Generic-compatible distinguishers for linear regression based attacks

Sana Boussam*

LIX, INRIA, CNRS, École polytechnique, Institut Polytechnique de Paris, Palaiseau, France
sana.boussam@inria.fr

Abstract. Non profiled attacks aim to recover secret information from a device without prior knowledge of its leakage model. However, most practical non profiled attacks, such as Differential Power Analysis, Correlation Power Analysis, and Linear Regression-based Attacks (LRA), still depend on *a priori* leakage assumptions. Designing a generic attack that does not rely on any such assumption therefore remains an open problem and has been actively investigated by the side-channel community for more than a decade. Although Whitnall *et al.* showed in [37] that LRA can be considered generic when all predictors are included, this is not feasible in practice due to inherent multicollinearity issues and the inadequacy of classical distinguishers, which lose discriminating power when targeting injective functions. In this work, we overcome these limitations and propose the first fully generic-compatible non profiled attack. We show that using a Walsh-Hadamard basis enables generic LRA by eliminating multicollinearity and allowing all predictors to be considered without loss of precision. We also introduce new generic-compatible distinguishers tailored to LRA and formally prove their soundness for both linear and non-linear cryptographic operations. Finally, we validate our approach experimentally using simulations and publicly available datasets.

Keywords: Side-Channel Attacks · Non profiled attacks · Linear Regression based Attacks · Generic attacks · Walsh-Hadamard basis · Sound distinguisher.

1 Introduction

Side-Channel Attacks (SCA) are a type of physical attacks in which an attacker exploits implementation weaknesses of a cryptographic algorithm in embedded systems in order to extract secret information, by targeting all types of physical leaks generated during the execution of this algorithm such as execution time [21], power consumption [20] or electromagnetic radiation [1]. SCA can be divided into two categories: *profiled* attacks and *non profiled* attacks. *Profiled*

* This work was carried out while the author was a PhD student in collaboration with Thales ITSEF, Toulouse, France.

attacks [9,34] are one of the most powerful SCA and require a device similar to the targeted one in order to build a learning database that models device behavior (*i.e.* leakage model) according to the values of sensitive variables that depend on the secret. *Non profiled attacks* [20,4,12] are less restrictive than *profiled attacks*. In this case, an attacker only has traces and possibly some details about the implementation of the targeted algorithm to carry out his attack. In an evaluation context, this type of attacks are more realistic as they are the closest to real attacks cases. There are different type of attacks which are based on various types of distinguishers.

Historically, one of the first *non profiled* attack introduced in SCA community by Kocher *et al.* is Differential Power Analysis (DPA) [20]. In their work, Kocher *et al.* showed that a difference of means exploitation leads to recover secret information. Since then, numerous attacks exploiting a difference of means appear in the literature using various distinguishers. The best-known and most widespread attacks in literature are Correlation Power Analysis (CPA) [4], Mutual Information Analysis (MIA) [12] and Linear Regression Analysis (LRA) [34,11], which use respectively as distinguishers Pearson's correlation coefficient, mutual information and Euclidean distance-based distinguishers (*e.g.* coefficient of determination R^2). However, a major limitation of the aforementioned *non profiled* attacks lies in the fact that, in practice, an attacker must have some prior knowledge about the leakage model when using any of these different attacks. For instance, the adversary must assume that the leakage model is defined by the Hamming weight, the Hamming distance, etc. Such prior knowledge is restrictive since, in practice, no one knows the exact leakage model that should be considered.

Hence, over at least the last ten years, the SCA community has been working to develop a 'generic' *non profiled* attack. Such an attack would allow recovering secret information from any device without any *a priori* about leakage characteristic of the targeted device. Throughout this paper, the term 'generic' refers thus to attacks that do not rely on explicit prior assumptions on the leakage model (see Definition 4). This notion of genericity has important implications for the evaluation of cryptographic devices: if such attacks exist, any implementation can in principle be evaluated without prior knowledge of its leakage characteristics, making profiled approaches mainly relevant in terms of efficiency rather than applicability. Conversely, if fully generic non profiled attacks are not achievable, some devices may only be reliably assessed using profiled methods. In [37, Proposition 4], Whitnall *et al.* show that LRA can be regarded as generic when all predictors are included. However, in practice, such an attack cannot be conducted in a full-basis setting when targeting an injective function, since the classical distinguishers used to conduct LRA fail to discriminate the correct key [37]. As a consequence, LRA is usually performed using only a restricted subset of predictors, which inherently prevents it from achieving its full generic potential. To circumvent this issue, Whitnall *et al.* propose a 'generic-emulating' LRA distinguisher: stepwise regression based on a non-device-specific intuition, namely that the regression coefficients under the correct key tend to be simple and structured *i.e.* only a few coefficients carry meaningful information and

they typically correspond to low-degree bits' interactions. However, as pointed out in [27], stepwise regression tends to overfit and is sensitive to predictors order, which leads to an unreliable and non-deterministic final model. In particular, changing the order in which predictors are considered may result in a different selected model, so that convergence toward a unique solution is not guaranteed. Hence, the problem of designing generic attacks in the sense of full-basis LRA [37] remains still open to this day. Until now, it has generally been believed that fully generic non profiled attacks are not achievable in practice, and that relying solely on non profiled evaluations may overestimate device security. We nuance this view by showing that generic non profiled attacks can be meaningfully approached using LRA in a full-basis setting with appropriate representations and distinguishers, suggesting that the gap between profiled and non profiled attacks should primarily be understood in terms of efficiency rather than applicability. In this work, we propose new theoretically grounded and generic-compatible distinguishers specifically designed for full-basis LRA.

Contributions. The main contribution of our paper is twofold. First, we propose a new basis to perform linear regression based attacks (LRA), thereby mitigating the limitations of the classical monomial basis and enabling non profiled generic attacks in the sense of full-basis predictor modeling without requiring explicit leakage model assumptions. Second, to better exploit the genericity offered by this basis in full-basis settings, we introduce two new sound and generic-compatible distinguishers, well suited for LRA (see Definition 9 and Definition 10). Our goal is not to claim systematic superiority over existing distinguishers in all settings, but rather to design sound distinguishers that remain effective in settings where classical distinguishers can fail. An overview of our main contributions is presented in Figure 1 (this summary is not comprehensive).

To justify the need for new LRA distinguishers, we provide further theoretical contributions regarding generic LRA and the current state-of-the-art LRA distinguishers. First, we propose to consider Walsh-Hadamard basis, which was introduced first in SCA context for profiled attacks [14], to perform LRA. We show that using this basis is more suitable for conducting generic LRA than traditional bases. Indeed, we prove in Theorem 1 and Corollary 1 that, unlike traditional bases, the use of Walsh-Hadamard basis, even when an attacker performs a generic LRA, doesn't lead to multicollinearity issues, which may cause a loss of precision of the estimated coefficients [16, Chapter 10]. Then, we highlight in Theorem 2, Corollary 2 and Corollary 3 the limitations encountered when performing both generic and implementation-optimized LRA [22] using Walsh-Hadamard basis and the current state-of-the-art LRA distinguishers. Indeed, we show that in such a configuration the latter lose their discriminating ability. Based on these results, we highlight the need of new distinguishers that are suitable for LRA. To properly design these new distinguishers, we provide additional theoretical contributions. We first prove in Proposition 3 that the underlying goal of all the current state-of-the-art LRA distinguishers consists in a minimization of a difference between the true determinisitic part and an

estimated one. This philosophy is thus derived to design our new distinguishers.

Moreover, we highlight some interesting properties when using Walsh-Hadamard basis to perform LRA, namely the bias and the uninformative samples do not provide information during the key recovery (see Proposition 4 and Proposition 5 respectively). These properties are exploited to design and justify our distinguishers.

Finally, by combining side-channel attack theory and Boolean function theory, we formally establish the soundness of each proposed distinguisher under well-defined conditions. We prove their soundness both when LRA is performed with a full basis (see Theorem 3 and Theorem 4) and when LRA is performed with a non-full basis (see Corollary 4 and Corollary 5). Moreover, we also prove in Proposition 6 and Proposition 7 that our distinguishers remain sound when using implementation-optimized versions of LRA. To the best of our knowledge, this work is the first SCA work that provides theoretical proofs of soundness for LRA distinguishers.

We experimentally assess and validate all theoretical results provided by our distinguishers and all theoretical claims about state-of-the-art LRA distinguishers limitations using both simulations and publicly available datasets, hence covering a wide range of targeted cryptographic functions. The whole implementation as well as notebooks containing all the experiments carried out in this article, are available in the following GitHub repository: `https://github.com/Sbsm157/LRA_generic_distinguishers`.

Paper organization. Our paper is organized as follow: in Section 2, we define our notation and recall important concepts that will be reused throughout the paper. Then, we highlight in Section 3 limitations of the state-of-the-art linear regression attacks metrics (*i.e.* bases projections and distinguishers). In Section 4, we propose novel distinguishers for linear regression based attacks that address limitations highlighted in the previous section. We experimentally validate in Section 5 and Section 6 all our theoretical claims and results using both simulations and publicly available datasets covering a wide range of use-cases. Finally, we conclude in Section 7.

2 Background

In this section we define notations, recall definitions and summarize notions that will be useful throughout the reading of this article.

2.1 Notations and definitions

Probability theory. Calligraphic letters such as $\mathcal{X}$ denote sets. The cardinality of the set $\mathcal{X}$ -if finite- is denoted $|\mathcal{X}|$. The corresponding bold capital letter $\mathbf{X}$

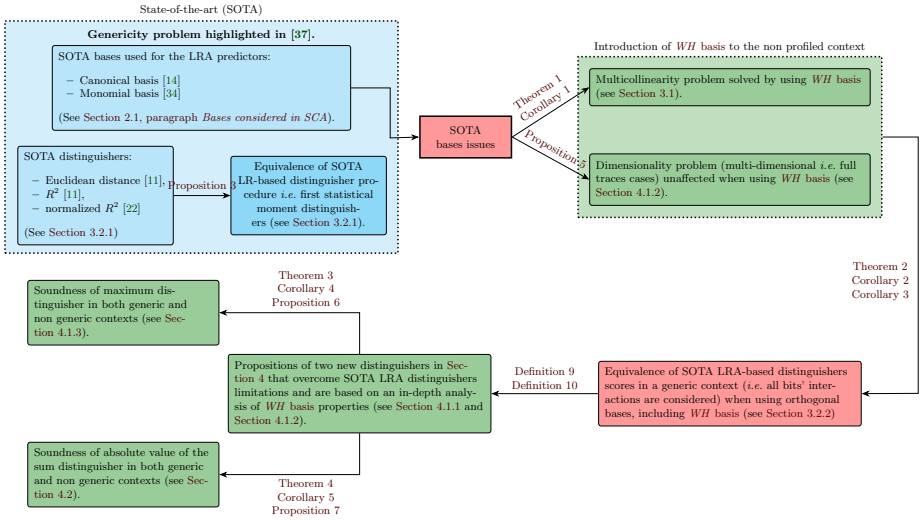


Fig. 1: Overview of our main theoretical contributions. Blue: latest advancements in the LRA state-of-the-art. Red: issues/limitations encountered. Green: proposed solutions to address these limitations.

denotes a random variable over $\mathcal{X}$. Its corresponding bold lowercase $\mathbf{x}$ refers to a realization of $\mathbf{X}$. The notation $\mathbf{X} \sim \mathcal{N}_D(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ (resp. $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$) refers to a random vector $\mathbf{X}$ of dimension D (resp. a random scalar) that follows a multivariate Gaussian distribution (resp. Gaussian distribution) of parameters $\boldsymbol{\mu} \in \mathbb{R}^D$ (resp. $\boldsymbol{\mu} \in \mathbb{R}$) and $\boldsymbol{\Sigma} \in \mathbb{R}^{D \times D}$ (resp. $\boldsymbol{\sigma^2} \in \mathbb{R}$). The expectation (resp. variance) of a random variable is denoted $\mathbb{E}[\cdot]$ (resp. $\mathbb{V}[\cdot]$). We denote $\mathbf{a} \xleftarrow{\$} \mathcal{A}$ when $\mathbf{a}$ is chosen uniformly in $\mathcal{A}$. Let $(\mathbf{X}_Q)_{Q \geqslant 1}$ be a sequence of random variables and $(a_Q)_{Q \geqslant 1}$ constants. We write $\mathbf{X}_Q = O_p(a_Q)$ if for every $\varepsilon > 0$ there exists $M > 0$ such that $\sup_{Q \geqslant 1} \Pr(|\mathbf{X}_Q| > M a_Q) < \varepsilon$. If $\mathbf{X}_Q = O_p(a_Q)$ and $a_Q \to 0$, then $\mathbf{X}_Q \to 0$ in probability, meaning that the claimed bound holds for all but a negligible fraction of possible realizations.

Linear algebra. The ith entry of a vector A is denoted A_i. $\mathcal{M}_{n,m}(\pm 1)$ denotes the matrix space of n rows and m columns so that all coefficients of matrices belonging to this space are in $\{-1, 1\}$. For a given matrix M, the entry of the ith row and j^{th} column is denoted $M_{i,j}$. We denote $M_{i,:}$ (resp. $M_{:,j}$) the ith row (resp. j^{th} column) of the matrix M. The transpose and inverse operator will be denoted $(\cdot)^\top$ and $(\cdot)^{-1}$ respectively. The identity matrix of size $D \times D$ is denoted I_D. Given a vector A, we denote its infinity norm and Euclidean norm as $\|A\|_\infty$ and $\|A\|_2$ respectively.

Recall on Boolean functions theory. Considering $a, b \in \mathbb{F}_2^n$, we define the usual dot product in $\mathbb{F}_2^n$ as $a \cdot b = \oplus_{i=1}^n a_i \wedge b_i$ where $\oplus$ and $\wedge$ corresponds

respectively to addition and multiplication in $\mathbb{F}_2$. A n-variable Boolean function g is a mapping from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ defined as

$$\begin{aligned} g : \mathbb{F}_2^n &\to \mathbb{F}_2 \\ X &\mapsto g(x_1, \ldots, x_n) \end{aligned}$$

Hence, based on this definition, a (n, m)-vectorial Boolean function h which corresponds to a mapping from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$ is defined as

$$\begin{aligned} h : \mathbb{F}_2^n &\to \mathbb{F}_2^m \\ X &\mapsto (g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n)) \end{aligned}$$

where $g_v(X)$ are Boolean functions $g_v(X) = v \cdot h(X)$, $v \in \mathbb{F}_2^m$ which are called the coordinate functions of h. The Walsh transform of h is defined by the Walsh transform of its coordinate functions g_v [8]. Given $u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m$, we have:

$$W_{u,v}^h = \sum_{x \in \mathbb{F}_2^n} (-1)^{u \cdot x \oplus v \cdot h(x)}.$$

The linear approximation table (LAT) of the vectorial Boolean function h denoted LAT^h is a $2^n \times 2^m$ table such that $LAT_{u,v}^h = W_{u,v}^h$. In the case of a bijective (n, n)-vectorial Boolean function, the LAT contains only even values and the expected number of zeros entries amounts to $\frac{4}{\sqrt{2\pi \times 2^n}}$ [28]. The linearity of a given vectorial Boolean function, denoted $l(h)$, corresponds to the maximum absolute entry of LAT^h, excluding the first line and column *i.e.* :

$$l(h) = \max_{\substack{u \in \mathbb{F}_2^n \setminus \{0\}, \\ v \in \mathbb{F}_2^m \setminus \{0\}}} |LAT_{u,v}^h| = \max_{\substack{u \in \mathbb{F}_2^n \setminus \{0\}, \\ v \in \mathbb{F}_2^m \setminus \{0\}}} \Big| \sum_{x \in \mathbb{F}_2^n} (-1)^{u \cdot x \oplus v \cdot h(x)} \Big|.$$

The nonlinearity of a vectorial Boolean function denoted $nl(h)$ [8] is thus defined as:

$$nl(h) = 2^{n-1} - \frac{1}{2} l(h).$$

It plays a key role in the design of cryptographic functions such as Substitution Boxes (SBoxes) as it quantifies the resistance of a function against linear cryptanalysis [26]. According to the covering radius bound, the nonlinearity of a vectorial Boolean function can be bounded by $nl(h) \leqslant 2^{n-1} - 2^{\frac{n}{2}-1}$ [18,8]. In the rest of this paper we consider a subset of vectorial Boolean function, namely the bijective (n, n)-vectorial Boolean functions.

SCA terminology. In SCA domain, the secret key recovery is often performed using a divide-and-conquer strategy. It consists in targeting separately each secret key chunk $k_i^* \in \mathcal{K}$. In this paper, we focus on targeting a secret key chunk where $\mathcal{K} = \mathbb{F}_2^n$ with $n \geqslant 1$. To ease the reading, we denote a secret key chunk $k^* \in \mathbb{F}_2^n$ and refers to it as the correct secret key in the rest of the paper. A leakage trace $\mathbf{T} \in \mathbb{R}^D$ is a random vector that consists in a deterministic part Ψ and an independent noise $\mathbf{N}$ such that $\mathbf{T} = \Psi\big(Y^{k^*}\big) + \mathbf{N}$, where $Y^{k^*} = f(X, k^*) \in \mathbb{F}_2^n$

is a sensitive value that corresponds to the image of a cryptographic function f (*e.g.* SBox function) given a public input (*e.g.* plaintext, ciphertext) chunk $X \in \mathbb{F}_2^n$ and the secret key chunk $k^* \in \mathbb{F}_2^n$. In this paper, the deterministic part is defined as a pseudo-boolean function [6] (*i.e.* the most generic definition). Hence, Ψ takes the form of a multi-linear polynomial $\Gamma(Y^k)$ of maximal degree n, composed of 2^n monomials which represent the contribution of bits' interactions of Y^k.

In the rest of the paper, to ease the reading, this multi-linear polynomial will be considered as vector in which each entry corresponds to a monomial *i.e.* $\Gamma(Y^k) \in \mathbb{F}_2^{2^n}$. Hence, we define the leakage model Ψ as:

$$\begin{aligned} \Psi : \mathbb{F}_2^n &\to \mathbb{R}^D \\ Y^k &\mapsto \mathcal{B}(\Gamma(Y^k)) \cdot \alpha \end{aligned} \tag{1}$$

where $\mathcal{B}(\Gamma(Y^k)) \in \mathbb{F}_2^{2^n}$ and $\alpha \in \mathbb{R}^{2^n \times D}$ is a matrix of coefficients representing contribution of each monomial/bits' interaction for all D time samples. We note $\mathcal{B}$ the basis on which Γ is projected. Following the independent Gaussian noise assumption, we assume that the noise $\mathbf{N}$ follows a centered Gaussian distribution $\mathcal{N}_D(\mathbf{0^D}, \boldsymbol{\Sigma})$ and is independent at each time sample from the sensitive variable. Hence, this implies that $\mathbf{T} \sim \mathcal{N}_D(\Psi(Y^{k^*}), \boldsymbol{\Sigma})$.

Usual bases considered in SCA. Bases $\mathcal{B}$ usually considered in SCA to project Γ are either the canonical basis or the monomial basis, which was introduced into SCA context by Schindler *et al.* through profiled stochastic attacks [34]. However, as pointed out in [14], these bases suffer from intrinsic limitations that prevent a clear and comprehensible reading of bits and bits' interaction contributions. Indeed, since the monomial basis is neither orthonormal nor even orthogonal, the monomials of this basis (*i.e.* bits and bit interactions), also called predictors, are correlated. This correlation makes it difficult to interpret the true contribution of each individual bit or bits' interaction. The canonical basis does not suffer from this problem, as it is orthonormal by construction. However, when considering this basis, the degree of the Boolean function characterizing bits' interaction is not preserved (unlike monomial basis) and equals n for all bits' interactions. Consequently, this makes bits' interactions not easily interpretable. To circumvent limitations of these two bases, Guilley *et al.* propose to consider a new basis to conduct profiled SCA: the Fourier basis [14]. This basis corresponds to a scaled Walsh-Hadamard matrix that has been permuted in such a way as to be ordered by increasing Hamming Weight. Thanks to its orthonormality and degree preservation proprieties, this basis allows a clear and comprehensible reading of the contribution of each monomial/bits' interaction.

Definition 1. ***Walsh-Hadamard (WH) basis.*** *Given a variable $x \in \mathbb{F}_2^n$ and $u \in \mathbb{F}_2^n$ the support of a monomial which characterizes bits' interactions,*

$$(WH)_{x,u} = \frac{1}{2^{\frac{n}{2}}}(H_{2^n})_{x,u} = \frac{1}{2^{\frac{n}{2}}}(-1)^{u \cdot x}$$

where $\cdot$ denotes the usual inner product in $\mathbb{F}_2^n$ and $H_{2^n} \in \mathcal{M}_{2^n,2^n}(\pm 1)$ is a Sylvester constructed Hadamard matrix in which columns are sorted, according to their indexes, by increasing Hamming Weight order.

In the rest of the paper, we refer to this basis as *WH* basis. If the Hadamard matrix H_{2^n} must be built from scratch, the overall complexity of the projection onto *WH* is dominated by the explicit generation of H_{2^n}, which requires $\mathcal{O}(2^n)$ operations. In practice, H_{2^n} is often pre-computed or already available. As a result, the complexity of the projection onto *WH* is primarily determined by the computation of $u \cdot x$, which requires $\mathcal{O}(n)$ operations.

2.2 Classical non profiled SCA

Given a set of leakage traces $\mathcal{T}$ corresponding to the manipulation of a fixed secret key k^*, the goal of a side-channel attack is to recover k^* from $\mathcal{T}$. More formally, side-channel attacks exploit the statistical dependence between physical leakages and sensitive intermediate values computed during the execution of a cryptographic algorithm, and can be defined as follows.

Definition 2. ***Side-Channel attacks*** *Let $\mathcal{X} = \{x_1, \ldots, x_Q\}$ be a set of Q public inputs and $\mathcal{T} = \{\mathbf{t}_1, \ldots, \mathbf{t}_Q\}$ the associated leakage traces such that, for all $i \in [1, Q]$, $\mathbf{t}_i = \Psi(y_i^{k^*}) + \mathbf{n}_i \quad with \quad y_i^{k^*} = f(x_i, k^*)$.*
The goal of the attacker is to select a distinguisher $\mathcal{D}$ such that, by considering as few traces as possible,

$$\hat{k} = \arg\max_{k \in \mathcal{K}} \mathcal{D}^k(\mathcal{T}, \mathcal{X}), \tag{2}$$

with $\hat{k} = k^*$.

Depending on the attacker's capabilities, two main classes of SCAs are usually distinguished: *profiled* and *non profiled* attacks. *Profiled* attacks [9,34] are considered the most powerful class of SCAs, as they assume access to a device similar to the target, which allows the attacker to characterize the leakage model during a dedicated profiling phase. This leakage model can then be exploited through the log-likelihood distinguisher, which is optimal when the leakage distributions are correctly estimated [19]. *Non profiled* attacks [20,4,12] are less restrictive and more realistic, as they only require leakage traces, public data such as plaintexts or ciphertexts, and possibly some knowledge about the implementation. In this case, the leakage model is unknown and the attacker must rely on alternative statistical distinguishers.
In the remainder of this paper, and without any loss of generality, we consider the standard setting where the attacker can choose or observe public data that are uniformly distributed, an assumption commonly made in the side-channel literature and verified in practice [25,31], and which we formalize below.

Assumption 1. *Public data such as plaintexts or ciphertexts are assumed to be uniformly distributed over $\mathbb{F}_2^n$. Furthermore, for a given set of Q public inputs, we assume that each possible public input value occurs at least $\left\lfloor \frac{Q}{2^n} \right\rfloor$ times.*

Consequently, Q can be expressed as $Q = A \times 2^n + R$ where $A = \left\lfloor \frac{Q}{2^n} \right\rfloor$ and $0 \leqslant R = Q \bmod 2^n < 2^n$.

This assumption is standard in the side-channel literature and is typically satisfied in practice [36,30,2] (see Section 6). In the following, we recall three representative non profiled attacks based on different distinguishers, namely correlation power analysis, mutual information analysis, and linear regression analysis.

Correlation Power Analysis. Introduced in 2003 by Clavier *et al.* as a generalization of Differential Power Analysis (DPA) [20], Correlation Power Analysis or CPA [4] exploits and assume that there exists a linear relationship between the leakage traces and a predicted leakage model $\hat{\Psi}^k$, when considering a good key hypothesis. This leakage model is built from a key hypothesis $k \in \mathcal{K}$ and a prediction of the leakage function which is usually the Hamming Weight function. This attack is considered as optimal when the true leakage model Ψ corresponds to the Hamming Weight model up to some affine transformation [11,19]. However, as the true leakage model is not known by the attacker in a *non profiled* attack context, the use of the CPA might prove insufficient, as it only detects linear dependence between the leakage traces and the predicted leakage model and does not capture the non-linear dependencies.

Mutual Information Analysis. To address CPA limitations, Gierlichs *et al.* propose the Mutual Information Analysis (MIA) [12]. This attack is based on using mutual information as an alternative distinguisher in order to exploit all possible linear and non-linear dependencies between the leakage traces and predicted leakage model. Although MIA performs a more generic attack than CPA as it captures linear and non-linear dependencies between the predicted leakage model and leakage traces, it suffers from two major limitations. The first limitation is related to mutual information computation. Indeed, the computation of the conditional entropy $H\big(\mathbf{T}|\hat{\Psi}^k(Y^k)\big)$ requires an estimation, for all key hypotheses k, of the probability density functions of the conditional leakage traces *i.e.* $\mathbb{P}\big(\mathbf{T}|\hat{\Psi}^k(Y^k)\big)$. While several works investigate some methods to circumvent this issue (*e.g.* histograms, kernel density function, or parametric estimation [32]), selecting the appropriate one can be very challenging for an attacker. The second limitation concerns leakage function selection. As the distinguisher used is the mutual information, the attacker has a constraint on the leakage model assumption. Indeed, $\hat{\Psi}^k$ must be non-injective [32]. As a result, the choice of the leakage function is restricted for the attacker, who usually ends up choosing the Hamming weight as leakage function.

Linear Regression Analysis. To be generic regarding the choice of leakage model and thus to circumvent MIA limitations, an attacker can consider the Linear Regression Analysis (LRA) [34,11]. LRA is a *non profiled* adaptation of a *profiled* attack called stochastic attacks [34]. This type of attacks allows the attacker to consider any form of leakage model as the latter is seen as a weighted sum of bits and possibly bits interactions of the sensitive value. Indeed, given

a key hypothesis k, a sensitive value Y^k and $\Gamma_m(Y^k) \in \mathbb{F}_2^{\sum_{i=0}^{m}\binom{n}{i}}$ the binary decomposition of Y^k in which one considers up to the mth bits interaction degree with $m \leqslant n$ (see Section 2.1), LRA consists in computing for all k the predicted leakage model $\hat{\Psi}^k$ by estimating coefficients $\widehat{\alpha}^k$. This coefficients estimation is done by performing a linear regression (ordinary least squares). The correct key is then retrieved by minimizing an Euclidean distance based distinguisher that can be formally described as follows:

Definition 3. ***Euclidean distance based LRA distinguisher [34,11].*** *Given a D-dimensional leakage trace $\mathbf{T}$, its associated public input X, a key hypothesis $k \in \mathcal{K}$, the sensitive value $Y^k = f(X,k)$ and an hypothetical leakage model $\hat{\Psi}^k$ such that $\hat{\Psi}^k(Y^k) = \Gamma_m(Y^k) \cdot \widehat{\alpha}^k$ with $\widehat{\alpha}^k \in \mathbb{R}^{\sum_{i=0}^{m}\binom{n}{i} \times D}$ and $\Gamma_m \in \mathbb{F}_2^{\sum_{i=0}^{m}\binom{n}{i}}$ with $m \leqslant n$, the LRA distinguisher based on Euclidean distance is defined as*

$$\mathcal{D}_{ED}^k(\mathbf{T}, X) = \min_{j \in [1,D]} ED_j^k \tag{3}$$

$$= \min_{j \in [1,D]} \|\mathbf{T}_j - \hat{\Psi}_j^k(Y^k)\|_2^2 \tag{4}$$

where $\|\cdot\|_2$ denotes the Euclidean distance. Based on this distinguisher, the distinguishing rule applied is then

$$\hat{k} = \arg\min_{k \in \mathcal{K}} \mathcal{D}_{ED}^k(\mathbf{T}, X)$$

where $\hat{k}$ is the key candidate.

Other distinguishers can be considered to perform a LRA such as the coefficient of determination R^2 that we define in Section 3.2.1. Hence this attack, in contrast to MIA, offers the attacker greater flexibility regarding the choice of leakage models. This flexibility makes LRA a particularly appealing candidate for the design of generic SCAs, *i.e.* , attacks that do not rely on any a priori assumption on the leakage behavior. This question has been formally investigated by Whitnall *et al.* in [37], leading to the notion of generic attacks.

Generic attacks. In [37], Whitnall *et al.* formalize the concept of generic side-channel attacks as follows.

Definition 4. ***Generic Attacks [37, Definition 4].*** *A generic strategy consists in performing a standard univariate Differential Power Analysis using:*

- *a generic power model, namely a nominal labeling of the equivalence classes induced by the key-hypothesised target function,*
- *a generic-compatible distinguisher operating on nominal-scale measurements.*

In other words, such attacks aim at recovering secret information from physically observable device leakage without any a priori knowledge about the device's leakage characteristics. In practice, the generic power model typically corresponds

to the identity labeling of the targeted sensitive variable, interpreted nominally only [37]. Moreover, it should be noted, as discussed in [37], that this notion of genericity concerns the absence of an explicit leakage model for the data-dependent leakage, but does not exclude additional assumptions on the statistical setting or noise behavior. In particular, in [37, Proposition 4], Whitnall *et al.* show that LRA can be regarded as a generic non profiled SCA when considering a leakage model with a full basis, *i.e.* when all bit interactions are taken into account.

Proposition 1. ***Genericity of LRA [37, Proposition 4].*** *Conducting a LRA with full basis constitutes a generic strategy.*

However, they also establish a fundamental limitation of the generic attack framework.

Proposition 2. ***Impossibility of generic SCA [37, Proposition 1].*** *There is no generic side-channel attack able to distinguish the correct key from an other key hypothesis if the composition of the function connecting the secret key to the public input (e.g. plaintext, ciphertext) is injective.*

In particular, when the targeted function is injective (*e.g.* the AES S-box), standard LRA distinguishers such as R^2 are unable to discriminate and thus retrieve the correct key hypothesis, when considering a leakage model with a full basis $\mathcal{B}(\Gamma_m)$ where $m = n$ [10,37]. Nevertheless, Whitnall *et al.* observed in [37] that, in contrast to classical non profiled attacks such as CPA or MIA which only output a scalar distinguisher score per key hypothesis, LRA uniquely provides, in addition, key-dependent sets of estimated coefficients. These coefficients contain structural information about the relationship between the sensitive variables and the leakage, and can therefore be seen as a form of *additional knowledge* that does not rely on any explicit additional assumption on the leakage model.
While such coefficient sets may, in principle, reveal the correct key through non-device-specific intuition, no fully automatic and generic way of exploiting them was proposed. In particular, Whitnall *et al.* showed that they could not be transformed into a sound generic distinguisher able to systematically identify the correct hypothesis. As a consequence, in order to exploit this additional information carried by the coefficients, the authors turned to a so-called *generic-emulating* approach based on stepwise linear regression, which incrementally selects predictors to build a reduced model capable of attacking injective targets, relying on a non-device-specific intuition that the regression coefficients corresponding to the correct key yield a significantly simpler and more structured model than those obtained under incorrect hypotheses. However, this strategy departs from true genericity, as it relies on heuristic model-selection procedures and implicitly introduces additional assumptions, even though these are non-device-specific. Moreover, stepwise regression is known to be unstable, prone to overfitting, and highly sensitive to the order of predictors, which results in unreliable final models [27]. Therefore, despite these efforts, the problem of designing generic non profiled attacks remains open.

The goal of this paper is to take a substantial step toward the first fully generic-compatible non profiled side-channel attack. To this end, we revisit LRA from a generic-attack perspective in Section 3 and first focus on the choice of the predictor basis, as it directly determines the estimated regression coefficients on which LRA relies. We show that standard bases suffer from fundamental limitations that prevent LRA from fully exploiting its generic potential. To overcome these issues, we propose a novel basis projection that removes key obstacles such as multicollinearity and enables the use of all predictors. This new basis changes the structure of the estimated coefficients. However, we then show that the use of classical LRA distinguishers, such as Euclidean distance or R^2, remain ineffective. This motivates the design in Section 4 of new coefficient-based distinguishers specifically tailored to this setting. Building on this new basis and the proposed distinguishers, we redesign LRA into a framework that enables efficient (*i.e.* implentation-optimized LRA) and generic non profiled attacks.

3 Limitations of LRA-based metrics

This section aims at highlighting current LRA limitations and questioning the relevance of state-of-the-art LRA distinguishers. First we briefly recall in Section 3.1 a major limitation namely the multicollinearity problem encountered when performing generic LRA (*i.e.* LRA in which an attacker considers all bits interactions), then we address this limitation by proposing to perform LRA using *WH* basis, introduced in [14]. Finally, we show in Section 3.2 practical limitations of state-of-the-art LRA distinguishers.

3.1 Multicollinearity problem in LRA

One of the major limitation of LRA, notably when considering a full basis, is the multicollinearity problem. This phenomenon, which is intrinsic to linear regression approach, occurs when at least two predictors are highly correlated. Indeed, as the basis commonly used for LRA attacks is the monomial basis [34,14], bits' interactions are thus highly correlated when projected in this basis. This has an impact on coefficients estimation in linear regression, leading to poor coefficients estimation. Moreover, when using the monomial basis, multicollinearity makes the interpretation of the estimated coefficients unreliable: it becomes impossible to distinguish the real individual contribution of each bit or bits' interaction to the leakage, since correlated predictors can arbitrarily share their explanatory power. To deal with this problem, Long *et al.* [23] recommend thus, to both consider a balanced (*i.e.* an equal number of traces for all public input chunk values) or large set of Q traces and restrict bits' interactions degree to $m = 1$ *i.e.* bits interactions of higher-order are not taken into account. Indeed, by considering $m = 1$, the multicollinearity problem is avoided as the columns of the regressors matrix $M_{\mathcal{B}} = \left(\mathcal{B}\left(\Gamma_1(y_1^k)\right), \ldots, \mathcal{B}\left(\Gamma_1(y_Q^k)\right)\right) \in \mathbb{R}^{Q \times n}$ are no longer correlated. Consequently, by only considering individual bits, LRA-based attacks are restricted to only estimate Hamming Weight and independent bit leakage

models. In other words, the presence of multicollinearity significantly undermines the genericity and flexibility of LRA-based attacks, which are meant to constitute LRA's main advantage over classical attacks such as CPA and MIA. Nevertheless, we prove in the following that such restrictions are not necessary and LRA-based attacks can fully exploit their genericity potential (*i.e.* exploiting the information provided by higher-order bits interactions without *a priori* assumption on the leakage model) when considering *WH* basis [14].
Indeed, an intuitive solution to address multicollinearity problem is the orthogonalization of the regressors matrix. This procedure allows then to decorrelate, from linear algebra point of view, predictors. Even though it represents a promising beginning, decorrelation in terms of linear algebra does not directly imply a statistical decorrelation of predictors pairwise (each predictor is treated as a random variable). Thus, to solve multicollinearity problem, the chosen basis into which bits' interactions are projected must meet the latter condition. By taking a closer look at *WH* basis properties, one can thus show and prove in the following theorem that using *WH* basis address multicollinearity problem. Hence, this result enables the use of LRA with a full basis while preserving the accuracy and stability of the estimated coefficients.

Theorem 1. *Given a set of Q traces and its associated set of Q public data, when performing a LRA under Assumption 1, the use of WH basis prevents multicollinearity problem as $Q \to \infty$.*

Proof. Let define $M_{WH} \in \mathbb{R}^{Q \times 2^n}$ such that $M_{WH} = \left(\mathcal{B}\left(\Gamma(y_1^k)\right), \ldots, \mathcal{B}\left(\Gamma(y_Q^k)\right)\right)$ for a given key hypothesis $k \in \mathcal{K}$, where $\mathcal{B} = WH$, this proof consists in showing that all columns of the regressors matrix M_{WH} become uncorrelated as $Q \to \infty$. Without loss of generality, let consider that M_{WH} is cyclically ordered *i.e.* its rows are permuted such that M_{WH} consists in the concatenation of $A = \lfloor \frac{Q}{2^n} \rfloor$ submatrices $m_{WH} \in \mathbb{R}^{2^n \times 2^n}$ defined as $m_{WH} = \left(\mathcal{B}\left(\Gamma(0)\right), \ldots, \mathcal{B}\left(\Gamma(2^n - 1)\right)\right)$ and an additional matrix $m'_{WH} \in \mathbb{R}^{R \times 2^n}$ where, as a reminder, $R = Q \bmod 2^n$ (see Assumption 1). By construction, as the projection onto *WH* basis consists in a projection onto the Fourier basis [14], M_{WH} can consequently be rewritten as a concatenation of A weighted Walsh-Hadamard matrices $H_{2^n} \in \mathcal{M}_{2^n}(\pm 1)$ and $H' \in \mathcal{M}_{R,2^n}(\pm 1)$.

$$M_{WH} = \frac{1}{2^{\frac{n}{2}}} \left.\begin{pmatrix} H_{2^n} \\ \vdots \\ H_{2^n} \\ H' \end{pmatrix}\right\} A \text{ times} \tag{5}$$

Let $\rho\left((M_{WH})_{:,i}, (M_{WH})_{:,j}\right)$ be the correlation between the ith and j^{th} $(i \neq j)$ columns of M_{WH} (*i.e.* two distinct predictors) such that

$$\rho\left((M_{WH})_{:,i}, (M_{WH})_{:,j}\right) = \frac{\mathbb{E}[(M_{WH})_{:,i} \times (M_{WH})_{:,j}] - \mathbb{E}[(M_{WH})_{:,i}]\,\mathbb{E}[(M_{WH})_{:,j}]}{\sigma_{(M_{WH})_{:,i}}\,\sigma_{(M_{WH})_{:,j}}}.$$

By linearity of expectation, $\mathbb{E}[(M_{WH})_{:,i} \times (M_{WH})_{:,j}] = \frac{1}{2^{\frac{n}{2}}}\mathbb{E}[H'_{:,i} \times H'_{:,j}]$ as H_{2^n} is orthogonal. Due to the dimensions and structure of H', one can deduce that:

$$-\frac{2^n}{Q} \leqslant \mathbb{E}[H'_{:,i} \times H'_{:,j}] \leqslant \frac{2^n}{Q}$$
$$\lim_{Q\to\infty} -\frac{2^n}{Q} = 0 \leqslant \lim_{Q\to\infty} \mathbb{E}[H'_{:,i} \times H'_{:,j}] \leqslant \lim_{Q\to\infty} \frac{2^n}{Q} = 0.$$

Consequently, by the squeeze theorem, $\lim_{Q\to\infty} \mathbb{E}[(M_{WH})_{:,i} \times (M_{WH})_{:,j}] = 0$. Moreover, considering the fact that a property of Walsh-Hadamard matrices is that −1 and 1 coefficients are evenly distributed in all rows and colums except for the first row and column which are only composed of 1 coefficients, one deduces by the linearity of expectation, that

$$\mathbb{E}[(M_{WH})_{:,i}] = \begin{cases} \frac{1}{2^{\frac{n}{2}}} & \text{if } i = 1 \\ \\ \frac{1}{2^{\frac{n}{2}}} \times \mathbb{E}[H'_{:,i}] & \text{otherwise.} \end{cases}$$

Similarly, by applying the same reasoning used for $\mathbb{E}[H'_{:,i} \times H'_{:,j}]$, one deduces that $\lim_{Q\to\infty} \mathbb{E}[H'_{:,i}] = 0$ using the squeeze theorem. Consequently, as $(M_{WH})_{:,1}$ is the intercept, this means that all regressors are uncorrelated *i.e.* $\forall\, i, j \in [2, 2^n]$ such that $i \neq j$, $\lim_{Q\to\infty} \rho\left((M_{WH})_{:,i}, (M_{WH})_{:,j}\right) = 0$. □

Along with works highlighting multicollinearity problem, numerous works identified and addressed another major limitation: LRA computation efficiency [22,29]. Indeed, inversion and products of $(Q \times \sum_{i=0}^{m} \binom{n}{i})$-sized prediction matrices, included in the LRA process (due to ordinary least square estimator) [22, Algorithm 3], are computationally expensive and can become difficult or even impossible to perform as the number of traces Q tends to ∞ and all bits' interactions are considered (full basis) *i.e.* $m = n$, $\sum_{i=0}^{m} \binom{n}{i} = 2^n$. To address this other limitation, several works propose implementation improvements [22,29] in which they consider a 'coalesced' LRA version.

Definition 5. ***'Coalesced' LRA [22,29].*** *A 'coalesced' LRA is an implementation-optimized LRA in which the Q leakage traces are averaged by class, according to their corresponding public input as a bijective relationship (e.g. SBox function) exists between sensitive values and public inputs.*

By considering this 'coalesced' LRA version, the computation complexity is reduced as matrices manipulated have a maximum size of $2^n \times 2^n$, when considering a full basis. Based on Theorem 1, one can thus deduce that when Q is a multiple of 2^n as in 'coalesced' LRA version for example, the use of *WH* basis overcomes the multicollinearity problem.

Corollary 1. *Let Q be a multiple of 2^n. Given a set of Q traces and its associated set of Q public data, when performing a LRA under Assumption 1 i.e. $R = 0$, the use of WH basis prevents multicollinearity problem.*

Proof. Following the notations of the proof of Theorem 1, M_{WH} can be seen as a concatenation of $\frac{Q}{2^n}$ Walsh-Hadamard matrices H_{2^n} up to a multiplicative constant factor $\frac{1}{2^{\frac{n}{2}}}$. Considering the orthogonality and structure properties of Walsh-Hadamard matrices, it is then straightforward to prove that for all regressors $i, j \in [2, 2^n]$, $\rho\left((M_{WH})_{:,i}, (M_{WH})_{:,j}\right) = 0$ (see proof of Theorem 1). □

Hence, since the *WH* basis is not affected by multicollinearity, its use allows LRA to fully exploit its generic potential: no *a priori* on the leakage model is required, while efficiency is preserved through the use of the 'coalesced' LRA version. In the next section, we review state-of-the-art LRA distinguishers and discuss their practical limitations.

3.2 Limitations of the state-of-the-art LRA distinguishers

3.2.1 Overview of state-of-the-art LRA distinguishers

Definition of R^2 based distinguishers. In addition to the Euclidean distance distinguisher (see Definition 3), Doget *et al.* introduced in [11] a second distinguisher that can be considered as an alternative when performing LRA: the R^2-based distinguisher. This distinguisher derives from statistics literature [16] and is based on the well-known coefficient of determination R^2, which measures the goodness-of-fit of a statistical model considering a set of observations.

Definition 6. $\mathbf{R^2}$-*based LRA distinguisher [11].* *Given a key hypothesis $k \in \mathcal{K}$, a set of Q leakage traces $\mathcal{T} = \{\mathbf{t}_1, \ldots, \mathbf{t}_Q\}$ of dimension D, its associated set of public inputs (resp. sensitive values) $\mathcal{X} = \{x_1, \ldots, x_Q\}$ (resp. $\mathcal{Y} = \{y_1^k, \ldots, y_Q^k\}$) and an hypothetical leakage model that can be decomposed, $\forall\, i \in [1, Q]$, as $\Gamma_m(y_i^k) \cdot \hat{\alpha}^k$ with $\hat{\alpha}^k \in \mathbb{R}^{\sum_{i=0}^{m} \binom{n}{i} \times D}$ and $\Gamma_m \in \mathbb{F}_2^{\sum_{i=0}^{m} \binom{n}{i}}$ the binary decomposition of sensitive values up to m^{th} bits interactions degree with $m \leqslant n$ projected into $\mathcal{B}$ basis, the R^2 distinguisher is defined for a given key hypothesis k as*

$$\mathcal{D}_{R^2}^k(\mathcal{T}, \mathcal{X}) = \max_{j \in [1,D]} r_j^k \tag{6}$$

$$= \max_{j \in [1,D]} 1 - \frac{\sum_{i=1}^{Q} \left(\mathbf{t}_{i,j} - \mathcal{B}(\Gamma_m(y_i^k)) \cdot \hat{\alpha}_j^k\right)^2}{Q \times \boldsymbol{\sigma}^2_{\boldsymbol{\mathcal{T}_j}}} \tag{7}$$

where $\boldsymbol{\sigma}^2_{\boldsymbol{\mathcal{T}_j}}$ corresponds to the empirical variance of leakage traces set $\mathcal{T}$ at sample j.

However, as pointed out in [22], the reliability of this distinguisher relies on the assumption that the R^2 score associated with the correct key hypothesis is the maximum over all key hypotheses and time samples. This, in turn, is based on the assumption that the distinguisher value obtained for an incorrect key hypothesis or an incorrect time sample tends toward a minimum value (*e.g.* zero) as the

number of traces increases [24]. In practice, these assumptions are not always satisfied, in particular in the presence of the so-called *ghost peaks* phenomenon [4]. This casts doubt on the effectiveness of R^2 as a reliable distinguisher. To mitigate this issue, Lomné *et al.* proposed in [22] a modified R^2-based distinguisher relying on a normalized score, which they validated experimentally.

Definition 7. ***Normalized* $\mathbf{R^2}$ *distinguisher [22].*** *Following notations defined in Definition 6, the normalized R^2 distinguisher for the sample $j \in [1, D]$ is a 2-tuple which can be computed as*

$$\mathcal{D}^j_{R^2_{norm}}(\mathcal{T}, \mathcal{P}) = \left(\underset{k \in \mathcal{K}}{\arg\max}\ r^k_j\ , \frac{\max\limits_k (r^k_j) - \boldsymbol{\mu_k}[r^k_j]}{\boldsymbol{\sigma^2_k}[r^k_j]} \right) \tag{8}$$

where $r^k_j = 1 - \dfrac{\sum_{i=1}^{Q} \left(\mathbf{t}_{i,j} - \hat{\alpha}^k_j \, \Gamma_m(y^k_i) \right)^2}{Q \times \boldsymbol{\sigma^2_{T_j}}}$ *corresponds to the R^2 score obtained for the jth sample $\boldsymbol{\mu_k}[r^k_j] \in \mathbb{R}$ and (resp. $\boldsymbol{\sigma^2_k}[r^k_j] \in \mathbb{R}$) is the empirical mean over k (resp. empirical variance over k) of r^k_j.*
Hence, for the jth sample, the first element of the 2-tuple corresponds to the key hypothesis k that maximizes the R^2 score while the second element corresponds to its normalized R^2 score.
Based on this distinguisher, the distinguishing rule applied is then

$$\hat{k} = \texttt{arg1max2}\left(\mathcal{D}^j_{R^2_{norm}} \right) \tag{9}$$

where $\hat{k}$ corresponds to the key candidate and `arg1max2` *is a function that returns the key candidate (i.e. the first element) corresponding to the maximum normalized R^2 score (i.e. the second element of the 2-tuple).*

The normalization of the R^2 scores makes the proposed distinguisher experimentally more robust than the classical R^2-based distinguisher, as demonstrated in [22].

State-of-the-art LRA distinguishers and their interpretation. LRA targets the first statistical moment of the leakage, *i.e.* the conditional mean, which corresponds in the standard side-channel model to the deterministic part of the leakage. Consequently, state-of-the-art LRA distinguishers such as the R^2-based, normalized R^2-based and Euclidean-distance-based distinguishers operate by comparing, explicitly or implicitly, the estimated deterministic part $\hat{\Psi}^k$ to the true deterministic leakage Ψ, and select the key hypothesis that minimizes their discrepancy (as formalized in Proposition 3).

Proposition 3. *The state-of-the-art LRA distinguishers asymptotically consist, when the number of traces $Q \to \infty$, in minimizing the discrepancy between the true deterministic part Ψ and the hypothetical deterministic part $\hat{\Psi}^k$ associated with a key hypothesis $k \in \mathcal{K}$.*

Proof. This result is obtained by expanding Equation 6. Let $\mathcal{T} = \{\mathbf{t}_1, \ldots, \mathbf{t}_Q\}$ be a set of Q leakage traces of dimension D, with associated public inputs $\mathcal{X}$ and sensitive values $\mathcal{Y}^k = \{y_1^k, \ldots, y_Q^k\}$ for a key hypothesis $k \in \mathcal{K}$. Let $\hat{\Psi}^k \in \mathbb{R}^D$ denote the deterministic part estimated by regression. For a given sample j, discarding k-independent constants and denoting $\stackrel{c}{=}$ equality up to an additive constant, we have

$$\begin{aligned}\mathcal{D}^k_{R^2}(\mathcal{T}, \mathcal{X}) &\stackrel{c}{=} \min_{j\in[1,D]} \sum_{i=1}^{Q} (\mathbf{t}_{i,j} - \hat{\Psi}^k_j)^2 \quad \left(= \mathcal{D}^k_{ED}(\mathcal{T}, \mathcal{X})\right) \\ &\stackrel{c}{=} \min_{j\in[1,D]} \sum_{i=1}^{Q} (\Psi_j + \mathbf{n}_{i,j} - \hat{\Psi}^k_j)^2,\end{aligned}$$

where for fixed j the noise variables $(\mathbf{n}_{i,j})_{i=1}^{Q}$ are assumed independent and identically distributed with $\mathbb{E}[\mathbf{n}_{i,j}] = 0$ and $\mathbb{V}(\mathbf{n}_{i,j}) = \boldsymbol{\sigma}_j^2 < \infty$. Let $\Delta_j^k := \Psi_j - \hat{\Psi}^k_j$. Then

$$\sum_{i=1}^{Q} (\Psi_j + \mathbf{n}_{i,j} - \hat{\Psi}^k_j)^2 = \sum_{i=1}^{Q} (\Delta_j^k + \mathbf{n}_{i,j})^2 = Q(\Delta_j^k)^2 + 2\Delta_j^k \sum_{i=1}^{Q} \mathbf{n}_{i,j} + \sum_{i=1}^{Q} \mathbf{n}_{i,j}^2.$$

Dividing by $Q > 0$ (which does not affect the min over j) and discarding k-independent terms yields

$$\frac{1}{Q} \sum_{i=1}^{Q} (\Psi_j + \mathbf{n}_{i,j} - \hat{\Psi}^k_j)^2 \stackrel{c}{=} (\Delta_j^k)^2 + 2\Delta_j^k \left(\frac{1}{Q} \sum_{i=1}^{Q} \mathbf{n}_{i,j} \right).$$

By the law of large numbers, for each fixed j, $\frac{1}{Q}\sum_{i=1}^{Q} \mathbf{n}_{i,j} \to 0$ as $Q \to \infty$, so the cross term vanishes asymptotically and the criterion reduces to minimizing $(\Psi_j - \hat{\Psi}^k_j)^2$. Thus, the R^2 and Euclidean-distance based distinguishers minimize the discrepancy between Ψ and $\hat{\Psi}^k$, and the optimum is reached when $\hat{\Psi}^k = \Psi$. The result for the normalized R^2 distinguisher $\mathcal{D}^k_{R^2_{\text{norm}}}$ is immediate, as it coincides with $\mathcal{D}^k_{R^2}$ up to additive and multiplicative k-dependent constants $\boldsymbol{\mu_k}[r_j^k]$ and $\boldsymbol{\sigma_k^2}[r_j^k]$ respectively (see Equation 8), which do not affect the min over j. □

Proposition 3 provides a clear interpretation of state-of-the-art LRA distinguishers: although linear regression fundamentally produces a set of coefficients $\hat{\boldsymbol{\alpha}}^k$, existing distinguishers do not exploit these coefficients directly. Indeed, the core output of LRA is the coefficient matrix $\hat{\boldsymbol{\alpha}}^k$, while the reconstructed deterministic leakage $\hat{\Psi}^k$ is obtained only by projecting these coefficients onto the chosen basis. Consequently, state-of-the-art distinguishers rely on the coefficients only indirectly, via their contribution to the estimated mean leakage.

In a generic setting with a full basis, all estimated coefficient sets $\hat{\boldsymbol{\alpha}}^k$ are computed via a least-squares minimization that ensures the reconstructed deterministic leakage $\hat{\Psi}^k$ is as close as possible to the leakage traces [16,13,17]. When the

number of predictors matches the number of distinct sensitive values *i.e.* 2^n, this minimization can be achieved for any key hypothesis $k \in \mathcal{K}$. As a result, classical LRA distinguishers, which rely on the discrepancy between $\hat{\Psi}^k$ and the observed leakage, tend to return nearly identical scores across all key hypotheses. In practice, due to measurement noise and imperfect modeling, these scores may not be exactly equal, particularly when performing a classical LRA (*i.e.* not 'coalesced' LRA), but they are generally too close to allow reliable discrimination of the correct key. This renders them ineffective for discriminating the correct key in a fully generic context. We discuss and formalize these practical limitations in detail in the following section.

3.2.2 Limitations of state-of-the-art LRA distinguishers

So far, we see how to prevent the multicollinearity problem using *WH* basis and thus, resulting in the exploitation of a generic LRA (*i.e.* considering a full basis) without any risk of losing precision in the coefficient estimates, which could lead to falsely high R^2 scores for wrong key hypotheses [16]. Unfortunately, in the following, we show that these improvements with regard to distinguishers proposed in the literature do not allow to conduct successful attacks, especially when considering 'coalesced' LRA implementations. Indeed, using *WH* basis considering a full basis deprives the Euclidean distance distinguisher of its ability to discriminate the correct key hypothesis from the others.

Theorem 2. *The Euclidean distance based distinguisher returns the same score for all key hypothesis $k \in \mathcal{K}$ when considering 'coalesced' LRA implementations and a full orthogonal basis.*

Proof. Considering $\mathbf{C}_j = (\mathbf{c}_{j,0}, \ldots, \mathbf{c}_{j,2^n-1})$ as a vector of class-averaged traces for a given time sample $j \in [1, D]$ *i.e.* $\mathbf{c}_{j,i} = \frac{1}{|\{a:p_a=i\}|} \sum_{a,p_a=i} \mathbf{t}_{a,j}$ with $\mathbf{t}_a$ a D-dimensional leakage trace and p_a a public input such as a plaintext, the Euclidean distance score ED can be computed as follows:

$$\begin{aligned} ED_j^k &= \|\mathbf{C}_j - \hat{\Psi}^k(Y^k)\|_2^2 \\ &= \|\mathbf{C}_j - B^k\, \hat{\alpha}_j^k\|_2^2 \end{aligned}$$

where $\hat{\Psi}^k$ is an hypothetical leakage model, $B^k = \mathcal{B}(\Gamma(Y^k))$ correspond to the projection onto the orthogonal basis $\mathcal{B}$ of the binary decomposition Γ of sensitive values Y^k given a key hypothesis k and $\hat{\alpha}_j^k$ are estimated coefficients. By performing the ordinary least square and as B^k is orthogonal, $\hat{\alpha}_j^k = \left((B^k)^\top B^k\right)^{-1} (B^k)^\top \mathbf{C}_j = (B^k)^\top \mathbf{C}_j$. Moreover, B^k orthogonality also implies that $ED_j^k = 0$ for all time sample $j \in [1, D]$ and key hypothesis $k \in \mathcal{K}$. □

As seen in Section 3.2.1, since R^2 distinguisher corresponds to Euclidean distance distinguisher up to constants, one can deduce the following result.

Corollary 2. *Given an orthogonal basis, 'coalesced' LRA implementations and a full basis, R^2 distinguisher returns the same score for all key hypothesis $k \in \mathcal{K}$.*

Proof. The result is straightforward since the R^2 score at the jth is defined according to Equation 6 as

$$(R^2)_j^k = 1 - \frac{ED_j^k}{2^n \times \boldsymbol{\sigma^2_{\mathcal{T}_j}}}.$$

where $\boldsymbol{\sigma^2_{\mathcal{T}_j}}$ corresponds to the empirical variance of leakage traces set $\mathcal{T}$ at the jth sample and ED_j^k is the Euclidean distance obtained at the jth sample considering $k \in \mathcal{K}$ as key hypothesis. □

By extension, the same phenomenon occurs also for the normalized R^2 distinguisher.

Corollary 3. *The normalized R^2 distinguisher failed when considering 'coalesced' LRA implementations and a full orthogonal basis.*

Proof. Without loss of generality, let consider a fixed sample j and its associated R^2 score r_j^k for a given key hypothesis k. The normalized R^2 is a 2-tuple computed as

$$\mathcal{D}^j_{R^2_{norm}} = \left(\arg\max_k r_j^k \, , \, \frac{\max\limits_k \left(r_j^k\right) - \boldsymbol{\mu_k}[r_j^k]}{\boldsymbol{\sigma^2_k}[r_j^k]} \right)$$

where $\boldsymbol{\mu_k}[r_j^k]$ (resp. $\boldsymbol{\sigma^2_k}[(R^2)_j^k]$) corresponds to the empirical mean (resp. empirical variance) over k of r_j^k. Considering the 'coalesced' LRA implementations and a full orthogonal basis it appears that $\forall\, k$, $r_j^k = 1$ as mentioned in Corollary 2. Consequently, a division by zero occurs in the computation of $\mathcal{D}^j_{R^2_{norm}}$ second element and thus making the use of the normalized R^2 distinguisher unsuitable. □

Hence, through this section, we show that the state-of-the-art works on LRA distinguishers are deprived of their ability to discriminate the correct key hypothesis from the others using a full orthogonal basis decomposition. As a result, these distinguishers are incompatible with efforts aimed at improving the LRA process to make it generic. In the following section, we propose sound distinguishers that overcomes all practical limitations previously highlighted and thus allowing one to perform a generic LRA and a generic 'coalesced' LRA.

4 Proposition of new sound distinguishers for LRA

This section presents new LRA distinguishers that are sound (*i.e.* , can retrieve the correct key even with a full orthogonal basis) and allow one to perform both generic and 'coalesced' LRA. The first distinguisher is presented in Section 4.1, then an extension is proposed in Section 4.2 to target particular cryptographic functions. We emphasize that the proposed distinguishers are designed to be generic-compatible rather than optimal in a classical sense. Their purpose is to guarantee key discrimination without relying on model selection or device-specific tuning. The search for optimal LRA metrics is left for future work.

4.1 Proposition of a distinguisher for LRA-based attacks

To propose a suitable LRA distinguisher, a certain number of requirements need to be met. Indeed, since LRA exploits the first statistical moment, a suitable LRA distinguisher must in some way include a comparison of deterministic parts as shown in Proposition 3. Since the comparison of deterministic parts fails at finding the correct key as shown in previous sections, a solution is to focus on estimated coefficients of the deterministic parts and design a distinguisher that discriminates keys according to estimated coefficients values as also noticed in [37]. Since the comparison of deterministic parts fails at identifying the correct key, as shown in previous sections, a natural alternative is to focus directly on the estimated coefficients $\hat{\alpha}^k$ of the deterministic part $\hat{\Psi}^k$. This direction was already pointed out by Whitnall *et al.* in [37], who observed that the regression coefficients carry additional information beyond classical goodness-of-fit scores. Designing distinguishers that discriminate key hypotheses based on these coefficients therefore appears as a promising strategy. Nevertheless, turning this idea into a sound and fully automated distinguisher remained challenging in previous works. Indeed, LRA is performed using the ordinary least squares estimator, the estimation of these coefficients $\hat{\alpha}^k$ necessarily involves a choice of basis such as monomial basis or *WH* basis. Hence, a major difficulty in previous works stems from the use of the monomial basis for coefficient estimation. Indeed, as explained in Section 3.1, this basis does not provide a faithful representation of the true contributions of individual bits and bit interactions due to multicollinearity problem. This prevents the estimated coefficients $\hat{\alpha}^k$ from accurately reflecting the underlying leakage structure. Consequently, the potential of coefficient-based discrimination could not be fully exploited.
In contrast, the *WH* basis offers an orthogonal representation that captures the real contributions of each bit and bits' interaction. Estimating the coefficients $\hat{\alpha}^k$ in this basis therefore provides a more meaningful and stable characterization of the leakage, enabling effective discrimination between key hypotheses. The foundation of our following approach and proposed distinguishers is based on this observation. Consequently, this section will be divided into two subsections. We first establish key properties of the *WH* basis when estimating coefficients in Sections 4.1.1 and 4.1.2. Based on these properties, we then derive and propose a first LRA distinguisher, whose presentation is given in Section 4.1.3. As a reminder, in this paper we follow the commonly adopted definition of soundness in the SCA community [15].

Definition 8. ***Soundness [15, Proposition 1].*** *A distinguisher $\mathcal{D}$ is said to be sound if the success of the attack tends to* 100% *as the number of traces Q increases:*

$$SR(\mathcal{D}) \to 1, \text{ as } Q \to \infty,$$

where SR denotes the success rate.

Preliminary lemmas. Before proceeding to the next section, we introduce four lemmas that will be used in the proofs of all subsequent results of this paper.

Lemma 1. *The matrix multiplication of two projections onto WH basis* $B^{k_1} \in \mathbb{R}^{Q\times 2^n}$ *and* $B^{k_2} \in \mathbb{R}^{Q\times 2^n}$ *considering two distinct key hypotheses* $k_1 \in \mathbb{F}_2^n$ *and* $k_2 \in \mathbb{F}_2^n$ *results in a matrix such that its entry in the* i^{th} *row and* j^{th} *column equals:*

$$\left((B^{k_1})^\top B^{k_2}\right)_{u,v} = \frac{1}{2^n}\sum_{q=1}^{Q}(-1)^{u\cdot y_{p_q}^{k_1} \oplus v\cdot y_{p_q}^{k_2}},$$

where " $\cdot$ " *corresponds to the usual inner product in* $\mathbb{F}_2^n$ *and* $y_{p_q}^k = f(p_q, k)$ *for a given public input (e.g. plaintext)* $p_q \in \mathbb{F}_2^n$*, a key* $k \in \mathbb{F}_2^n$ *and a bijective cryptographic vectorial Boolean function* $f : \mathbb{F}_2^n \times \mathbb{F}_2^n \mapsto \mathbb{F}_2^n$*.*

Proof. The result is straightforward as by definition $(B^k)_{q,u} = \frac{1}{2^{\frac{n}{2}}}(H_{2^n})_{y_{p_q}^k,u} = \frac{1}{2^{\frac{n}{2}}}(-1)^{u\cdot y_{p_q}^k}$ where H_{2^n} is a Walsh-Hadamard matrix. □

Lemma 2. *Given a projection onto WH basis* $B^k \in \mathbb{R}^{Q\times 2^n}$ *such that* B^k *can be rewritten as the concatenation of* $\left\lfloor \frac{Q}{2^n} \right\rfloor$ *scaled Walsh-Hadamard matrices* $\frac{1}{2^{\frac{n}{2}}}H_{2^n}$ *and a scaled matrix* $\frac{1}{2^{\frac{n}{2}}}H'$ *where* $H_{2^n} \in \mathcal{M}_{2^n}(\pm 1)$ *and* $H' \in \mathcal{M}_{R,2^n}(\pm 1)$*,*

$$(B^k)^\top B^k = \left\lfloor \tfrac{Q}{2^n} \right\rfloor \times I_{2^n} + M_R,$$

where $M_R \in \mathbb{R}^{2^n\times 2^n}$*. In the case where* Q *is a multiple of* 2^n*, the equality holds i.e.* $(B^k)^\top B^k = \frac{Q}{2^n} \times I_{2^n}$*.*

Proof. Without loss of generality, let consider that B^k is cyclically ordered *i.e.* its rows are permuted such that B^k consists in the concatenation of $A = \left\lfloor \frac{Q}{2^n} \right\rfloor$ scaled Walsh-Hadamard matrices $\frac{1}{2^{\frac{n}{2}}}H_{2^n}$ and a scaled matrix $\frac{1}{2^{\frac{n}{2}}}H'$ (see Equation 5). As a reminder, following Assumption 1, Q can be decomposed as $Q = A\times 2^n + R$ with $A = \left\lfloor \frac{Q}{2^n} \right\rfloor$. Given $B^k \in \mathbb{R}^{Q\times 2^n}$, by applying Lemma 1, one deduces $\forall\, u, v \in [1, 2^n]$,

$$\begin{aligned}
\left((B^k)^\top B^k\right)_{u,v} &= \frac{1}{2^n}\sum_{q=1}^{Q}(-1)^{(u\oplus v)\cdot y_{p_q}^k} \\
&= A \times \frac{1}{2^n}\sum_{p_q\in\mathbb{F}_2^n}(-1)^{(u\oplus v)\cdot y_{p_q}^k} + \frac{1}{2^n}\sum_{q=A\times 2^n+1}^{Q}(-1)^{(u\oplus v)\cdot y_{p_q}^k} \\
&= A \times (I_{2^n})_{u,v} + \frac{1}{2^n}\sum_{q=A\times 2^n+1}^{Q}(-1)^{(u\oplus v)\cdot y_{p_q}^k} \\
&= A \times (I_{2^n})_{u,v} + (M_R)_{u,v}
\end{aligned}$$

where $M_R \in \mathbb{R}^{2^n\times 2^n}$ such that $(M_R)_{u,v} = \frac{R}{2^n}$ if $u = v$, $(M_R)_{u,v} \leqslant \left|\frac{R}{2^n}\right|$ otherwise. Regarding the case where Q is a multiple of 2^n, the equality is reached *i.e.* $(B^k)^\top B^k = \left\lfloor \frac{Q}{2^n} \right\rfloor \times I_{2^n}$ as $R = 0$. □

Lemma 3. *Following Lemma 2, given a projection onto WH basis* $B^k \in \mathbb{R}^{Q\times 2^n}$ *and* $(B^k)^\top \in \mathbb{R}^{2^n\times Q}$ *its transpose,*

$$\left((B^k)^\top B^k\right)^{-1} = \frac{1}{\left\lfloor \frac{Q}{2^n}\right\rfloor} I_{2^n} + O \text{ as } R \to 0$$

where $O \in \mathbb{R}^{2^n\times 2^n}$ *such that its* ∞*- matrix norm* $||O||_\infty \leqslant \frac{1}{\left\lfloor \frac{Q}{2^n}\right\rfloor + R}$*. In case where* Q *is a multiple of* 2^n*, the equality holds i.e.* $\left((B^k)^\top B^k\right)^{-1} = \frac{2^n}{Q} I_{2^n}$*.*

Proof. For the sake of proof readability, it should be noted that, following Assumption 1, Q can be decomposed as $Q = A \times 2^n + R$ with $A = \left\lfloor \frac{Q}{2^n}\right\rfloor$. Using the result from Lemma 2 $(B^k)^\top B^k = A \times I_{2^n} + M_R$, where $M_R \in \mathbb{R}^{2^n\times 2^n}$, an estimation of the inverse matrix can be provided. Indeed, since $Q \geqslant 2^n$, $\frac{1}{A}\left((B^k)^\top B^k\right) = I_{2^n} + M'_R$ where $M'_R = \frac{1}{A} \times M_R$. Then, denoting $||\cdot||_\infty$ as the ∞- matrix norm, $||M'_R||_\infty \leqslant \frac{R}{A} < 1$ as $R \to 0$. It is now possible to use the Neumann series identity to estimate the inverse of $(I_{2^n} + M'_R)$:

$$\left(I_{2^n} + M'_R\right)^{-1} = \sum_{i=0}^{\infty} (-1)^i (M'_R)^i.$$

Hence $(I_{2^n} + M'_R)^{-1} = I_{2^n} + O'$ where $O' \in \mathbb{R}^{2^n\times 2^n}$ such that $||O'||_\infty \leqslant \sum_{i=1}^{\infty} \left(\frac{(-1)\times R}{A}\right)^i$, which converges as the sum of a geometric series, since $\left|\frac{-R}{A}\right| < 1$ because for a given Q, R is fixed such that $R \to 0$. Consequently, $||O'||_\infty \leqslant \frac{A}{A+R} \leqslant 1$ as $R \to 0$. All in all, we obtain the following equalities:

$$\left(\frac{1}{A}((B^k)^\top B^k)\right)^{-1} = \left(I_{2^n} + M'_R\right)^{-1} = I_{2^n} + O',$$

which lead to

$$\left((B^k)^\top B^k\right)^{-1} = \frac{1}{A} I_{2^n} + O,$$

where $O = \frac{1}{A} O'$ such that $||O||_\infty \leqslant \frac{1}{A+R}$ as $R \to 0$. In case where $R = 0$, Lemma 2 states that $(B^k)^\top B^k = \frac{Q}{2^n} I_{2^n}$. Thus, applying the inverse operator leads to $\left((B^k)^\top B^k\right)^{-1} = \frac{2^n}{Q} I_{2^n}$. □

Remark 1. As $Q \to \infty$, we approximate $\left((B^k)^\top B^k\right)^{-1} = \frac{1}{\left\lfloor \frac{Q}{2^n}\right\rfloor} I_{2^n} + O \to \frac{1}{\left\lfloor \frac{Q}{2^n}\right\rfloor} I_{2^n}$ since $||O||_\infty \leqslant \frac{1}{\left\lfloor \frac{Q}{2^n}\right\rfloor + R} \to 0$.

Lemma 4. *Sum of independent centered noise Let* $(\mathbf{N}_q)_{q\geqslant 1}$ *be independent random variables with* $\mathbb{E}[\mathbf{N}_q] = 0$ *and* $\mathbb{V}(\mathbf{N}_q) \leqslant \boldsymbol{\sigma^2} < \infty$*. Then*

$$\sum_{q=1}^{Q} \mathbf{N}_q = O_p(\sqrt{Q}).$$

Proof. Let $\mathbf{S}_Q = \sum_{q=1}^{Q} \mathbf{N}_q$. By independence, $\mathbb{E}[\mathbf{S}_Q] = 0$ and $\mathbb{V}[\mathbf{S}_Q] = \sum_{q=1}^{Q} \mathbb{V}[\mathbf{N}_q] \leqslant Q\boldsymbol{\sigma^2}$. For any $\varepsilon > 0$, set $M = \boldsymbol{\sigma}/\sqrt{\varepsilon}$. Chebyshev's inequality gives

$$\Pr(|\mathbf{S}_Q| > M\sqrt{Q}) \leqslant \frac{\mathbb{V}(\mathbf{S}_Q)}{M^2 Q} \leqslant \varepsilon,$$

which matches the definition of $O_p(\sqrt{Q})$. □

4.1.1 Indiscriminability of the bias with WH basis

We present in Proposition 4 a first property of the *WH* basis when estimating coefficients $\hat{\alpha}^k$ in LRA. In particular, it can be shown that the estimated bias converges toward the true unknown bias independently of the key hypothesis $k \in \mathbb{F}_2^n$ considered.

$$\hat{\alpha}^k_{0,j} \to \alpha_{0,j} \text{ as } Q \to \infty \text{ and } R \to 0.$$

Proposition 4. *Let consider Q traces with associated public inputs and assume Assumption 1. When using the WH basis and considering the maximum degree of bits interactions n, the coefficient corresponding to the constant Walsh-Hadamard component (i.e. the bias term) is correctly recovered: for all key hypotheses $k \in \mathcal{K}$ and all time samples $j \in [1, D]$,*

$$\hat{\alpha}^k_{0,j} \xrightarrow[Q\to\infty,\, R\to 0]{\textit{in probability}} \alpha_{0,j},$$

such that the estimation error satisfies

$$\hat{\alpha}^k_{0,j} - \alpha_{0,j} = O_p\left(\frac{2^{n/2}}{\sqrt{Q}}\right).$$

If Q is a multiple of 2^n ($R = 0$), there is no error due to class imbalance, the remaining estimation error is solely caused by the trace noise and is $O_p(2^{n/2}/\sqrt{Q})$.

Proof. By definition,

$$\hat{\alpha}^k_{0,j} = \left(\left((B^k)^\top B^k \right)^{-1} (B^k)^\top \right)_{0,:} \mathbf{T}_{:,j},$$

where $B^k \in \mathbb{R}^{Q \times 2^n}$ and $\mathbf{T}_{:,j} \in \mathbb{R}^Q$. Applying Lemma 3 and Remark 1 yields, as $R \to 0$,

$$\hat{\alpha}^k_{0,j} \to \frac{1}{\lfloor Q/2^n \rfloor} \left((B^k)^\top \right)_{0,:} \mathbf{T}_{:,j}.$$

Expand this term using $\mathbf{T}_{:,j} = B^{k^*} \alpha_{:,j} + \mathbf{N}_{:,j}$:

$$\begin{aligned}
\left((B^k)^\top \right)_{0,:} \mathbf{T}_{:,j} &= \sum_{v \in \mathbb{F}_2^n} \frac{1}{2^n} \sum_{q=1}^{Q} (-1)^{0 \cdot y^k_{p_q} \oplus v \cdot y^{k^*}_{p_q}} \alpha_{v,j} + \frac{1}{2^{\frac{n}{2}}} \sum_{q=1}^{Q} (-1)^{0 \cdot y^k_{p_q}} \mathbf{N}_{q,j} \\
&= \sum_{v \in \mathbb{F}_2^n} \frac{1}{2^n} \sum_{q=1}^{Q} (-1)^{v \cdot y^{k^*}_{p_q}} \alpha_{v,j} + \frac{1}{2^{n/2}} \sum_{q=1}^{Q} \mathbf{N}_{q,j},
\end{aligned}$$

where $(\mathbf{N}_{q,j})_{q=1}^{Q}$ are independent, centered, with finite variance (*i.e.* $\boldsymbol{\sigma}_j^2$).

Noise contribution. By Lemma 4, $\sum_{q=1}^{Q} \mathbf{N}_{q,j} = O_p(\sqrt{Q})$. Hence the noise term in $\hat{\alpha}_{0,j}^k$ satisfies

$$\frac{1}{\lfloor Q/2^n \rfloor} \cdot \frac{1}{2^{n/2}} \sum_{q=1}^{Q} \mathbf{N}_{q,j} = O_p\left(\frac{2^{n/2}}{\sqrt{Q}}\right) \to 0 \quad \text{as } Q \to \infty.$$

Deterministic contribution. Since f is bijective over $\mathbb{F}_2^n$ and each public input (*e.g.* plaintext) value occurs at least $\lfloor Q/2^n \rfloor$ times,

$$\sum_{q=1}^{Q} (-1)^{v \cdot y_{pq}^{k^*}} = \begin{cases} Q, & v = 0, \\ \epsilon_v, \ |\epsilon_v| \leqslant R, & v \neq 0. \end{cases}$$

Therefore, as $R \to 0$,

$$\sum_{v \in \mathbb{F}_2^n} \frac{1}{2^n} \sum_{q=1}^{Q} (-1)^{v \cdot y_{pq}^{k^*}} \alpha_{v,j} \to \frac{Q}{2^n} \alpha_{0,j}.$$

Combining both parts gives

$$\hat{\alpha}_{0,j}^k \to \frac{\frac{Q}{2^n}}{\lfloor Q/2^n \rfloor} \alpha_{0,j} \to \alpha_{0,j} \quad \text{as } R \to 0,$$

and the equality is exact when Q is a multiple of 2^n. □

Hence, Proposition 4 implies that the correct key $k^* \in \mathbb{F}_2^n$ cannot be discriminated on the basis of the estimated biases. This is consistent with the nature of the bias term: unlike the other coefficients, it is not associated with any bit or bit interaction of the sensitive variable, but rather captures global effects such as trace offset and amplitude modulations. Consequently, the bias component does not convey any cryptographic information related to the key and should not be involved in the decision process of an LRA distinguisher. This observation guides our contribution by restricting the design of coefficient-based distinguishers to the informative coefficients only, while discarding the bias that merely reflects implementation-dependent signal modulations.

4.1.2 Estimation of coefficients of uninformative samples with WH basis

Another important property of the *WH* basis is that it alleviates the dimensionality issue inherent to LRA. This issue stems from the univariate nature of LRA: a separate regression is performed for each time sample of the trace, so that each time sample yields its own vector of coefficients and predictions. Consequently, the outcome of an LRA attack depends not only on the key hypothesis

k but also on the considered time sample, which makes the comparison of results across samples particularly difficult. This difficulty has been repeatedly observed in practice, where classical assumptions stating that the distinguisher value for wrong hypotheses tends toward its minimum as the number of traces increases are often violated, notably due to phenomena such as ghost peaks [4,22]. To mitigate this problem, Lomné *et al.* introduced the normalized R^2 distinguisher [22] in order to enable a fair comparison between instantaneous attacks, but this normalization remains tied to the classical distinguishers.
In contrast, we show in the following proposition that performing LRA and 'coalesced' LRA with the *WH* basis on whole traces, without restricting the analysis to selected points of interest (PoIs), does not degrade the attack performance. Indeed, the coefficients estimated on uninformative samples converge toward 0 for all key hypotheses when using the *WH* basis. Since LRA models the leakage of each sample individually, this property allows the comparison of leakage models over all samples without any prior dimensionality reduction or POI selection, a behavior that does not hold with canonical or monomial bases [22, Section 3.2], where uninformative samples may introduce spurious contributions.

Proposition 5. *Let f be a cryptographic (n,n)-vectorial Boolean function (e.g. an S-box), and consider a set of Q traces with corresponding public inputs. Under Assumption 1, if the time sample $j \in [1, D]$ is not a point of interest, then for all key hypotheses $k \in \mathcal{K}$, the estimated coefficients in the WH basis satisfy*

$$\hat{\alpha}^k_{u,j} \xrightarrow[Q\to\infty,\, R\to 0]{\text{in probability}} 0 \quad \text{for all } u \neq 0,$$

with rate

$$\hat{\alpha}^k_{u,j} = O_p\left(\frac{2^{n/2}}{\sqrt{Q}}\right).$$

Proof. According to Lemma 3 and Remark 1, $\hat{\alpha}^k_{:,j} \to \frac{1}{\lfloor Q/2^n \rfloor}(B^k)^\top \mathbf{T}_{:,j}$ as $R \to 0$, where $B^k \in \mathbb{R}^{Q\times 2^n}$ and $\mathbf{T}_{:,j} \in \mathbb{R}^Q$. Consequently, for all $u \in \mathbb{F}_2^n \backslash \{0\}$,

$$\begin{aligned}\hat{\alpha}^k_{u,j} &\to \frac{1}{\lfloor Q/2^n \rfloor}\left((B^k)^\top\right)_{u,:}\mathbf{N}_{:,j} \\ &= \frac{1}{\lfloor Q/2^n \rfloor}\sum_{q=1}^{Q}(-1)^{u\cdot f(p_q\oplus k)}\frac{1}{2^{n/2}}\mathbf{N}_{q,j}.\end{aligned}$$

By reordering the sum as in [22,14], we obtain

$$\hat{\alpha}^k_{u,j} \to \frac{1}{\lfloor Q/2^n \rfloor}\sum_{t\in\mathbb{F}_2^n}\left(\sum_{q\,:\,p_q=t}\mathbf{N}_{q,j}\right)\frac{1}{2^{n/2}}(-1)^{u\cdot f(t\oplus k)}.$$

The sign $(-1)^{u\cdot f(t\oplus k)}$ has modulus one and therefore does not affect the $O_p(\cdot)$ bound.

For a fixed t, let $m_t := |\{q : p_q = t\}|$. The variables $(\mathbf{N}_{q,j})_{q:p_q=t}$ are independent, centered, with finite variance (*i.e.* $\boldsymbol{\sigma}_j^2$), hence by Lemma 4,

$$\sum_{q\,:\,p_q=t} \mathbf{N}_{q,j} = O_p(\sqrt{m_t}).$$

Since $m_t \geqslant \lfloor Q/2^n \rfloor$ (see Assumption 1) and $\sum_t m_t = Q$, we get

$$\sum_{t\in\mathbb{F}_2^n} \sqrt{m_t} \leqslant 2^n \sqrt{\max_t m_t} \leqslant 2^n \sqrt{Q}.$$

Therefore,

$$\hat{\alpha}_{u,j}^k = O_p\left(\frac{1}{\lfloor Q/2^n \rfloor} \cdot \frac{1}{2^{n/2}} \cdot 2^n \sqrt{Q}\right) = O_p\left(\frac{2^{n/2}}{\sqrt{Q}}\right) \xrightarrow[Q\to\infty]{} 0,$$

for fixed n. This concludes the proof. □

4.1.3 A sound distinguisher for non-linear operations: Maximum distinguisher

Based on the two aforementioned properties, one can now design a suitable distinguisher for LRA-based attacks. Unlike state-of-the-art approaches, the proposed distinguisher does not rely on the comparison of reconstructed deterministic parts $\hat{\Psi}^k$, but directly exploits the estimated coefficients $\hat{\alpha}^k$ obtained through regression. More precisely, the comparison is performed only on the maximal coefficient for a given key hypothesis, excluding the bias term, which is not discriminative (see Proposition 4). This choice is motivated by the fact that the largest coefficient captures the strongest bit or bit-interaction contribution for a given time sample. According to Proposition 5, coefficients associated with uninformative samples converge toward zero; hence the maximal coefficient is expected to correspond to a point of interest and to carry the most discriminating information.

The use of the Walsh-Hadamard basis provides a natural interpretation of this strategy. In this basis, the estimated coefficients represent orthogonal and independent contributions of bits and bit interactions, in contrast to the monomial basis where multicollinearity entangles these contributions. Parseval's identity [6] further ensures that the information contained in the deterministic leakage is preserved in the coefficient domain, so that analyzing $\hat{\alpha}^k$ is equivalent to analyzing $\hat{\Psi}^k$ for the correct key hypothesis *i.e.* $k = k^*$. Moreover, for an incorrect key hypothesis, the mismatch between the true and guessed sensitive variables induces a permutation whose representation in the Walsh domain generally mixes the coefficient structure, so that dominant components tend to be spread over several bits/bits' contributions rather than concentrated on a single one. Consequently, the coefficient with the largest magnitude identifies the Walsh component that is most aligned with the leakage for the correct key, *i.e.* the dominant contribution of the sensitive variable. This provides a sound and intuitive justification for designing a distinguisher based on the maximal estimated coefficient, as formally defined below.

Definition 9. ***Maximum distinguisher based attack.*** *Given a set of D-dimensional traces $\mathcal{T}$, its associated set of public inputs $\mathcal{X}$ and the matrix $D_M \in \mathbb{R}^{|\mathcal{K}|\times D}$ for which an entry $(D_M)_{k,j} = \|\hat{\alpha}^k_{1:,j}\|_\infty$ with $\hat{\alpha}^k_{1:,j} \in \mathbb{R}^{\sum_{i=1}^{m}\binom{n}{i}}$ the estimated coefficients up to m^{th} bits interactions ($m \leqslant n$) excluding the bias and considering a key hypothesis k and a time sample $j \in [1, D]$. The maximum distinguisher is defined as*

$$\mathcal{D}^k_{max}(\mathcal{T}, \mathcal{X}) = \|(D_M)_{k,:}\|_\infty \tag{10}$$

where $\|\cdot\|_\infty$ is the infinity norm. Based on this distinguisher, the key candidate $\hat{k}$ is selected such that

$$\hat{k} = \arg\max_{k\in\mathcal{K}} \mathcal{D}^k_{max}. \tag{11}$$

Let f be the targeted cryptographic function (*e.g.* an S-box) and define, for any $\Delta \in \mathbb{F}_2^n$, the function $F_f^\Delta : x \mapsto f\big(f^{-1}(x) \oplus \Delta\big)$. We also define $nl_{\min}(F_f^\Delta) = \min_{\Delta\in\mathbb{F}_2^n} nl(F_f^\Delta)$ as the minimal nonlinearity reached by F_f^Δ over all Δ.
The soundness analysis of the proposed distinguisher relies on the study of f and F_f^Δ as vectorial Boolean functions and on the impact of its nonlinearity of F_f^Δ on the Walsh coefficients of the deterministic leakage. The key intuition is that an incorrect key hypothesis induces the transform F_f^Δ, whose nonlinearity governs the evolution of the Walsh coefficients of the true deterministic leakage: when this transform is sufficiently nonlinear, a dominant Walsh component tends to be dispersed across several bits and bits' interactions, whereas low nonlinearity may preserve such a structure and make key hypotheses harder to separate.
At the level of coefficient estimation, this means that under a wrong key hypothesis the estimated vector $\hat{\alpha}^k$ results from a distorted version of the true coefficients, where the action of F_f^Δ may generate spurious large components unrelated to any real point of interest. Under a condition on the true coefficients, this framework enables us to relate the behavior of the maximal estimated coefficient for the correct key hypothesis to the perturbation affecting $\hat{\alpha}^k$ in the case of an incorrect key hypothesis, which is controlled by the nonlinearity of F_f^Δ. In the following theorem, we establish the soundness of $\mathcal{D}^k_{\max}$ for conventional LRA. The extension to 'coalesced' LRA implementations is provided in Proposition 6.
The soundness guarantees established in the following results rely on structural assumptions on the true Walsh-Hadamard coefficients of the leakage representation, rather than on any *a priori* leakage model. For the maximum distinguisher, the assumption requires the existence of a sufficiently dominant coefficient. Importantly, this assumption does not restrict the expressiveness of the leakage model. Moreover, unlike leakage model assumptions used in classical non profiled approaches, such properties can be verified *a posteriori* from the estimated coefficients. Our experiments suggest that such dominant-coefficient structures consistently appear on the evaluated real-world datasets (see Section 6.2).

Theorem 3 (Soundness of the maximum distinguisher). *Consider a set of Q traces and its associated public inputs satisfying Assumption 1. For each*

sample $j \in [1, D]$, let $\alpha_{u,j}$ denote the true Walsh-Hadamard coefficients of the deterministic part $\Psi(Y^{k^*})$, and define

$$|\alpha_{\max,j}| = \max_{u \in \mathbb{F}_2^n \setminus \{0\}} |\alpha_{u,j}|, \qquad |\alpha_{\max}| = \max_{1 \leqslant j \leqslant D} |\alpha_{\max,j}|.$$

If, for every sample $j \in [1, D]$,

$$|\alpha_{\max}| \geqslant \left(1 - \frac{nl_{\min}(F_f^{\Delta})}{2^{n-1}}\right) \sum_{v \in \mathbb{F}_2^n \setminus \{0\}} |\alpha_{v,j}|,$$

then the linear regression based attack using the maximum distinguisher $\mathcal{D}^k_{\max}$ with the full WH basis is sound as $Q \to \infty$ and $R \to 0$.

Proof. This proof consists in showing that using the maximum distinguisher when conducting a LRA, using a full *WH* basis, on D-dimensional traces is sound and return as key candidate the true key k^*. Given $\alpha \in \mathbb{R}^{2^n \times D}$ the matrix of true coefficients for all samples, let define $\left|\alpha_{max,j}\right| = \max_{u \in \mathbb{F}_2^n \setminus \{0\}} \left|\alpha_{u,j}\right|$ and $\left|\alpha_{max}\right| = \max_{1 \leqslant j \leqslant D} \left|\alpha_{max,j}\right|$ such that $\left|\alpha_{max}\right| > 0$ as by definition there is at least one point of interest in side-channel traces.
The first step of this proof is to demonstrate that for a given time sample $j \in [1, D]$, such that j is not a point of interest, $\alpha_{max} > \hat{\alpha}^k_{:,j}$. Since $Q \to \infty$, $R \to 0$ and $\left|\alpha_{max}\right| > 0$, this result is straightforward by applying Proposition 5.
Now, the second step consists in showing that α_{max} is reached if and only if a correct key hypothesis is assumed (*i.e.* $k = k^*$) for all j such that j are PoIs. By definition, since j is a PoI, $\hat{\alpha}^k_{:,j} = \left((B^k)^\top B^k\right)^{-1} (B^k)^\top \mathbf{T}_{:,j}$. According to Lemma 3 and Remark 1, $\hat{\alpha}^k_{:,j} \to \frac{1}{\left\lfloor \frac{Q}{2^n} \right\rfloor} \times (B^k)^\top \mathbf{T}_{:,j}$ where $B^k \in \mathbb{R}^{Q \times 2^n}$ and $\mathbf{T}_{:,j} \in \mathbb{R}^Q$, as $R \to 0$. Consequently, $\forall\, u \in \mathbb{F}_2^n \setminus \{0\}$

$$\begin{aligned}
\hat{\alpha}^k_{u,j} &\to \frac{1}{\left\lfloor \frac{Q}{2^n} \right\rfloor} \times \left((B^k)^\top\right)_{u,:} \mathbf{T}_{:,j} \\
&\to \frac{1}{\left\lfloor \frac{Q}{2^n} \right\rfloor} \times \left(\left((B^k)^\top\right)_{u,:} B^{k^*} \alpha_{:,j} + \left((B^k)^\top\right)_{u,:} \mathbf{N}_{:,j}\right).
\end{aligned}$$

By expanding the right term, one deduces

$$\hat{\alpha}^k_{u,j} \to \frac{1}{\left\lfloor \frac{Q}{2^n} \right\rfloor} \times \sum_{q=1}^{Q} \frac{1}{2^n} \sum_{v \in \mathbb{F}_2^n \setminus \{0\}} \alpha_{v,j} (-1)^{u \cdot f(p_q \oplus k) \oplus v \cdot f(p_q \oplus k^*)} + (-1)^{u \cdot f(p_q \oplus k)} \frac{1}{2^{\frac{n}{2}}} \mathbf{N}_{q,j}.$$

Reordering the sum leads to:

$$\begin{aligned}
\hat{\alpha}^k_{u,j} &\to \frac{1}{\left\lfloor \frac{Q}{2^n} \right\rfloor} \times \sum_{t \in \mathbb{F}_2^n} \left(\sum_{q | p_q = t} \frac{1}{2^{\frac{n}{2}}} \sum_{v \in \mathbb{F}_2^n \setminus \{0\}} \alpha_{v,j} (-1)^{v \cdot f(t \oplus k^*)} + \mathbf{N}_{q,j} \right) \times \frac{1}{2^{\frac{n}{2}}} (-1)^{u \cdot f(t \oplus k)} \\
&\to \frac{1}{2^n} \sum_{v \in \mathbb{F}_2^n \setminus \{0\}} \alpha_{v,j} \sum_{t \in \mathbb{F}_2^n} (-1)^{u \cdot f(t \oplus k) \oplus v \cdot f(t \oplus k^*)}. \qquad (12)
\end{aligned}$$

The last line follows from two arguments. First, by Assumption 1, the public inputs (*e.g.* plaintexts) are uniformly represented over $\mathbb{F}_2^n$, so that for $m_t := |\{q : p_q = t\}|$ we have

$$\frac{m_t}{Q} = \frac{1}{Q}\sum_{q=1}^{Q} \mathbb{1}_{p_q=t} \xrightarrow[Q\to\infty]{\text{in probability}} \frac{1}{2^n},$$

which implies $\frac{1}{\lfloor Q/2^n \rfloor}\sum_{q:p_q=t} 1 \to 1$ as $R \to 0$.
Second, by Lemma 4, $\sum_{q:p_q=t} \mathbf{N}_{q,j} = O_p(\sqrt{m_t})$. Hence

$$\frac{1}{Q}\sum_{q:p_q=t} \mathbf{n}_{q,j} = O_p\left(\frac{\sqrt{m_t}}{Q}\right) = O_p\left(\frac{1}{\sqrt{Q}}\right) \xrightarrow[Q\to\infty]{} 0.$$

Combining these two facts yields Equation 12.
Based on this expression of $\hat{\alpha}_{u,j}^k$, $\forall\, u \in \mathbb{F}_2^n\backslash\{0\}$, one can now deduce which value will be returned by the maximum based distinguisher, assuming a given key hypothesis k. Let define Δ_k as the distance between a key hypothesis and the true unknown key k^*. The rest of the proof will be divided into two sub-proofs *i.e.* when a correct and an incorrect key hypothesis is considered.
Sub-proof 1: $k = k^*$ (*i.e.* $\Delta_k = 0$)
By definition, for a given time sample j that corresponds to a PoI, the maximum based distinguisher can be rewritten as:

$$\begin{aligned}
\mathcal{D}_{\max}^{k^*}(\mathcal{T},\mathcal{P}) &= \max_{1\leqslant j\leqslant D}\left[\max_{u\in\mathbb{F}_2^n\backslash\{0\}}\left|\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\backslash\{0\}}\alpha_{v,j}\sum_{t\in\mathbb{F}_2^n}(-1)^{u\cdot f(t\oplus k^*)\oplus v\cdot f(t\oplus k^*)}\right|\right]\\
&= \max_{1\leqslant j\leqslant D}\left[\max_{u\in\mathbb{F}_2^n\backslash\{0\}}\left|\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\backslash\{0\}}\alpha_{v,j}\sum_{t\in\mathbb{F}_2^n}(-1)^{(u\oplus v)\cdot f(t\oplus k^*)}\right|\right]\\
&= \max_{1\leqslant j\leqslant D}\left[\max_{u\in\mathbb{F}_2^n\backslash\{0\}}\left|\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\backslash\{0\}}\alpha_{v,j}\,2^n\,(I_{2^n})_{u,v}\right|\right]\\
&= \max_{1\leqslant j\leqslant D}\left[\max_{u\in\mathbb{F}_2^n\backslash\{0\}}\left|\sum_{v\in\mathbb{F}_2^n\backslash\{0\}}\alpha_{v,j}\,(I_{2^n})_{u,v}\right|\right] \quad = \max_{1\leqslant j\leqslant D}\left[\max_{u\in\mathbb{F}_2^n\backslash\{0\}}\left|\alpha_{u,j}\right|\right]\\
&= \max_{1\leqslant j\leqslant D}\left|\alpha_{max,j}\right| \leqslant \left|\alpha_{max}\right|.
\end{aligned}$$

By definition of $\left|\alpha_{max}\right|$, this value is reached for a certain j. Therefore, the equality holds for a given $j \in [1, D]$. Now, let's prove that for an incorrect key hypothesis, the value of $\left|\alpha_{max}\right|$ will never be reached.
Sub-proof 2: $k \neq k^*$ (*i.e.* $\Delta_k = k \oplus k^* \neq 0$)
For an incorrect key hypothesis and a given time sample j, the maximum distinguisher takes the form of:

$$\mathcal{D}_{\max}^{k}(\mathcal{T},\mathcal{P}) = \max_{1\leqslant j\leqslant D}\left[\max_{u\in\mathbb{F}_2^n\backslash\{0\}}\left|\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\backslash\{0\}}\alpha_{v,j}\sum_{t\in\mathbb{F}_2^n}(-1)^{u\cdot f(t\oplus k^*\oplus\Delta_k)\oplus v\cdot f(t\oplus k^*)}\right|\right].$$

Using the change of variable $z = f(t \oplus k^* \oplus \Delta_k)$, as f is bijective, allows one to rewrite the previous equation as:

$$\begin{aligned}\mathcal{D}^k_{\max}(\mathcal{T},\mathcal{P}) &= \max_{1\leqslant j\leqslant D}\left[\max_{u\in\mathbb{F}_2^n\setminus\{0\}}\left|\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\alpha_{v,j}\sum_{z\in\mathbb{F}_2^n}(-1)^{u\cdot z\oplus v\cdot F_f^{\Delta}(z)}\right|\right]\\ &\leqslant \max_{1\leqslant j\leqslant D}\left[\max_{u\in\mathbb{F}_2^n\setminus\{0\}}\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\left|\alpha_{v,j}\right|\left|\sum_{z\in\mathbb{F}_2^n}(-1)^{u\cdot z\oplus v\cdot F_f^{\Delta}(z)}\right|\right]\\ &< \max_{1\leqslant j\leqslant D}\left[\frac{\max\limits_{u,v\in\mathbb{F}_2^n\setminus\{0\}}\left|\sum\limits_{z\in\mathbb{F}_2^n}(-1)^{u\cdot z\oplus v\cdot F_f^{\Delta}(z)}\right|}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\left|\alpha_{v,j}\right|\right]\end{aligned}$$

The strict inequality comes from the fact that, by construction, excluding the first row and column *i.e.* $u, v \in \mathbb{F}_2^n\setminus\{0\}$, all entries of the linear table approximation of a bijective vectorial Boolean function (see Section 2.1) like F_f^{Δ} don't reach the maximum value. Indeed, for instance, the expected number of zeros entries in the table equals $\frac{4}{\sqrt{2\pi\times 2^n}}$ using Stirling's approximation [28]. Given that $nl(F_f^{\Delta}) = 2^{n-1} - \frac{1}{2}\max\limits_{u,v\in\mathbb{F}_2^n\setminus\{0\}}\left|\sum_{z\in\mathbb{F}_2^n}(-1)^{u\cdot z\oplus v\cdot F_f^{\Delta}(z)}\right|$ [5,7] (see Section 2.1), one can rewrite the latest equation, which leads to:

$$\mathcal{D}^k_{\max}(\mathcal{T},\mathcal{P}) < \max_{1\leqslant j\leqslant D}\left[\frac{2^n - 2\,nl(F_f^{\Delta})}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\left|\alpha_{v,j}\right|\right].$$

As $nl(F_f^{\Delta}) \geqslant nl_{min}(F_f^{\Delta})$, one deduces thus:

$$\mathcal{D}^k_{\max}(\mathcal{T},\mathcal{P}) < \max_{1\leqslant j\leqslant D}\left[\frac{2^n - 2\,nl_{min}(F_f^{\Delta})}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\left|\alpha_{v,j}\right|\right].$$

As it is assumed in Theorem 3 that $|\alpha_{max}| \geqslant \left(1 - \frac{nl_{min}(F_f^{\Delta})}{2^{n-1}}\right)\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\left|\alpha_{v,j}\right|$, one can finally find:

$$\mathcal{D}^k_{\max}(\mathcal{T},\mathcal{P}) < \left|\alpha_{max}\right|.$$

Hence, the value of $\left|\alpha_{max}\right|$ for all j is never reached when considering an incorrect key hypothesis. □

Based on these results, one can deduce the following corollary that states that the maximum distinguisher is also sound when considering a non-full *WH* basis *i.e.* when considering bits interactions up to the mth degree such that $m \leqslant n$.

Corollary 4. *Given a set of Q traces and its associated set of Q public inputs, f and F_f^Δ which are the bijective (n,n)-vectorial Boolean functions previously defined. Under Assumption 1 and for all samples $j \in [1, D]$, if the maximum coefficient of the true deterministic part $|\alpha_{max}| \geqslant \left(1 - \frac{nl_{min}(F_f^\Delta)}{2^{n-1}}\right) \sum_v |\alpha_{v,j}|$, $\forall\, v \in \mathbb{F}_2^n \backslash \{0\}$ such that $w_H(v) \leqslant m$ with $w_H(\cdot)$ and m being respectively the Hamming Weight function and the maximal degree of bits interactions considered to conduct the attack, then linear regression based attacks using the maximum distinguisher are sound as the number of traces $Q \to \infty$ and $R \to 0$, when considering WH basis up to the mth degree of bits interactions.*

Proof. This proof is a straightforward application of Theorem 3 proof, assuming that $\forall\, u, v \in \mathbb{F}_2^n \backslash \{0\}$, $w_H(u) \leqslant m$ and $w_H(v) \leqslant m$. □

We can now extend results obtained in Theorem 3 and in Corollary 4 and prove in the following proposition the soundness of our proposed distinguisher when 'coalesced' LRA versions are considered.

Proposition 6. *Following Theorem 3 and Corollary 4 configurations, linear regression based attacks using the maximum distinguisher recover the true key k^* asymptotically as $Q \to \infty$, and considering (full or non-full) WH basis and 'coalesced' linear regression attacks.*

Proof. The proof is a straightforward application of Theorem 3 and Corollary 4 proofs. Indeed, when considered 'coalesced' LRA, $\forall\, j \in [1, D]$, $\mathbf{T}_{:,j} \in \mathbb{R}^Q$ is averaged by class so as to end up only manipulating $\bar{\mathbf{T}}_{:,j} \in \mathbb{R}^{2^n}$ with $\bar{\mathbf{T}}_{i,j} = \frac{1}{|\{q|p_q = i\}|} \sum_{q|p_q=i} \mathbf{t}_{q,j}$ where $|\{q|p_q = i\}|$ denotes the cardinal of the set $\{q|p_q = i\}$, $i \in \mathbb{F}_2^n$, $q \in [1, Q]$ and p_q is the public input (*e.g.* plaintext) associated with the trace $\mathbf{t}_q \in \mathbb{R}^D$, to compute estimated coefficients $\hat{\alpha}_{:,j}^k$ for all key hypotheses $k \in \mathcal{K}$. □

So far, we have introduced a distinguisher and established its soundness under a condition on the true coefficients. A major limitation of this result is that the condition explicitly depends on the minimal nonlinearity of F_f^Δ. When $nl_{\min}(F_f^\Delta)$ is low, the transform induced by an incorrect key hypothesis may preserve part of the Walsh structure of the true leakage, which can artificially amplify some estimated coefficients. In such cases, the estimation of $\hat{\alpha}^k$ may contain spurious dominant components originating from F_f^Δ rather than from a true point of interest.

Consequently, the soundness condition requires the genuine dominant coefficient $|\alpha_{\max}|$ to be sufficiently large compared to this "structural noise", occurring in the coefficients estimation under wrong key hypotheses. If $nl_{\min}(F_f^\Delta) = 0$, this perturbation can be arbitrarily strong, so that the condition on $\alpha_{\max}$ either cannot be satisfied or becomes unrealistic in practical leakage scenarios. This situation occurs in particular for some nonlinear cryptographic functions f (see Table 2), and systematically when the targeted operation is linear, *e.g.* $f(x, k) = x \oplus k$.

To address the latter case, we introduce in the next section an alternative distinguisher whose soundness does not rely on the nonlinearity of F_f^Δ and which is specifically suited to linear cryptographic functions.

4.2 A sound distinguisher for linear operations: Absolute value of the sum distinguisher

This second distinguisher is derived from Definition 9 and is specifically designed to enable sound attacks on linear cryptographic functions, such as the XOR operation, for which the previous distinguisher may fail. It consists in selecting the sample that maximizes the absolute value of the sum of the estimated coefficients, excluding the bias term which is not discriminative (see Proposition 4). The intuition behind this approach is similar to that of the maximum distinguisher: for a given key hypothesis, the sample maximizing this quantity corresponds to the point where the total contribution of all bits and bit interactions is the largest. While the maximum distinguisher focuses on a single dominant component, the sum-based criterion aggregates all estimated contributions and is therefore more robust when the information is spread over several coefficients. In particular, for linear functions such as XOR, wrong key hypotheses tend to produce estimated coefficients that coincide with the true ones up to sign changes (see Proof of Theorem 4). In this situation, the absolute maximum may no longer discriminate the correct key, whereas the signed sum remains sensitive to the global alignment of the coefficients and allows the correct hypothesis to stand out. Consequently, the key hypothesis for which this aggregated contribution is maximal is expected to correspond to the correct key. The distinguisher is formally defined below.

Definition 10. ***Absolute value of the sum distinguisher based attacks.*** *Given a set of D-dimensional traces $\mathcal{T}$, its associated set of public inputs $\mathcal{X}$ and the matrix $D_{abs_sum} \in \mathbb{R}^{|\mathcal{K}|\times D}$ for which an entry $(D_{abs_sum})_{k,j} = |\hat{\alpha}_j^k \mathbf{1}|$ with $\hat{\alpha}_{1:,j}^k \in \mathbb{R}^{\sum_{i=1}^m \binom{n}{i}}$ and $\mathbf{1}$ being a column vector of ones of size $\sum_{i=1}^m \binom{n}{i}$, the estimated coefficients up to m^{th} bits interactions ($m \leqslant n$) excluding the bias and considering a key hypothesis k and a time sample j, the absolute value of the sum distinguisher is defined as*

$$\mathcal{D}_{abs_sum}^k(\mathcal{T}, \mathcal{X}) = \|(D_{abs_sum})_{k,:}\|_\infty \tag{13}$$

where $\|\cdot\|_\infty$ is the infinity norm. Based on this distinguisher, the key candidate $\hat{k}$ is selected such that

$$\hat{k} = \arg\max_{k\in\mathcal{K}} \mathcal{D}_{abs_sum}^k. \tag{14}$$

Similarly to the maximum distinguisher (see Definition 9), we can also establish the soundness of the absolute-sum distinguisher using Boolean function theory under an appropriate condition on the true coefficients, when targeting a linear cryptographic function f, namely a XOR operation. In the following theorem, we show the soundness of $\mathcal{D}_{\text{abs_sum}}^k$ for conventional LRA; the extension to 'coalesced' LRA implementations is provided in Proposition 7.

The intuition of the proof is the following. For linear functions, an incorrect key hypothesis induces a transformation that preserves the magnitude of the true coefficients but may flip their signs. As a consequence, the estimated coefficients under a wrong key are essentially the true ones up to sign changes. If the genuine coefficients share a globally consistent sign pattern (for instance, they are predominantly positive or predominantly negative), these sign inversions tend to cancel out in the signed sum, whereas the correct key preserves the constructive accumulation of contributions. This makes the absolute value of the sum a discriminating criterion even when the maximum coefficient alone would fail to separate key hypotheses. As for the previous proposed distinguisher, the soundness of the absolute-sum distinguisher relies on a structural assumption on the true coefficients, namely a globally consistent sign pattern. As previously discussed, this assumption concerns only the key-discrimination phase and does not impose any restriction on the leakage estimation itself. Unlike classical leakage model assumptions, it does not constrain the form of the leakage *a priori* and can instead be verified *a posteriori* from the estimated coefficients. Our experiments indicate that such sign-consistency behaviors are observed on the evaluated real-world datasets (see Section 6.2).

Theorem 4. ***Soundness of the absolute value of the sum distinguisher.*** *Let consider a set of Q traces and its associated set of public inputs satisfying Assumption 1. Let also consider that the targeted cryptographic function f is linear i.e. the non linearity of f equals 0 (e.g. XOR). If the true coefficients of the true deterministic part $\alpha \in \mathbb{R}^{2^n \times D}$ have all the same sign (i.e. all coefficients are either positive or negative), excluding the bias, then linear regression based attack using the absolute value of the sum distinguisher $\mathcal{D}_{abs_sum}$ with the full WH basis is sound as $Q \to \infty$ and $R \to 0$.*

Proof. This proof consists in showing that using the absolute value of the sum distinguisher when targeting a linear operation *i.e.* a XOR operation on D-dimensional traces returns as key candidate the true key k^*. Given $\alpha \in \mathbb{R}^{2^n \times D}$ the matrix of true coefficients such that all coefficients have the same sign (*i.e.* all coefficients are either positive or negative), let define $|\alpha_{as,j}| = \left|\sum_{u \in \mathbb{F}_2^n \setminus \{0\}} \alpha_{u,j}\right|$ and $|\alpha_{max_as}| = \max\limits_{1 \leqslant j \leqslant D} |\alpha_{as,j}|$. By definition $|\alpha_{max_as}| > 0$, as it is assumed that side-channel traces contain at least one PoI. Thus, for a given time sample $j \in [1, D]$ such that j is not a PoI, $|\alpha_{max_as}| > \left|\sum_{u \in \mathbb{F}_2^n \setminus \{0\}} \hat{\alpha}^k_{u,j}\right|$ since according to Proposition 5, $\forall\, u \in \mathbb{F}_2^n$, $\hat{\alpha}^k_{u,j} \to 0$ as $Q \to \infty$ and $R \to 0$. Now, the rest of the proof will focus on time samples j such that j is a PoI. It will consist in showing that $|\alpha_{max_as}|$ is reached only if a correct key hypothesis is assumed. As j is a PoI, by expanding the expression of $\hat{\alpha}^k_{u,j}$, one deduces that:

$$\forall\, u \in \mathbb{F}_2^n \setminus \{0\},\ \hat{\alpha}^k_{u,j} \to \frac{1}{2^n} \sum_{v \in \mathbb{F}_2^n \setminus \{0\}} \alpha_{v,j} \sum_{t \in \mathbb{F}_2^n} (-1)^{u \cdot (t \oplus k) \oplus v \cdot (t \oplus k^*)}$$

as $Q \to \infty$ and $R \to 0$ (see proof of Theorem 3, Equation 12). Based on this expression, one can now deduce values returned by the absolute value of the sum

distinguisher for a given key hypothesis $k \in \mathcal{K}$. Let define Δ_k as the distance between the key hypothesis and the true key k^*, the rest of the proof will be divided into two sub-proofs: one for the case $\Delta_k = 0$ and the other for the case $\Delta_k \neq 0$.

Sub-proof 1: $k = k^*$ (*i.e.* $\Delta_k = 0$)

$$\begin{aligned}
\mathcal{D}^{k^*}_{\text{abs_sum}}(\mathcal{T},\mathcal{P}) &= \max_{1\leqslant j\leqslant D}\left[\Big|\sum_{u\in\mathbb{F}_2^n\setminus\{0\}}\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\alpha_{v,j}\sum_{t\in\mathbb{F}_2^n}(-1)^{u\cdot(t\oplus k^*)\,\oplus\, v\cdot(t\oplus k^*)}\Big|\right]\\
&= \max_{1\leqslant j\leqslant D}\left[\Big|\sum_{u\in\mathbb{F}_2^n\setminus\{0\}}\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\alpha_{v,j}\sum_{t\in\mathbb{F}_2^n}(-1)^{(u\oplus v)\cdot(t\oplus k^*)}\Big|\right]\\
&= \max_{1\leqslant j\leqslant D}\left[\Big|\sum_{u\in\mathbb{F}_2^n\setminus\{0\}}\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\alpha_{v,j}\,2^n\,(I_{2^n})_{u,v}\Big|\right]\\
&= \max_{1\leqslant j\leqslant D}\left[\Big|\sum_{u\in\mathbb{F}_2^n\setminus\{0\}}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\alpha_{v,j}\,(I_{2^n})_{u,v}\Big|\right]\\
&= \max_{1\leqslant j\leqslant D}|\alpha_{as,j}| \leqslant |\alpha_{max_as}|.
\end{aligned}$$

By definition of $|\alpha_{max_as}|$, this value is reached for a certain j. Therefore, the equality holds for a given $j \in [1, D]$. The second sub-proof will then consist in showing that $|\alpha_{max_as}|$ will not be reached for all j and a wrong key hypothesis.

Sub-proof 2: $k \neq k^*$ (*i.e.* $\Delta_k = k \oplus k^* \neq 0$)

$$\mathcal{D}^{k}_{\text{abs_sum}}(\mathcal{T},\mathcal{P}) = \max_{1\leqslant j\leqslant D}\left[\Big|\sum_{u\in\mathbb{F}_2^n\setminus\{0\}}\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\alpha_{v,j}\sum_{t\in\mathbb{F}_2^n}(-1)^{u\cdot(t\oplus k^*\oplus\Delta_k)\,\oplus\, v\cdot(t\oplus k^*)}\Big|\right].$$

Using the change of variable $z = t \oplus k^*$ allows one to rewrite the last equation as:

$$\mathcal{D}^{k}_{\text{abs_sum}}(\mathcal{T},\mathcal{P}) = \max_{1\leqslant j\leqslant D}\left[\Big|\sum_{u\in\mathbb{F}_2^n\setminus\{0\}}\frac{1}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\alpha_{v,j}\sum_{z\in\mathbb{F}_2^n}(-1)^{u\cdot(z\oplus\Delta_k)\,\oplus\, v\cdot z}\Big|\right].$$

By the linearity of dot product, one deduces:

$$\begin{aligned}
\mathcal{D}^k_{\text{abs_sum}}(\mathcal{T},\mathcal{P}) &= \max_{1\leqslant j\leqslant D}\left[\Big|\sum_{u\in\mathbb{F}_2^n\setminus\{0\}}\frac{(-1)^{u\cdot\Delta_k}}{2^n}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\alpha_{v,j}\sum_{z\in\mathbb{F}_2^n}(-1)^{(u\oplus v)\cdot z}\Big|\right]\\
&= \max_{1\leqslant j\leqslant D}\left[\Big|\sum_{u\in\mathbb{F}_2^n\setminus\{0\}}(-1)^{u\cdot\Delta_k}\sum_{v\in\mathbb{F}_2^n\setminus\{0\}}\alpha_{v,j}\,(I_{2^n})_{u,v}\Big|\right]\\
&= \max_{1\leqslant j\leqslant D}\left[\Big|\sum_{u\in\mathbb{F}_2^n\setminus\{0\}}(-1)^{u\cdot\Delta_k}\,\alpha_{u,j}\Big|\right] \qquad (15)\\
&< \max_{1\leqslant j\leqslant D}|\alpha_{as,j}| \;\leqslant |\alpha_{max_as}|. \qquad (16)
\end{aligned}$$

The strict inequality comes from the uniform distribution over $\{-1,1\}$ induced by the term $(-1)^{u\cdot\Delta_k}, \forall\, u\in\mathbb{F}_2^n$. Hence, for all j, the value of $|\alpha_{max_as}|$ is not reached when considering an incorrect key hypothesis. □

One can also extend these results in the following corollary when considering a non-full *WH* basis.

Corollary 5. *Let consider a set of Q traces and its associated set of Q public inputs and that the targeted cryptographic function f is linear i.e. the nonlinearity of f equals 0 (e.g. XOR). Under Assumption 1, if the true coefficients of the true deterministic part $\alpha\in\mathbb{R}^{\sum_{i=0}^{m}\binom{n}{i}\times D}$ have all the same sign (i.e. all coefficients are either positive or negative), excluding the bias, with m being the maximal degree of bits interactions considered to conduct the attack, then linear regression based attacks using the absolute value of the sum distinguisher recover the true key k^* asymptotically as $Q\to\infty$ and $R\to 0$, when considering WH basis up to the mth degree of bits interactions.*

Proof. This proof is a straightforward application of Theorem 4 proof, assuming that $\forall\, u,v\in\mathbb{F}_2^n\setminus\{0\}$, $w_H(u)\leqslant m$ and $w_H(v)\leqslant m$ where $w_H(\cdot)$ is the Hamming Weight function. In the case of a wrong key assumption *i.e.* $\Delta_k\neq 0$, the strict inequality in Equation 16 is obtained by noticing that for all values of $\Delta_k\in\mathbb{F}_2^n$, there exists at least one $u\in\mathbb{F}_2^n\setminus\{0\}$, $w_H(u)=1$ verifying $u\cdot\Delta_k=1$ in Equation 15. □

Based on Theorem 4 and Corollary 5 results, we can prove in the following proposition the soundness of our distinguisher when 'coalesced' LRA versions are considered.

Proposition 7. *Following Theorem 4 and Corollary 5 configurations, linear regression based attacks using the absolute value of the sum distinguisher recover the true key k^* asymptotically as $Q\to\infty$, and considering (full or non-full) WH basis and 'coalesced' linear regression attacks.*

Proof. The proof is a direct application of Theorem 4 and Corollary 5 proofs considering 'coalesced' LRA *i.e.* in this case $\bar{\mathbf{T}}_{:,j} \in \mathbb{R}^{2^n}$ is used to compute estimated coefficients $\hat{\alpha}^k_{:,j}$ for all key hypotheses $k \in \mathcal{K}$. Indeed, $\bar{\mathbf{T}}_{:,j}$ corresponds to the classes-averaged version of $\mathbf{T}_{:,j} \in \mathbb{R}^Q$ *i.e.* $\bar{\mathbf{T}}_{i,j} = \frac{1}{|\{q|p_q=i\}|} \sum_{q|p_q=i} \mathbf{t}_{q,j}$ where $|\{q|p_q = i\}|$ denotes the cardinal of the set $\{q|p_q = i\}$, $i \in \mathbb{F}_2^n, q \in [\![1, Q]\!]$ and p_q is the public input (*e.g.* plaintext) associated with the trace $\mathbf{t}_q \in \mathbb{R}^D$. □

Hence, in this section, we introduce a second distinguisher intended for cases where the targeted cryptographic function f is linear, *i.e.* a XOR operation, or more generally when the minimal nonlinearity of F_f^Δ equals 0 even though f is nonlinear. This distinguisher can be applied to both classical and 'coalesced' LRA, with either full or non-full *WH* basis. A current limitation is that the soundness proof is established only for linear functions f. Although the distinguisher can still be applied in situations where f is nonlinear but $nl_{\min}(F_f^\Delta) = 0$, its soundness in this setting remains an open question. An interesting direction for future work is therefore to extend the analysis to such nonlinear functions, in order to obtain soundness guarantees for all possible targeted cryptographic operations. Overall, the distinguishers proposed in Section 4 aim at overcoming the limitations of state-of-the-art LRA metrics highlighted in Section 3 and at covering a wide range of practical use-cases.
The effectiveness of the proposed distinguishers depends on the targeted cryptographic function. For injective targets, classical metrics such as R^2 lose discriminating power in the full-basis setting (see Section 3.2), whereas our coefficient-based distinguishers remain effective under the stated assumptions. For non-injective targets, classical distinguishers may also remain effective. Hence, our distinguishers should be viewed as a principled and generic-compatible alternative for regimes where classical full-basis LRA fails. In the next section, we experimentally evaluate these distinguishers on several scenarios.

5 Experimental results: impact of the basis choice

In this section, we experimentally assess, through simulations, the benefit of using the *WH* basis for LRA compared to the monomial basis. To do so, we first present our experimental setup in Section 5.1. Then, in Section 5.2, we analyze the impact of the basis choice on the estimation of LRA coefficients. Finally, we assess in Section 5.3 how the choice between the monomial and Walsh-Hadamard bases affects the overall performance of LRA attacks.

5.1 Experimental setup

To assess our theoretical claims, we simulate an AES-128 unprotected dataset. This dataset contain $100,000$ traces of dimension $D = 1,000$. Among these $1,000$ time sample, only 10 are *points of interest* (PoIs) *i.e.* informative samples. Let $\mathbf{T}$ be a leakage trace, Y^{k^*} the unprotected targeted variable and $\{l_1, \ldots, l_{10}\}$ the

set corresponding to the indexes of the 10 PoIs, the ith sample $\mathbf{T}_i$ is defined as:

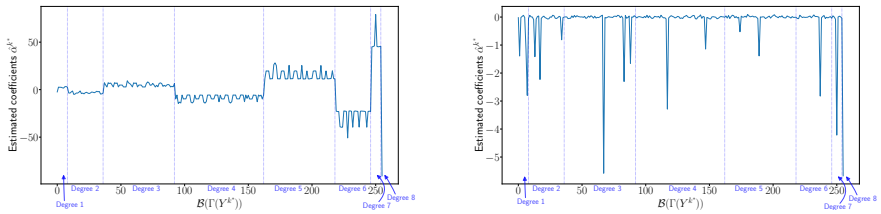
$$\mathbf{T}_i = \begin{cases} \Psi_{ML}(Y^{k^*}) & \text{if } i \in \{l_1, \ldots, l_{10}\} \\ \mathbf{N}_i \sim \mathcal{N}(0, \boldsymbol{\sigma}_i^2) \; s.t. \; \boldsymbol{\sigma}_i \xleftarrow{\$} \{1, 5, 10\} & \text{otherwise.} \end{cases}$$

Ψ_{ML} corresponds to a multivariate leakage model, which is more complex than the Hamming Weight (HW) and Independent Bit Leakage (IBL) models, as it involves high-degree interactions between bits. The highest degree considered in Ψ_{ML} is 8, which is the maximal possible degree and corresponds to the joint interaction of all bits. The coefficients and monomials used to define Ψ_{ML}, representing both individual bits and their interactions, are provided in Appendix A. All our experiments and code are available at the following GitHub repository: https://github.com/Sbsm157/LRA_generic_distinguishers.

5.2 Monomial vs Walsh-Hadamard bases: coefficients estimation

Figure 2 shows the estimation of the coefficients of Ψ_{ML}, for the correct key hypothesis, when using the monomial and the Walsh-Hadamard basis. The comparison between these two set of coeffcients $\hat{\alpha}^{k^*}$ highlights significant differences in terms of stability and interpretability between the two bases. Indeed, when using the monomial basis, the estimated coefficients exhibit very large amplitudes and strong variability across all degrees of interaction (see Figure 2a). In particular, the magnitude of the coefficients tends to increase with the interaction degree, with high-degree monomials reaching values comparable to, or even larger than, those of low-degree terms. This increasing magnitude with the interaction degree is a typical symptom of multicollinearity, which severely deteriorates the conditioning of the Gram matrix (*i.e.* $(B^k)^\top B^k$) associated with the LRA system [13,17]. As a consequence, the regression becomes numerically unstable, and the leakage information is spread over numerous correlated monomials rather than being concentrated on meaningful components [13,17].
In contrast, with the Walsh–Hadamard basis, the coefficients are much better structured and remain within a narrow range (see Figure 2b). Most of the significant contributions are concentrated on a limited number of basis elements. Importantly, the estimated non-zero coefficients correspond precisely to the bits and bit interactions defined in Ψ_{ML}, showing that the regression correctly identifies the true leakage components. On the other hand, the coefficients associated with terms that do not contribute to Ψ_{ML} remain close to zero, consistently with their true null contribution. This reflects the orthogonality property of the *WH* basis, which suppresses interactions between predictors and leads to a more accurate and reliable estimation of the underlying leakage model, as detailed in Theorem 1 and Corollary 1.
Overall, these observations confirm that the *WH* basis provides a more reliable representation for LRA, yielding stable and interpretable coefficients. In contrast, the monomial basis introduces numerical instability and can distort the estimation of the deterministic leakage component. Such an inaccurate estimation may severely degrade the effectiveness of the attack, especially in the case of

complex leakage models involving high-degree bit interactions, and this remains true even for non-generic LRA attacks, *i.e.* when the basis is restricted and not full.



(a) Monomial basis (b) *WH* basis

Fig. 2: Estimated coefficients of Ψ_{ML} using the monomial and *WH* bases for the correct key hypothesis *i.e.* $k = k^*$.

5.3 Monomial vs Walsh-Hadamard bases: LRA performance

Table 1 compares classical LRA based on the R^2 distinguisher with the monomial basis to our approach using the $\mathcal{D}^k_{max}$ distinguisher with the *WH* basis, for various maximum interaction degrees m. All reported results are averaged over 10 independent attacks. It should be noted that, for LRA with the monomial basis and the R^2 distinguisher, the results reported in Table 1 correspond to attacks performed with $10,000$ traces and averaged over 10 independent attacks. To ensure that the failure was not due to an insufficient number of traces, we also conducted experiments with $50,000$ traces averaged over 10 independent attacks, which led to globally similar average rank values. In addition, we performed a single attack using up to $100,000$ traces. Despite this substantial increase in data, the attack still did not converge and the average rank remained close to the values reported in the table, confirming that the issue is not related to a lack of data.

The results reported in Table 1 clearly confirm the practical impact of the coefficient estimation issues previously highlighted in Section 5.2. When LRA is performed with the R^2 distinguisher and a restricted monomial basis, the attack systematically fails, with final average ranks close to the random baseline (*i.e.* AR≈ 128) for almost all values of m. This shows that restricting the monomial basis does not solve the fundamental problem: even low-degree representations lead to an unreliable model that is unable to discriminate the correct key hypothesis, although, in theory, the R^2 distinguisher is expected to be effective when the basis is sufficiently restricted [37,23].

In contrast, our approach based on the *WH* basis and the $\mathcal{D}^k_{max}$ distinguisher succeeds for all configurations and requires only a moderate number of traces,

even when the equivalent interaction degree is low. This difference directly echoes the observations made in Figure 2. As explained previously, the monomial basis produces highly unstable coefficient estimates, whose magnitude increases with the interaction degree due to severe multicollinearity and poor conditioning of the Gram matrix $(B^k)^\top B^k$. Such a distorted estimation of the leakage model coefficients prevents the distinguisher from correctly separating key hypotheses, which explains the quasi-random average rank values observed in the first row of the table.

Conversely, the *WH* basis provides well-structured coefficients that accurately reflect the true contributions of Ψ_{ML}, while non-contributing terms remain close to zero. Thanks to the orthogonality of the basis, the regression remains numerically stable and the leakage information is concentrated on the true meaningful components, leading to a reliable key recovery. These results therefore demonstrate that the degradation of classical LRA (*i.e.* LRA with monomial basis and R^2 distinguisher) is not only due to the use of a full basis but is inherent to the monomial representation itself. In particular, when the leakage model is complex, as in the case of Ψ_{ML} which involves high-degree bit interactions, the inaccurate estimation of the deterministic part can prevent classical LRA from functioning at all. Even in non-generic settings with restricted bases, such modeling errors may render the attack ineffective, whereas our *WH*-based approach remains robust. Hence, as the *WH* basis has proven to be the most reliable representation for LRA, we restrict the remainder of our study to this basis. The following section is therefore dedicated to assessing our proposed generic-compatible distinguishers.

	Maximum interaction degree m (restricted monomial basis)							
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$
LRA + R^2 + monomial basis [11]	$AR:123$	$AR:149$	$AR:132$	$AR:135$	$AR:106$	$AR:133$	$AR:151$	$AR:22$
LRA + $\mathcal{D}^k_{max}$ + WH basis (ours)	5 000	5 000	750	750	750	1 500	2 000	2 500

Table 1: Comparison between classical LRA using R^2 with a restricted monomial basis and our approach using $\mathcal{D}^k_{max}$ with the WH basis. For the classical approach, AR:X denotes the final average rank when the key is not recovered. For our approach, the number indicates the traces required to reach a constant rank of 1.

6 Experimental results: proposed generic-compatible distinguishers

In this section, we experimentally assess using simulations in Section 6.1 and publicly available datasets in Section 6.2 theoretical claims made on previous sections for both our proposed generic-compatible distinguishers.

Regarding the validity of Assumption 1 made throughout this work, this setting is standard in side-channel analysis [25,31] and is widely observed in practice. During evaluation campaigns, public inputs are typically generated uniformly at random, and numerous public datasets show that each value occurs approximately $Q/2^n$ times [36,30,2].

6.1 Attack results on simulations

6.1.1 Experimental setup

Presentation of datasets. To illustrate our theoretical claims, we simulate 4 unprotected datasets to cover large cases encountered when performing SCA *i.e.* differents targeted operations, various specificities related to attacked cryptographic functions (input dimension n of the function, minimal nonlinearity reached $nl_{min}(F_f)$, see Table 2 for more details). These datasets target respectively AES SBox output, a XOR operation, PRINCE SBox output and Clyde-128 SBox output. These datasets contain 100 000 traces of dimension $D = 1\,000$, including both informative and uninformative samples, with leakage models that vary across the PoIs. Among these 1 000 samples, only about 1% correspond to PoIs, *i.e.* 10 samples. For all datasets, we consider unprotected targeted variables Y^{k^*}. For all SBox functions considered, we target the first round. Hence, Y^{k^*} takes the form of:

- For AES and XOR simulations, $Y^{k^*} = Sbox(P_0 \oplus k^*)$ and $Y^{k^*} = P_0 \oplus k^*$ respectively, given the first plaintext byte P_0 and $k^* \in \mathbb{F}_2^8$.
- For PRINCE simulations, $Y^{k^*} = Sbox(P_0 \oplus k_0^* \oplus k_1^*)$ with $P_0, k_0^*, k_1^* \in \mathbb{F}_2^4$. In our experiments, we consider the whitening key k_0^* known.
- For Clyde-128 simulations, $Y^{k^*} = Sbox(P_0 \oplus k^* \oplus TK)$ with $P_0, k^*, TK \in \mathbb{F}_2^4$. In our experiments, we consider the tweakey TK known.

Consequently, given a leakage trace $\mathbf{T}$ composed of 10 PoIs and $\{l_1, \ldots, l_{10}\}$ the set corresponding to the indexes of the 10 PoIs, the ith sample of $\mathbf{T}$ for all datasets is defined as:

$$\mathbf{T}_i = \begin{cases} \Psi_{HW}(Y^{k^*}) & \text{if } i \in \{l_1, \ldots, l_{10}\} \text{ and } i \equiv 0 \mod 3 \\ \Psi_{IBL}(Y^{k^*}) & \text{if } i \in \{l_1, \ldots, l_{10}\} \text{ and } i \equiv 1 \mod 3 \\ \Psi_{ML}(Y^{k^*}) & \text{if } i \in \{l_1, \ldots, l_{10}\} \text{ and } i \equiv 2 \mod 3 \\ \mathbf{N}_i \sim \mathcal{N}(0, \boldsymbol{\sigma}_i^2) \; s.t. \; \boldsymbol{\sigma}_i \xleftarrow{\$} \{1, 5, 10\} & \text{otherwise.} \end{cases}$$

where Ψ_{HW} corresponds to the Hamming Weight leakage model, Ψ_{IBL} and Ψ_{ML} correspond respectively to the independent bit leakage and multivariate leakage (*i.e.* a leakage model that also depends on high degree bits' interactions) models. Coefficients for Ψ_{IBL} and Ψ_{ML} are arbitrarily chosen such that they vary across different bits and their interactions. All detailed coefficients are respectively presented in Appendix A and Appendix B for $n = 8$ and 4-bit cryptographic functions. In our experiments, we consider a non-isotropic white Gaussian noise *i.e.* variances vary accross all samples. Resulting Signal-to-Noise ratios (SNR) for all datasets are summarized in Table 2.

	f	n	$nl_{min}(F_f)$	$\max_{i\in[1,D]}(SNR)_i$	Proposed distinguishers
AES Simulation	SBox	8	112	2.01	$\mathcal{D}^k_{max}$
PRINCE Simulation	SBox	4	2	2	
XOR Simulation	XOR	8	0	2.04	$\mathcal{D}^k_{abs_sum}$
Clyde-128 Simulation	SBox	4	0	2.01	

Table 2: Simulations − Specificities of cryptographic functions targeted.

Configuration. We implement our experiments using Python. Our code is available at the following GitHub repository : https://github.com/Sbsm157/LRA_generic_distinguishers. We conduct attacks using *WH* basis, considering the highest degree of bits' interactions *i.e.* the full basis, for all datasets. Attacks considered are LRA and 'coalesced' LRA. The latter are performed considering whole traces. We use as attack metric the guessing entropy *i.e.* expected rank of the true key k^* [35]. Regarding traces preprocessing, we standardize all traces sets. The attack results obtained by averaging 10 independent shuffled experiments for all simulations are reported in Table 3. The following section presents results obtained for simulations.

6.1.2 Results obtained on simulations

		AES Simulation	PRINCE Simulation	XOR Simulation	Clyde-128 Simulation
R^2 distinguisher	LRA	AR:139	AR:5	AR:22	AR:9
	'coalesced' LRA	AR:22	AR:6	AR:22	AR:6
Proposed distinguishers	LRA	**2,500**	**300**	**5,000**	**150**
	'coalesced' LRA	**2,500**	**300**	**5,000**	**150**

Table 3: Simulations − Number of traces needed to retrieve the secret key considering LRA and 'coalesced' LRA. AR:X denotes the final average rank in cases where the rank of the secret key does not converge to 1.

According to reported results in Table 3, we observe that both our proposed distinguishers lead, for LRA and 'coalesced' LRA based attack, to retrieve the secret key for all simulations, which is not the case when attacks are carried out considering R^2 distinguisher. Indeed, unlike 'coalesced' LRA, for which we know the attack won't succeed with state-of-the-art distinguishers (see Theorem 2, Corollary 2 and Corollary 3), it interestingly appears that even with conventional LRA the attack is also inconclusive when considering R^2 distinguisher. These results can be explained by the fact that intuitively when considering only deterministic part of traces (*i.e.* informative part of traces) the system of

equations resolved by the linear regression is fully determined for all key hypotheses as there is 2^n equations for 2^n unknowns [37]. This therefore leads to a solution that maximizes R^2 distinguisher for all key hypotheses. Hence, these preliminary empirical observations highlight the practical limitations of using R^2 as a distinguisher to carry out SCA *i.e.* R^2 is not suitable to conduct SCA in a full basis setting. Thus, these observations highlight practical limitations of the R^2 distinguisher in full-basis settings, whereas our proposed distinguishers remain effective in all considered simulations. In particular, regarding $\mathcal{D}^k_{max}$ distinguisher, especially when considering PRINCE simulation dataset for which conditions required to guarantee $\mathcal{D}^k_{max}$ soundness are not met *i.e.* the condition $\forall\, j \in [1, D]$, $|\alpha_{max}| \geqslant \frac{3}{4} \sum_{v \in \mathbb{F}_2^n \setminus \{0\}} |\alpha_{j,v}|$ is not fulfilled (see Theorem 3), we notice that the secret key is still retrieved using $\mathcal{D}^k_{max}$ for our complex leakage model. By conducting supplementary simulations for which there is no guarantee of soundness, we observe that when only considering respectively Hamming Weight leakage model and independent bit leakage model (as defined in Appendix B), our distinguisher is no longer able to find the key as the attack converges towards an averaged rank of 8 against 6 for R^2 distinguisher when considering Hamming Weight leakage model and an averaged rank of 2 against 6 for R^2 distinguisher, in the case of independent bit leakage model. Hence, for the independent bit leakage model, our distinguisher still allows one to complete the attack, whereas this is not the case for the R^2 distinguisher. Regarding $\mathcal{D}^k_{abs_sum}$ distinguisher, we notice that even if it hasn't been proven that our distinguisher works for Clyde-128 simulation dataset, $\mathcal{D}^k_{abs_sum}$ distinguisher is still able to recover the secret key. Here again, we conduct supplementary simulations for which there is no guarantee about the soundness of $\mathcal{D}^k_{abs_sum}$. We remark that when only considering respectively Hamming Weight leakage model and independent bit leakage model (as defined in Appendix B), LRA achieves an average rank of 2, using $\mathcal{D}^k_{abs_sum}$, for Clyde-128 simulations. This still allows one to complete the attack, which is not the case when LRA is carried out using R^2 distinguisher as it achieves an averaged rank of 6 for both datasets.
Overall, these simulations suggest that the proposed distinguishers provide a more robust alternative in full-basis settings where the discriminative power of R^2 deteriorates. In the following section, we assess our proposed distinguishers using publicly available datasets.

6.2 Attack results on real datasets

6.2.1 Setup

Presentation of datasets. To assess our proposed distinguishers and confirm results obtained in simulations (see Section 6.1), we select 4 publicly available datasets that implement AES (as this algorithm is one of the only cryptographic primitive used in publicly available databases) considering a wide range of use cases: unprotected software implementation on smart cards, unprotected hardware implementation on FPGA and protected software implementations with first-order masking.

- **DPA contest v4.2** is a dataset from fourth edition of DPA contest [36,3] that implements an AES-128 software protected with first-order masking and shuffling countermeasures on an Atmel ATMega-163 smart-card. The dataset is available at https://dpacontest.telecom-paris.fr/v4. In our experiments, we consider this implementation as unprotected because we assume that an attacker has at its disposal mask and shuffling offset values. The targeted variable is the first byte of the first round SBox operation. Consequently, given a known mask value M and the shuffling offset value *offset*, the targeted variable associated to each pair of traces $\mathbf{T}$ and 16-byte plaintexts P is $SBox(P_0 \oplus k^*) \oplus M[\mathit{offset} + 1]$ with P_0 being the first byte of the plaintext P.
- **AES_HD_Ext** is an extension of **AES_HD** dataset introduced in [30] and consists in an unprotected AES-128 hardware implementation on Xilinx Virtex-5 FPGA. In this dataset, the targeted operation is the register writing in the last round. Consequently, for a given leakage trace $\mathbf{T}$ and its associated 16-byte ciphertext C, the targeted variable is $SBox^{-1}(C_{12} \oplus k^*) \oplus C_8$ with C_8 and C_{12} being respectively the eighth and twelfth byte of C. The relation between C_8 and C_{12} is given through the inverse ShiftRows of AES. The dataset is available at https://github.com/AISyLab/AES_HD_Ext.
- **ASCAD v1** is a dataset introduced in [2] that consists in a protected AES-128 software implementation which runs over an 8-bit AVR microcontroller ATMega8515. The countermeasure implemented is a first-order boolean masking. For each leakage trace $\mathbf{T}$ and its associated 16-byte plaintext P, the targeted operation is the third byte of the first round SBox operation *i.e.* the sensitive value is $SBox(P_3 \oplus k^*)$ with P_3 being the third byte of the plaintext P. However, as this implementation is protected, the sensitive variable is blinded by considering a 16-byte random mask R *i.e.* $SBox(P_3 \oplus k^*) \oplus R_3$. We consider two versions to conduct our experiments: **ASCAD v1-F** and **ASCAD v1-R**. The whole database is available at https://github.com/ANSSI-FR/ASCAD.

Configuration. We adopt the same configuration as described in Section 6.1.1. Regarding **ASCAD v1** datasets, as both datasets are masked implementations, a preliminary preprocessing step is necessary to conduct an attack. To do so, we recombine 13 samples that maximizes the SNR associated to the masked value $SBox(P_3 \oplus k^*) \oplus R_3$ and the mask value R_3. The combination function used in our case is the optimal product combining *i.e.* centered product [33]. This results in recombined traces of dimension $D = 169$. Regarding traces preprocessing, we consider for **DPA contest v4.2** and **ASCAD v1-F** datasets maximum absolute value scaling. We apply no scaling for **ASCAD v1-R** dataset. Finally, for **AES_HD_Ext** dataset, we apply on traces the bias suppression preprocessing method and then add an arbitrary constant that we set, in our case, to 5. The bias suppression preprocessing consists in removing traces modulation. It is done by subtracting the empirical mean $\hat{\boldsymbol{\mu}} \in \mathbb{R}^{\boldsymbol{D}}$ from traces, with D being traces dimension. The attack results are obtained for all datasets by averaging 10 independent shuffled experiments and are reported in Table 4. As for simulations,

we use as attack metric the guessing entropy *i.e.* expected rank of the true key k^* and conduct attacks considering whole traces, using $\mathcal{D}^k_{max}$ as distinguisher since targeted cryptographic function is AES Sbox. As a reminder, conventional and 'coalesced' LRA are conducted using *WH* basis and considering full basis *i.e.* all bits' interactions. The following section presents results obtained for all datasets.

6.2.2 Results obtained on real datasets

		DPA contest v4.2	AES_HD_Ext	ASCAD v1-F	ASCAD v1-R
R^2 distinguisher	LRA	**2,000**	15,000	AR:119	AR:162
	'coalesced' LRA	N/A	N/A	AR:161	AR:126
$\mathcal{D}^k_{max}$ distinguisher	LRA	3,500	**12,500**	**3,000**	**2,500**
	'coalesced' LRA	N/A	N/A	**3,000**	**2,500**

Table 4: Number of traces needed to retrieve the secret key considering LRA and 'coalesced' LRA for all datasets. AR:X denotes the final average rank in cases where the rank of the secret key does not converge to 1.

The experimental results highlight two distinct regimes depending on the structure of the targeted cryptographic function. For non-injective target functions, such as **DPA contest v4.2** and **AES_HD_Ext**, classical distinguishers such as R^2 may remain effective even in a full-basis setting, as the additional known mask/ciphertext values increase the effective number of leakage configurations. In contrast, for injective target functions, such as the simulations and **ASCAD** datasets, the theoretical limitations established in Section 3.2.2 apply and may cause the discriminative power of R^2 to deteriorate. Referring to Table 4, which reports results obtained for LRA only as 'coalesced' LRA is not applicable for **DPA contest v4.2** dataset (sensitive value targeted cannot be partitioned considering only the 256 possible plaintexts values), we observe that the R^2 distinguisher enables a successful attack. This apparent contradiction with the theoretical limitations of R^2 in a full *WH* basis (see Section 3.2.2) is explained by the masking countermeasure. The random mask (known in our context) introduces additional variability so that, for each plaintext class, multiple distinct leakage realizations are observed. Consequently, the regression no longer involves only 256 deterministic unknowns associated with 256 plaintext classes, but effectively about 256×256 mask–plaintext configurations. This increases the effective rank of the regression matrix B^k, which is no longer close to a 256-rank structure but approaches a rank of roughly 256×256. As a result, wrong keys can no longer perfectly interpolate the leakage, restoring contrast between key hypotheses and thus the discriminative power of the R^2 distinguisher. Indeed, LRA using R^2 as distinguisher converges towards a stabilized rank of 1 within 2,000 traces. Regarding our proposed distinguisher, 3,500 traces are required to achieve a

stabilized rank of 1. This gap in performance can be explained by the fact that the SNR for this dataset amounts to a maximum value of 6. Indeed, the sample selected by the R^2 distinguisher is the one that maximizes SNR. This may allow the R^2 distinguisher to retrieve the correct key k^* more easily (even when considering a full basis). This gap also highlights the fact that our distinguisher $\mathcal{D}^k_{max}$, although proven to be sound (see Theorem 3, Proposition 6), is not necessarily optimal. For this dataset, R^2 achieves convergence with fewer traces than our proposed distinguisher. This indicates that our distinguishers are not uniformly optimal in terms of trace complexity, even though they remain theoretically sound. More generally, in such non-injective settings, both classical and proposed distinguishers may successfully recover the key, with no clear criterion allowing one to predict which distinguisher will be the most trace-efficient. Indeed, regarding **AES_HD_Ext** dataset for which results obtained only concerns LRA since 'coalesced' LRA is not applicable for the same reasons as those mentioned for **DPA contest v4.2** dataset, one can observe that both distinguishers lead to a successful attacks. The same explanation as above applies: the final XOR with the ciphertext byte C_8 increases the effective number of configurations and the rank of the regression matrix B^k, so that wrong keys can no longer perfectly interpolate the leakage, which restores the discriminative power of R^2 despite the use of a full WH basis. However, in contrast to results achieved for **DPA contest v4.2** dataset, results for **AES_HD_Ext** dataset suggest to rather favor our distinguisher when considering a low-SNR hardware implementation. Indeed, $12,500$ traces are required to retrieve the correct key when considering our proposed distinguisher against $15,000$ for R^2-distinguisher. Finally, for **ASCAD v1-F** and **ASCAD v1-R** datasets, we notice that there is no successful attacks when using R^2 as distinguisher whether the dataset considered. This is not the case when considering our proposed distinguisher $\mathcal{D}^k_{max}$ since the correct key k^* is retrieved within $3,000$ and $2,500$ traces for **ASCAD v1-F** and **ASCAD v1-R** datasets respectively. These experimental results are in line with those reported in Section 6.1 as they highlight again practical limitations of using R^2 as a distinguisher for SCA and seem to suggest that R^2 is not a suitable and reliable distinguisher to conduct SCA, unlike our proposed distinguisher as stated in Theorem 3 and Proposition 6. We nevertheless note an additional limitation when interpreting the **ASCAD v1** results. Since traces are obtained after centered-product recombination in order to conduct a higher-order attack, the resulting noise component may deviate from the Gaussian noise assumption introduced in Section 2.1. It thus remains unclear whether Euclidean distance-based distinguishers such as R^2, are still suitable under such preprocessed leakages.

All in all, the experiments suggest that the proposed distinguishers provide a robust alternative to classical R^2-based distinguishers in full-basis LRA settings. In particular, they remain effective in injective regimes where the discriminative power of R^2 deteriorates, while remaining competitive in more favorable non-injective settings. Therefore, our distinguishers should be viewed as a new, more

robust baseline for LRA, as they address limitations of classical ones, especially without leakage assumptions.

7 Discussion and conclusion

In this paper, we address a long-standing problem in side-channel analysis: the design of generic non profiled attacks, namely attacks that do not rely on explicit prior assumptions on the leakage model.
To this end, we revisit Linear Regression Analysis from this perspective and show that its previous limitations were not inherent to LRA itself, but to the interaction between classical distinguishers, the full predictor space, and the monomial basis.
We show that the monomial representation, even in a restricted setting, can lead to misleading estimations, explaining why classical LRA distinguishers behave inconsistently in practice. In parallel, we demonstrate that the Walsh-Hadamard basis removes multicollinearity and provides a structured coefficient space consistent with Boolean function theory, enabling the use of all predictors and a truly model-agnostic estimation of the leakage. However, when LRA is performed in this full-basis configuration, state-of-the-art metrics such as R^2 still lose their discriminating ability, as the regression becomes nearly determined for every key hypothesis. Consequently, the degradation of R^2 is not only a consequence of the full basis, but can also stem from the inadequacy of the monomial representation in the presence of complex leakage models involving high-degree interactions.
To overcome this issue, we introduce new coefficient-based and generic-compatible distinguishers in the Walsh-Hadamard basis and provide formal soundness results under explicit conditions using Boolean function theory. These results are supported by experiments on simulations and publicly available datasets, covering a wide range of practical scenarios. Our approach differs fundamentally from previous *generic-emulating* LRA based on stepwise or ridge regression: instead of selecting or penalizing predictors through optimization heuristics, we provide a deterministic and reproducible procedure that preserves the full expressiveness of the leakage model. While such heuristic approaches could also benefit from the use of the Walsh-Hadamard basis, they remain inherently experimental, as their success depends on empirical tuning and model-selection choices. Investigating these combinations is left for future work.
Despite these advances, several aspects deserve further investigation. The current soundness proofs, especially when targeting non-linear cryptographic functions, rely on the assumption that, for the correct key, the dominant coefficient contribution sufficiently exceeds the noise-induced contributions that may arise under wrong key hypotheses. Relaxing this requirement would broaden the applicability of the method and constitutes an important direction for future work. Moreover, for cryptographic functions with extremely low nonlinearity (*i.e.* when $F_f = 0$), extensions of the distinguisher may be necessary. Finally, the optimality of the proposed metrics within the LRA framework was not addressed in this paper, studying this question remains an open problem for future research.

Overall, this work shows that generic non profiled LRA in the sense of full-basis predictor modeling can be performed in a principled way and that sound distinguishers can be designed under explicit structural assumptions on the regression coefficients, without relying on predefined leakage models.

Acknowledgments

The author thanks Gabriel Zaid, Mathieu Carbone, Benoît Gérard, and Guénaël Renault for their supervision, insightful discussions, and valuable feedback that improved this work. The author is also grateful to François-Xavier Standaert for the fruitful discussions on generic LRA and for suggesting the inclusion of Section 5. Finally, the author thanks the anonymous reviewers for their constructive comments.

A Leakage models considered in simulations for $n = 8$-bit cryptographic functions

To simulate traces manipulating a $n = 8$-bit cryptographic functions such as an AES SBox or a XOR operation, we generate traces $\mathbf{T}$ composed of 10 PoIs such that the ith sample is defined as:

$$\mathbf{T}_i = \begin{cases} \Psi_{HW}(Y^{k^*}) & \text{if } i \in \{l_1, \ldots, l_{10}\} \text{ and } i \equiv 0 \mod 3 \\ \Psi_{IBL}(Y^{k^*}) & \text{if } i \in \{l_1, \ldots, l_{10}\} \text{ and } i \equiv 1 \mod 3 \\ \Psi_{ML}(Y^{k^*}) & \text{if } i \in \{l_1, \ldots, l_{10}\} \text{ and } i \equiv 2 \mod 3 \\ \mathbf{N}_i \sim \mathcal{N}(0, \sigma_i^2) \; s.t. \; \sigma_i \xleftarrow{\$} \{1, 5, 10\} & \text{otherwise.} \end{cases}$$

where $\{l_1, \ldots, l_{10}\}$ corresponds to the set of indexes of the 10 PoIs, Ψ_{HW}, Ψ_{IBL} and Ψ_{ML} corresponds respectively to the Hamming Weight leakage model, independent bit leakage model and multivariate leakage model. In our experiments, coefficients of independent bit leakage model and multivariate leakage model are arbitrarily chosen such that they vary across bits and their interactions. Hence, these two leakage models are defined as follow:

- **Independent bit leakage model**: we assume that each bit leaks independently from the others.

$$\Psi_{IBL} = \sum_{b=0}^{7} \alpha_b \, (Y^{k^*})_b + \mathbf{N}_i,$$

  where $\alpha = [1, 0.5, 1, 0, 2, 0.5, 0.75, 0.25]$ and $\mathbf{N}_i \sim \mathcal{N}(0, \boldsymbol{\sigma}_i^2)$ such that $\boldsymbol{\sigma}_i \xleftarrow{\$} \{1, 5, 10\}$.
- **Multivariate leakage model**: we assume that the leakage also depends on high degree of bits' interaction. To simplify the equations, we exceptionally

denote here $(Y^{k^*})_b$ *i.e.* the bth bit of the sensitive value Y^{k^*} as Y_b.

$$\begin{aligned}\Psi_{ML} =& 0.25\,Y_0 + 0.2\,Y_5 + 0.5\,Y_6 + 0.25\,r(Y_1Y_3) + 0.4\,r(Y_2Y_4) + 0.15\,r(Y_4Y_7)\\ &+ 1\,r(Y_0Y_5Y_6) + 0.4\,r(Y_1Y_5Y_7) + 0.3\,r(Y_1Y_6Y_7) + 0.6\,r(Y_2Y_3Y_4Y_6)\\ &+ 0.2\,r(Y_3Y_4Y_5Y_7) + 0.1\,r(Y_0Y_1Y_2Y_5Y_6) + 0.25\,r(Y_0Y_1Y_2Y_5Y_7)\\ &+ 0.5\,r(Y_0Y_1Y_2Y_5Y_6Y_7) + 0.75\,r(Y_0Y_1Y_2Y_3Y_5Y_6Y_7)\\ &+ 1\,r(Y_0Y_1Y_2Y_3Y_4Y_5Y_6Y_7) + \mathbf{N}_i,\end{aligned}$$

where r is a function that characterizes bits' interaction (*e.g.* XOR or AND operation) and $\mathbf{N}_i \sim \mathcal{N}(0, \boldsymbol{\sigma}_i^2)$ such that $\boldsymbol{\sigma}_i \xleftarrow{\$} \{1, 5, 10\}$.

B Leakage models considered in simulations for $n = 4$-bit cryptographic functions

To simulate traces manipulating a $n = 4$-bit cryptographic functions such as an AES SBox or a XOR operation, we generate traces $\mathbf{T}$ composed of 10 PoIs such that the ith sample is defined as:

$$\mathbf{T}_i = \begin{cases} \Psi_{HW}(Y^{k^*}) & \text{if } i \in \{l_1, \ldots, l_{10}\} \text{ and } i \equiv 0 \mod 3 \\ \Psi_{IBL}(Y^{k^*}) & \text{if } i \in \{l_1, \ldots, l_{10}\} \text{ and } i \equiv 1 \mod 3 \\ \Psi_{ML}(Y^{k^*}) & \text{if } i \in \{l_1, \ldots, l_{10}\} \text{ and } i \equiv 2 \mod 3 \\ \mathbf{N}_i \sim \mathcal{N}(0, \boldsymbol{\sigma}_i^2) \; s.t. \; \boldsymbol{\sigma}_i \xleftarrow{\$} \{1, 5, 10\} & \text{otherwise.} \end{cases}$$

where $\{l_1, \ldots, l_{10}\}$ corresponds to the set of indexes of the 10 PoIs, Ψ_{HW}, Ψ_{IBL} and Ψ_{ML} corresponds respectively to the Hamming Weight leakage model, independent bit leakage model and multivariate leakage model. In our experiments, coefficients of independent bit leakage model and multivariate leakage model are arbitrarily chosen such that they vary across bits and their interactions. Hence, these two leakage models are defined as follow:

- **Independent bit leakage model**: we assume that each bit leaks independently from the others.

$$\Psi_{IBL} = \sum_{b=0}^{3} \alpha_b\,(Y^{k^*})_b + \mathbf{N}_i,$$

  where $\alpha = [1,\ 0.5,\ 1,\ 0]$ and $\mathbf{N}_i \sim \mathcal{N}(0, \boldsymbol{\sigma}_i^2)$ such that $\boldsymbol{\sigma}_i \xleftarrow{\$} \{1, 5, 10\}$.
- **Multivariate leakage model**: we assume that the leakage also depends on high degree of bits' interaction. To simplify the equations, we exceptionally denote here $(Y^{k^*})_b$ *i.e.* the bth bit of the sensitive value Y^{k^*} as Y_b.

$$\begin{aligned}\Psi_{ML} =& 1\,Y_0 + 2\,Y_1 + 0.5\,Y_3 + 0.25\,r(Y_0Y_1) + 1\,r(Y_0Y_3)\\ &+ 0.15\,r(Y_1Y_2) + 1\,r(Y_1Y_3) + 0.4\,r(Y_2Y_3) + 0.3\,r(Y_0Y_1Y_2)\\ &+ 0.6\,r(Y_0Y_2Y_3) + 0.2\,r(Y_1Y_2Y_3) + 0.1\,r(Y_0Y_1Y_2Y_3) + \mathbf{N}_i,\end{aligned}$$

  where r is a function that characterizes bits' interaction (*e.g.* XOR or AND operation) and $\mathbf{N}_i \sim \mathcal{N}(0, \boldsymbol{\sigma}_i^2)$ such that $\boldsymbol{\sigma}_i \xleftarrow{\$} \{1, 5, 10\}$.

References

1. Agrawal, D., Archambeault, B., Rao, J.R., Rohatgi, P.: The em side—channel(s). In: Kaliski, B.S., Koç, ç.K., Paar, C. (eds.) Cryptographic Hardware and Embedded Systems - CHES 2002. pp. 29–45. Springer Berlin Heidelberg, Berlin, Heidelberg (2003)
2. Benadjila, R., Prouff, E., Strullu, R., Cagli, E., Dumas, C.: Deep learning for side-channel analysis and introduction to ascad database. Journal of Cryptographic Engineering **10**(2), 163–188 (2020)
3. Bhasin, S., Bruneau, N., Danger, J.L., Guilley, S., Najm, Z.: Analysis and improvements of the dpa contest v4 implementation. In: Chakraborty, R.S., Matyas, V., Schaumont, P. (eds.) Security, Privacy, and Applied Cryptography Engineering. pp. 201–218. Springer International Publishing, Cham (2014)
4. Brier, E., Clavier, C., Olivier, F.: Correlation power analysis with a leakage model. In: Cryptographic Hardware and Embedded Systems - CHES 2004: 6th International Workshop Cambridge, MA, USA, August 11-13, 2004. Proceedings. Lecture Notes in Computer Science, vol. 3156, pp. 16–29. Springer (2004). https://doi.org/10.1007/978-3-540-28632-5_2, https://iacr.org/archive/ches2004/31560016/31560016.pdf
5. Canteaut, A.: Lecture notes on cryptographic boolean functions. Inria, Paris, France **3** (2016)
6. Carlet, C.: Boolean Functions for Cryptography and Error-Correcting Codes, p. 257–397. Encyclopedia of Mathematics and its Applications, Cambridge University Press (2010)
7. Carlet, C. (ed.): Boolean Functions for Cryptography and Coding Theory. Cambridge University Press (2020). https://doi.org/10.1017/9781108606806, https://doi.org/10.1017/9781108606806
8. Carlet, C., Crama, Y., Hammer, P.L.: Vectorial boolean functions for cryptography. (2010)
9. Chari, S., Rao, J.R., Rohatgi, P.: Template attacks. In: Kaliski, B.S., Koç, ç.K., Paar, C. (eds.) Cryptographic Hardware and Embedded Systems - CHES 2002. pp. 13–28. Springer Berlin Heidelberg, Berlin, Heidelberg (2003)
10. Dabosville, G., Doget, J., Prouff, E.: A new second-order side channel attack based on linear regression. IEEE Transactions on Computers **62**(8), 1629–1640 (2012)
11. Doget, J., Prouff, E., Rivain, M., Standaert, F.X.: Univariate side channel attacks and leakage modeling. Journal of Cryptographic Engineering **1**, 123–144 (2011)
12. Gierlichs, B., Batina, L., Tuyls, P., Preneel, B.: Mutual information analysis. In: Cryptographic Hardware and Embedded Systems - CHES 2008, 10th International Workshop, Washington, D.C., USA, August 10-13, 2008. Proceedings. Lecture Notes in Computer Science, vol. 5154, pp. 426–442. Springer (2008). https://doi.org/10.1007/978-3-540-85053-3_27, https://www.iacr.org/archive/ches2008/51540423/51540423.pdf
13. Greene, W.: Econometric Analysis Global Edition - 8th Edition. Pearson Deutschland (2019), https://elibrary.pearson.de/book/99.150005/9781292231150
14. Guilley, S., Heuser, A., Ming, T., Rioul, O.: Stochastic side-channel leakage analysis via orthonormal decomposition. In: Innovative Security Solutions for Information Technology and Communications: 10th International Conference, SecITC 2017, Bucharest, Romania, June 8–9, 2017, Revised Selected Papers 10. pp. 12–27. Springer (2017)

15. Guilley, S., Heuser, A., Rioul, O.: A key to success - success exponents for side-channel distinguishers. In: Biryukov, A., Goyal, V. (eds.) Progress in Cryptology - INDOCRYPT 2015 - 16th International Conference on Cryptology in India, Bangalore, India, December 6-9, 2015, Proceedings. Lecture Notes in Computer Science, vol. 9462, pp. 270–290. Springer (2015). https://doi.org/10.1007/978-3-319-26617-6_15, https://doi.org/10.1007/978-3-319-26617-6_15
16. Gujarati, D.: Basic Econometrics. Economic series, McGraw Hill, fourth edn. (2003)
17. Hansen, B.E.: Econometrics (March 2021 version). University of Wisconsin, Department of Economics (2021), http://home.ustc.edu.cn/~matheming/Econometrics.pdf
18. Helleseth, T., Kløve, T., Mykkeltveit, J.: On the covering radius of binary codes (corresp.). IEEE Trans. Inf. Theory **24**, 627–628 (1978), https://api.semanticscholar.org/CorpusID:206733497
19. Heuser, A., Rioul, O., Guilley, S.: Good is not good enough: Deriving optimal distinguishers from communication theory (2014), https://eprint.iacr.org/2014/527, https://eprint.iacr.org/2014/527
20. Kocher, P., Jaffe, J., Jun, B.: Differential power analysis. In: Wiener, M. (ed.) Advances in Cryptology — CRYPTO' 99. pp. 388–397. Springer Berlin Heidelberg, Berlin, Heidelberg (1999)
21. Kocher, P.C.: Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems pp. 104–113 (1996), uRL: http://www.cryptography.com/timingattack/paper.html
22. Lomné, V., Prouff, E., Roche, T.: Behind the scene of side channel attacks. In: Sako, K., Sarkar, P. (eds.) Advances in Cryptology - ASIACRYPT 2013 - 19th International Conference on the Theory and Application of Cryptology and Information Security, Bengaluru, India, December 1-5, 2013, Proceedings, Part I. Lecture Notes in Computer Science, vol. 8269, pp. 506–525. Springer (2013). https://doi.org/10.1007/978-3-642-42033-7_26, https://doi.org/10.1007/978-3-642-42033-7_26
23. Long, J., Ou, C., Wang, Z., Zhang, F.: The mysteries of LRA: Roots and progresses in side-channel applications. Cryptology ePrint Archive, Paper 2024/1735 (2024), https://eprint.iacr.org/2024/1735
24. Mangard, S.: Hardware countermeasures against dpa–a statistical analysis of their effectiveness. In: Topics in Cryptology–CT-RSA 2004: The Cryptographers' Track at the RSA Conference 2004, San Francisco, CA, USA, February 23-27, 2004, Proceedings. pp. 222–235. Springer (2004)
25. Mangard, S., Oswald, E., Popp, T.: Power Analysis Attacks: Revealing the Secrets of Smart Cards. Springer, United States (2007), other identifier: 0387308571
26. Matsui, M.: Linear cryptanalysis method for des cipher. In: Helleseth, T. (ed.) Advances in Cryptology — EUROCRYPT '93. pp. 386–397. Springer Berlin Heidelberg, Berlin, Heidelberg (1994)
27. McCann, D., Oswald, E., Whitnall, C.: Towards practical tools for side channel aware software engineering: 'grey box' modelling for instruction leakages. In: Kirda, E., Ristenpart, T. (eds.) 26th USENIX Security Symposium, USENIX Security 2017, Vancouver, BC, Canada, August 16-18, 2017. pp. 199–216. USENIX Association (2017), https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/mccann
28. O'Connor, L.: Properties of linear approximation tables. In: Preneel, B. (ed.) Fast Software Encryption. pp. 131–136. Springer Berlin Heidelberg, Berlin, Heidelberg (1995)

29. Ouladj, M., Guilley, S., Prouff, E.: On the implementation efficiency of linear regression-based side-channel attacks. In: Bertoni, G.M., Regazzoni, F. (eds.) Constructive Side-Channel Analysis and Secure Design. pp. 147–172. Springer International Publishing, Cham (2021)
30. Picek, S., Heuser, A., Jovic, A., Bhasin, S., Regazzoni, F.: The curse of class imbalance and conflicting metrics with machine learning for side-channel evaluations. IACR Transactions on Cryptographic Hardware and Embedded Systems **2019**(1), 209–237 (Nov 2018). https://doi.org/10.13154/tches.v2019.i1.209-237, https://tches.iacr.org/index.php/TCHES/article/view/7339
31. Prouff, E., Renault, G., Rivain, M., O'Flynn, C.: Embedded Cryptography 1. ISTE Invoiced, Wiley (2025)
32. Prouff, E., Rivain, M.: Theoretical and practical aspects of mutual information based side channel analysis. In: Applied Cryptography and Network Security: 7th International Conference, ACNS 2009, Paris-Rocquencourt, France, June 2-5, 2009. Proceedings 7. pp. 499–518. Springer (2009)
33. Prouff, E., Rivain, M., Bevan, R.: Statistical analysis of second order differential power analysis. IEEE Transactions on computers **58**(6), 799–811 (2009)
34. Schindler, W., Lemke, K., Paar, C.: A stochastic model for differential side channel cryptanalysis. In: Rao, J.R., Sunar, B. (eds.) Cryptographic Hardware and Embedded Systems – CHES 2005. pp. 30–46. Springer Berlin Heidelberg, Berlin, Heidelberg (2005)
35. Standaert, F.X., Malkin, T.G., Yung, M.: A unified framework for the analysis of side-channel key recovery attacks. In: Advances in Cryptology-EUROCRYPT 2009: 28th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cologne, Germany, April 26-30, 2009. Proceedings 28. pp. 443–461. Springer (2009)
36. Telecom ParisTech, R.G.o.: Dpa contest. https://dpacontest.telecom-paris.fr/home/ (2008)
37. Whitnall, C., Oswald, E., Standaert, F.: The myth of generic dpa...and the magic of learning. In: Benaloh, J. (ed.) Topics in Cryptology - CT-RSA 2014 - The Cryptographer's Track at the RSA Conference 2014, San Francisco, CA, USA, February 25-28, 2014. Proceedings. Lecture Notes in Computer Science, vol. 8366, pp. 183–205. Springer (2014). https://doi.org/10.1007/978-3-319-04852-9_10, https://doi.org/10.1007/978-3-319-04852-9_10