Generic Side-Channel Distinguishers: Improvements and Limitations

Nicolas Veyrat-Charvillon* and François-Xavier Standaert**

UCL Crypto Group, Université catholique de Louvain
Place du Levant 3, B-1348, Louvain-la-Neuve, Belgium
{nicolas.veyrat,fstandae}@uclouvain.be

Abstract. The goal of generic side-channel distinguishers is to allow key recoveries against any type of implementation, under minimum assumptions on the underlying hardware. Such distinguishers are particularly interesting in view of recent technological advances. Indeed, the traditional leakage models used in side-channel attacks, based on the Hamming weight or distance of the data contained in an implementation, are progressively invalidated by the increased variability in nanoscale electronic devices. In this paper, we consequently provide two contributions related to the application of side-channel analysis against emerging cryptographic implementations. First, we describe a new statistical test that is aimed to be generic and efficient when exploiting high-dimensional leakages. The proposed distinguisher is fully non-parametric. It formulates the leakage distributions using a copula and discriminates keys based on the detection of an "outlier behavior". Next, we provide experiments putting forward the limitations of generic side-channel analysis in advanced scenarios, where leaking devices are protected with countermeasures. Our results exhibit that all non-profiled attacks published so far can sometimes give a false sense of security, due to incorrect leakage models. That is, there exists settings in which an implementation is secure against such non-profiled attacks and can be defeated with profiling. This confirms that the evaluations of cryptographic implementations should always consider profiling, as a worst case scenario.

1 Introduction

Since the introduction of differential power analysis by Kocher, Jaffe and Jun in the late 1990s [13], physical attacks have become an important issue for the security of cryptographic devices. On the academic side, it gave rise to many exciting developments of new attacks, countermeasures and models for the evaluation (or, recently, proof) of physical security. On the industrial side, security against such attacks is now required to reach high certification levels for cryptographic products. Roughly speaking, side-channel attacks are usually classified

* Postdoctoral researcher supported by the Walloon region SCEPTIC project.
** Associate researcher of the Belgian Fund for Scientific Research (FNRS-F.R.S.).

P. Rogaway (Ed.): CRYPTO 2011, LNCS 6841, pp. 354–372, 2011.

as profiled or non-profiled [15]. Profiled attacks are typically useful in an evaluation context, where one can exploit devices with known keys in order to build precise leakage models, e.g. templates [5]. They can then be used in order to estimate evaluation metrics such as the mutual information, in order to capture a worst-case scenario [28]. By contrast, non-profiled attacks, on which we will focus in this paper, rather aim to capture the behavior of actual adversaries, who do not have a precise prior characterization of the devices they target. They are usually suboptimal from a data complexity point of view, and exploit the "on-the-fly" estimation of the leakage probability distributions (or their moments) in order to recover secret information. In general, the gap between these two scenarios can be large. Hence, profiled and non-profiled attacks are complementary and shed a different light on the security of embedded devices.

In brief, a non-profiled side-channel attack generally works by comparing key-dependent leakage models with actual measurements. If the attack is successful, the key candidate giving rise to the best comparison is the one manipulated by the target device. As a consequence, evaluating the performances of such a distinguisher can typically be done along two axes. On the one hand, attacks are expected to be efficient, meaning that they allow recovering keys with limited data (i.e. measurements), time and memory. On the other hand, attacks should also be generic, i.e. applicable against any type of device and (if possible) unsensitive to unprecise leakage models. A brief look at the state-of-the-art suggests that most previous works can be viewed as exploring the tradeoff between these two (usually contradictory) goals. For example, in the seminal paper of Crypto 1999, a "single-bit DPA" is performed using a simple difference-of-means test. This approach has limited efficiency, because the single-bit model implies a low SNR, and it cannot take advantage of any knowledge of the target device that may be available to the adversary. As a consequence, many "multiple bit" extensions have been proposed in the literature. Most prominently, the Correlation Power Analysis (CPA) introduced in 2004 works very efficiently in contexts where a device leaks according to a well known and linear (e.g. Hamming weight) model [2]. But it is also quite specific, and can end up to be completely ineffective if an unprecise leakage model is used. In order to avoid this model-dependency, a powerful solution is to build stochastic models "on-the-fly", as suggested by Schindler, Lemke and Paar in 2005 [26]. A possible drawback of stochastic models is their parametric nature. But as discussed in [6], such a linear regression-based approach gives excellent results in the context of first-order side-channel attacks against unprotected devices. Alternatives to stochastic models include the Mutual Information Analysis, introduced in 2008 [8], and tests such as the Cramér-von-Mises one, discussed in [31]. These last two distinguishers are quite generic, and can capture any type of leakage dependency. They are also non-parametric in the case of MIA (which still implies pdf estimation, e.g. with histograms or Kernels, hence requiring to select number of bins or Kernel bandwidth adequately [19]) and completely free of parameters for the Cramér-von-Mises test.

Interestingly, recent technological advances suggest that the genericity of side-channel distinguishers could be an important feature in the evaluation of future

cryptographic devices. Indeed, as discussed in [24], the move towards nanoscale electronic circuits implies the apparition of new leakage functions, that strongly deviate from the traditional (Hamming weight, distance) assumptions. Also, the increasing device variability implies that each target implementation can be characterized by a different leakage model. Quite naturally, the situation turns out to be even nastier when moving to higher-order attacks against devices protected with masking [4,10]. Indeed, straightforward extensions of DPA and CPA require the introduction of a heuristic dimensionality reduction technique, usually denoted as the combination function in the literature [17]. But as discussed in [20,29], the selection of a good combination function is inherently dependent on the leakage function (i.e. the target device), and can only degrade the amount of information exploited by the distinguisher. Similarly, extensions of the stochastic model are direct in the profiled context [14], when masks are available during the profiling, but their application in a non-profiled scenario requires either to estimate pdf mixtures in an unsupervised manner (i.e. a problem for which we do not have systematic and efficient solutions), or to take advantage of some heuristic assumptions (e.g. using a combination function). In fact, only MIA directly generalizes to multivariate side-channel attacks, without requiring a combination function [1,7,19]. Given this attractive feature, it appears natural to investigate how such distinguishers can deal with advanced scenarios, mixing non-linear leakage functions and countermeasures like masking.

This paper presents two contributions in this direction. First, we propose a new test for side-channel analysis, aimed to be generic and efficient when exploiting high-dimensional leakages. For this purpose, we start from the observation that, in order to be generic, MIA selects the key candidate that maximizes the mutual information between an adversary's key-dependent leakage models and actual measurements. Our new distinguisher uses the alternative criterion to select the key candidate for which these leakage models deviate the most from a reference (e.g. uniform) distribution. Next, it has been observed in experiments on MIA that the generalization to multivariate attacks can be less efficient, because of the difficulty of estimating a multivariate pdf "on-the-fly", without specific assumptions [29]. Also, previously considered estimation methods such as using histograms, Kernels, or splines [30], generally require to tune a parameter, e.g. the number of bins in histograms, that directly impacts the efficiency of the attack. Hence, although MIA aims at genericity more than efficiency, it would be interesting to avoid such parameters, or to make their tuning as easy as possible. Our new distinguisher takes advantage of advanced statistical tools in order to mitigate these issues. For this purpose, we first apply a leakage transformation, exploiting copulas [18]. It projects the samples into a new space where their distribution (among each dimension) is uniform. Thanks to this transform, we base our distinguisher on the generic criterion of selecting the key candidate for which the model maximizes the deviation from uniform. Afterwards, we exploit distance sampling, i.e. we evaluate the distribution of the distance between two samples (conditioned on a leakage model), rather than the distribution of single samples. Distance sampling has interesting features in a multivariate setting,

as it allows to avoid dealing with multivariate distributions directly (we rather evaluate the univariate distribution of a distance taken over several dimensions). As a result, our test is completely free of parameters and mainly requires to compute empirical cumulative distributions, for each dimension taken independently. Summarizing, it is pointed out in [34] that the efficiency loss of MIA is due to the problem of estimating the leakage distributions. The present paper complements this view and aims at making this estimation step easier.

Second, we propose different experiments in order to evaluate this new distinguisher. Namely, we investigated attacks against unprotected and masked S-box computations, in three different settings: naive Hamming weight simulations, real measurements of a 65 nanometer CMOS chip, Spice simulations of a dual-rail logic style. These examples are expected to be reflective of the variety of leakage functions that one can find in side-channel analysis. They highlight that the proposed generic test compares favorably with MIA, in all investigated scenarios. They also underline that, despite the generic nature of MIA and the new distinguisher, their application against modern electronic devices can be strongly affected by inaccurate leakage models, and that the impact of such imprecisions is amplified with countermeasures such as masking. In other words, even generic tests can be unsuccessful in certain contexts, unless preliminary assumptions (similar to profiling) are available to the adversary. We note that this last observation holds for all non-profiled distinguishers published so far. Hence, our results raise the question whether non-profiled attacks could be improved in order to deal with such critical contexts, or alternatively, whether these contexts can be formalized and used as a design criteria for new countermeasures. For now on, they at least confirm the importance of a profiled information theoretic analysis in the evaluation of leaking cryptographic devices.

2 Side-Channel Analysis

The next sections of the paper analyze the attack depicted in Figure 1, as described in [1]. That is, we consider a device performing several cryptographic computations $\mathsf{E}_k(p)$ on different plaintexts p drawn uniformly from the text space $\mathcal{P}$, using some fixed key k drawn uniformly from the key space $\mathcal{K}$. While computing $\mathsf{E}_k(P)$ (where P is a random variable over $\mathcal{P}$), the device will handle some intermediate values (defined as sensitive variables in [25]) that depend on the known input P and the unknown key k. In practice, the interesting sensitive variables in a DPA attack are the ones that only depend on an enumerable subkey s: we denote them as $V_{s,P}$. Anytime such a sensitive intermediate value is computed, the device generates some physical leakage, denoted as $Y_{k,P}$.

In order to perform a key recovery, an adversary first has to select a sensitive value. Given that this value only depends on a subkey s, he can then evaluate its result for the same plaintexts that have been used to generate $Y_{k,P}$ and all the possible subkey candidates $j \in \mathcal{S}$. It gives rise to different hypothetical values $V_{j,P}$. Afterwards, he uses a leakage model to map these values from their original space $\mathcal{V}$ towards a hypothetical leakage space $\mathcal{X}$. It is usually in this step that

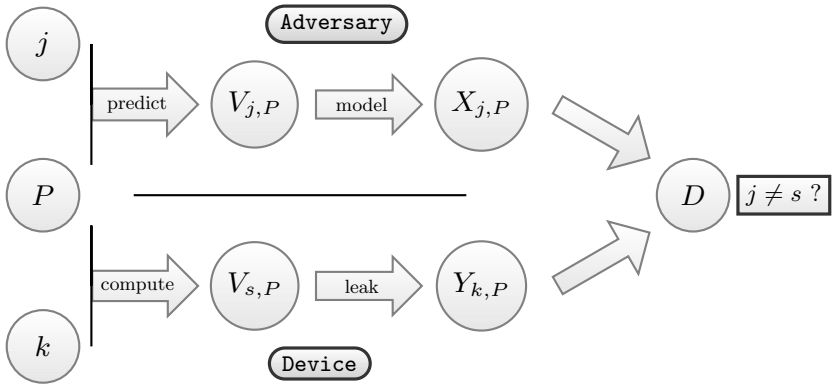

Fig. 1. Schematic illustration of a side-channel key recovery attack

engineering intuition can be exploited, if available. For example, a usual model that has been experimentally confirmed, e.g. in [15], is to take the Hamming weight of the values $V_{j,P}$. Such a model is justified by the dominating dynamic part of the power consumption in certain microelectronic devices. As a result, the adversary obtains $|\mathcal{S}|$ different models denoted as $X_{j,P}$, again corresponding to the different subkey candidates. Eventually, he uses a distinguisher D to compare the different models $X_{j,P}$ with the actual leakages $Y_{k,P}$. If the attack is successful, the best comparison result (*i.e.* the highest value of the distinguisher) should be obtained for the correct subkey candidate $j = s$. This procedure can then be repeated for different subkeys in order to eventually recover the full key.

3 The New Generic Test

A central problem in non-profiled side-channel analysis is to properly estimate the leakage distribution "on the fly" during an attack. Previous methods for this purpose typically range in two categories. A first class of distinguishers exploits specific assumptions about the target implementation, resulting in efficient attacks provided that these assumptions are fulfilled. Another class of distinguishers aims to avoid relying on assumptions, at the cost of a (hopefully small) efficiency loss. Quite naturally, the genericity of this second class of distinguishers essentially comes from that they try to completely characterize the leakage distribution (rather than its first- or second-order moments, typically). And its efficiency loss comes from the difficulty of the density estimation problem. For example, tools for estimating pdf usually rely on a good choice of parameters: number of modes in Gaussian mixtures, number of bins in histograms, bandwidth in Kernel estimators, to name a few. Also, these estimators strongly suffer from the curse of dimensionality: a 9-bin univariate histogram will typically require 81 bins in two dimensions. In the remainder of the section,

we first describe a distinguisher, illustrated in Figure 2, that aims to limit these drawbacks. Next, we discuss its advantages and limitations.

3.1 Specification

The distinguisher is based on six main steps (one being optional).

Leakage space transform. In order to circumvent the problem of estimating the leakage distribution, our method first transforms the samples by means of copula. A copula simply applies the probability integral transform to every marginal variable, which renders the distribution of samples along each axis uniform. More precisely, for $Y = (Y_1, Y_2, \ldots, Y_d)$ a d-dimension random variable, the copula transformation gives a derived variable $Z = (Z_1, Z_2, \ldots, Z_d) = (F_1(Y_1), F_2(Y_2), \ldots, F_d(Y_d))$, where F_i is the cumulative distribution function of Y_i, defined by $F_i(y_i) = \Pr[Y_i \leq y_i]$. Interestingly, the Z_i have a uniform distribution by definition of the probability integral transform, but any dependency among the Y_i variables implies a corresponding dependency among the Z_i's. This is illustrated in Figure 2, where it is clearly seen that the marginal distribution $\Pr[Z = z]$ is uniform after reduction, while the conditional distributions remain easily distinguishable. For illustration, the figure represents the simple case of a single-bit leakage model. In practice, since the leakage density and its cumulative function are both unknown, we compute an empirical copula, i.e. a copula where the cumulative distributions are approximated by empirical distributions. For an n-sample set $y_1, \cdots y_n$ drawn from the distribution of a one-dimension random variable Y, the empirical cumulative distribution of a value y is given by $\hat{F}(y) = \frac{1}{n}\sum_i \mathrm{I}(y \leq y_i)$, where I is the indicator function. That is, it only requires to sort the samples and is equivalent to computing the quantile of y for that sample set. Thanks to the Glivenko-Cantelli theorem [3,9], we know that the empirical cumulative distribution function converges *almost surely* towards the common cumulative distribution function, uniformly over all y. That is:

$$\|\hat{F}_n - F\|_\infty \equiv \sup_{y \in \mathbb{R}} \left|\hat{F}_n(y) - F(y)\right| \xrightarrow{a.s.} 0. \tag{1}$$

An advantage of the empirical copula is that reduced samples only take quotient values: $0, \frac{1}{n}, \ldots, \frac{i}{n}, \ldots 1$, built from n real-valued leakages. Hence, it allows dealing with probability mass functions instead of densities, which simplifies computations. In general, the transform amounts to working on modified values $Z_{k,P}$ instead of the leakages $Y_{k,P}$, which have by construction a simple uniform distribution, but still retain the information contained in the original leakages.

Leakage partitioning. This step is common to all standard side-channel attacks. After transformation, the leakage samples are classified, based on predictions $X_{j,P}$ made up from the plaintexts and key hypotheses. Intuitively, the expectation is that predictions obtained from the correct key candidate will lead to a meaningful leakage partition, i.e. there will be a dependency between the categories x and the samples y they contain. By contrast, a wrong key hypothesis should give rise to random predictions, so that the categories only correspond to

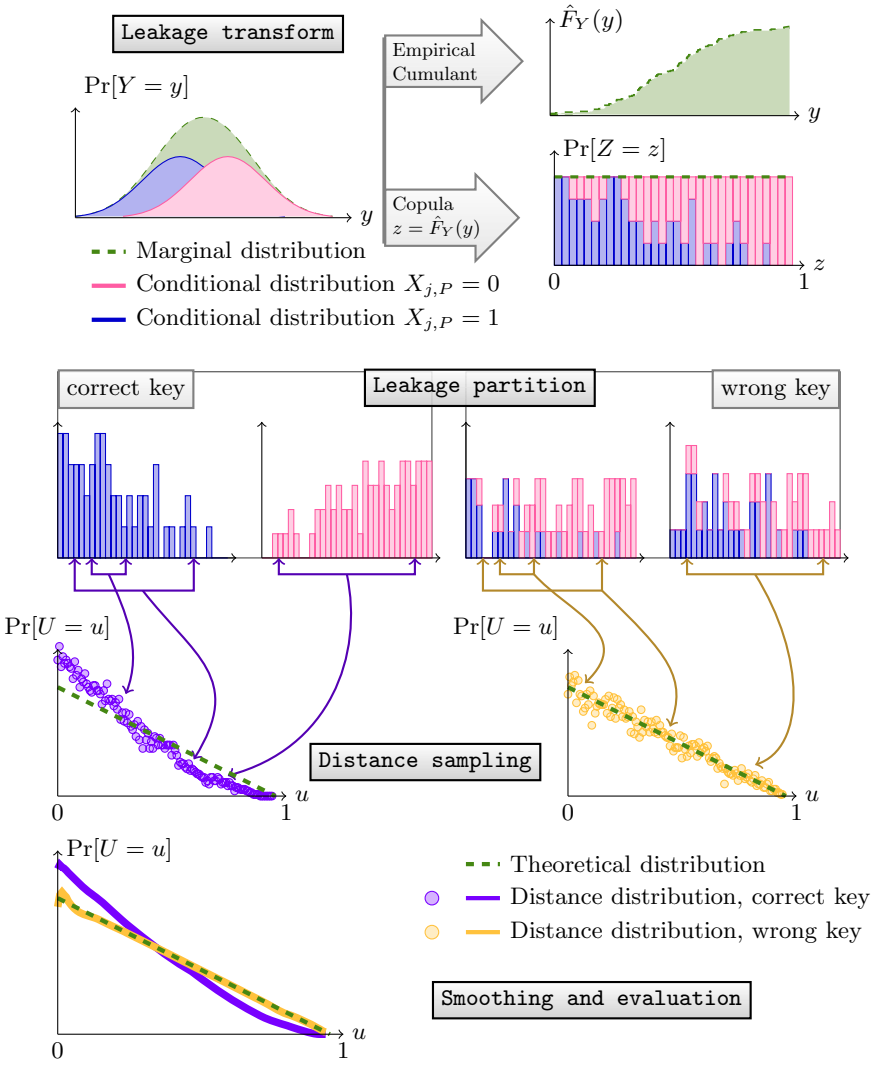

Fig. 2. Illustrated process of the distinguisher (1D)

a random shuffling of leakages from the sample set. Interestingly, thanks to the transformation step, a random sampling should tend towards a uniform distribution, as illustrated in the leakage partition step in Figure 2.

Feature selection and template building. Given the (uniform or not) distributions representing the key candidates, one can select a feature of these distributions, that properly captures possible non-uniformities. In the following, we will consider the distance between couples of samples for this purpose, which corresponds to the notion of spacings in statistics [21]. When dealing with multivariate leakages, we define spacings via the L_1 or Manhattan distance (instead of

the Euclidean one), which avoids dealing with irrational values. The Manhattan distance between two samples z and z' is given by $d_1(z, z') = \sum_i |z_i - z'_i|$, where the sum is taken over the different dimensions of the samples. Next, since the marginal distribution of the leakages is known to be always uniform, we simply build a template for the distribution of the distance between two uniform samples. This template is precomputed once for all attacks, independently of the target implementation. Its shape is actually a spline, which ensure a degree of smoothness (see appendix A for details). We note that the interest of feature selection is not obvious in the univariate attack context of Figure 2. But it implies a dimensionality reduction that becomes convenient in a multivariate setting[1]. It also leads to an efficient solution for the estimation problem, as we now detail.

Estimation. This is the central step of the attack, in which we try to model the distributions corresponding to the different key candidates. For this purpose, one limitation of MIA was the need to estimate one conditional distribution per model value, for each key candidate. In other words, each of these distributions has to be estimated from only a part of the available samples. By contrast, our new test allows to estimate only one distribution, from all the available samples. This is possible because we know (again thanks to the copula) that the marginal distribution should be uniform for a wrong key candidate. Hence, we can sample the Manhattan distance for all partitions, and combine the results into a single probability mass function, which has to be consistent with uniformly drawn samples. Combined with the previous feature selection, it implies that one can estimate the distributions for each key candidate by iterating the next steps:

- Pick a random leakage z from the complete set of samples.
- Pick a different random leakage z', from the same model category as z.
- Compute their Manhattan distance $d_1(z, z')$.

Finally, we obtain the sampled probability mass function of the distance. This distance can only take $n \cdot d$ distinct values, with n the number of samples available and d the dimensionality of the leakages. This is possible because we use the Manhattan distance (i.e. there are as many possible distances as there are samples). Note that if there are m leakages corresponding to a model value $X = x$, the number of possible couples to sample scales as m^2. In our following experiments, it was always possible to test these couples exhaustively. But in attacks requiring millions of traces, exploiting Monte Carlo sampling would of course be an alternative to reduce the time complexity to more tractable values.

Smoothing. As can be seen from Figure 2, the distance histograms can be used to discriminate the different partitions, but they remain quite noisy. In order

[1] Distance sampling is reminiscent of the use of an absolute distance combining function in higher-order side-channel attacks [17]. The prior application of a copula allows us to give it a stronger foundation and to remove its device-dependent flavor.

to improve the signal-to-noise ratio of the pdf estimations, one straightforward solution is to apply a lowpass filter. Different solutions can be used for this purpose. A very generic one, that we applied in our experiments, is to use Kernel smoothing with an Epanechnikov function [12] (which is both the most efficient Kernel and the least costly to compute). The advantage of Kernel smoothing is to generalize easily to distributions with any number of dimensions. Its drawback is to introduce a parameter (the window size used in the smoothing). However, we note that this parameter is significantly easier to select than, e.g. the number of bins or Kernel bandwidth in the case of MIA, since we know exactly how the distribution of the wrong key candidates should look like. Namely, this distribution should correspond to the previously described uniform distance template. How to select this parameter adequately will be explained in the next subsection.

Evaluation. Finally, we compute the integrated square distance between the sample distributions of all the key candidates and the theoretical distance distribution (i.e. the uniform ditance template construted in step 3). The attack is successful if the correct key candidate corresponds to the largest deviation[2].

3.2 Pros and Cons

One important advantage of this new test is that it nicely extends to multivariate attacks. As illustrated in Figure 3 for a bivariate example (i.e. $d = 2$) and a Hamming weight leakage model, the leakage transformation is performed independently for each dimension. That is, we just have to compute d univariate empirical cumulative functions (rather than one d-dimensional distribution when applying MIA). The empirical cumulative functions are straightforward to compute and require no parameter at all. By applying the copula, the leakage samples are transformed in such a way that their marginal distribution along each dimension becomes uniform over $[0, 1]$. As a result, we can again discriminate key candidates by simply assuming that the leakage model generated with the good key candidate should lead to a partitioning for which the distance to uniform is large. On the negative side, the estimation of the distance distribution follows a multinomial distribution (a generalization of the binomial distribution to more than two categories), which implies that the sampled distribution tends to be noisy. The smoothing part that aims at reducing this noise requires to set a parameter that can be seen as the counterpart of the number of bins or Kernel bandwidth when directly trying to estimate the leakage pdf in MIA. However, this task is arguably easier in our new distinguisher, since we only need to detect departures from a well-characterized distribution. In practice, a Kernel smoothing with window size equal to 1 allows us to only retain the general features of the distance distribution, which is enough to detect departures from the uniform template. This is only suboptimal when dealing with very low-noise scenarios, where a smaller window size is enough to smooth out the estimated densities.

[2] As detailed in [31], different alternatives could be considered. A robust one would be to use the smoothed median of the distributions for all key hypotheses as a template.

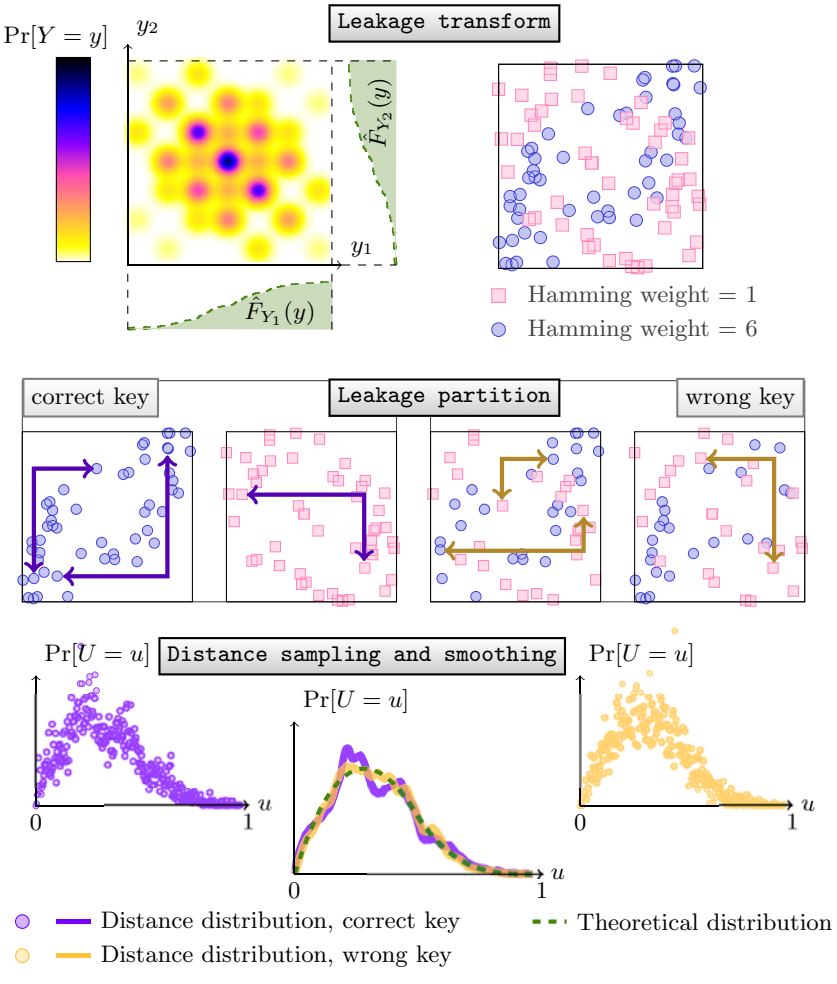

Fig. 3. Illustrated process of the distinguisher (2D)

4 Experiments

We now provide experiments, in order to verify the relevance of the previously introduced generic test and to compare its efficiency with other distinguishers used in side-channel analysis. For this purpose, we consider the following contexts.

1. Different target computations. We investigated three possible cases. In the first one, we target the leakage corresponding to the execution of an unprotected AES S-box, denoted as $v_{k,p} = \mathsf{S}(p \oplus k) \rightsquigarrow y_{k,p}$, where the adversary is provided with the plaintext p and leakage $y_{k,p}$. In the second one, we consider the execution of a masked AES S-box, denoted as $v^1_{k,p} = \mathsf{S}(p \oplus k) \oplus m \rightsquigarrow y^1_{k,p}$, $v^2_{k,p} = m \rightsquigarrow y^2_{k,p}$. In this case, the mask m is a uniformly random value and

the adversary is provided with the plaintext p and leakages $y^1_{k,p}, y^2_{k,p}$. Finally, we consider the execution of a masked S-box where the the adversary only receives the plaintext p and a combination of the two leakage samples. Following the previous analyzes in [20,29], we used the normalized product between the samples, i.e. $\mathsf{C}(y^1_{k,p}, y^2_{k,p}) = (y^1_{k,p} - \hat{\mathbf{E}}(Y^1_{k,P})) \cdot (y^2_{k,p} - \hat{\mathbf{E}}(Y^1_{k,P}))$, where $\hat{\mathbf{E}}$ denotes the sample mean operator.

2. Different leakage functions and target devices. We again analyzed three possible cases. In the first one, the leakages are simulated with a Hamming weight function, to which we added a Gaussian noise, with mean 0 and variance σ_n^2. Although not always realistic, the investigation of this setting is justified by the numerous works carried out under this assumption, as a reference. In the second case, we use the leakage measured from an S-box implemented in a 65 nanometer CMOS technology, running at 2MHz and sampled with a 1 Gsample/sec digital oscilloscope), previously analyzed in [24]. In the third case, the leakage of a dual-rail pre-charged S-box, implemented in the same 65 nanometer technology and using the logic style described in [11], was simulated with Spice. The details of this S-box are out of the scope of this paper, but it was selected as an example of leakage function that strongly deviates from the Hamming weight model. Both for the CMOS and the dual-rail S-boxes, we selected one single leakage sample per target operation. This selection is not supposed to be optimal, but was the same for all the investigated attacks, in order to allow fair comparisons.

3. Different distinguishers. First, correlation attacks based on Pearson's coefficient were applied, following the descriptions in [2]. Next, we performed MIA[3] with histogram-based pdf estimation, following [8]. The number of bins in histograms was selected according to Scott's rule of thumb [27]. Third, we used the stochastic approach first described in [26] and analyzed in the non-profiled setting by Doget et al. [6]. The goal of the stochastic approach is to approximate the S-box leakages with a linear function $\hat{\mathsf{L}}(j,p) = \sum \alpha_i \; g_i(j,p)$, where the coefficients α_i are determined by regression, and the $g_i(j,p)$'s correspond to the base functions used in the attack. Finally, we experimented our new generic test, with the window size in the Kernel smoothing step systematically set to one in the attacks (i.e. a version of the distinguisher completely free of parameters).

4. Different leakage assumptions. As detailed in Section 2, the non-profiled distinguishers studied in this paper need to rely on some preliminary assumptions on the leakage. For correlation attacks, MIA and the new test, we first evaluated the Hamming weight and identity leakage models (where one takes $x_{j,p}$ as the Hamming weight, or the 7 least significant bits of $v_{j,p}$, respectively). These are usual assumptions when performing a side-channel attack. However, as will be discussed in the remaining of the section, these models were not sufficient to perform successful key recoveries in all the investigated contexts. Hence, we additionally used a profiled leakage model in some experiments. Different solutions

[3] To avoid issues like described in [32], we performed a robust variant of MIA, by selecting the subkeys according to their mutual information bias, i.e. the distance between $\hat{\mathrm{I}}(X_{j,P}, Y_{k,P})$ and the median of this quantity, computed over all key candidates.

are possible for this purpose, e.g. exploiting templates [5]. In the following, we built model classes by grouping together transitions leading to similar leakage values (e.g. 9 such groups would appear for an 8-bit Hamming weight model, corresponding to the 9 possible weights). The model groups were built using a K-means clustering algorithm. Again, this selection is not supposed to be optimal, but to serve as a background to discuss generic distinguishers. As for the stochastic approach, one just needs to select the base vectors used in the adversary's predictions. We followed the classical strategy and used the target S-box output bits for this purpose (i.e. a 9-element basis, with 8 bits and a constant).

As our goal is to compare distinguishers, the evaluations we performed followed the security metrics in [28]. For each investigated context, we computed the success rate of the attacks, over a set of 100 to 500 independent experiments. The figures corresponding to these experiments have been reported in appendix. They allow a number of interesting observations that we now detail.

Observation 1. On different types of leakage functions. The different leakage functions considered imply very different constraints for the non-profiled adversaries. In the case of Hamming weight leakages (Figures 4 and 5 left), this function is purely linear, i.e. $\mathsf{L}(k,p) = \sum \alpha_i \; v_{j,p}[i]$, with $v_{j,p}[i]$ the ith bit of the S-box output and all α_i coefficients set to 1. In addition, as shown in [29], the bivariate distribution of the leakages conditioned on the key, for a masked S-box, is accurately characterized by the correlation between the samples $L^1_{k,P}$ and $L^2_{k,P}$ in this case. As a result, all investigated attacks are very efficient. By contrast, when moving to the analysis of real measurements on a 65nm chip (Figures 5 right and 6), the Hamming weight assumption becomes invalid, and the quadratic, cubic, ... terms in $\mathsf{L}(k,p) = \sum \alpha_i \; v_{j,p}[i] + \sum \beta_{i,j} \; v_{j,p}[i] \cdot v_{j,p}[j] + \ldots$ are non-negligible. As a result, attacks using this model are not successful anymore. Interestingly, the stochastic attack using a linear basis is still efficient, confirming the analysis in [24] that the linear terms of the leakage function are still significant. Surprisingly, the correlation attacks in Figure 6 suggests that a masked S-box may be easier to attack than an unmasked one, under a Hamming weight assumption. This is explained by the fact that actual leakages are not accurately predicted by Hamming weights in the unprotected case, whereas their inter-sample correlation remains informative in a second-order attack against a masked S-box. Eventually, the (simulated) dual-rail S-box is an example of implementation with completely non-linear leakages, as witnessed by the impossibility to perform a successful stochastic attack using a linear basis (see Figure 7).

Observation 2. On the limits of generic distinguishers and models. Generic distinguishers are expected to capture any type of leakage dependency. Still, they are dependent on the leakage model used to build the partitions in a side-channel attack. An interesting outcome of our experiments is that these distinguishers are in fact strongly affected by incorrect assumptions. For example, the Hamming weight leakage model does not lead to successful attacks, neither against the 65nm CMOS chip, nor against the dual-rail pre-charged one. More critically, the identity leakage model that is supposed to provide a generic way to target any implementation in [8], is not successful either in certain cases (Fig-

ures 5 right, 7). For Figures 5 right and 7 left, only models obtained through profiling lead to successful key recoveries with the new distinguisher. For Figure 7 right, only template attacks are successful. These limitations are due to the lack of relevance of the leakage models used by the adversary. They are in fact not new: already in 2005, Mangard et al. observed an implementation for which even single-bit leakage models were not accurate enough to perform a successful DPA [16]. More generally, and as also emphasized in [33], MIA-like distinguishers can naturally exploit identity models when applied to non-bijective S-boxes (as in the DES), because such S-boxes imply a meaningful partition by design. But the genericity of this model does not extend to bijective S-boxes (or other block cipher components), excepted if justified by specific implementation choices. In other words, there is no generic leakage model. As discussed in [31], even MIA requires a partitioning such that $\hat{\mathrm{I}}(X_{j,P}; Y_{k,P})$ is maximized for the correct key candidate. The extension of MIA in this paper faces a very similar requirement.

Observation 3. On the limits of non-profiled side-channel attacks. Another consequence of our experiments is to emphasize that, in the context of the dual-rail S-box (Figure 7), none of the non-profiled side-channel attacks could lead to successful key recoveries. Interestingly, the "on-the-fly" stochastic approach fails in this context, even when increasing the size of the basis (e.g. using not only linear, but quadratic, cubic, ... terms). The failure of a stochastic model using linear base vectors is easily explained by the strongly non-linear nature of the simulated leakages produced by the dual-rail S-box. The unsuccessful results with larger bases just derive from the fact that these large bases allow refining the model for all key candidates (i.e. not only the correct one). In fact, also in this case, it is important that the base vectors are justified by a reasonable physical intuition. For example, in case of linear leakage functions, the regression is easy for the correct key candidate (because provided with a good basis) and difficult for the wrong key candidates (because the regression essentially has to capture the non-linearity of a modified S-box S' such that $\mathsf{S}(x \oplus k_w) = \mathsf{S}'(x \oplus k_g)$, with k_g and k_w the good and a wrong key candidate). But as soon as the base vectors do not have a connection with the actual leakages, the advantage of the correct key candidate in building a good stochastic model vanishes[4]. When combining this dual-rail logic style with masking (in Figure 7 right), we see that only a bivariate template attack, similar to the ones in [29], allows recovering keys. In this respect, it is worth noting that a leakage model that is sound in an unprotected setting (e.g. the one based on clustering in Figure 7 left) does not translate into a sound model for the corresponding masked implementation (in Figure 7 right).

Observation 4. MIA versus the new test. Finally, our new distinguisher compares favorably to MIA in all the investigated experiments. We note that the efficiency of MIA could possibly be improved, by exploiting pdf estimation based on

[4] This limitation is only due to the application of the stochastic approach in a non-profiled scenario. In profiled attacks, the stochastic approach remains perfectly sound, as soon as provided with enough base vectors, just as a template attack.

Kernels, splines or parametric techniques [19]. However, more than the efficiency of the distinguisher, it is worth to notice that in certain settings, e.g. the masked 65nm CMOS S-boxes in Figures 5 right and 6 right, it allows exploiting a leakage assumption while MIA cannot. It is an open question to determine whether these experiments can be formally confirmed (e.g. are they due to identified limitations as in the example given in [32], Section 3) or are the result of measurement artifacts that would vanish with more intensive measurement efforts.

5 Conclusion and Open Problems

Generic distinguishers are a useful tool for evaluating leaking devices. In this paper, we first proposed a new and efficient generic test that is fully non-parametric, based on a natural discriminating criterion, and exploits state-of-the-art statistical tools. It can be useful in all scenarios where the previously introduced MIA shows significant advantages over other non-profiled distinguishers.

Next, we discussed the relevance of generic distinguishers in general. Based on experimental validation in different contexts, we put forward that such non-profiled attacks do not get rid of the need of sound assumptions during the partitioning step in a side-channel analysis. Similarly, when applied "on-the-fly", the stochastic approach of Schindler et al. is only sound when provided with meaningful based vectors, that may not be easy to guess for practical adversaries. This observation suggests that the gap between profiled and non-profiled side-channel attacks can be huge, when such assumptions are are not available. It re-emphasizes the need to consider two aspects in the security analysis of a leaking device, as advocated in [28]. First, the worst case security can only be evaluated with a profiled attack (e.g. using templates), and quantified with an information theoretic metric. Second, different types of non-profiled distinguishers can be compared with a security metric, in order to measure how efficiently they can take advantage of the available information. In this respect, generic tests bring an interesting alternative to more specific (e.g. correlation-based) statistical tools. But they are not immune to model imprecisions, and resisting such attacks is not sufficient to conclude that an implementation is secure.

Admittedly, the most critical examples we exhibit in this paper are based on simulations and practical implementations usually show linear dependencies in their leakages. Nevertheless, this discussion underlines that the theoretical limits of present non-profiled attacks have to be properly understood. It also leads to interesting questions for the design and analysis of secure implementations. First, as non-profiled distinguishers published in the literature seem specially affected by non-linear leakage functions, designing hardware logic styles with this criterion in mind appears as an interesting scope for further research. Preliminary experiments reported in [23] suggest that the DDSLL logic style may not be the most suitable for this purpose. Second, the partitioning step of generic distinguishers is made specially easy when non-bijective S-boxes (or components) are used. Hence, such S-boxes should be avoided when designing ciphers to be secured against physical attacks. In the same lines, targeting the

output of MixColumn in an AES implementation could be an interesting topic for further investigation. Depending on the architectures (e.g. 8-bit software or 32-bit hardware), it could also lead to leakage models that are simple to exploit. Eventually, non-profiled security evaluations are typically misleading when randomization-based countermeasures such as masking are combined with the difficulty to make sound assumptions on the leakage model. Hence, developing new tools to get rid of this limitation, or showing that no such tools actually exist, is an important challenge for evaluating the security of future embedded cryptographic devices.

References

1. Batina, L., Gierlichs, B., Prouff, E., Rivain, M., Standaert, F.-X., Veyrat-Charvillon, N.: Mutual information analysis: a comprehensive study. J. Cryptology 24(2), 269–291 (2011)
2. Brier, E., Clavier, C., Olivier, F.: Correlation Power Analysis with a Leakage Model. In: Joye, M., Quisquater, J.-J. (eds.) CHES 2004. LNCS, vol. 3156, pp. 16–29. Springer, Heidelberg (2004)
3. Cantelli, F.P.: Sulla determinazione empirica della legge di probabilita. Giorn. Ist. Ital. 4, 421–424 (1933)
4. Chari, S., Jutla, C.S., Rao, J.R., Rohatgi, P.: Towards Sound Approaches to Counteract Power-Analysis Attacks. In: Wiener, M.J. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 398–412. Springer, Heidelberg (1999)
5. Chari, S., Rao, J.R., Rohatgi, P.: Template attacks. In: Kaliski Jr., B.S., Koç, Ç.K., Paar, C. (eds.) CHES 2002. LNCS, vol. 2523, pp. 13–28. Springer, Heidelberg (2003)
6. Doget, J., Prouff, E., Rivain, M., Standaert, F.-X.: Univariate side channel attacks and leakage modeling. In: COSADE, Darmstadt, Germany, pp. 1–15 (February 2011)
7. Gierlichs, B., Batina, L., Preneel, B., Verbauwhede, I.: Revisiting higher-order DPA attacks. In: Pieprzyk, J. (ed.) CT-RSA 2010. LNCS, vol. 5985, pp. 221–234. Springer, Heidelberg (2010)
8. Gierlichs, B., Batina, L., Tuyls, P., Preneel, B.: Mutual Information Analysis. In: Oswald, E., Rohatgi, P. (eds.) CHES 2008. LNCS, vol. 5154, pp. 426–442. Springer, Heidelberg (2008)
9. Glivenko, V.: Sulla determinazione empirica della legge di probabilita. Giorn. Ist. Ital. 4, 92–99 (1933)
10. Goubin, L., Patarin, J.: DES and differential power analysis (the "duplication" method). In: Koç, Ç.K., Paar, C. (eds.) CHES 1999. LNCS, vol. 1717, pp. 158–172. Springer, Heidelberg (1999)
11. Hassoune, I., Macé, F., Flandre, D., Legat, J.-D.: Dynamic differential self-timed logic families for robust and low-power security ics. Integration, the VLSI Journal 40(3), 355–364 (2007)
12. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning, ch.6. Springer Series in Statistics. Springer New York Inc., New York (2001)

13. Kocher, P.C., Jaffe, J., Jun, B.: Differential power analysis. In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 388–397. Springer, Heidelberg (1999)
14. Lemke-Rust, K., Paar, C.: Analyzing side channel leakage of masked implementations with stochastic methods. In: Biskup, J., López, J. (eds.) ESORICS 2007. LNCS, vol. 4734, pp. 454–468. Springer, Heidelberg (2007)
15. Mangard, S., Oswald, E., Popp, T.: Power Analysis Attacks: Revealing the Secrets of Smart Cards. Springer, Heidelberg (2007)
16. Mangard, S., Pramstaller, N., Oswald, E.: Successfully attacking masked aes hardware implementations. In: Rao, Sunar [22], pp. 157–171
17. Messerges, T.S.: Using second-order power analysis to attack dpa resistant software. In: Koç, Ç.K., Paar, C. (eds.) CHES 2000. LNCS, vol. 1965, pp. 238–251. Springer, Heidelberg (2000)
18. Nelsen, R.B.: An Introduction to Copulas, 1st edn. Lecture Notes in Statistics. Springer, Heidelberg (1998)
19. Prouff, E., Rivain, M.: Theoretical and practical aspects of mutual information based side channel analysis. In: Abdalla, M., Pointcheval, D., Fouque, P.-A., Vergnaud, D. (eds.) ACNS 2009. LNCS, vol. 5536, pp. 499–518. Springer, Heidelberg (2009)
20. Prouff, E., Rivain, M., Bevan, R.: Statistical analysis of second order differential power analysis. IEEE Trans. Computers 58(6), 799–811 (2009)
21. Pyke, R.: Spacings revisited. In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of statistics, pp. 417–427. Univ. California Press, Berkeley (1972)
22. Rao, J.R., Sunar, B. (eds.): CHES 2005. LNCS, vol. 3659. Springer, Heidelberg (2005)
23. Renauld, M., Kamel, D., Standaert, F.-X., Flandre, D.: Scaling trends for dual rail logic styles (2011) (preprint)
24. Renauld, M., Standaert, F.-X., Veyrat-Charvillon, N., Kamel, D., Flandre, D.: A formal study of power variability issues and side-channel attacks for nanoscale devices. In: Paterson, K.G. (ed.) EUROCRYPT 2011. LNCS, vol. 6632, pp. 109–128. Springer, Heidelberg (2011)
25. Rivain, M., Dottax, E., Prouff, E.: Block ciphers implementations provably secure against second order side channel analysis. In: Nyberg, K. (ed.) FSE 2008. LNCS, vol. 5086, pp. 127–143. Springer, Heidelberg (2008)
26. Schindler, W., Lemke, K., Paar, C.: A stochastic model for differential side channel cryptanalysis. In: Rao, Sunar [22], pp. 30–46
27. Scott, D.W.: On optimal and data-based histograms. Biometrika 66(3), 605–610 (1979)
28. Standaert, F.-X., Malkin, T., Yung, M.: A unified framework for the analysis of side-channel key recovery attacks. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 443–461. Springer, Heidelberg (2009)
29. Standaert, F.-X., Veyrat-Charvillon, N., Oswald, E., Gierlichs, B., Medwed, M., Kasper, M., Mangard, S.: The world is not enough: Another look on second-order dpa. In: Abe, M. (ed.) ASIACRYPT 2010. LNCS, vol. 6477, pp. 112–129. Springer, Heidelberg (2010)
30. Venelli, A.: Efficient entropy estimation for mutual information analysis using b-splines. In: Samarati, P., Tunstall, M., Posegga, J., Markantonakis, K., Sauveron, D. (eds.) WISTP 2010. LNCS, vol. 6033, pp. 17–30. Springer, Heidelberg (2010)

31. Veyrat-Charvillon, N., Standaert, F.-X.: Mutual information analysis: How, when and why? In: Clavier, C., Gaj, K. (eds.) CHES 2009. LNCS, vol. 5747, pp. 429–443. Springer, Heidelberg (2009)
32. Veyrat-Charvillon, N., Standaert, F.-X.: Generic side channel distinguishers: Improvements and limitations. Cryptology ePrint Archive, Report 2011/149 (2011), http://eprint.iacr.org/
33. Whitnall, C.: An information theoretic assessment of first-order mia. First year PhD report, University of Bristol (2010)
34. Whitnall, C., Oswald, E.: A comprehensive evaluation of mutual information analysis using a fair evaluation framework. In: To Appear In The Proceedings Of Crypto 2011, Santa Barbara, California, USA, August 2011, vol. xxxx, pp. yyy–zzz (2011)

A Results of Our Experiments

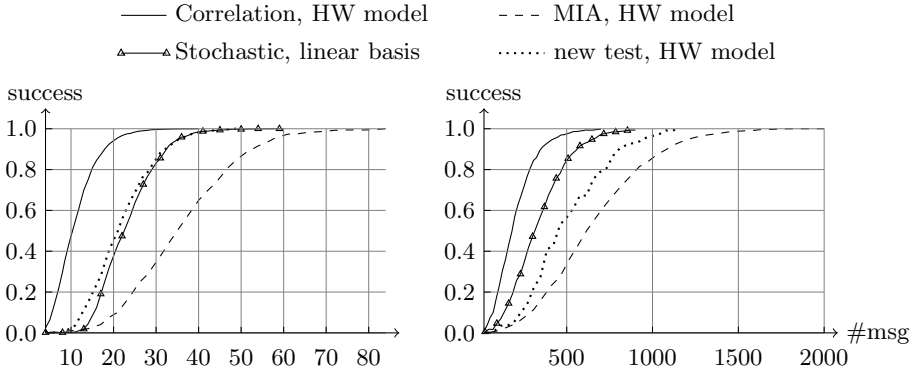

Fig. 4. Simulations with a Hamming weight leakage function, attacks against an unprotected S-box (left) and a masked S-box with combined samples (right)

B Building the Distance Templates

The distance templates are built by convolution of independent uniform random variables. That is, for one dimension, the probability of a spacing of length u is the probability that two random variables X and X' drawn from the n-valued discrete uniform distribution on the interval $[0, 1]$ (each value has probability $\frac{1}{n}$) will differ by an amount u. That is:

$$\Pr[U = u] = \sum_x \Pr[X = x] \cdot \Pr[X' = x \pm u]$$

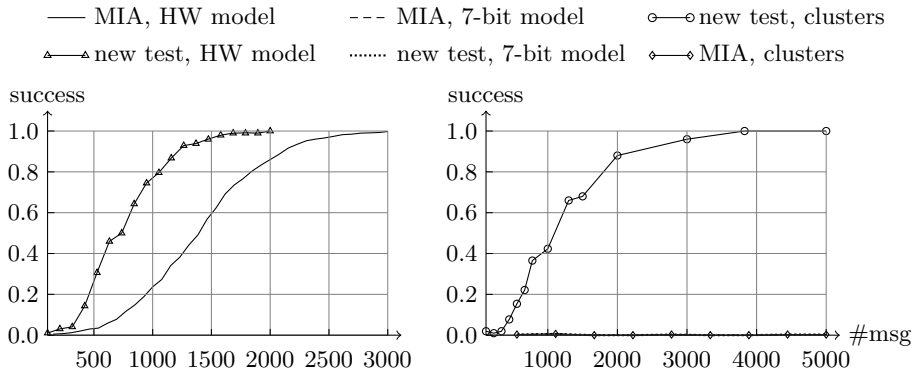

Fig. 5. Masked S-box with bivariate leakages. Attacks with simulated Hamming weight leakage function (left) and actual measurements on a 65nm CMOS chip (right).

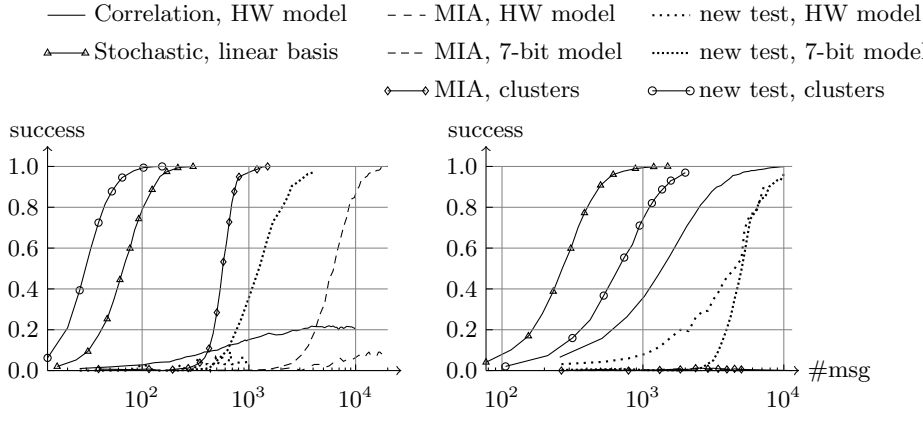

Fig. 6. Measurements on a 65nm CMOS chip, attacks against an unprotected S-box (left) and a masked S-box with combined samples (right)

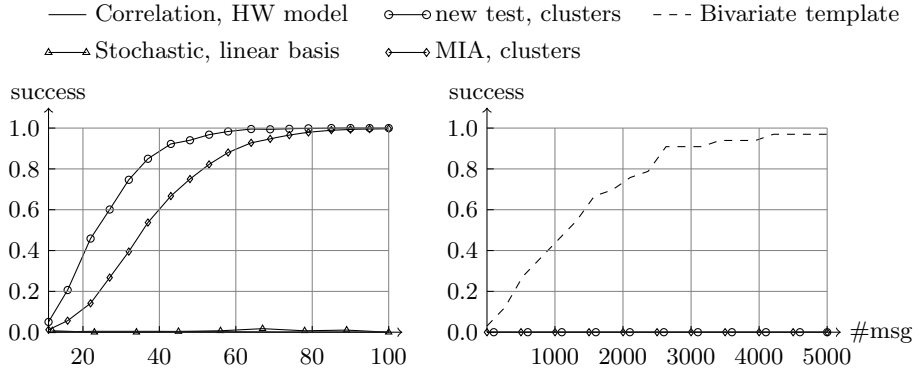

Fig. 7. Spice simulations of a 65nm dual-rail logic style, attacks against an unprotected S-box (left) and a masked S-box with bivariate samples (right)

In our specific case, since the discrete uniform distribution results from an integral transform of continuous variables, it is impossible to have a spacing of 0. The spacing distribution is therefore $\Pr[U = u] = \frac{2}{n-1} \times (1 - u)$ for $u > 0$, 0 otherwise. As one considers higher dimensions, using the Manhattan distance to extend the notion of spacings, the distribution becomes a convolution of spacings on one dimension:

$$\Pr[U = u] = \sum_{u_i} \prod_{i=1}^{d-1} \Pr[U_i = u_i] \cdot \Pr\left[U_d = \left(u - \sum_{i=1}^{d-1} u_i\right)\right],$$

where the U_i are spacings taken along dimension i. In the case of two dimensions, this formula simply gives :

$$\Pr[U = u] = \sum_{u_1} \Pr[U_1 = u_1] \cdot \Pr[U_2 = u - u_1],$$

which is the integral of two affine functions, therefore a piecewise cubic polynomial since U_1 and U_2 are only defined on the interval $[0, 1]$ while U ranges over $[0, 2]$. While it is possible to compute the sampling distribution analytically for higher dimensions, it quickly becomes cumbersome, and it is much more practical to build the distribution by composing lower-dimension distance histograms. This is done very efficient by following a method similar to the square-and-multiply algorithm, where for example the template for dimension 4 is built by convoluting the template for dimension 2 with himself.