no more $\hat{\mathrm{PI}}(X;L)$ but $\hat{\mathrm{PI}}(K;X,L)$, with conditional entropy matrix:

$$\hat{\mathbf{H}}_{k,k^*} = -\sum_{x\in\mathcal{X}} \Pr[x] \sum_{l\in\mathcal{L}} \hat{\Pr}_{\mathtt{chip}}[l|x,k] \log_2 \hat{\Pr}_{\mathtt{model}}[k^*|l,x],$$

where k and k^* denote the correct key and a key candidate. As each line of the matrix is computed by averaging over 256 possible inputs x, even some misclassified traces do not always prevent a successful DPA. In other words, DPA-soundness, corresponding to a matrix $\hat{\mathbf{H}}_{k,k^*}$ with minimum diagonal, is a much weaker requirement than model soundness. In our setting, even the identity mapping $\delta(x) = x$ gave rise to successful DPA attacks using templates. The analysis of this scenario will be investigated in the next section.

3 Security analysis

The previous section provided an extensive evaluation of the information leakage of an AES S-box implemented in a 65-nanometer CMOS technology. It shows that inter-chip variability makes the straightforward application of profiled attacks (such as using templates, or stochastic models) less efficient than when variability can be neglected. In this section, we perform the second part of the evaluation framework in [28], i.e. security analysis. For this purpose, we analyze the success rates of various distinguishers in a standard DPA setting, in order to determine the impact of variability in this context. Namely, we performed:

- Template attacks, using exactly the profiling described in Section 2.3.
- Correlation attacks [2] with a Hamming weight leakage model[3].
- Mutual Information Analysis (MIA) attacks, using a Hamming weight leakage model and an identity leakage model (corresponding to 7 out of 8 S-box output bits). Our implementation of MIA followed the guidelines given in [8]: we used histogram-based pdf estimation, with 32 (linearly-spaced) bins, which allowed us to deal with the weak accuracy of our leakage models.
- Non-profiled attacks using stochastic models generated "on-the-fly" , with the linear bases described in Section 2.2. That is, we used exactly the profiling techniques proposed in [27], but this profiling was done for each key candidate separately. The attack then proceeds as carefully described in [12]: each time the adversary gains a new trace, he repeats the profiling and tests his key dependent models directly on the set of available traces.

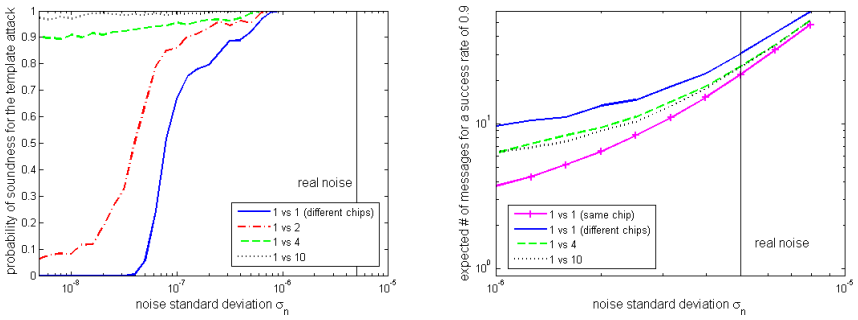

Fig. 7. Left: Probability of soundness for a template attack. Right: Expected number of messages to reach a 0.9 success rate for a sound template attack (1.2V supply).

Figure 7 illustrates the effect of variability on template attacks, for a 1.2V supply voltage. Its left part shows that attacks may not work at all, when the profiling done for a chip is used to attack a significantly different chip. But as

[3] This is equivalent to a Hamming distance model, corresponding to the transitions between a pre-charge of the S-box input to zero and its evaluation on input x.

discussed in the previous section, one reaches DPA-soundness easily, by profiling on more than 4 chips. The right part of the figure shows that the number of traces to attack is low once a sound model is available, and lower bounded by the worst case curve corresponding to a perfect model. Moving to 0.5V supply would lead to similar conclusions, with both curves slightly translated on the right.

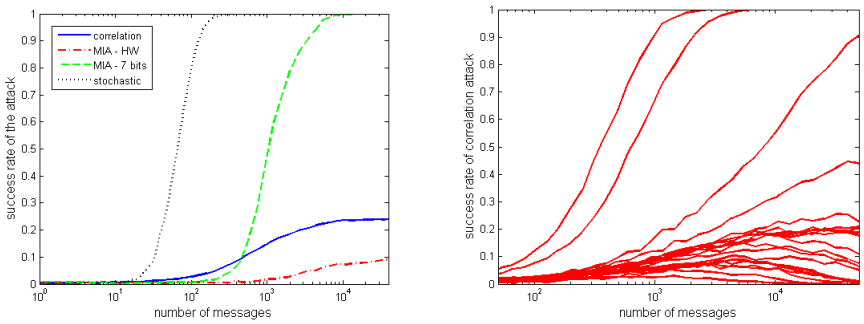

Fig. 8. Left: success rates of a various non profiled attacks averaged over 20 chips. Right: success rates of the correlation attacks for each of the 20 chips (1.2V supply).

Figure 8 shows the results of various non-profiled attacks. Its right part contains results of a correlation attack against each of our 20 chips. It underlines that the Hamming weight leakage model only allows to reach a 100% success rate for a few of these chips. Hence, it cannot be used to evaluate the security of a cryptographic implementation in this case. The left part of the figure illustrates the average success rates (over the 20 chips) of our different non-profiled attacks. It suggests that attacks performing "on-the-fly" model estimation, such as using stochastic models or MIA, are a promising approach for dealing with variability.

4 Conclusions and open problems

Process variability in nanoscale devices raises new challenges for the theory and practice of side-channel attacks. Experiments performed on 20 prototype chips show that former DPA attacks are not perfectly adequate to evaluate the security of an implementation in this context. In the absence of variability, adversaries could first profile (or assume) a leakage model, and then exploit this model in an online attack. With the increase of process variability, it becomes necessary to infer the correct leakage model for each target implementation. The deep integration of recent microelectronic circuits also increases the coupling between their interconnected parts, with consequences on the "independent leakage" assumption that is frequently required in formal works on leakage resilience. Hence, developing new techniques to deal with this new attacks scenario is essential, in order to avoid overestimating the security levels of actual products.

Acknowledgements. M. Renauld is a PhD student funded by the Walloon region SCEPTIC project. F.-X. Standaert is an associate researcher of the Belgian fund for scientific research (FNRS - F.R.S). N. Veyrat-Charvillon is a post-doctoral researcher funded by the Walloon region SCEPTIC project. D. Kamel is a PhD student funded by the Walloon region E.USER project.

References

1. K.A. Bowman, X. Tang, J.C. Eble, J.D. Menldl, *Impact of extrinsic and intrinsic parameter fluctuations on CMOS circuit performance*, IEEE Journal of Solid State Circuits, vol 35, issues 8, pp 1186-1193, August 2002.
2. E. Brier, C. Clavier, F. Olivier, *Correlation Power Analysis with a Leakage Model*, in the proceedings of CHES 2004, Lecture Notes in Computer Science, vol 3156, pp 16-29, Boston, Massachussets, USA, August 2004.
3. S. Chari, J. Rao, P. Rohatgi, *Template Attacks*, in the proceedings of CHES 2002, Lecture Notes in Computer Science, vol 2523, pp 13-28, Redwood Shores, California, USA, August 2002.
4. S. Dziembowski, K. Pietrzak, *Leakage-Resilient Cryptography*, in the proceedings of FOCS 2008, pp 293-302, Philadelphia, Pennsylvania, USA, October 2008.
5. S. Faust, T. Rabin, L. Reyzin, E. Tromer, V. Vaikuntanathan, *Protecting Circuits from Leakage: the Computationally-Bounded and Noisy Cases*, in the proceedings of Eurocrypt 2010, Lecture Notes in Computer Science, vol 6110, pp 135-156, Nice, France, May 2010.
6. K. Gandolfi, C. Mourtel, F. Olivier, *Electromagnetic analysis: Concrete results*, in the proceedings of CHES 2001, Lecture Notes in Computer Science, vol 2162, pp 251-261, Paris, France, May 2001.
7. S. Ghosh, K. Roy, *Parameter Variation Tolerance and Error Resiliency: New Design Paragigm for the Nanoscale Era*, in the Proceedings of the IEEE, vol 98, num 10, pp 1718 - 1751, October 2010.
8. B. Gierlichs, L. Batina, P. Tuyls, B. Preneel, *Mutual Information Analysis*, in the proceedings of CHES 2008, Lecture Notes in Computer Science, vol 5154, pp 426-442, Washington DC, USA, August 2008.
9. D. Kamel, F.-X. Standaert, D. Flandre, *Scaling Trends of the AES S-box Lower Power Consumption in 130 and 65 nm CMOS Technology Nodes*, in the proceedings of ISCAS 2009, Taipei, Taiwan, May 2009.
10. D. Kamel, C. Hocquet, F.-X. Standaert, D. Flandre, D. Bol, *Glicth-Induced Within Die Variations of Dynamic Energy in Voltage-Scaled Nano-CMOS Circuits*, in the proceedings of ESSCIRC 2010, Seville, Spain, September 2010.
11. P.C. Kocher, J. Jaffe, B. Jun, *Differential Power Analysis*, in the proceedings of Crypto 1999, Lecture Notes in Computer Science, vol 1666, pp 388-397, Santa Barbara, California, USA, August 1999.
12. K. Lemke-Rust, *Models and Algorithms for Physical Cryptanalysis*, PhD Thesis, Ruhr University Bochum, Germany, June 2007.
13. L. Lin, W. Burleson, *Analysis and Mitigation of Process Variation Impacts on Power-Attack Tolerance*, in the proceedings of DAC 2009, pp 238-243, San Francisco, CA, USA, July 2009.
14. S. Mangard, E. Oswald, T. Popp, *Power Analysis Attacks*, Springer, 2007.
15. S. Mangard, E. Oswald, F.-X. Standaert, *One for All - All for One: Unifying Standard DPA Attacks*, Cryptology ePrint Archive: Report 2009/449.

16. N. Mentens, L. Batina, B. Preneel, and I. Verbauwhede, *A Systematic Evaluation of Compact Hardware Implementations for the Rijndael SBOX*, in the proceedings of CT-RSA 2005, Lecture Notes in Computer Science, vol 3376 pp 323-333, San Francisco, California, USA, March 2005.
17. S. Nassif, K. Bernstein, D.J. Frank, A. Gattiker, W. Haensch, B.L. Ji, E. Nowak, D. Pearson, N.J. Rohrer, *High Performance CMOS Variability in the 65nm Regime and Beyond*, in the proceedings of IEDM 2007, pp 569-571, Washington DC, USA, December 2007.
18. A.K. Nieuwland, A. Katoch, M. Meijer, *Reducing Cross-Talk Induced Power Consumption and Delay*, in the proceedings of PATMOS 2004, Lecture Notes in Computer Science, vol 3254, pp 179-188, Santorini, Greece, September 2004.
19. K. Pietrzak, *A Leakage-Resilient Mode of Operation*, in the proceedings of EUROCRYPT 2009, Lecture Notes in Computer Science, vol 5479, pp 462-482, Cologne, Germany, April 2009.
20. K. Pietrzak, *Provable Security for Physical Cryptography*, in the proceedings of WEWORC 2009, Graz, Austria, July 2009.
21. J.-J. Quisquater, D. Samyde, *ElectroMagnetic Analysis (EMA): Measures and Counter-Measures for Smart Cards*, in the proceedings of E-smart 2001, Lecture Notes in Computer Science, vol 2140, pp 200-210, Cannes, France, September 2001.
22. F. Regazzoni, A. Cevrero, F.-X. Standaert, S. Badel, T. Kluter, P. Brisk, Y. Leblebici, P. Ienne, *A Design Flow and Evaluation Framework for DPA-Resistant Instruction Set Extensions*, in the proceedings of CHES 2009, Lecture Notes in Computer Science, vol 5747, pp 205-219, Lausanne, Switzerland, September 2009.
23. M. Renauld, F.-X. Standaert, *Algebraic Side-Channel Attacks*, in the proceedings of Inscrypt 2009, Lecture Notes in Computer Science, vol 6151, pp 393-410, Beijing, China, December 2009.
24. M. Renauld, F.-X. Standaert, N. Veyrat-Charvillon, *Algebraic Attacks on the AES: Why Time also Matters in DPA*, in the proceedings of CHES 2009, Lecture Notes in Computer Science, vol 5747, pp 97-111, Lausanne, Switzerland, September 2009.
25. K. Roy, S. Mukhopadhyay, H. Mahmoodi-Meimand, *Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits*, Proceedings of the IEEE, vol 91, issue 2, pp 305-327, April 2003.
26. Y. Oren, M. Kirschbaum, T. Popp, A. Wool, *Algebraic Side-Channel Analysis in the Presence of Errors*, in the proceedings of CHES 2010, Lecture Notes in Computer Science, vol6225, pp 428-442, Santa Barbara, California, USA, August 2010.
27. W. Schindler, K. Lemke, C. Paar, *A Stochastic Model for Differential Side Channel Cryptanalysis*, in the proceedings of CHES 2005, Lecture Notes in Computer Science, vol 3659, pp 30-46, Edinburgh, Scotland, August 2005.
28. F.-X. Standaert, T.G. Malkin, M. Yung, *A Unified Framework for the Analysis of Side-Channel Key Recovery Attacks*, in the proceedings of Eurocrypt 2009, LNCS, vol. 5479, pp. 443-461, Cologne, Germany, April 2009.
29. F.-X. Standaert, C. Archambeau, *Using Subspace-Based Template Attacks to Compare and Combine Power and Electromagnetic Information Leakages*, in the proceedings of CHES 2008, Lecture Notes in Computer Science, vol 5154, pp 411-425, Washington DC, USA, August 2008.
30. F.-X. Standaert, N. Veyrat-Charvillon, E. Oswald, B. Gierlichs, M. Medwed, M. Kasper, S. Mangard, *The World is Not Enough: Another Look on Second-Order DPA*, to appear in the proceedings of Asiacrypt 2010, Lecture Notes in Computer Science, vol 6477, pp 112-129, Singapore, December 2010.

31. K. Tiri, I. Verbauwhede, *Securing encryption algorithms against dpa at the logic level: Next generation smart card technology*, in the proceedings of CHES 2003, Lecture Notes in Computer Science, vol 2779, pp 125-136, Cologne, Germany, September 2003.

A Exemplary leakage traces

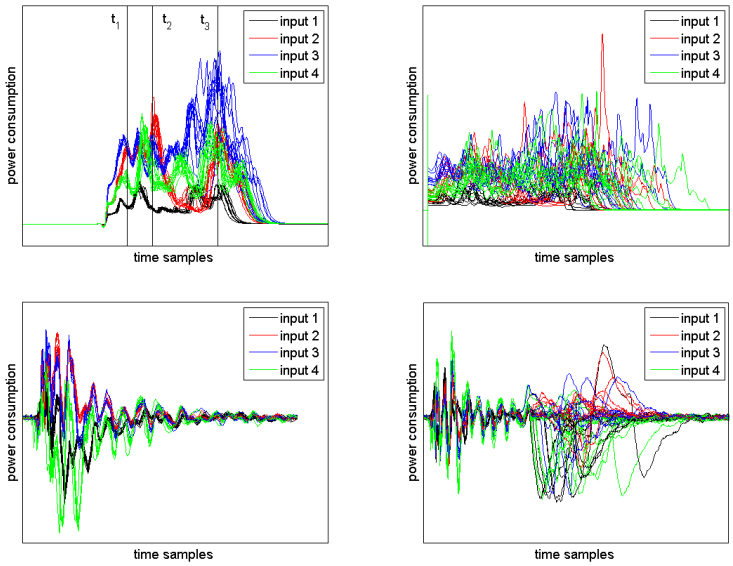

Fig. 9. Illustrative noise-free leakage traces corresponding to 4 inputs and 10 chips. Left: 1.2V supply, Right: 0.5V supply / Up: simulations, Down: actual measurements.

B Preprocessing of the traces

Side-channel attacks exploit information about the internal state of a computing device that is leaked, e.g. through power consumption traces. These power traces also contain some noise, either due to unpredictable physical effects, or to other forms of perturbations such as measurement artifacts (outliers) or parasitic signals (interference). The influence of this second source of noise can sometimes be reduced by processing the power traces prior to the actual attack.

The top curve in the left part of Figure 10 illustrates the autocorrelation of a power trace. It measures the linear correlation between the trace and its shift by τ time samples. The autocorrelation function shows two mixed components: regularly spaced peaks (around $\tau = 1000$, 2000, etc.) and some periodic sinusoidal component with period close to 2000. Roughly speaking, the correlation

peaks which correspond to successive clock cycles of the chip, correspond to the useful signal. By contrast, the periodic sinusoidal component is a parasitic that can be filtered. An example of filtered trace is given in the bottom curve of the left part of Figure 10. And the impact of this preprocessing on the distribution of the residual noise (estimated with histograms) is in the right part of the figure. It clearly illustrates the gain in terms of noise standard deviation.

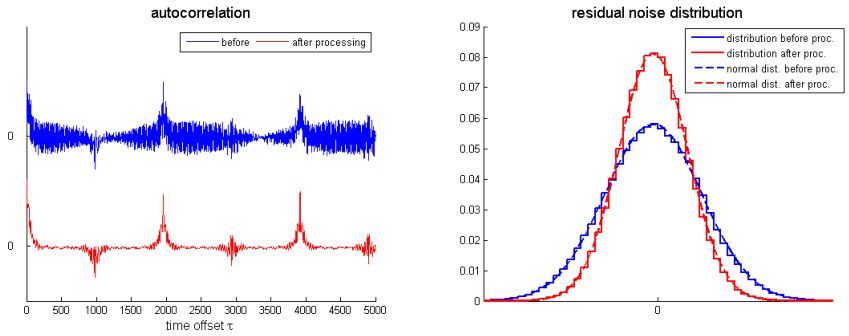

Fig. 10. Left: autocorrelation of the signal before (top) and after (bottom) preprocessing of the traces. Right: residual noise distribution with/without preprocessing.

C PCA and LDA eigenvectors

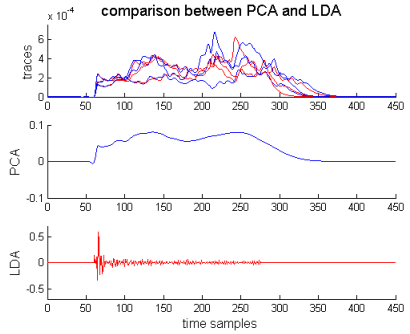

Fig. 11. PCA (middle) and LDA (below) applied to simulated traces at 1.2V (above).