IACR Transactions on Cryptographic Hardware and Embedded Systems
ISSN 2569-2925, Vol. 2026, No. 1, pp. 402–424. DOI:10.46586/tches.v2026.i1.402-424

Improving the Selection Rule of Correlation Attacks for Remote Power Analysis

Oriol Farràs[1], Vincent Grosso[2], Miquel Guiot[1] and Carlos Andres Lara-Nino[1]

[1] Universitat Rovira i Virgili, Tarragona, Spain
[2] CNRS/Laboratoire Hubert Curien, Saint-Étienne, France

Abstract. Remote power analysis is a novel threat to information systems. Under this attack model, the adversary does not require direct physical access to the platform or specialized sensing equipment. Most of the literature in this field deals with advanced acquisition methods and adversarial models. In contrast, side-channel analysis techniques for remote attacks have not been sufficiently explored. We bridge this gap by taking a look at the characteristics of the data recovered from remote power analysis. We use these insights to propose a novel selection rule for correlation-based attacks that boosts success confidence. This improvement comes from the observation that the samples in a power trace are not independent. We show that adjacent samples can also provide useful information by proposing a post-processing step that capitalizes on these additional leakages. In contrast to previous work, the proposed technique does not rely on the selection of points of interest within the power traces. We further investigate the characteristics of "remote" power traces and their effect on the proposed selection rule through experiments with real (TDC, ChipWhisperer) and synthetic data sets. To assess the advantage of the proposed improvement, we also introduce novel performance metrics that divert from known-key evaluation techniques.

Keywords: Attack performance evaluation · Correlation power analysis · Hypothesis discrimination · Remote power analysis · Side channel attacks · Statistical attacks

1 Introduction

Side channel analysis attacks (SCA) exploit the dependencies between the logical and physical states of a computing system. The "logical state" of a circuit is constituted by program data, constants, inputs, outputs, register values, etc. The physical state consists of characteristics that can be monitored through an acquisition process. These include phenomena such as sound, voltage, current, power, temperature, and electromagnetism. By studying fluctuations in the physical state, an adversary might gain insight into its logical state. And potentially into its sensitive information [CKN01, MOP08].

The usual adversarial models in SCAs confer the adversary with several advantages. We assume that the attacker can query a "physical" oracle an unrestricted number of times and that they have the capability to register, store and operate over these responses. This physical oracle $\mathcal{E}[k]$ (depicted in Fig. 1) contains some secret information (k) whose recovery is the ultimate goal of the adversary. The rule of the game is that the adversary cannot directly request k, but can provide y queries (m) and the oracle will respond with a cryptographic output (c) plus additional information (ℓ). If the algorithm $\mathcal{E}$ is cryptographically sound, we can assume that it is computationally difficult to infer k from $\{m, c\}$. However, due to the physical nature of the oracle, some information "is leaking" from k through ℓ. In practical terms, the physical oracle is the device that performs sensitive operations over m and produces c. This process involves internal secrets k, which

Licensed under Creative Commons License CC-BY 4.0.
Received: 2025-07-15 Accepted: 2025-09-15 Published: 2026-01-16

can be some data or the operation itself. The additional information $\boldsymbol{\ell}$ reveled by the oracle is embedded in the physical magnitudes associated with the device: current draw, power dissipation, heat dissipation, electromagnetic emanation, etc. Therefore, the adversary must have some control over the device and the sensing capabilities to retrieve the second oracle output ($\boldsymbol{\ell}$).

However, the physical transients produced by computing platforms are challenging to observe. Their fluctuation rate is, in most cases, proportional to the operating frequency of the device. Additionally, their magnitude is quite small and easily obfuscated by noise. Historically, this has been understood to pose two main challenges for adversaries. First, they would require sophisticated sensing equipment (high monetary costs). And second, they would need physical access to the device in order to prepare a proper experimental setup [MOP08]. However, advances in the development of sensing technologies now allow us to instantiate internal sensors in ordinary computing systems [ZS18, GDTL21]. In fact, it is now possible to leverage hardware components within the device to retrieve measurements that correlate with the electrical activity of the circuit. These sensors can be introduced during the manufacturing step or by software means, enabling an attacker to mount SCAs remotely. This paradigm change can lead to the introduction of more practical adversarial scenarios and the definition of *remote power analysis* (RPA).

Internal sensors such as time-to-digital converters (TDC) and ring oscillators are not capable of capturing physical phenomena *per se*. Instead, they are designed to register logical values which may fluctuate as a function of the propagation delay within the circuit. That is, when the system is in repose, their output is more-or-less constant. However, when some computing operation is performed, the charge/discharge of components within the circuit affects its current draw. In turn, this produces an instantaneous power drop, which leaves the system in a transitory state. During this phase, anything connected to the same power source is bound to be affected. The most notable outcome of this process is that the speed of the stimuli propagating throughout the platform is reduced. Thus, the logical values registered by the sensor circuit are expected to divert from their steady states. This phenomenon is demonstrated in [FTBLN24]. Multiple quantification and decoding strategies can then be applied to obtain estimations of the fluctuation in the propagation delay. Ideally, there would be some linearity in the sensor output. In such a case, the output of the internal sensor would be similar to the observations registered by a conventional capture device, such as an oscilloscope or a spectrum analyzer. However, using internal sensors as capture systems posses new challenges.

They are meant to be small to improve their chances of remaining undetected. This impacts their resource availability and thus their quantification step, which yields low resolution traces. A common solution has been to capture multiple traces under the same $\{m, k\}$ tuple and then to "compose" these traces to improve their resolution. This increases the number of observations that must be captured, however. Another important issue is that trigger mechanisms are not ideal. On one hand, some works assume that the adversary can probe the control signals of the victim circuit to start and stop the acquisition. This limits the scope of RPA attacks. To solve this issue, some works have proposed to detect large fluctuations in the sensor output [SGMT21] or to use covert channels [FTBLN23] to synchronize the capture setup. These methods come with the disadvantage that some traces are not properly captured. These observations not only do not contribute to the quality of the dataset, but effectively act as noise which impacts it. Additional steps are needed to re-synchronize the traces.

In this work, we introduce a post-processing step for correlation-based attacks that improves the divergence of the correct hypothesis. Instead of simply selecting the maximum absolute value from the correlation matrix, we combine multiple Pearson coefficients for consecutive time points into a single coefficient. Then, we select the maximum value among these combined coefficients. This approach takes into account multiple time points

near the instant of interest, which is maximized for the correct hypothesis. Some works in the literature [LCSL07] have proposed to first create a composition of samples and then apply the statistical distinguisher. Our approach first computes the distinguisher and then composes a selection function. We base the soundness of this method on the analysis of Pearson's correlation coefficient and its composability through Fisher's transform. The proposed selection rule can help to discern the correct hypothesis with a smaller number of traces. In some cases, this is evidenced by an increased divergence (as a function of the number of observations) between the correct hypothesis and the corpus of other hypotheses. In other cases, the proposed improvement can discriminate the correct hypothesis with fewer observations. To assess its performance, we introduce three indicators based on the divergence of the correct hypothesis (separation from the second-highest hypothesis, the ratio between the top hypothesis and the second-best, the ratio between the top hypothesis and the cluster mean) as the number of observations increases.

We investigate the characteristics of "remote" power traces and their effect on the proposed method. We study the effects of noise and misalignment in the traces through experiments with real (TDC, ChipWhisperer) and synthetic data sets. These data capture (TDC, ChipWhisperer) or model (synthetic) the behavior of the encryption function of AES-128, allowing us to adjust the noise (synthetic) and misalignment (ChipWhisperer) rates of the dataset to present a comprehensive evaluation of the method. As shown, our findings can be useful for attacking any type of dataset. Indeed, a similar method has concurrently been proposed to mount template attacks [DJJW25]. In our case, we focus on RPA to address the aforementioned open challenges.

The remainder of the paper is structured as follows. In Section 2, we describe some mathematical methods employed in our work. Section 3 provides an overview of the previous work and the positioning of our study with regard to the literature. In Section 4, we describe the experimental conditions that led to the proposed method, as well as the improvements introduced in this paper. Section 5 features an experiment with synthetic data that aims to demonstrate the reproducibility of our findings. In Section 6, we show the applicability of the proposed improvement to conventional data sets. In Section 7 we show how the proposed method contributes to improve the literature results. Lastly, Section 8 expands our discussion and concludes the manuscript.

2 Preliminaries

We denote datasets X using capital letters, their i-th observation as X_i, and their mean as $\bar{X}$. We denote vectors $\boldsymbol{x}$ using bold symbols, and their i-th coordinates as x_i. We denote matrices M using capital letters, their i-th column as M_i, their j-th row as M^j, and their entry in the i-th row and j-th column as M_i^j.

2.1 Pearson's correlation coefficient

Pearson's correlation coefficient, in this work denoted as ρ, is a statistical measure used to quantify the linear relationship between two continuous variables [PG95]. It serves as a foundational tool in many scientific disciplines to assess how closely two variables are linearly related.

Definition 1. Let $X = \{X_1, X_2, \ldots, X_n\}$ and $Y = \{Y_1, Y_2, \ldots, Y_n\}$ represent two datasets, each of size n. The Pearson correlation coefficient, ρ, is given by

$$\rho(X,Y) = \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 \sum_{i=1}^{n} \left(Y_i - \bar{Y}\right)^2}}. \tag{1}$$

The formula of Pearson's correlation coefficient is equivalent to the covariance of X and Y divided by the product of their standard deviations. In this way, we ensure that ρ is dimensionless and invariant to the scale and units of the variables being studied. Several key properties of Pearson's correlation coefficient make it a valuable measure. Firstly, it is symmetric, meaning that the correlation between X and Y is the same as that between Y and X. Furthermore, the coefficient ranges from -1 to 1, and its interpretation depends on both its magnitude and direction, with values closer to ± 1 indicating a stronger linear relationship. More specifically, a value of $\rho \approx 1$ indicates a positive linear relationship, $\rho \approx -1$ signifies a negative linear relationship, and $\rho \approx 0$ implies no linear relationship between variables. However, it is important to note that ρ captures only linear relationships and alternative methods are required to evaluate non-linear relationships. Furthermore, Pearson's correlation coefficient is sensitive to outliers, which can disproportionately affect its value.

2.2 Fisher's transform

Fisher's transform, in this work denoted as z, is a mathematical function used to stabilize the variance of the Pearson's correlation coefficient and make it suitable for parametric statistical methods [Fis15]. This transform is particularly useful when combining or comparing correlations, as the distribution of ρ is not normally distributed, especially for values close to -1 or 1. By applying the Fisher transform, the correlation coefficients are converted into a form that more closely approximates a normal distribution.

Definition 2. Let $X = \{X_1, X_2, \ldots, X_n\}$ and $Y = \{Y_1, Y_2, \ldots, Y_n\}$ represent two datasets of size n, and let ρ represent their Pearson correlation coefficient. The Fisher transform is given by

$$z(X, Y) = \frac{1}{2} \ln \left(\frac{1+\rho}{1-\rho} \right). \tag{2}$$

Fisher's transform z is approximately normally distributed. It is symmetric and ensures that small sample variances and skewness in the distribution of r are reduced, allowing for more reliable inferential procedures. When analyzing data from multiple studies or groups, it is often necessary to compute the mean of several Pearson correlation coefficients. Direct averaging of them is inappropriate because ρ does not have additive properties. Instead, the Fisher transform is used to compute the mean more accurately [DMCB98].

3 Positioning in the literature

Let an observation of the physical state of a characteristic of the device ($\boldsymbol{\ell}$) be composed of an ordered sequence of x samples ℓ^j, uniformly spaced at intervals t. Let the system $\mathcal{E}$ instantiated by $k \in \mathcal{K}$ (its sensitive information) be noted as $\mathcal{E}[k]$ and receive an input $m \in \mathcal{M}$. The system produces the output $c \in \mathcal{C}$ as $c = f(m, k)$ and $\boldsymbol{\ell} = \mu(m, k, t)$. We assume that the adversary can choose m or c, query $\mathcal{E}$ a number of times y, and select the start and end of $\boldsymbol{\ell}$ (and thus implicitly define x). In general terms, the function f is the procedure under attack and may represent a cipher, a hash function, a digital signature algorithm, a neural network, etc. Whereas μ is known as the leakage model and can be understood as the linear combination of f and a stochastic process:

$$\mu(m, k, t) = f(m, k) + \{\mathcal{R}_t\}. \tag{3}$$

In the literature, $\{\mathcal{R}_t\}$ is generally assumed to be a signal with normal distribution (Gaussian noise). In contrast, much research has been done on the nature of $f(m, k)$. Different authors have proposed to model this component as the Hamming weight or the Hamming distance, for example [BCO04].

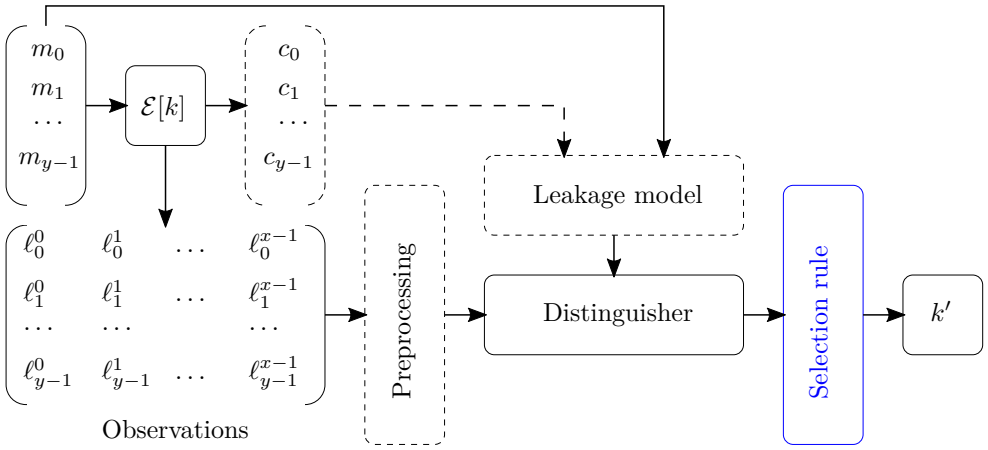


Figure 1: The general composition of a vertical SCA. A full description of this graphic is available in [FGGLN25].

SCAs were first introduced by Kocher et al. observing that the logical state influences the power consumption of the device [KJJ99]. The authors first described the *simple power analysis* attack (SPA), which studies the direct relationship between power traces and runtime [Koc96]. Subsequently, they proposed the *differential power analysis* attack (DPA), which studies the statistical dependency of the power traces on the secret data on the platform. Their seminal work has been adopted as a reference for the design of attacks and countermeasures. However, it also established a framework for the classification of SCAs into two broad families: horizontal and vertical. The *correlation power analysis* (CPA) proposed by [BCO04] belongs to the latter group.

As illustrated in Fig. 1, vertical SCAs follow a well-defined sequence of steps, an idea partially developed by [DPRS11]. A full description of these steps may be found in [FGGLN25]. The final step of attacks like CPA is the selection rule used to discriminate between the different hypotheses. This process is often overlooked, as—historically—most works dictate the selection of the hypothesis that maximizes the absolute magnitude of the distinguisher [KJJ99, BCO04, GHT05, MOP08, PRB09, MOS09, DPRS11, OP13, LPR+14]. To our knowledge, only one work in the literature has elaborated on the selection rule that best adapts to a particular kind of attack [LPR13]. In their work, Lomné et al. proposed an improvement in hypotheses discrimination by capitalizing on the characteristics of linear regression attacks [DPRS11]. Similarly, we observe the characteristics of the Pearson correlation test to propose an improvement to the selection rule of correlation attacks, particularly CPA. Instead of simply selecting the maximum absolute value from the correlation matrix, we first calculate a weighted mean of the absolute values over a defined interval of x. Then, we identify the maximum value among these weighted means. This approach takes into account all correlation coefficients near the instant of interest, which is maximized for the correct hypothesis.

Our proposed improvement comes from the observation that the samples in a power trace are not independent. Indeed, CPA was designed to select the sample with the highest correlation coefficient. However, this ignores adjacent samples, which can also provide useful information. Some works in the literature [LCSL07] have proposed to first create a composition of samples and then apply the statistical distinguisher. Our approach first computes the distinguisher and then composes a selection function. We base the soundness of this method on the analysis of Pearson's correlation coefficient and its composability through Fisher's transform.

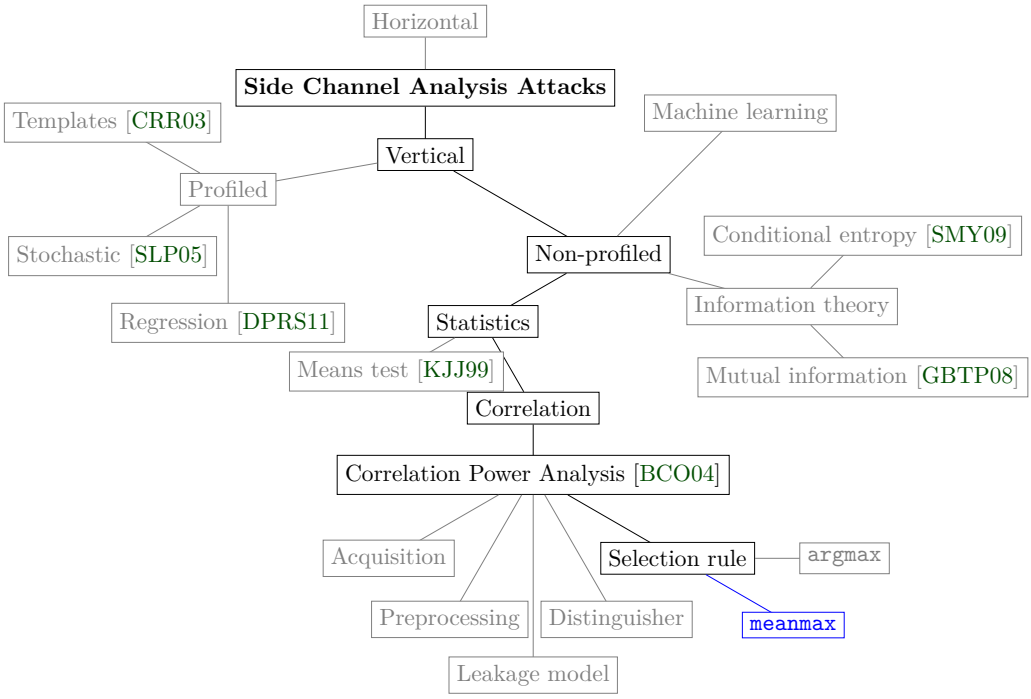


Figure 2: Taxonomy of side channel analysis attacks.

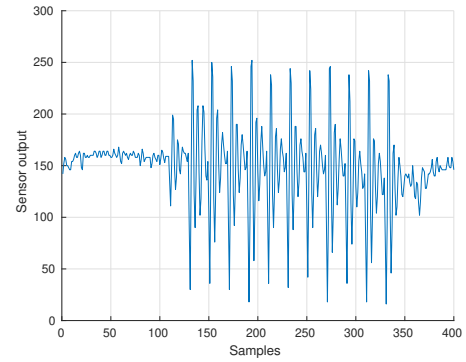
In Fig. 2, we show the positioning of our work in relation to the state of the art. We focus on the selection rule for correlation attacks, using as a case study the correlation power analysis attack proposed by Brier et al. [BCO04]. Instead of using the conventional `argmax` function, we propose a post-processing + selection step called `meanmax`. This has implications for attacks that use the Pearson correlation coefficient as a distinguisher.

4 Correlation attacks in practice

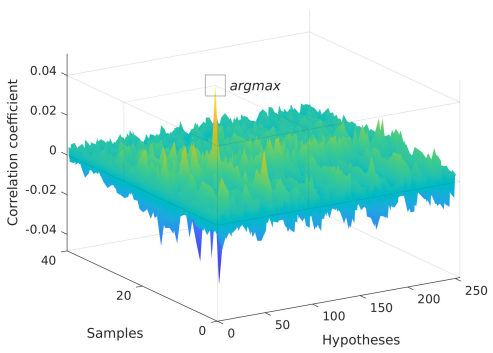
Remote power analysis is an emerging threat to computing systems. It allows us to reduce the adversarial assumptions required to mount an attack. Indeed, the versatility and proficiency of these attacks cannot be underestimated. However, as a recent field of study, it has many open questions and offers areas for improvement. In Fig. 3, we illustrate a basic RPA attack scenario. The adversary implements a sensor circuit in a heterogeneous SoC and uses side-channel leakage to carry out an attack. This allows them to retrieve the secret key of an encryption algorithm, in this case AES-128.

4.1 Correlation power analysis

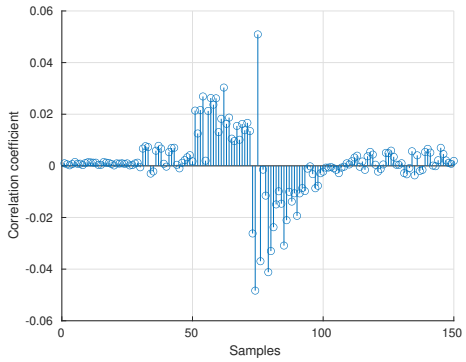
Correlation power analysis improved the original DPA attack in multiple ways. It introduced the use of leakage models, extended the number of bits to be "guessed" in a single pass, and used Pearson's correlation coefficient as a distinguisher. A detailed overview of this method can be found in [FGGLN25]. If we look inside the algorithm, the result of CPA is not a single value for the suggested hypothesis. Rather, the algorithm constructs a two-dimensional array of correlation coefficients with dimensions 2^n-by-x. Recall that x is the number of samples in the power observations and let n represent the binary length of the secret k. In DPA, we have $n = 1$ since each bit is evaluated independently. But CPA allows us to test multiple bits simultaneously. When attacking AES-128 it is common to



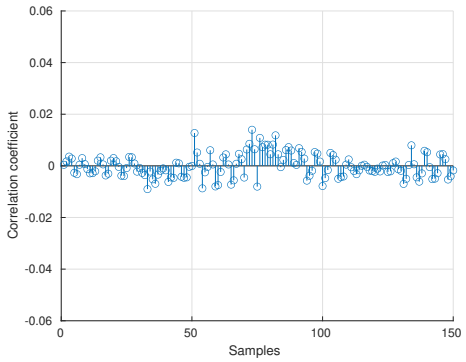
(a) Sample averaged power trace.



(b) Sample correlation matrix.



(c) Correct hypothesis.

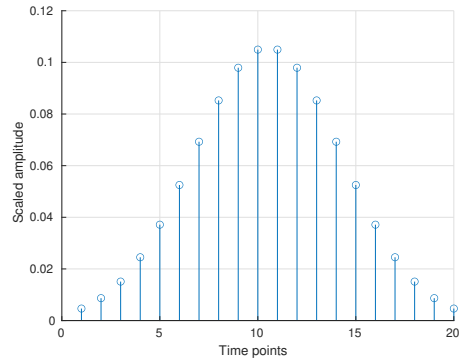


(d) Incorrect hypothesis.

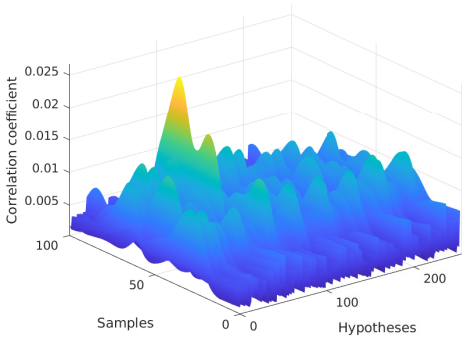
Figure 3: Remote power analysis results using an AMD-Xilinx Zynq-7000 SoC-FPGA (28nm). The sensor is an array of eight time-to-digital converters with 8-bit resolution and a sampling rate of 200MHz. The victim is a textbook, iterative implementation of AES-128 clocked at 10MHz. Both, the sensor and the target are implemented in the reconfigurable fabric of the device.

target its SBOX output, thus $n = 8$-bits are tested and $2^n = 256$. As all the hypotheses are tested for each sample in the power observations, this yields a correlation matrix with the dimensions aforementioned. An example of this correlation matrix is illustrated in Fig. 3b. Conventionally, the `argmax` selection rule is applied to identify the global maximum of this array. However, when y is not large enough, the correlation values for the correct hypothesis can be obfuscated by other local maxima. It is assumed that with increasing y the correct hypothesis will eventually "emerge." Nonetheless, a larger y means capturing and processing more power observations.

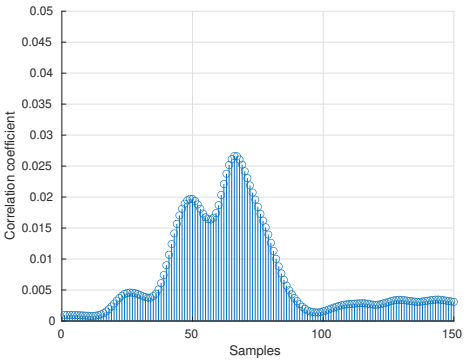
The main limitation with the use of `argmax` as a selection rule is that it only considers one dimension of the correlation matrix. Indeed, from the proposed definition of vertical attacks, CPA is most efficient with $x = 1$. That is, if we can sample the precise instant in which the information leakage takes place. But in practice, the adversary needs to capture multiple samples and test them. The attacker is supposed to know the instants in which the victim begins and ends their processing. This is a very generous assumption and yet produces a large number of samples, as can be observed in Fig. 3a. Based on this, we will operate on the understanding that $x \gg 1$ and therefore there is information that is not fully used by `argmax`. In fact, when we look at the correlation values for the correct hypothesis, it is possible to see that multiple samples contain useful information. In



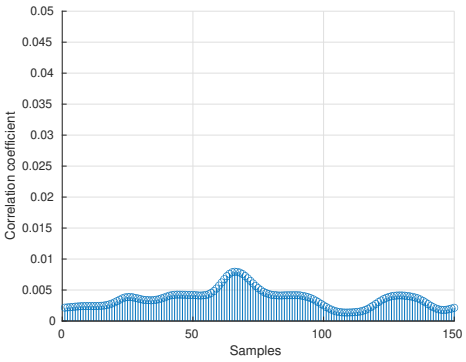
(a) Gaussian window with $s = 20$ coefficients, computed as in [OSB99].



(b) Improved correlation matrix



(c) Improved correct hypothesis



(d) Improved incorrect hypothesis

Figure 4: Application of a sliding-window post-processing step to improve the selection rule of CPA. These data correspond with the experiment in Fig. 3.

contrast, the behavior of the correlation coefficients for multiple samples over an incorrect is mostly uniform. This is shown in Fig. 3c and Fig. 3d, respectively.

It is likely that there are several factors behind the behavior of the correlation test on the samples of a correct hypothesis. First, positive and negative correlation values are interesting since they imply some dependence on the secret information. That is why we take the absolute value on the correlation matrix as noted in [BCO04]. We must then recall that the time samples retrieved from the system are not independent events. The behavior of the system at instant t will depend on its state at instant $t-1$, and both will influence its state at instant $t+1$. Thus, we expect information leakage to be present in several samples. This phenomenon can be amplified if the sampling frequency surpasses the frequency of the target. In fact, logical circuits are expected to complete a given operation within a cycle. However, the actual time response of the circuit depends on several implementation factors. This leads to a scenario where we cannot assume that we obtain a single sample of interest. Lastly, physical defects, such as parasitic currents and synchronization issues during acquisition, can "shift" the point of interest over time. The latter is of special interest for this work.

4.2 Our improvement

Trace desynchronization or misalignment can lead to observations that yield correlation maxima at different moments t. This affects the conventional `argmax` selection rule, which

Algorithm 1 The proposed `meanmax` selection rule for CPA.

Require: a vector $\boldsymbol{m}$ of known y inputs processed by $\mathcal{E}[k]$
Require: a matrix $\boldsymbol{L}$ of y power traces (of x samples each)
Require: n, the binary length of k
Require: a scalar s, the number of samples per cycle
Ensure: k', a hypothesis for the sensitive information k

$\boldsymbol{w} \leftarrow \omega(s)$ {Compute the window weights, as per [OSB99].}
for $i = 0$ **to** $2^n - 1$ **do**
  $\boldsymbol{u} = f(\boldsymbol{m}, i)$ {Compute the leakage model, as per (3).}
  for $j = 0$ **to** $x - 1$ **do**
    $q_j \leftarrow z\left(\left|\rho(\boldsymbol{u}, \boldsymbol{L}^j)\right|\right)$ {Fisher transform (z) is applied over the ρ, as per (1), (2).}
  end for
  for $j = 0$ **to** $x - s - 1$ **do**
    $R_i^j \leftarrow z^{-1}\left(\boldsymbol{w} \cdot (q_j, \ldots, q_{j+s})\right)$ {Sliding window, inverse Fisher transform.}
  end for
end for
$k' \leftarrow \text{argmax}\left(\boldsymbol{R}\right)$
return k'

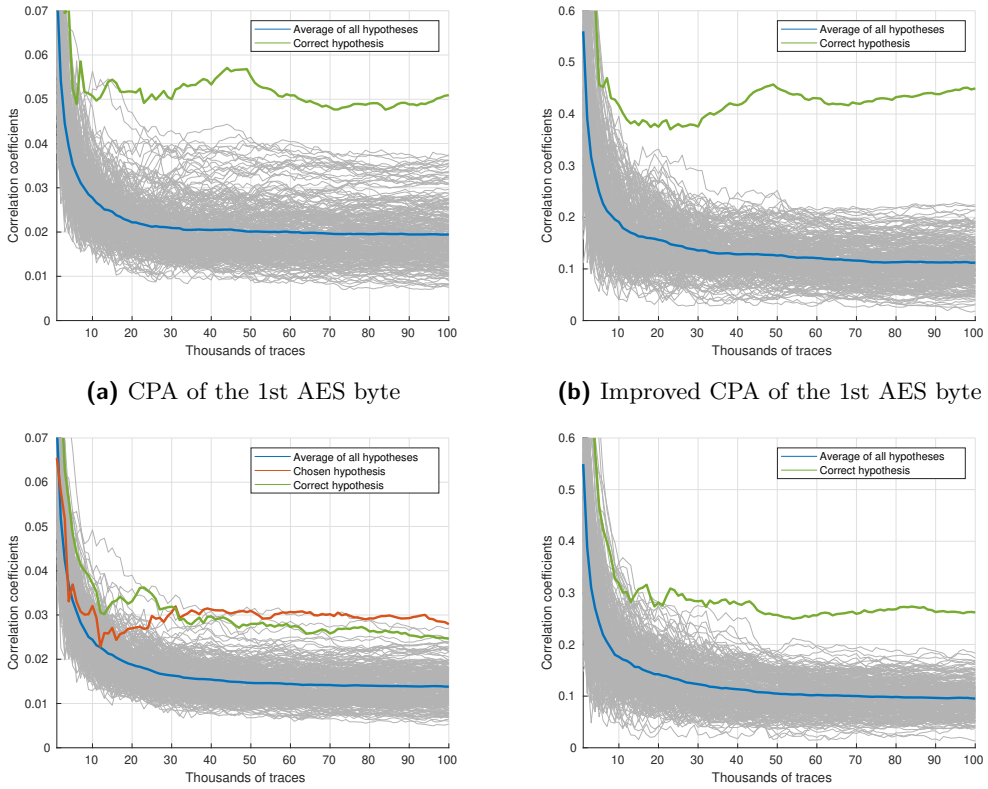
relies on a single instant to determine the best candidate hypothesis. In this regard, pre-processing techniques are normally suggested to deal with this issue. However, we can also leverage the characteristics of the correlation matrix to improve CPA.

To account for the information leakage that appears along multiple samples, we can compute a weighted mean over a finite interval of time points, where the size of this interval can be defined as the number of samples obtained from a cycle of the target (the number of samples per cycle). That is, the ratio between the sampling frequency and the target's frequency. The weights represent the probability of the maximum correlation spike to be found in any given t. Thus, we can employ a window function with a given distribution for the weights and perform a convolution over the correlation matrix along the x samples for all the $|\boldsymbol{h}|$ hypotheses. However, since Pearson's correlation coefficients do not have additive properties, we first apply the Fisher's transform over the correlation matrix. Then, we compute the mean of the resulting values and convert it back using the inverse Fisher's transform. Finally, we can test the maximum among all these means to choose a hypothesis. By applying this sliding-window technique, we mitigate the dependency on any techniques to detect points of interest. The effects of the proposed selection rule, called `meanmax` are shown in Fig. 4 and its application to CPA is detailed in Alg. 1.

The proposed selection rule can help to discern the correct hypothesis with a smaller number of traces. This result is illustrated in Fig. 5. In some cases, this is evidenced by an increased divergence, as a function of the number of observations, between the correct hypothesis and the corpus of other hypotheses. In other cases, the proposed improvement can separate the correct hypothesis with fewer observations than the `argmax` selection rule. However, questions remain that must be answered. How can we determine that the correct hypothesis becomes clearer? Can this improvement be quantified? What is the success rate of the proposed method over conventional CPA? Can this be used in other contexts? In the following section, we try to address these points.

5 Correlation attacks *in vitro*

In this section, we intend to reproduce the behavior observed in remote power analysis with synthetic power traces. This allowed us to control the noise and misalignment rates.



(a) CPA of the 1st AES byte

(b) Improved CPA of the 1st AES byte

(c) CPA of the 7th AES byte

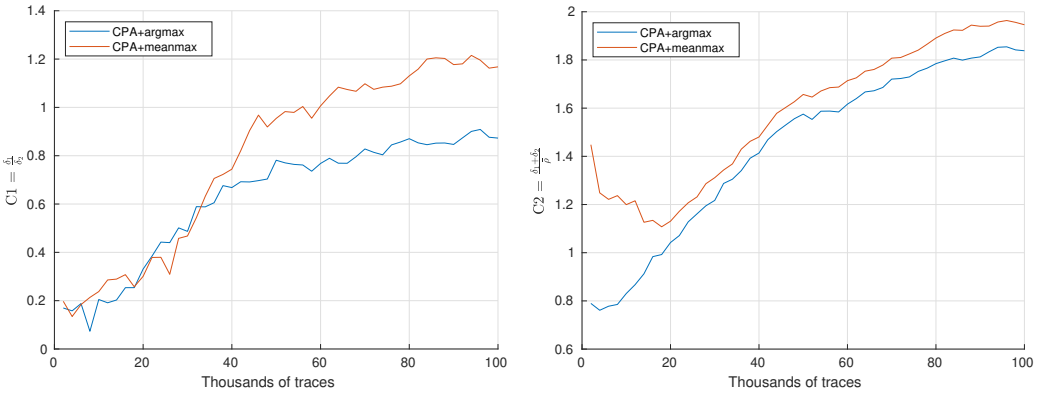
(d) Improved CPA of the 7th AES byte

Figure 5: Comparison of two CPA attacks using the conventional `argmax` selection rule (left column) and the introduced `meanmax` selection rule. The proposed approach increases the divergence of the correct hypothesis (top row) and in some cases can reveal the correct hypotheses using less power traces (bottom row). In this instance, we use a Gaussian window function with $s = 20$ coefficients and standard deviation $\sigma = 0.346$. These data correspond with the experiment in Fig. 3.

5.1 Divergence of the correct hypothesis

Quantifying the performance of SCAs has been a topic of much discussion. Usually, we can register the progression of the greatest correlation point for each hypothesis as a function of the number of observations. Nevertheless, by looking at these results we can draw different conclusions. The adversary will be inclined to report that the correct key was identified with a lower number of observations. So, they might stop the algorithm as soon as the correct hypothesis "comes on top." We refer to this point as the first divergence from the other hypotheses (the corpus). This notion has been complemented with the development of the key-rank metric commonly used in the literature [VGS13]. However, in some circumstances, an incorrect hypothesis may "remount" when more observations are added to the dataset. The adversary is only led to believe that the correct key was found because they know its value *a priori*. On the other hand, the victim may claim that the actual value of the key is not known until the correct hypothesis shows a significant divergence from the corpus. A similar challenge arises when trying to compare SCA methods, as in our case.

We propose that the rate of divergence for an attack can be estimated based on the Euclidean distances between the top hypothesis and the second-best hypothesis (δ_1) and



(a) using performance indicator C1 **(b)** using performance indicator C2

Figure 6: Comparison between CPA+`argmax` and CPA+`meanmax`. The computed performance indicators are averaged over 16 bytes. These data correspond with those in Fig. 3.

between the latter and the mean of the corpus (δ_2), for a fixed number of observations. Please refer to Fig. 6 of [FGGLN25] for a graphic description of these distances. Let $\bar{\rho}$ represent the mean correlation score of all the hypotheses for a given number of observations (y) and define the following *performance indicators*:

$$\mathrm{C1} = \frac{\delta_1}{\delta_2}, \qquad \mathrm{C2} = \frac{\delta_1 + \delta_2}{\bar{\rho}}. \tag{4}$$

As the magnitudes of these indicators are normalized, we can use them to assess the performance of the proposed selection rule compared to the `argmax` function. This is illustrated in Fig. 6. In this case, when looking at C1 we can see how the proposed selection rule improves the separation of the first hypothesis as the number of observation increases. As for C2, our improvement consistently increases the distance between the top hypothesis and the mean of the cluster. This indicates that the proposed technique has the effect of smothering the correlation scores of the incorrect hypotheses and thus reducing the number of traces reduced to conduct a successful attack.

There is room for discussion regarding the selection of the proposed performance indicators. As observed by [Man04], "the number of samples that is needed to see a [correlation] peak in practice, is mainly determined by the distance between the sampling distributions with $\rho = 0$ and $\rho = \rho_{max}$." These sampling distributions represent the corpus of all incorrect hypotheses which converges towards $\rho = 0$ and the correct hypothesis which converges towards $\rho = \rho_{max}$. It follows that we want to observe two indicators: C1 verifies that a single hypothesis is left in the sampling distribution with $\rho = \rho_{max}$ and C2 verifies that it indeed converges towards $\rho = \rho_{max}$ (as the number of observations increases). Note that we do not simply rely on the Euclidean distances δ_1, δ_2 as we also want to normalize the magnitude of the indicators as to allow for straightforward comparison between independent pairs of sampling distributions.

5.2 Synthetic power traces

Following the leakage model in (3), we generated a dataset of synthetic power traces. We used the AES-128 SubBytes function to define the procedure under attack

$$f(m, k) = \mathrm{HW}(\mathrm{SubBytes}(m \ \mathbf{xor} \ k)), \tag{5}$$

where HW() represents the calculation of Hamming weight normalized to the interval $[-0.5, 0.5]$. The values of k and m were chosen at random every time, note that $\log_2 m =$

$\log_2 k = 128$. The output of f was collected in a single sample of interest i. The precedent and subsequent samples in the trace were initialized with the Hamming weight of the output of the AddRoundKey and MixColumns functions of AES-128, respectively. We used this strategy to approach the behavior of a hardware AES-128 architecture. To model the influence of the stochastic process $\{\mathcal{R}_t\}$ we generated a sequence of pseudo-random bits $R(t)$ in the interval $[-0.5, 0.5]$ with a normal distribution. We used a scaling constant α to adjust the amplitude of this noise. It can be said that the noise amplitude is alpha times greater than the data. Each synthetic trace consisted of multiple samples t defined as

$$\mu(m,k,t) = \begin{cases} \mathrm{HW}(m \ \mathbf{xor} \ k) + \alpha R(t) & t < i \\ \mathrm{HW}(\mathrm{SubBytes}(m \ \mathbf{xor} \ k)) + \alpha R(t) & t = i \\ \mathrm{HW}(\mathrm{MixColumns}(\mathrm{SubBytes}(m \ \mathbf{xor} \ k))) + \alpha R(t) & t > i. \end{cases} \tag{6}$$

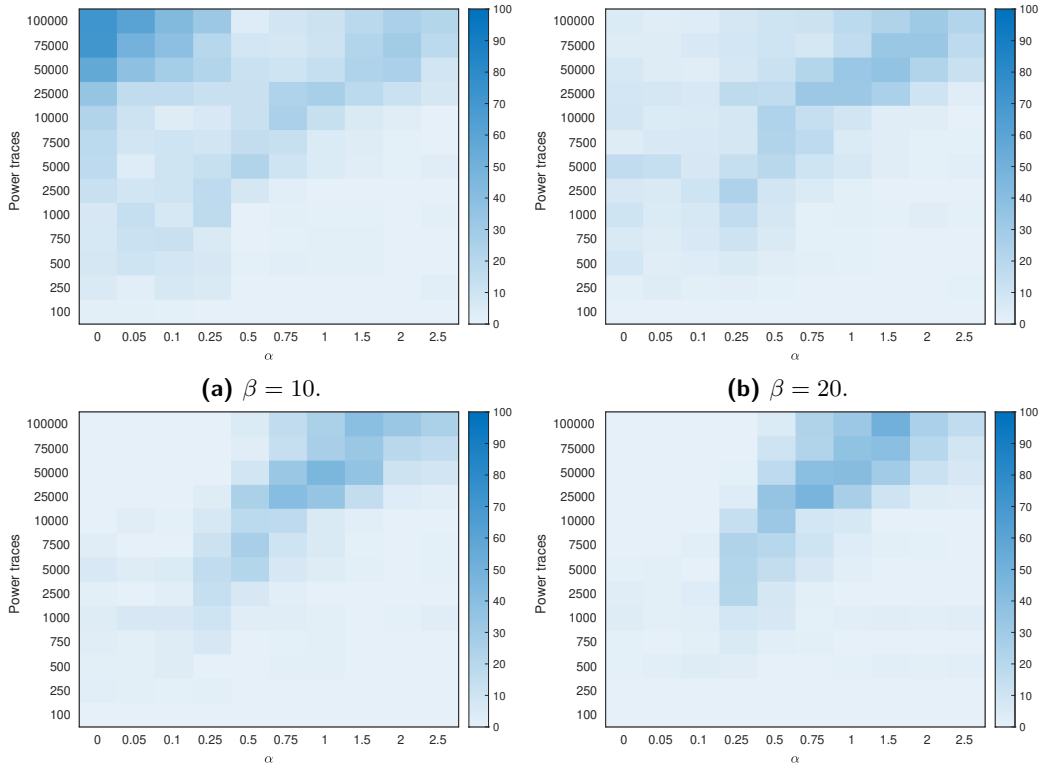
This strategy allowed us to produce a two-dimensional array of y-by-x synthetic samples with y as the number of observations (messages, synthetic traces) and $x \geq 3$ as the number of samples. Evidently, only the samples at $t = i$ contain leakage information. However, since we intend to study the characteristics of RPA (namely misalignment), additional containers are needed.

5.3 Effects of trace misalignment

The leakage model provided in (6) accounts for most of the noise generated by the acquisition process, if we assume that it is uniformly distributed. But it disregards noise sources with low periodicity (i.e. thermal noise and jitter), as their influence over small values of x can be considered negligible. This model is accurate for the study of power analysis attacks in conventional setups. However, remote power analysis attacks are susceptible to other perturbations, notably trace misalignment. This is especially noticeable in capture schemes that do not rely on logical trigger signals [FTBLN23]. Although it is true that conventional capture setups are also prone to trace misalignment, in such cases, the common recommendation is to perform a pre-processing step which considers resynchronization algorithms [MOP08]. But this process is complex in remote scenarios given the lower resolution of the capture setup and the limited processing capabilities of the sensor itself. Nonetheless, we show that the `meanmax` selection rule can contribute to improving the performance indicators of CPA (C1, C2) in the presence of moderate noise and trace misalignment rates.

We designed an experimental methodology inspired by the evaluation methods described in [DPRS11] (Fig. 10 in their paper). This statistical test was originally intended to observe the evolution of the success rate according to the number of messages (y) and the amplitude of the noise component (α). However, we added an additional dimension to this analysis by using datasets with different misalignment rates (β). We provide experimental results for $\alpha \leq 2.5$ and $\beta \leq 40$. These upper bounds were determined by the computing time. We observe that the proposed method is most effective for large α and β, if enough observations are provided. As we also perform multiple CPA runs per data point, the chosen values show this effect with reasonable computing time. We considered that each synthetic trace could be misaligned by ± 1 sample, with β% of the traces being randomly misaligned. This means that $\beta/2$ of every 100 traces will have the AddRoundKey data at instant i and $\beta/2$ of every 100 traces will have the MixColumns data at instant i. The reason behind choosing a single sample misalignment is that the temporal leakage dispersion can also be captured with the proposed sliding window. Thus, greater dispersion distributions can be addressed by adjusting the window function.

Since we wanted to evaluate the prowess of the proposed selection rule, each test reports only the percentage of experiments where CPA+`meanmax` is correct (i.e. the key is



(a) $\beta = 10$. **(b)** $\beta = 20$.

(c) $\beta = 30$. **(d)** $\beta = 40$.

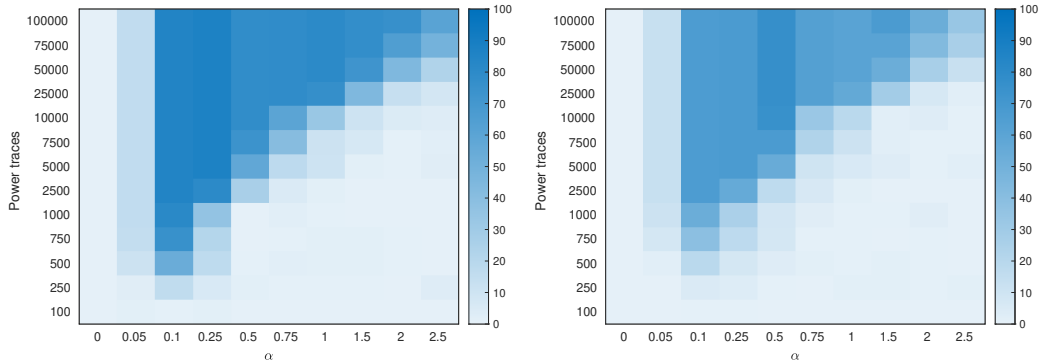
Figure 7: Success rate (W1) of CPA+`meanmax` in function to the number of messages (y), the noise gain (α), and the misalignment rate (β).

succesfully recovered) and a) improves indicator C1 (the results are shown in Fig. 7) or b) improves indicator C2 (the results are shown in Fig. 8). These indicators are compared against the same indicators calculated for CPA+`argmax`. Given that the synthetic traces produced from the model in (6) contain information about 16 different AES-128 key bytes, we repeated the experiments 10 times to estimate the success rate of attacking 160 different key bytes. To derive our comparison results we performed a byte-to-byte analysis of both methods. We first computed the distances δ_1, δ_2 and asserted their correctness. We then calculated the performance indicators C1 and C2 as in (4). Lastly, we computed the "success rate" (W) for CPA+`meanmax` as

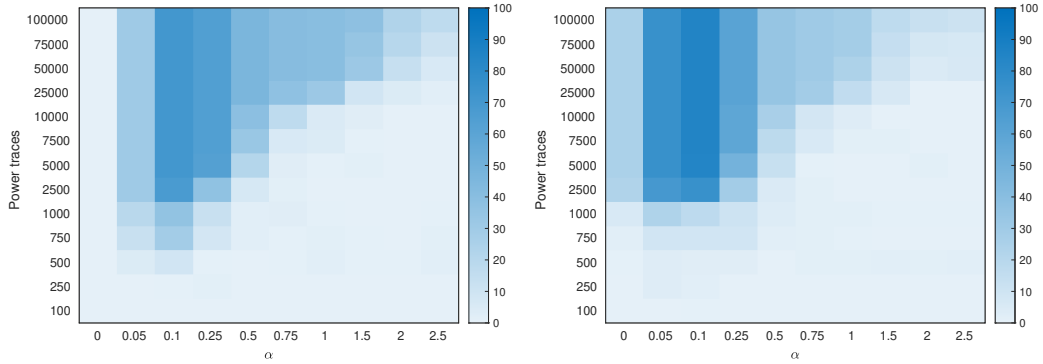
$$\mathrm{W1} = \frac{\sum_{byte=1}^{160} (\text{(C1'} > \text{C1) \& byte_found})}{160} \times 100. \tag{7}$$

In this case, C1 is computed for `argmax` and C1' is computed for `meanmax`. Note that we denote the "success rate according to the indicator C1" as W1, but an analogue formula can be derived for indicator C2.

To clarify, these experiments show that in the given percentage of experiments (how many bytes are recovered out of the 160) the proposed method (CPA+`meanmax`) contributes either to increase the divergence between the correct hypothesis and the second best (Fig. 7) or to segregate the correct hypothesis while a substantial portion of the corpus is led to converge towards low correlation values at a faster rate (Fig. 8). This implies that in many cases CPA+`meanmax` is correct, but it does not manage to improve either indicator. Indeed, as shown in Appendix A, the proposed method effectively improves the success rate of the baseline.



(a) $\beta = 10$. **(b)** $\beta = 20$.

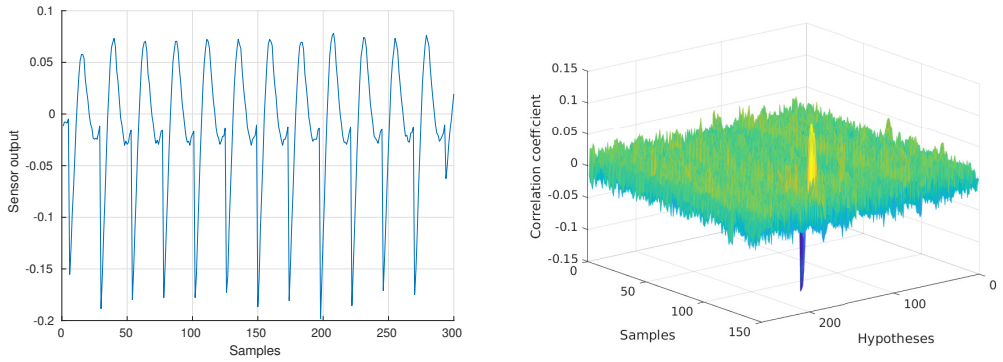


(c) $\beta = 30$. **(d)** $\beta = 40$.

Figure 8: Success rate (W2) of CPA+`meanmax` in function to the number of messages (y), the noise gain (α), and the misalignment rate (β).

In Fig. 7 we observe a moderate advantage for CPA+`meanmax` when looking at the indicator C2. The success rate appears to be somewhat affected by the misalignment rate (β), in particular for a large number of observations (y) and lower noise content (α). This is expected since misaligned traces introduce additional noise into the dataset. It is likely that by increasing the number of power traces we would observe a similar behavior for all values of β. More research is needed to understand the non-continuity of the success rate W1 in relation to the noise gain α and the number of observations. However, we remark that this phenomenon only affects the capacity of CPA+`meanmax` for improving the performance indicator C1. Indeed, from Appendix A it can be corroborated that with enough observations both methods are consistently correct.

In Fig. 8 we instead look at the advantage obtained for the indicator C2. In this case, we observe a greater advantage for CPA+`meanmax` which appears to be maintained even as the misalignment rate increases. This analysis highlights the effectiveness of the proposed selection rule for smothering the incorrect hypothesis. An interesting effect that can be observed in Fig.8 is that when the noise gain (α) is zero, the `argmax` selection rule has a significant advantage. In fact, the proposed `meanmax` is a post-processing technique that can mitigate spurious correlation spikes found in the correlation matrix. And these spikes are generated, in the synthetic model, from the Gaussian noise ($R(t)$) introduced in the traces. In that sense, we can consider the proposed method as a low-pass filter applied over the correlation matrix. The cut frequency is implicitly defined by the number of samples in the sliding window (s). As described above, a convenient way to define s is to take the relationship between the sampling frequency and the operating frequency of the target. Nevertheless, this simply represents an upper bound, and in some cases a



(a) Sample individual power trace **(b)** CPA+`argmax` correlation matrix

Figure 9: Power analysis results using the ChipWhisperer-Lite and ChipWhisperer-CW305 setup. In the sample power trace it is possible to appreciate eleven segments separated by power spikes, these correspond to the internal iterations of the AES-128 architecture.

smaller window could yield better results.

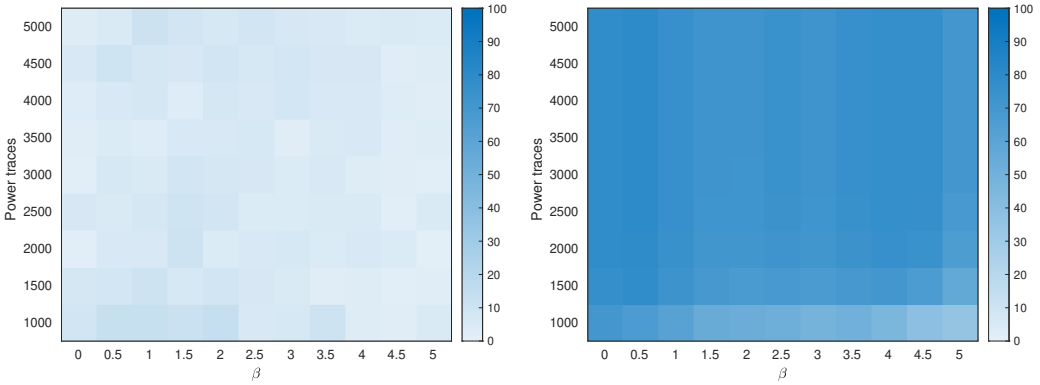
In the experiments reported in this section, we consider that the percentage of misaligned samples is known. Therefore, we can define a perfect mask instead of employing a window function as in the previous section. This represents the optimal window where the proposed selection function is immune to misalignment. However, as shown in Section 4, it is possible to employ window functions that approximate the probability density function of the misalignment in the correlation coefficients for the correct hypothesis. The width and shape of this window should approach the distribution of the dispersion of the misaligned traces (how many samples) and the misalignment rate (how many traces). Evidently, applying a sliding-window filter over the correlation matrix is equivalent to performing a convolution with a filter window. This suggests that applying a Fourier transform to the correlation matrix can yield useful information for discriminating the correct hypothesis. This idea is further explored in [FGGLN25].

6 Correlation attacks *in vivo*

In this section, we study the performance of the proposed selection rule when small misalignments are introduced into the traces. That is, we used a dataset of generic side-channel measurements and induced some controlled perturbations. This experiment differs from the results provided in the previous sections in the following aspects:

- The sensor is a ChipWhisperer-Lite device with an acquisition rate of 96 MSps.
- The victim is a textbook iterative implementation of AES-128 (a round is computed each cycle) clocked at 4 MHz. Implemented in a Chipwhisperer CW305, which embeds an AMD-Xilinx Artix 7 FPGA (28nm).
- This capture setup showed greater challenges to obtain positive CPA results; therefore, we used the last round differential leakage model (hence, a lower number of traces is required w.r.t. Section 4).

The traces used in this experiment included a full encryption of AES-128 with random plaintexts. We obtained ten sets of traces with random different keys. Therefore, we could attack a total of 160 different key bytes as in the previous Section. For the capture, we used the internal triggers of the victim to synchronize the acquisition, and thus it



(a) W1 **(b)** W2

Figure 10: Success rate (7) of CPA+`meanmax` according to the performance indicators from (4), in function to the number of messages (y) and the misalignment rate (β).

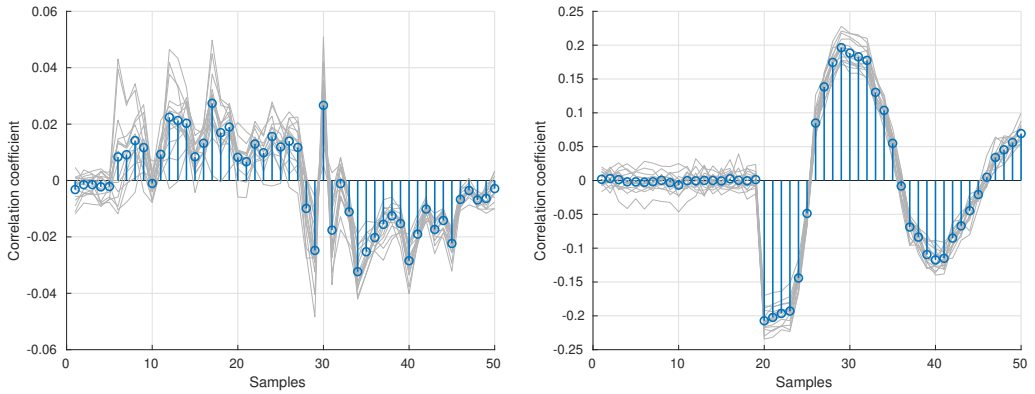
represents a scenario where $\beta \approx 0$. However, with this dataset we can introduce a controlled misalignment rate to study the behavior of the proposed selection rule when the noise (α) comes from a physical system. Fig. 9 shows a sample trace and classic CPA results.

Next, we applied the proposed selection rule to perform an evaluation of its improvement in conventional datasets. The attack attempted to retrieve the 160 secret bytes in the dataset for misalignment rates up to 5%. As in the previous section, we were interested in the cases where CPA+`meanmax` outperforms CPA+`argmax` according to indicators C1 (greater divergence between the second hypothesis and the mean of the cluster) and C2 (overall divergence from the mean of the cluster). The results are shown in Fig. 10.

This experiment shows that CPA+`meanmax` contributes moderately to improve the separation between the correct hypothesis and the second choice for small misalignment rates. In contrast, it significantly increases the separation between the correct hypothesis and the cluster's mean, even for small misalignment rates. This suggests that the proposed method overall accelerates the convergence of incorrect hypothesis towards a low correlation value, while in some cases not necessarily reducing the correlation score of some spurious hypotheses. Indeed, this can be observed in Fig. 5. This suggests that in cases where the segregation of the correct hypothesis is not conclusive, it may still reduce the number of potential candidates for an enumeration attack. These results can appear less favorable than the results given in the previous Section 5, where significant misalignment rates were used. However, we must note that, in this case the artificial misalignment is applied over a strong noise component derived from a real measurement system. Another relevant factor is the change in the acquisition setup. In particular, a closer look reveals that the dispersion first observed in Fig. 3 is less pronounced when using the ChipWhisperer capture setup, as illustrated in Fig. 11.

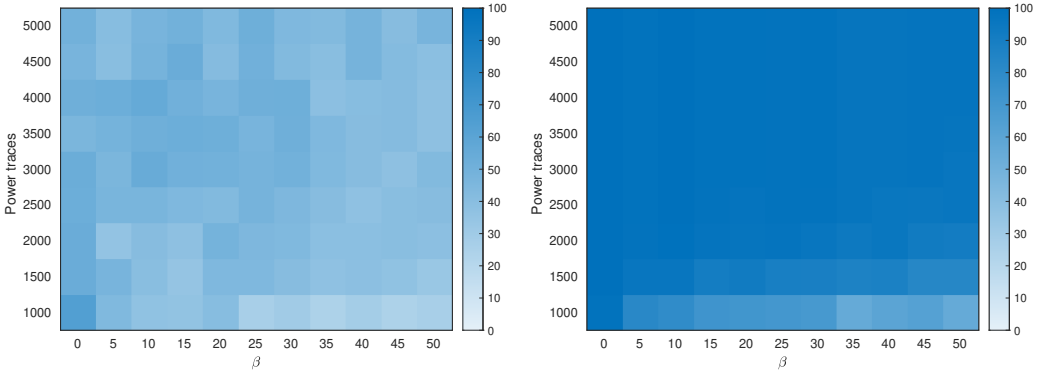
This poses the question of why the leakage is more dispersed over the RPA power traces. Initially, we formulated the hypothesis that trace misalignment could contribute to increasing this dispersion. But as shown in this section, a strong desynchronization is in fact detrimental for the proposed method. Nevertheless, the solution is evidently to adjust the number of samples used to define the sliding window in Alg. 1. For example, for the ChipWhisperer scenario, the results are improved if we set a window width of four samples. These results are provided in Fig. 12. With this adjustment, the improvement rates for the divergence between the second hypothesis and the cluster mean are less sensitive to large misalignment rates. This is consistent with the results observed in Section 5.

In Section 4.2 we noted that the window width could be computed as the ratio between the sampling and the victim's frequencies. This works well for RPA scenarios. However, as



(a) TDC power traces **(b)** ChipWhisperer power traces

Figure 11: Detail of the correlation coefficients for the correct hypotheses over multiple samples. The correlation coefficients are computed following the conventional CPA approach (CPA+`argmax`). In gray we show individual (correct) hypotheses from the correlation matrix and in blue their average values.



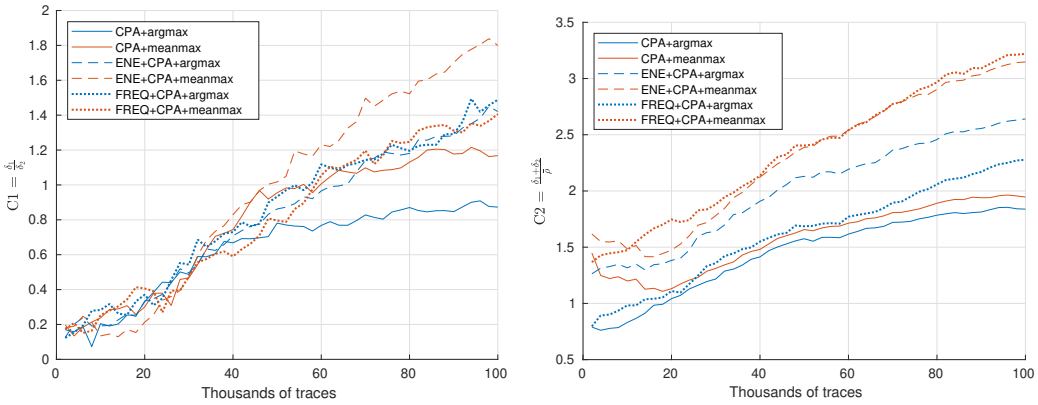
(a) W1 **(b)** W2

Figure 12: Evolution of the success rate (7) in function to the number of messages (y) and the misalignment rate (β). In this case the sliding window considers a leakage dispersion over four samples.

shown in this Section, this likely represents an upper bound in terms of finding the target computation within a given time frame (concretely a clock cycle of the victim). Therefore, for some cases it might be more interesting to focus on narrower intervals. This can be addressed either by reducing the window width or by adjusting the window model so that its coefficients better capture the behavior of the system.

7 Comparison with the literature

Works in the literature have proposed pre-processing techniques oriented to combine multiple samples of a trace into a single time instant [LCC+06, LCSL07]. This allows statistical attacks to capitalize on the amplified leakages to derive a result through `argmax`, which in this way can benefit from multiple time points. One of these approaches [LCC+06] relies on selecting multiple samples of interest to formulate a multiple-bit DPA, ultimately



(a) using performance indicator C1 **(b)** using performance indicator C2

Figure 13: Comparison of different pre-processing techniques and their interaction with the proposed selection rule. For each pre-processing method (None, ENE [LCSL07], FREQ [GHT05]) we perform CPA using both selection rules (`argmax`, `meanmax`) and compute the performance indicators from (4). For ENE+CPA+`meanmax` we used a pre-processing interval of 20 samples and a Gaussian window with five samples for `meanmax`. In the other instances we used a Gaussian window of 20 samples for `meanmax`.

concluding that CPA is an instance of their method. [LCSL07] proposed to divide the traces in small windows and to integrate over these intervals to create energy scores. This has the additional benefit of reducing the size of the input. Other works have instead propose to study the frequency contents of the input traces, which can solve the issue of misalignment [GHT05, JZ23]. We note that since these are pre-processing techniques they are compatible with our approach. Indeed, we can easily apply the proposed selection rule to the CPA output regardless of the nature of the input. This is demonstrated in Fig. 13.

From this experiment we can corroborate that the proposed selection rule can improve the performance indicators of CPA independently of the type of input. Nonetheless, further research is needed to understand the effect of these methods on the correlation matrix.

8 Conclusions

In this paper, we have introduced a new selection rule for correlation attacks that exhibits favorable characteristics in the presence of trace misalignment. The proposed improvement relies on the observation that the correlation matrix used as a distinguisher in certain vertical attacks is not made up of totally independent observations. Instead, the independence exists only between hypotheses, but the correlation values for consecutive samples are related. Indeed, for the vector associated with the correct hypothesis we can identify—under ideal conditions—not only the global maximum, but also a subset of samples with increased correlation values.

This observation was first made in the analysis of the behavior of remote power analysis attacks, where the proposed technique can be used in practice to improve the effectiveness of CPA. From this scenario, we set out to investigate the causes of the observed phenomenon. It is likely that multiple factors are at play. But in this paper, we focus on the idea that moderate desynchronization of the traces is one of the causes. This is a particular characteristic of remote power analysis, where the acquisition mechanism is sometimes decoupled from the logical triggers of the algorithm to simplify the adversarial assumptions. This produces, consequently, measurements that have to be resynchronized to be useful for standard analysis methodologies.

To test this hypothesis, we constructed a set of power traces according to well-understood power models and modified the noise and misalignment rates. This allowed us to study the performance of CPA under the classical and novel selection rules. We introduced two attack performance indicators based on the divergence of the correct hypothesis (C1: separation from the second highest hypothesis, C2: the separation between the top hypothesis and the cluster mean) as the number of observations increases. This allowed us to demonstrate that the proposed method improves CPA under certain scenarios.

We also evaluated our approach with a conventional power traces dataset. We used a ChipWhisperer capture setup and created a corpus of well-aligned power traces. We then introduced some artificial misalignment to these data and evaluated the performance of the proposed selection rule. In this case, we identified that the method is sensitive to the width of the window, which can be upper-bounded by the ratio between the sampling frequency and the frequency of the target.

The proposed selection rule requires applying a sliding window over the correlation matrix along the time dimension for each hypothesis. This can be perceived as an application of a low-pass filter on the correlation matrix. Indeed, smoothening high frequency components contributes to discern the correct hypothesis with greater certainty. As a side effect, the amplitude of the correlation coefficients decreases. But this is true for all hypotheses. Yet, this is not a drawback since the proposed evaluation metrics rely on the normalized Euclidean distances within the cluster and thus are agnostic to this effect. The window function used with the proposed method simply approaches the probability density function of the correlation coefficients. We have defined its width as the number of samples captured from a cycle of the target and the windows are normalized so that the sum of their coefficients is 1. Therefore, it should be noted that as the width of the window decreases, the proposed selection rule approaches the behavior of CPA+`argmax`.

Identifying temporal dependencies within the correlation matrix has interesting implications. For example, it is possible to compute its Fourier transform and perform the hypothesis selection over a defined frequency band. However, this requires additional assumptions which are not necessary for the proposed method. The proposed approach is of special interest when dimension reduction/projection is not an option. In fact, the proposed technique can be seen as a projection of the post-processing results.

In our work, we used a dataset of power traces from an actual RPA setup. We have also created a synthetic dataset of traces that approach the characteristics of RPA scenarios where misalignment is found. We have also used a dataset of ChipWhisperer traces. These data may be found in the associated artifact.

Acknowledgments

The first, third and fourth authors are supported by grant 2021 SGR 00115 from the Government of Catalonia, by the projects ACITHEC PID2021-124928NB-I00 and MATSE PID2024-156636NB-C22 funded by MCIN/AEI/10.13039/501100011033/FEDER, EU, and by the project HERMES, funded by the European Union NextGenerationEU/PRTR via INCIBE. The second author is supported by the French Agence Nationale de la Recherche through the grant ANR-22-CE39-0008 (Project PROPHY).

References

[BCO04] Eric Brier, Christophe Clavier, and Francis Olivier. Correlation power analysis with a leakage model. In Marc Joye and Jean-Jacques Quisquater, editors, *Cryptographic Hardware and Embedded Systems – CHES 2004*, volume 3156 of

Lecture Notes in Computer Science, pages 16–29, Cambridge, Massachusetts, USA, August 11–13, 2004. Springer, Heidelberg, Germany.

[CKN01] Jean-Sébastien Coron, Paul C. Kocher, and David Naccache. Statistics and secret leakage. In Yair Frankel, editor, *FC 2000: 4th International Conference on Financial Cryptography*, volume 1962 of *Lecture Notes in Computer Science*, pages 157–173, Anguilla, British West Indies, February 20–24, 2001. Springer, Heidelberg, Germany.

[CRR03] Suresh Chari, Josyula R. Rao, and Pankaj Rohatgi. Template attacks. In Burton S. Kaliski Jr., Çetin Kaya Koç, and Christof Paar, editors, *Cryptographic Hardware and Embedded Systems – CHES 2002*, volume 2523 of *Lecture Notes in Computer Science*, pages 13–28, Redwood Shores, CA, USA, August 13–15, 2003. Springer, Heidelberg, Germany.

[DJJW25] Elena Dubrova, Sönke Jendral, Yanning Ji, and Ruize Wang. The Weighted Sum Correlation Analysis. Cryptology ePrint Archive, Paper 2025/1250, 2025.

[DMCB98] William P. Dunlap David M. Corey and Michael J. Burke. Averaging correlations: Expected values and bias in combined pearson rs and fisher's z transformations. *The Journal of General Psychology*, 125(3):245–261, 1998.

[DPRS11] Julien Doget, Emmanuel Prouff, Matthieu Rivain, and François-Xavier Standaert. Univariate side channel attacks and leakage modeling. *Journal of Cryptographic Engineering*, 1(2):123–144, August 2011.

[FGGLN25] Oriol Farràs, Vincent Grosso, Miquel Guiot, and Carlos Andres Lara-Nino. Improving the selection rule of correlation attacks for remote power analysis. Cryptology ePrint Archive, Paper 2025/1290, 2025. https://eprint.iacr.org/2025/1290.

[Fis15] R. A. Fisher. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4):507–521, 1915.

[FTBLN23] Anis Fellah-Touta, Lilian Bossuet, and Carlos Andres Lara-Nino. Combined internal attacks on SoC-FPGAs: Breaking AES with remote power analysis and frequency-based covert channels. In *2023 IEEE European Symposium on Security and Privacy Workshops*, pages 281–286, Delft, Netherlands, July 03–07, 2023. IEEE Computer Society Press.

[FTBLN24] Anis Fellah-Touta, Lilian Bossuet, and Carlos Andres Lara-Nino. A lightweight non-oscillatory delay-sensor for remote power analysis. In *2024 IEEE International Symposium on Hardware Oriented Security and Trust*, pages 343–348, Tysons Corner, VA, USA, May 06–09, 2024. IEEE.

[GBTP08] Benedikt Gierlichs, Lejla Batina, Pim Tuyls, and Bart Preneel. Mutual information analysis. In Elisabeth Oswald and Pankaj Rohatgi, editors, *Cryptographic Hardware and Embedded Systems – CHES 2008*, volume 5154 of *Lecture Notes in Computer Science*, pages 426–442, Washington, DC, USA, August 10–13, 2008. Springer, Heidelberg, Germany.

[GDTL21] Joseph Gravellier, Jean-Max Dutertre, Yannick Teglia, and Philippe Loubet-Moundi. SideLine: How delay-lines (may) leak secrets from your SoC. In Shivam Bhasin and Fabrizio De Santis, editors, *COSADE 2021: 12th International Workshop on Constructive Side-Channel Analysis and Secure Design*,

volume 12910 of *Lecture Notes in Computer Science*, pages 3–30, Lugano, Switzerland, October 25–27, 2021. Springer, Heidelberg, Germany.

[GHT05] Catherine H. Gebotys, Simon Ho, and C. C. Tiu. EM analysis of Rijndael and ECC on a wireless Java-based PDA. In Josyula R. Rao and Berk Sunar, editors, *Cryptographic Hardware and Embedded Systems – CHES 2005*, volume 3659 of *Lecture Notes in Computer Science*, pages 250–264, Edinburgh, UK, August 29 – September 1, 2005. Springer, Heidelberg, Germany.

[JZ23] Chengbin Jin and Yongbin Zhou. Enhancing non-profiled side-channel attacks by time-frequency analysis. *Cybersecurity*, 6(15), August 2023.

[KJJ99] Paul C. Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In Michael J. Wiener, editor, *Advances in Cryptology – CRYPTO'99*, volume 1666 of *Lecture Notes in Computer Science*, pages 388–397, Santa Barbara, CA, USA, August 15–19, 1999. Springer, Heidelberg, Germany.

[Koc96] Paul C. Kocher. Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems. In Neal Koblitz, editor, *Advances in Cryptology – CRYPTO'96*, volume 1109 of *Lecture Notes in Computer Science*, pages 104–113, Santa Barbara, CA, USA, August 18–22, 1996. Springer, Heidelberg, Germany.

[LCC+06] Thanh-Ha Le, Jessy Clédière, Cécile Canovas, Bruno Robisson, Christine Servière, and Jean-Louis Lacoume. A proposition for correlation power analysis enhancement. In Louis Goubin and Mitsuru Matsui, editors, *Cryptographic Hardware and Embedded Systems – CHES 2006*, volume 4249 of *Lecture Notes in Computer Science*, pages 174–186, Yokohama, Japan, October 10–13, 2006. Springer, Heidelberg, Germany.

[LCSL07] Thanh-Ha Le, Jessy Clediere, Christine Serviere, and Jean-Louis Lacoume. Efficient solution for misalignment of signal in side channel analysis. In *IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP 2017*, pages II–257–II–260, Honolulu, HI, USA, April 15–20, 2007. IEEE.

[LPR13] Victor Lomné, Emmanuel Prouff, and Thomas Roche. Behind the scene of side channel attacks. In Kazue Sako and Palash Sarkar, editors, *Advances in Cryptology – ASIACRYPT 2013, Part I*, volume 8269 of *Lecture Notes in Computer Science*, pages 506–525, Bengalore, India, December 1–5, 2013. Springer, Heidelberg, Germany.

[LPR+14] Victor Lomné, Emmanuel Prouff, Matthieu Rivain, Thomas Roche, and Adrian Thillard. How to estimate the success rate of higher-order side-channel attacks. In Lejla Batina and Matthew Robshaw, editors, *Cryptographic Hardware and Embedded Systems – CHES 2014*, volume 8731 of *Lecture Notes in Computer Science*, pages 35–54, Busan, South Korea, September 23–26, 2014. Springer, Heidelberg, Germany.

[Man04] Stefan Mangard. Hardware Countermeasures against DPA – A Statistical Analysis of Their Effectiveness. In Tatsuaki Okamoto, editor, *Cryptographers' Track at the RSA Conference*, volume 2964 of *Lecture Notes in Computer Science*, pages 222–235, Heidelberg, Germany, February 2004. Springer.

[MOP08] Stefan Mangard, Elisabeth Oswald, and Thomas Popp. *Power analysis attacks: Revealing the secrets of smart cards*. Springer, Heidelberg, Germany, Berlin-Heidelberg, Germany, first edition, 2008.

[MOS09] Stefan Mangard, Elisabeth Oswald, and Francois-Xavier Standaert. One for all - all for one: Unifying standard DPA attacks. Cryptology ePrint Archive, Report 2009/449, 2009. https://eprint.iacr.org/2009/449.

[OP13] David Oswald and Christof Paar. Improving side-channel analysis with optimal linear transforms. In Stefan Mangard, editor, *International Conference on Smart Card Research and Advanced Applications - CARDIS 2013*, pages 219—-233, Graz, Austria, November 28–30, 2013. Springer Cham.

[OSB99] Alan V. Oppenheim, Ronald W. Schafer, and John R. Buck. *Discrete-Time Signal Processing*. Prentice Hall, Upper Saddle River, NJ, 1999.

[PG95] Karl Pearson and Francis Galton. Vii. note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58(347-352):240–242, 1895.

[PRB09] Emmanuel Prouff, Matthieu Rivain, and Regis Bevan. Statistical analysis of second order differential power analysis. *IEEE Transactions on Computers*, 58(6):799–811, January 2009.

[SGMT21] Falk Schellenberg, Dennis R. E. Gnad, Amir Moradi, and Mehdi B. Tahoori. An inside job: Remote power analysis attacks on FPGAs. *IEEE Design & Test*, 38(3):58–66, 2021.

[SLP05] Werner Schindler, Kerstin Lemke, and Christof Paar. A stochastic model for differential side channel cryptanalysis. In Josyula R. Rao and Berk Sunar, editors, *Cryptographic Hardware and Embedded Systems – CHES 2005*, volume 3659 of *Lecture Notes in Computer Science*, pages 30–46, Edinburgh, UK, August 29 – September 1, 2005. Springer, Heidelberg, Germany.

[SMY09] François-Xavier Standaert, Tal Malkin, and Moti Yung. A unified framework for the analysis of side-channel key recovery attacks. In Antoine Joux, editor, *Advances in Cryptology – EUROCRYPT 2009*, volume 5479 of *Lecture Notes in Computer Science*, pages 443–461, Cologne, Germany, April 26–30, 2009. Springer, Heidelberg, Germany.

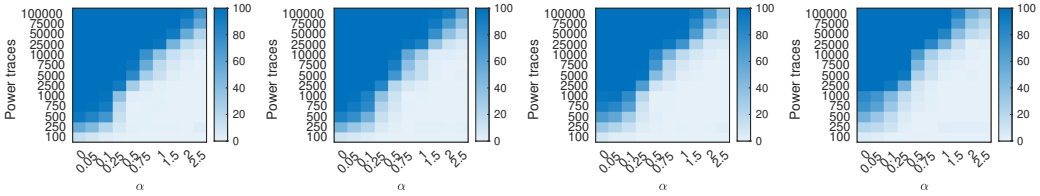
[VGS13] Nicolas Veyrat-Charvillon, Benoît Gérard, and François-Xavier Standaert. Security evaluations beyond computing power. In Thomas Johansson and Phong Q. Nguyen, editors, *Advances in Cryptology – EUROCRYPT 2013*, volume 7881 of *Lecture Notes in Computer Science*, pages 126–141, Athens, Greece, May 26–30, 2013. Springer, Heidelberg, Germany.

[ZS18] Mark Zhao and G. Edward Suh. FPGA-based remote power side-channel attacks. In *2018 IEEE Symposium on Security and Privacy*, pages 229–244, San Francisco, CA, USA, May 21–23, 2018. IEEE Computer Society Press.

A Attacks' correctness

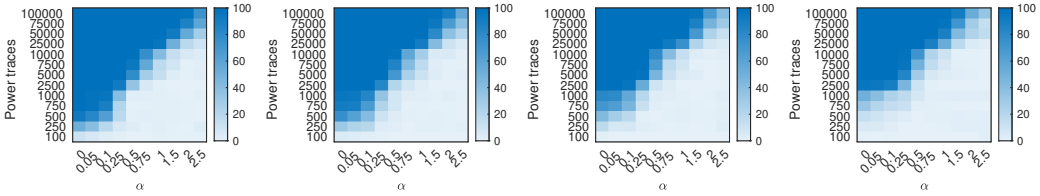
In Figs. 14 and 15 we provide results for the success rates of CPA on the synthetic traces introduced in Section 5. In these experiments we do not use any performance indicators. This complicates conducting any comparison, beyond intuitive visual analyses. However, we provide these data to illustrate the actual success rates of the attacks.

From these results, some readers might appreciate a slight advantage of the `meanmax` selection rule. We have quantified this improvement in Section 5 by measuring the divergence of the correct hypothesis. We note that counterintuitively in some cases CPA+`meanmax` is correct when CPA+`argmax` is not.



(a) $\beta = 10$. **(b)** $\beta = 20$. **(c)** $\beta = 30$. **(d)** $\beta = 40$.

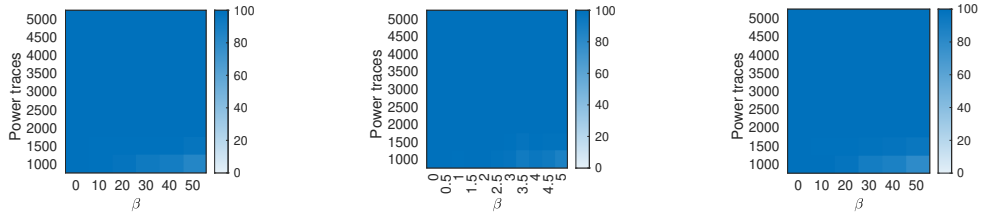
Figure 14: Percentage of bytes retrieved correctly by CPA+`argmax` in function to the number of traces (y), the noise gain (α), and the misalignment rate (β).



(a) $\beta = 10$. **(b)** $\beta = 20$. **(c)** $\beta = 30$. **(d)** $\beta = 40$.

Figure 15: Percentage of bytes retrieved correctly by CPA+`meanmax` in function to the number of traces (y), the noise gain (α), and the misalignment rate (β).

Similarly, we provide attack-correctness results for the experiments in Section 6.



(a) CPA+`argmax`. **(b)** CPA+`meanmax`, $s = 24$. **(c)** CPA+`meanmax`, $s = 4$.

Figure 16: Percentage of bytes retrieved correctly in function to the number of traces (y) and the misalignment (β). For CPA+`meanmax` we provide results using windows of different widths (s).