HAL
open science

Intrinsic Resiliency of S-boxes Against Side-Channel Attacks -Best And Worst Scenarios

Claude Carlet, Eloi de Chérisey, Sylvain Guilley, Selçuk Kavut, Deng Tang

▶ **To cite this version:**

Claude Carlet, Eloi de Chérisey, Sylvain Guilley, Selçuk Kavut, Deng Tang. Intrinsic Resiliency of S-boxes Against Side-Channel Attacks -Best And Worst Scenarios. IEEE Transactions on Information Forensics and Security, 2021, 16, pp.203-218. 10.1109/TIFS.2020.3006399 . hal-02915674

HAL Id: hal-02915674

https://cnrs.hal.science/hal-02915674v1

Submitted on 15 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Intrinsic Resiliency of S-boxes Against Side-Channel Attacks – Best And Worst Scenarios

Claude Carlet, Éloi de Chérisey, Sylvain Guilley, Selçuk Kavut, and Deng Tang

***Abstract*—Constructing S-boxes that are inherently resistant against side-channel attacks is an important problem in cryptography. By using an optimal distinguisher under an additive Gaussian noise assumption, we clarify how a defender (resp., an attacker) can make side-channel attacks as difficult (resp., easy) as possible, in relation with the auto-correlation spectrum of Boolean functions. We then construct balanced Boolean functions that are optimal for each of these two scenarios. Generalizing the objectives for an S-box, we analyze the auto-correlation spectra of some well-known S-box constructions in dimensions at most 8 and compare their intrinsic resiliency against side-channel attacks. Finally, we perform several simulations of side-channel attacks against the aforementioned constructions, which confirm our theoretical approach.**

***Index Terms*—Substitution boxes (S-boxes), Cryptography, Side-channel analysis, Constructions.**

I. Introduction

S-BOXES are prominent targets for side-channel attacks, because they allow, from an attacker standpoint, to distinguish clearly between correct and incorrect hypotheses on key guesses. It has already been underlined in early papers [22], [47], [7], [23] that a notion of correlation for S-box coordinate functions relates to the side-channel efficiency.

Recently, the article [12] revisited from a mathematical point of view the link between S-box properties and side-channel attacks. However, the scope of this analysis is limited, since it targets a particular attack (namely the differential power analysis [28]) and a particular kind of attacked device (namely a hardware implementation with precharge logic which leaks in the Hamming weight model).

In this article, we generalize the analysis by leveraging on the *optimal* side-channel attack. The optimal side-channel attack consists in the maximum likelihood distinguisher, considering that the leakage model is known by the attacker. In particular, there is no issue of possible misinterpretation of the output with this distinguisher, as opposed to particular attacks, such as differential power analysis, where either the largest or the smallest bias (positive or negative peak) betrays the correct key. Therefore, in this paper, the criterion does not need to resort to absolute values (as is the case in [12]). Moreover, we aim to be independent of specific leakage models, hence we consider the simple *mono-bit* leakage model. As noticed in the seminal paper about side-channel attacks [28], the mono-bit leakage model allows for a direct connection between the target algorithmic properties of the S-box and the side-channel attack.

a) Contributions: In this paper, we show that, in the case of mono-bit side-channel attacks, the attack outcome is determined by the auto-correlation of the targeted S-box coordinates. This criterion is not usually considered when analyzing S-boxes. Therefore, we study both best and worst cases of S-boxes which optimize also the auto-correlation parameter. General constructions are studied, which are primarily optimizing the auto-correlation, considering the set of the other classical robustness metrics as a second improvement factor. As an interesting byproduct, the value of autocorrelation for Dobbertin's iterative construction is provided. Besides, some particular constructions, leveraging rotation-symmetric S-boxes, also reveal new S-boxes.

b) Outline: The rest of the paper is structured as follows. Section II provides the necessary mathematical tools useful for the subsequent analyses. The next section III explains how side-channel attacks relate to the auto-correlation function of the S-box coordinates. S-boxes taking into consideration the optimization of this new parameter are constructed in Sec. IV. Specific rotation-symmetric S-boxes are analyzed under the same prism in Sec. V (and some truth tables are listed in Appendix A. Practical evaluation using simulated side-channel is carried out in Sec. VI. The same section also lists open-issues not resolved in this paper. Eventually, Sec. VII concludes the paper.

II. Preliminaries

A. Mathematical Definition of S-boxes

We denote by $\mathbb{F}_2 = \{0,1\}$ the finite field with two elements; $\mathbb{F}_2^n$ is the n-dimensional vectorspace over $\mathbb{F}_2$. The (canonical) inner product over $\mathbb{F}_2^n$ is the $\mathbb{F}_2$-bilinear operation: $(a,b) \mapsto a \cdot b = \bigoplus_{i=1}^n a_i b_i$. A linear hyperplane is a vector subspace whose dimension is one less than that of its ambient space. In $\mathbb{F}_2^n$, the linear hyperplanes are the sets of equation $a \cdot x = 0$ where $a \neq 0$. An $n \times m$ S-box (or equivalently,

The work of S. Guilley was supported in part by the European Commission under H2020 grant TeamPlay no 779882.

The work of D. Tang was supported in part by the National Natural Science Foundation of China (grants 61872435 and 61602394).

C. Carlet is with the LAGA, University of Paris 8, France and University of Bergen, Norway (e-mail: claude.carlet@gmail.com).

É. Chérisey is with the TELECOM-Paris, Institut Polytechnique de Paris, Paris Cedex 13, France (e-mail: eloi.de.cherisey@pm.me).

S. Guilley is with the Secure-IC S.A.S. Rennes, France, and TELECOM-Paris, Institut Polytechnique de Paris, Paris Cedex 14, France (e-mail: sylvain.guilley@TELECOM-ParisTech.fr).

S. Kavut is with the Department of Computer Engineering, Balıkesir University, Balıkesir, Turkey (e-mail: skavut@balikesir.edu.tr).

D. Tang is with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China and the State Key Laboratory of Cryptology, P.O. Box 5159, Beijing, 100878, China (e-mail: dtang@foxmail.com).

an (n,m)-function) $F:\mathbb{F}_2^n \to \mathbb{F}_2^m$ can be considered as the parallelization of m Boolean functions sharing the same input:

$$f_i : \mathbb{F}_2^n \to \mathbb{F}_2, \text{ where } 1 \le i \le m,$$

so that $F(x) = (f_1(x), f_2(x), \ldots, f_m(x))$ for all $x \in \mathbb{F}_2^n$. The functions $(f_i)_{1\le i\le m}$ are called the coordinate functions of F, and their linear combinations $c \cdot F = \bigoplus_{i=1}^m c_i f_i$ with non-all-zero coefficient vectors $c = (c_1, c_2, \ldots, c_m) \in \mathbb{F}_2^{m*}$ are called the component functions of F.

For any $x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n$, we define a cyclic permutation ρ on x by $\rho(x) = (x_2, x_3, \ldots, x_{n-1}, x_1)$. Then an $n \times m$ S-box F is said to be rotation symmetric (RSSB) if $F(\rho(x)) = \rho(F(x))$ for all $x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n$. For instance, any power function $F(x) = x^d$ over $\mathbb{F}_{2^n}$ gives a rotation symmetric (n,n)-function when $\mathbb{F}_{2^n}$ is decomposed over a normal basis.

We denote the set of all n-variable Boolean functions by $\mathcal{B}_n$. Any Boolean function $f(x_1, x_2, \ldots, x_n) \in \mathcal{B}_n$ can be expressed uniquely in the form of a multivariate polynomial over $\mathbb{F}_2$, called its algebraic normal form (ANF):

$$\bigoplus_{u\in\mathbb{F}_2^n} a_u \left(\prod_{i=1}^n x_i^{u_i}\right),$$

where the coefficients a_u belong to $\mathbb{F}_2$. The summation variable u is a dummy variable running over the universe $\mathbb{F}_2^n$. The maximum Hamming weight of u with non-zero a_u is called the algebraic degree of f, which is denoted by $deg(f)$.

In some cases we shall identify $\mathbb{F}_{2^n}$ with the field $\mathbb{F}_{2^n}$ (after this field being an n-dimensional vector space over $\mathbb{F}_2$)

B. Cryptographic Properties of S-boxes

We now briefly review the basic definitions regarding the cryptographic properties of Boolean functions and extend them to S-boxes by using component functions.

Cryptographic functions must have high algebraic degree to achieve good confusion properties (the notion of confusion has been originally introduced by Shannon [51], as well as that of diffusion). The affine functions are those Boolean functions with algebraic degree at most 1. An affine function having constant term equal to zero is called a linear function.

The Walsh-Hadamard transform of an n-variable Boolean function f is the even integer-valued function W_f defined as

$$W_f : \mathbb{F}_2^n \to [-2^n, 2^n]\ ; \quad \omega \mapsto W_f(\omega) = \sum_{x\in\mathbb{F}_2^n} (-1)^{\omega\cdot x \oplus f(x)}.$$

We call f balanced if its Hamming weight is equal to 2^{n-1}, which is cryptographically desirable to avoid the statistical imbalance in the output of f. Notice that f is balanced if and only if $W_f(\mathbf{0}_n) = 0$ (we use $\mathbf{0}_n$ to denote the all-zero vector of length n).

For n even, bent functions are those Boolean functions achieving optimal Hamming distance $2^{n-1} - 2^{\frac{n}{2}-1}$ to the vector space of affine Boolean functions. Equivalently, they have their Walsh spectrum taking only the two values $\pm 2^{\frac{n}{2}}$. It is known that any bent function has algebraic degree at most $\frac{n}{2}$, see e.g. [8]. Semi-bent functions have, by definition, their Walsh spectrum taking the three values 0 and $\pm 2^{\frac{n}{2}+1}$. For n odd, semi-bent functions (or near-bent[1]since there are two names for the same notion) have their Walsh spectrum taking the three values 0 and $\pm 2^{\frac{n+1}{2}}$. The notions of bent and semi-bent functions extend to any S-boxes: such function F is bent (resp. semi-bent) if all its component functions $c \cdot F$, $c \neq 0$, are also bent (resp. semi-bent).

The nonlinearity of f is defined as the minimum Hamming distance between f and n-variable affine functions. It can be expressed in terms of the Walsh-Hadamard transform as follows:

$$NL_f = 2^{n-1} - \frac{1}{2}\max_{\omega\in\mathbb{F}_2^n} |W_f(\omega)|. \tag{1}$$

Boolean functions with high nonlinearity are required in a cryptosystem to resist linear cryptanalysis and to achieve good confusion properties.

- For even n, Boolean functions attaining the maximum nonlinearity of $2^{n-1} - 2^{\frac{n}{2}-1}$ are precisely the bent functions [16], [32], [38]. Bent functions are not balanced, however they can be used to construct balanced functions with high nonlinearity.
- For odd n, the nonlinearity value $2^{n-1} - 2^{\frac{n-1}{2}}$, which is attainable for any odd n, is known as the bent concatenation bound (the concatenation of two $(n-1)$-variable bent functions achieves this nonlinearity).

The auto-correlation function of f is given by:

$$r_f(d) = \sum_{x\in\mathbb{F}_2^n} (-1)^{f(x)\oplus f(x\oplus d)} = \sum_{x\in\mathbb{F}_2^n} (-1)^{D_d f(x)},$$

where $d \in \mathbb{F}_2^n$. It is clear that, for any n-variable Boolean function f, we have $r_f(\mathbf{0}) = \sum_{x\in\mathbb{F}_2^n}(-1)^0 = 2^n$. One can see that, for balanced Boolean functions, the sum of all the auto-correlation values is null. Indeed, we have (according to the so-called Wiener-Khintchine theorem):

$$\sum_{d\in\mathbb{F}_2^n} r_f(d)(-1)^{\omega\cdot d} = W_f^2(\omega)$$

for every ω (and we have the result by taking $\omega = \mathbf{0}_n$). There are two important cryptographic criteria called global avalanche characteristics (GAC) [57] related to the autocorrelation spectrum, which are used to quantify the level of diffusion ensured by a function. The maximum absolute value in the autocorrelation spectrum (except at the origin—this value is uninteresting because it does not depend on f) is referred to as the absolute indicator, denoted by

$$\Delta_f = \max_{d\in\mathbb{F}_2^{n*}} |r_f(d)|,$$

where $\mathbb{F}_2^{n*} = \mathbb{F}_2^n \setminus \{0\}$ and the other one is known as the sum-of-squares indicator, given by

$$\sigma_f = \sum_{d\in\mathbb{F}_2^n} r_f^2(d).$$

[1]We shall call "near-bent" the semi-bent functions in odd dimension, and keep the term "semi-bent" for when the parity of n will not be specified.

For the purpose of this paper, we define the *non-absolute indicator* as

$$\Gamma_f = \max_{d \in \mathbb{F}_2^{n*}} r_f(d),$$

which is used as a measure of side-channel resiliency (as we shall see in Section III-A). The higher the value of Γ_f, the better resistance against side-channel attacks.

Let us now consider the case of S-boxes. The nonlinearity and absolute indicator of an S-box are determined by the component function(s) having the worst measure. In other words, the nonlinearity (resp., the absolute indicator) equals the lowest (resp., the highest) nonlinearity (resp., absolute indicator) of all the component functions of the S-box. A nonlinearity is considered as good if it is not too far from the optimum, which is $2^{n-1} - 2^{\frac{n-1}{2}}$ for (n,n)-functions, according to the Sidelnikov-Chabaud-Vaudenay bound (see e.g. [9]). The algebraic degree of an S-box is defined as the maximum algebraic degree of the coordinate functions and it is also the maximum algebraic degree of the component functions.

An $n \times m$ S-box F is called differentially δ-uniform [45] if the equation $F(x) \oplus F(x \oplus \gamma) = \beta$ has at most δ solutions for all $\gamma \in \mathbb{F}_2^{n*}$ and $\beta \in \mathbb{F}_2^m$. Accordingly, δ is called the differential uniformity of F. The values of δ are always even since if x is a solution of equation $F(x) \oplus F(x \oplus \gamma) = \beta$ then $x \oplus \gamma$ is also a solution. This implies that the smallest possible value of δ for an (n,m)-functions is 2; the functions achieving this value are called *almost perfect nonlinear* (APN). A cryptographically desirable S-box is required to have low differential uniformity ($\delta = 2$ is optimal, $\delta = 4$ is good), which makes the probability of occurrence of a particular pair of input and output differences (γ, β) low, and hence provides resistance against differential cryptanalysis.

Given two (n,m)-functions G and H, we say that they are affine equivalent if $G(x) = A_1(H(A_2(x)))$, where A_1 is an affine permutation on $\mathbb{F}_2^m$ and A_2 is an affine permutation on $\mathbb{F}_2^n$. It is known that the nonlinearity, algebraic degree, and differential uniformity are invariant under affine equivalence.

III. Side-Channel Problem Statement

A. Monobit Case

Let $t, k \in \mathbb{F}_2^n$ respectively be a plaintext and a key used in a cryptographic algorithm, such as a block cipher, which starts by a key addition (`xor` operation) followed by a confusion function (an S-box F). In this section, the attacker targets one bit of the S-box, that is the output of $f : \mathbb{F}_2^n \to \mathbb{F}_2$, applied to $t \oplus k$. So here, f is a balanced Boolean function (e.g., a coordinate function of the S-box). In realistic scenarios, the measurements are noisy; hence, not only one but several of them need to be captured by the attacker, so as to make a statistical attack. We resort to vectorial notation, where measurements $\vec{x} = (x_1, \ldots, x_Q)$ consist of a collection of Q queries. Let us denote the correct key by k^*. It is unknown and shall be guessed by exhaustive search over all keys $k \in \mathbb{F}_2^n$. The observable leakage is thus $\vec{x} = \vec{y}(k^*) + \vec{n}$, where the model is $\vec{y}(k) = f(\vec{t} \oplus k)$, that is $\vec{y}(k) = (f(t_1 \oplus k), \ldots, f(t_Q \oplus k))$. In case the measurements feature additive Gaussian noise, $\vec{n} \sim \mathcal{N}(\vec{0}, \Sigma)$ is a Gaussian noise, with $\Sigma = \sigma^2 \mathsf{Id}_{Q \times Q}$, where

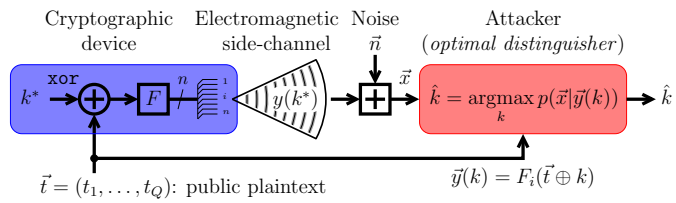

Fig. 1. Attack setup on the leakage function $y = f(t \oplus k^*)$, where t is one known plaintext, k^* is the secret key, F is the substitution box and i is the leaking coordinate of F (hence $f = F_i$)

$\mathsf{Id}_{Q \times Q}$ is the $Q \times Q$ identity matrix. Each value x_q (where $1 \le q \le Q$) is also called a *trace*, and the observable leakage $\vec{x}$ is altogether referred to as the traces *acquisition campaign*.

The adversary optimizes its probability of success to recover the correct key thanks to the *optimal distinguisher* [24]. It consists in guessing the correct key k^* with the maximum likelihood rule

$$\hat{k} = \operatorname{argmax}_k p(\vec{x}|\vec{y}(k)). \tag{2}$$

The setup we consider is depicted in Fig. 1, where "vectorial values" are represented as fat arrows, whereas "single bits" are represented as thin wires. In this figure, the attack target is the S-box coordinate i (where $1 \le i \le n$). We have that $p(\vec{x}|\vec{y}(k))$ is equal to:

$$\begin{aligned} &\frac{1}{(2\pi|\Sigma|)^{\frac{D}{2}}} \exp\left[-\frac{1}{2}(\vec{x} - \vec{y}(k))^\top \Sigma^{-1} (\vec{x} - \vec{y}(k))\right] \qquad (3) \\ = \ &\text{constant} \times \exp\left[-\frac{1}{2\sigma^2} \sum_{q=1}^{Q} (x_q - y(t_q, k))^2\right], \end{aligned}$$

hence the attacker aims at minimizing

$$\frac{1}{Q} \sum_{q=1}^{Q} (x_q - y(t_q, k))^2, \tag{4}$$

which, by the law of large numbers when $Q \to +\infty$, tends to:

$$\begin{aligned} \mathbb{E}(X - Y(T,k))^2 &= \mathbb{E}(Y(T,k^*) + N - Y(T,k))^2 \\ &= \mathbb{E}(f(T \oplus k^*) - f(T \oplus k))^2 + \sigma^2. \end{aligned}$$

Notice that in (3), the notation $z^\top$ stands for *transposition* of column z (hence $z^\top$ is a row), and that Σ^{-1} stands for the inverse of Σ, namely $\Sigma^{-1} = \frac{1}{\sigma^2} \mathsf{Id}_{Q \times Q}$.

So, assuming that the plaintexts T are uniformly distributed over $\mathbb{F}_2^n$ (which is a fair assumption in cryptography), the attack is equivalent to minimizing over all k the value:

$$\sum_{t \in \mathbb{F}_2^n} (f(t \oplus k^*) - f(t \oplus k))^2. \tag{5}$$

This quantity is classical in cryptography, namely:

$$\sum_t (f(t \oplus k^*) - f(t \oplus k))^2 = 2^{n+2} \kappa(k, k^*), \tag{6}$$

where $\kappa(k, k^*)$ bears the name of *confusion coefficient* [19]. Now, we know that difference square $(f(t \oplus k^*) - f(t \oplus k))^2$ equals

$$\begin{aligned} & \left(\frac{1}{2}\left(1-(-1)^{f(t\oplus k^*)}\right) - \frac{1}{2}\left(1-(-1)^{f(t\oplus k)}\right)\right)^2 \\ = & \frac{1}{4}\left((-1)^{f(t\oplus k)} - (-1)^{f(t\oplus k^*)}\right)^2 \\ = & \frac{1}{2}\left(1-(-1)^{f(t\oplus k^*)\oplus f(t\oplus k)}\right). \end{aligned}$$

Hence, minimizing (5) amounts to maximizing

$$\sum_t (-1)^{f(t\oplus k^*)\oplus f(t\oplus k)}. \tag{7}$$

This value which depends only on $d = k^* \oplus k$ is maximized when $d = 0$. For the sake of clarity, the value in Eqn. (7) is also referred to as the *autocorrelation* of f at input difference d and is denoted as $r_f(d)$.

To increase the success of the attack, one aims at having the nearest rival[2] ($k \neq k^*$) be as far as possible from the correct guess. Hence the goal is to make as small as possible the maximum value within $\{\sum_t (-1)^{f(t)\oplus f(t\oplus d)}, d \neq 0\}$. Notice that this objective holds for comparisons between S-boxes of the same input bitwidth (called n). Otherwise, when S-boxes f and f' have different numbers of input bits ($n \neq n'$), the comparison would hold on $\max_{d\neq 0} 2^n - r_f(d)$ vs $\max_{d'\neq 0} 2^{n'} - r_{f'}(d')$, or alternatively $\max_{d\neq 0} \frac{1}{2^n} r_f(d)$ vs $\max_{d'\neq 0} \frac{1}{2^{n'}} r_{f'}(d')$. We do not consider comparison between S-boxes of different sizes in the sequel.

Also notice that when $r_f(d) = 2^n$ for a nonzero d, then the possible keys are the correct key k^* or the challenger $k^* \oplus d$. For example, the least significant bit f of the PRESENT [3] S-box F features such a tie because one has

$$\forall z \in \mathbb{F}_2^4, \quad f(z) = f(z \oplus \texttt{0x9}),$$

where $\texttt{0x9}$ (in hexadecimal) represents $(1001)_2$ in binary notation. Hence, it is possible to distinguish by side-channel analysis only between pairs of key candidates k^* and $k^* \oplus \texttt{0x9}$.

B. Multi-bit S-boxes

1) Attacks on coordinate functions: We now consider that the S-box is vectorial ($m > 1$). we need that the S-box be balanced (i.e. with uniformly distributed output). As explained in e.g. [9], one simply requires that the S-box number of output bits be bounded above by its number of input bits. A special case is that the S-box is a permutation of $\mathbb{F}_2^n$.

The situation of S-boxes regarding side-channel analysis is depicted in Fig. 2 for 2 coordinates (out of $m = n$). This figure represents the auto-correlation $r_f(d)$ as a function of the difference $d = k \oplus k^* \in \mathbb{F}_2^n$ between candidate key k and actual secret key k^*. The origin value is $\mathbf{0}_n$ and other values on the abscissa axis represent the vectorspace $\mathbb{F}_2^n$. The left-hand side graph represents the situation of a coordinate f where the nearest rival $d_{\max}$ (relative to the correct key k^*) features an auto-correlation value close to $2^n = r_f(\mathbf{0})$.

[2]The *nearest rival*, in the context of side-channel distinguishers, is a term coined by Whitnall and Oswald, e.g., in [55].

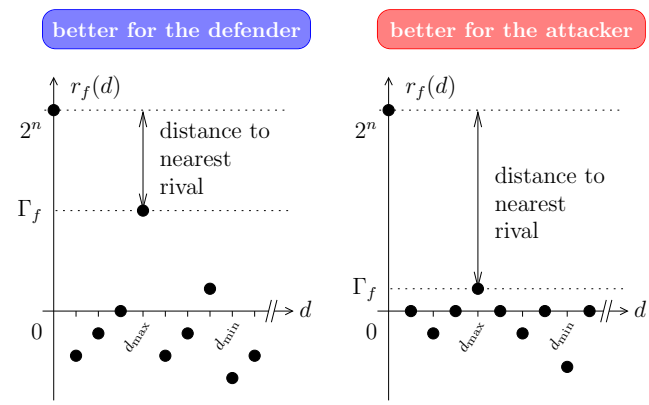

Fig. 2. Illustration of two side-channel situations, optimal for the defender (**in blue**) and for the attacker (**in red**)

This configuration favors the defender, as the attacker has hard time distinguishing between $d = 0$ and $d = d_{\max}$ (recall that in real side-channel, some noise blurs the values of the auto-correlations). At the opposite, the right-hand side graph highlights the situation of another coordinate for which maximum auto-correlation over incorrect key differences is inferior to the former case. The attacker can distinguish more clearly between the correct and the nearest rival key.

The relevant metrics are then deduced from the following analysis:

- From the attacker point of view, the attacker chooses the coordinate which is the most favorable for his key guess in the presence of noise. Hence, the attacker would like (if it was possible for him to devise an S-box) to select a coordinate i, $1 \leq i \leq n$, which increases the "*distance to the nearest rival*", i.e., his objective is to minimize $\min_{1\leq i\leq n} \max_{d\neq 0} r_{F_i}(d) = \min_{1\leq i\leq n} \Gamma_{F_i}$.
- From the defender (or designer) point of view, the goal is to avoid any weak coordinate in the S-box, because it is clear that it is the one which would be targeted by the most powerful attacker. Hence the objective is to maximize $\min_{1\leq i\leq n} \Gamma_{F_i}$.

Remark 1. *[Signedness of $r_f(d)$] The important parameter is $\max_{d\neq 0} r_f(d)$ and not $\max_{d\neq 0} |r_f(d)|$. Indeed, the criterion for the side-channel attacker to succeed* hardly *(resp.* easily*) is that the auto-correlation of the nearest rival is* close to *(resp.* far from*) 2^n. For example, in the conceptual figure 2, the largest value of r_f (for $d \neq 0$) occurs at $d = d_{max} = \operatorname{argmax}_{d\neq 0} r_f(d)$ whilst its smallest value occurs at $d = d_{min} = \operatorname{argmin}_{d\neq 0} r_f(d)$. It can be seen that in the left case, $\max_{d\neq 0} |r_f(d)|$ is same as $r_f(d_{max})$. However, in the right case, $\max_{d\neq 0} |r_f(d)|$ is same as $|r_f(d_{min})|$.*

Therefore, r_f shall be considered without absolute values. For the sake of illustration, in the DES block cipher (NIST FIPS PUB 46-3, which features $n = 6$ and $m = 4$), the extreme values for $r_f(d), d \neq 0$ in S-box 2 (denoted as F) are given in Table I. It is easy to see that values of $\max_{d\neq 0} r_f(d)$ would all be incorrect by considering $|r_f|$ instead of r_f (for

TABLE I. Extremal values of $r_f(d), d \neq 0$, where f is any component function of the DES second S-box $F : \mathbb{F}_2^6 \to \mathbb{F}_2^4$

	$f = F_1$	$f = F_2$	$f = F_3$	$f = F_4$
$\max_{d\neq 0} r_f(d)$	32	24	40	32
$\mathrm{argmax}_{d\neq 0} r_f(d)$	13	11	32	19
$\min_{d\neq 0} r_f(d)$	-40	-32	-56	-40
$\mathrm{argmin}_{d\neq 0} r_f(d)$	5	7	31	53

$f \in \{F_1, F_2, F_3, F_4\}$). Indeed, for this S-box, we have that

$$|\min_{d\neq 0} r_f(d)| > \max_{d\neq 0} r_f(d).$$

Therefore, in the remainder of the paper, we never study $|r_f|$ but stick to r_f (signed).

2) Attacks on component functions: In the previous section III-B1, we argued that, often, attackers base their guess on single bits. However, these bits can be extended from *coordinate* to *component* functions of the S-box[3]. We recall that a component function is a linear combination (with coefficients in $\mathbb{F}_2$) of all the coordinate functions. Let $c \in \mathbb{F}_2^m$. A component function of an S-box $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is $f : \mathbb{F}_2^n \to \mathbb{F}_2$, $t \mapsto f(t) = c \cdot F(t)$, where " $\cdot$ " is the canonical scalar product in $\mathbb{F}_2^m$. Indeed, modern block ciphers consist in the iterative alternation of a confusion layer (e.g., made up of S-boxes) and a diffusion layer (e.g., a linear mapping). For example, in substitution-permutation networks (such as the AES), the S-box is fed into a linear bijection (e.g., MixColumns in the case of AES) computing linear combination of bits. All of those bits leak their values through a side-channel, hence it is safe to imagine that an attacker will combine bits (in $\mathbb{F}_2$) to find the most favorable linear combination, as might show up in the diffusion layer (such as MixColumns). Notice that MixColumns is made up of `xor` gates (additions in $\mathbb{F}_2$), which are known to be very glitchy. Now glitches do contribute significantly to the overall leakage of the cryptographic function (see attack [35], [31], defense [20], [42], [43], and analysis [2] papers).

Hence, we pursue in the sequel of this article the following goals:

- From the attacker point of view: minimize $\min_{c\neq 0} \max_{d\neq 0} r_{c\cdot F}(d)$.
- From the defender (or designer) point of view: maximize $\min_{c\neq 0} \max_{d\neq 0} r_{c\cdot F}(d)$.

C. Positioning of our Work with respect to the State-of-the-Art

The theoretical study of side-channel analysis allows to grasp the impact of several factors on the outcome of attacks. Historically, Whitnall and Oswald [56] suggested the distance to nearest rival for distinguishers, which they studied in different scenarios (noise, model discrepancy with respect to actual side-channel, etc.). However, their criterion has consistency issues, because it is not invariant by the scaling of the distinguisher. Therefore, it fails to be fair when comparing distinguishers of different kinds. Such flaws relative to the unfairness of the attack outcome predictability were reported for instance in [49]. An analysis based on success probability (the focus of which is not on the distinguisher value but on the attack outcome) is proposed in [21]. The analysis reveals that the relevant parameter, called *success exponent*, is a normalized quantity of the asymptotical distinguisher. In [14], it is analyzed that only two factors impact the success exponent, namely the confusion coefficient and the noise variance. The confusion coefficient gathers cryptographic properties (typically, of the S-box) and the leakage model (non-injective function, such as the Hamming weight or the UWSB).

In our paper, we consider a mono-bit leakage model, so that we focus only on the impact of the S-box. Therefore, we assume that the attacker targets one bit of the S-box output, and selects it so as to maximize his advantage (that is: improve the distance of the autocorrelation regarding the correct key guess to its nearest rival). The motivation to select a bit at the output of the S-box arises from the goal of best distinguishing the correct key from others, which (as already mentioned) is also captured by the notion of *confusion coefficient* κ (recall (6)). It is now well known that the confusion coefficient is favorable to the attacker at the output of S-boxes [22], [47], [7], [23], [46], [10], [12].

However, it is unclear how to use output bits to devise an attack. Historically, Kocher introduced side-channel distinguishers targeting one bit [28], in attacks now referred to as "difference of means": one bit of the S-box output is selected, side-channel traces are partitioned in two groups according to this bit, and the difference of means in each partition constitutes the distinguisher. Later it has been noticed that usual devices leak all the bits at once, because processors or application specific circuits manipulate words (e.g., bytes when $n = 8$). Since the actual *leakage function* is hard to characterize, the assumption is often made that the leakage is the sum (in $\mathbb{Z}$) of the bits. This yields the so-called Hamming weight leakage model, as analyzed typically in [4]. Still, in practice, all the bits in a register do not have the same leaking characteristics. For instance, the LSB (least significant bit) can receive or not an input carry when performing arithmetic computations. For this reason, a refined Unevenly Weighted Sum of Bits (UWSB) model has been introduced [54], [58], [24]. It can be noticed that for imperfect masking schemes (see e.g., low-entropy masking schemes such as Rotating Substitution-box Masking, also known as RSM [41]), the effect of masking can be to have a leakage model which is a UWSB (see equation (4) of [39]). In the first article about the confusion coefficient [19], the performance of a distinguisher was based on the computation of true/false positive/negative matrix (also known as a confusion matrix), based on a binary outcome of the prediction. Later on, the confusion coefficient has been extended to real-valued leakage models [21], such as the UWSB model. The coefficients in this combination (in $\mathbb{R}$, now; indeed, they model as accurately as possible the physical leakage arising from analog logic implementing the cryptographic computation under analysis) are unknown, and possibly of opposite signs. Therefore it is still considered

[3]Recall that paper [12] handles multi-bit differential power analysis, but simply assuming that the leakage is impacted by the coordinates alone, and not the component functions of the S-box.

a safe practice to attack on a bit-by-bit basis: the $(n-1)$ remaining bits are considered unknown, and thus *de facto* integrated amongst the noise sources.

Constructively combining all the n bits requires profiling. In side-channel analysis, this method is referred to as *stochastic attacks* [52]. However, building a model requires many training traces. Depending on the operational constraints, this training set might not be available, since the attacker needs an open copy of the device he can manipulate freely to generate traces of his choice. Besides, owing to miniaturization of silicon technologies, the dispersion increases, making two instances of the same device fairly different. This is studied in papers about *template attacks* (profiling on one device and attacking on another) [18], [40]. Because of those limitations and the complexity of the learning stage (and of resulting estimation errors, creating so-called *epistemic* noise), many practical attacks remain based on mono-bit models.

IV. Construction of Optimal Boolean Functions

Recall that the non-absolute indicator Γ_f has been defined in Subsection II-B. In this section, we first show that $\Gamma_f < 0$ is impossible for $n > 1$ and that the values of the Walsh transform of a balanced Boolean function f such that $\Gamma_f = 0$ (i.e. which is optimal in terms of the objective of an attacker) all belong to a set that we determine. We deduce that there is no 4-variable or 6-variable balanced Boolean function with $\Gamma_f = 0$. We also deduce that the minimum possible nonlinearity of f with an odd number n of variables and such that $\Gamma_f = 0$ is $2^{n-1} - 2^{\frac{n-1}{2}}$. We compute the possible nonlinearities of those functions satisfying $\Gamma_f = 0$ for even $n \leq 16$. This shows that the nonlinearity requirement while allowing $\Gamma_f = 0$ is less demanding while increasing n. Secondly, we construct balanced Boolean functions with $\Gamma_f = 0$ (resp. $\Gamma_f = 2^n$) for the case of odd (resp. even) number of variables. Further, we present a construction, obtained by modifying the class of Maiorana-McFarland (M-M) bent functions [16], [32] and employing the balanced Boolean functions generated by Dobbertin's iterative construction [17], whose auto-correlation spectrum is completely characterized (hence, it can be utilized by search algorithms to construct optimal Boolean functions for best and worst scenarios).

A. Impossibility of having $\Gamma_f < 0$ for $n > 1$

We start by recalling the following lemma, which is subsequently used to prove that no Boolean function f exists with a number of variables greater than 1 and such that $\Gamma_f < 0$.

Lemma 1 [13]. *Let $f \in \mathcal{B}_n$, where $n > 1$. Then*

$$r_f(d) \equiv 0 \pmod 4 \text{ for any } d \in \mathbb{F}_2^n.$$

Moreover, if f is balanced, then $r_f(d)$ is a multiple of 8.

Theorem 1. *There is no Boolean function $f \in \mathcal{B}_n$ (where $n > 1$) with $\Gamma_f < 0$.*

Proof. We have already recalled the Wiener-Khintchine theorem, whose statement is that the Fourier transform of $r_f(d)$ coincides with the squared Walsh transform of f, i.e.,

$$W_f^2(a) = \sum_{d \in \mathbb{F}_2^n} r_f(d)(-1)^{a \cdot d} \text{ for all } a \in \mathbb{F}_2^n.$$

Hence, substituting $a = \mathbf{0}_n$ into this equation, we have $\sum_{d \in \mathbb{F}_2^n} r_f(d) \geq 0$. Suppose there exists f with $\Gamma_f < 0$. From Lemma 1, it is then clear that

$$\sum_{d \in \mathbb{F}_2^n} r_f(d) = 2^n + \sum_{d \in \mathbb{F}_2^{n*}} r_f(d) \leq 2^n - 4(2^n - 1) < 0,$$

which is a contradiction. □

B. Relating Γ_f with Nonlinearity

We now study the possible values of the Walsh transform of a balanced function such that $\Gamma_f = 0$.

Theorem 2. *Let $n > 3$ and f be a balanced n-variable Boolean function such that $\Gamma_f = 0$ (i.e. having only non-positive auto-correlation values, except for the one at all-zero point). Then all the values of the Walsh transform of f belong to the set*

$$S_n = \{\omega \in 4\mathbb{Z};\ \exists k \in \{0, 1, \ldots, 2^{n-4}\};\ \omega^2 = 2^n \pm 16k\}.$$

Proof. Since f is balanced, we have

$$\sum_{d \in \mathbb{F}_2^n} r_f(d) = 0, \tag{8}$$

where $r_f(d)$ is the auto-correlation function of f.

Let $D = \{d \in \mathbb{F}_2^{n*} \mid r_f(d) \neq 0\}$ and M be the multi-set of all the elements of D, each of which with multiplicity $\frac{|r_f(d)|}{8}$. Then, as $r_f(d) \leq 0$ for every $d \in \mathbb{F}_2^{n*}$, the left-hand side of the above sum can be rewritten as follows:

$$2^n + \sum_{d \in D} r_f(d) = 2^n - 8|M| = 0,$$

and so, $|M| = 2^{n-3}$.
On the other hand, by Fourier transform on the auto-correlation function, we have:

$$W_f^2(a) \quad = \quad \sum_{d \in \mathbb{F}_2^n} r_f(d)(-1)^{a \cdot d} = 2^n - 8 \sum_{d \in M} (-1)^{a \cdot d}.$$

Since $|M|$ is even, we have $\sum_{d \in M} (-1)^{a \cdot d} \equiv 0 \pmod 2$ and the proof is complete. □

As a consequence, we have:

Corollary 1. *For $n = 4$ and $n = 6$, there is no n-variable balanced Boolean function f with $\Gamma_f = 0$.*

Proof. We find that $S_4 = \{0, \pm 4\}$ and $S_6 = \{0, \pm 4, \pm 8\}$.

So, the minimum nonlinearity can be $2^{4-1} - \frac{4}{2} = 6$ and $2^{6-1} - \frac{8}{2} = 28$ for $n = 4$ and 6, respectively. However, these are the nonlinearities of the bent functions for both cases, which cannot be attained by balanced functions. □

For both $n = 4$ and 6, it is easy to find by a computer search that the minimum achievable value of Γ_f for a balanced n-variable Boolean function f is equal to 8. For $n = 4$,

an exhaustive search yields that there exist 12000 balanced functions with Γ_f=8. For $n = 6$, we have performed a heuristic search and found that there exist balanced functions with $\Gamma_f = 8$. Note that by Theorem 1 and Corollary 1, we have that $\Gamma_f > 0$ for 4- and 6-variable balanced functions. Then, as a consequence of Lemma 1, it is clear that Γ_f can be at least 8 for these functions, which is achieved by our search results. But our heuristic search could not find an 8-variable balanced function with $\Gamma_f = 8$. Note that in that case, there exist in the literature a few examples of balanced Boolean functions with absolute indicator $\Delta_f = 16$. An example is given in [26] as follows in hexadecimal:

```
18CA9ED8BC4EC1AFE2F4C023FA63E789 \
49455BC59DB873BE79409BAE4B289029
```
(9)

and we know then that the minimum of Γ_f can be at most 16 for $n = 8$. It is difficult to have more precise information but a little more insight can be obtained through the study of the nonlinearity. In Theorem 2, the maximum value of ω equals $4 \cdot \lfloor 2^{\frac{n-3}{2}} \rfloor$. This implies that any balanced Boolean function $f \in \mathcal{B}_n$ with $\Gamma_f = 0$ has nonlinearity at least $2^{n-1} - 2 \cdot \lfloor 2^{\frac{n-3}{2}} \rfloor$. In Table II, for n even between 8 and 16, we have displayed this value $2^{n-1} - 2 \cdot \lfloor 2^{\frac{n-3}{2}} \rfloor$ and the value that Dobbertin obtained in [17] as the nonlinearity of a balanced function that he constructed with his iterative construction[4] (this latter value was conjectured by him as the best possible nonlinearity of any balanced function). Notice that the existence question for 8-variables balanced Boolean functions with nonlinearity 118 is still open. (A negative answer to this question, that is, a positive answer to Dobbertin's conjecture for $n = 8$, would also rule out the possibility of having balanced functions of 8 variables with $\Gamma_f = 0$: such functions would not exist).

TABLE II. Comparison of the minimum possible nonlinearities required to have $\Gamma_f = 0$ with the maximum nonlinearities conjectured by Dobbertin [17] for n-variable balanced functions.

n	$2^{n-1} - 2 \cdot \lfloor 2^{\frac{n-3}{2}} \rfloor$	Dobbertin's conjecture [17]
8	118	116
10	490	492
12	2004	2010
14	8102	8120
16	32588	32628

For n odd, we shall see that functions with $\Gamma_f = 0$ exist. The maximum value of ω in Theorem 2 equals $\sqrt{2^n + 2^n}$ (i.e., $2^{\frac{n+1}{2}}$) and we have then:

Corollary 2. *For odd $n > 3$, the nonlinearity of any n-variable balanced Boolean function f such that $\Gamma_f = 0$ is bounded below by $2^{n-1} - 2^{\frac{n-1}{2}}$ (i.e. by the value of the bent concatenation bound).*

The bent concatenation bound is here a lower bound. Thanks to Construction 1, we shall be able to design, for every odd $n > 3$, functions having $\Gamma_f = 0$ and nonlinearity equal to $2^{n-1} - 2^{\frac{n-1}{2}}$. We present these functions in Subsection IV-C below, which is devoted to constructions of functions for the attacker. There may also exist functions with strictly better nonlinearity, but we could not find any. We leave this as an open problem.

[4]See Subsection IV-D1.

C. Constructions of Boolean and vectorial functions for the attacker

1) Boolean functions: Recall that the lower is Γ_f, the better it is for the attacker.

For small values of n: we have seen that for $n = 4$ and $n = 6$, there is no n-variable balanced Boolean function f with $\Gamma_f = 0$. By performing an exhaustive search for Boolean functions in 3 and 5 variables, it can be found that the number of balanced functions with $\Gamma_f = 0$ (i.e. which are optimal in terms of the objective of an attacker) is 56 and 27776, respectively. We have computationally checked for each case that the functions are in fact affine equivalent (and represent then one function, only, up to equivalence).

Constructions for odd n: Let us recall the class of Maiorana-McFarland (M-M) functions (introduced originally for designing bent functions, see [16], [32], and later extended to other kinds of functions in [6]), and which is defined as:

$$f(x, y) = x \cdot \phi(y) \oplus g(y)$$

where $x \in \mathbb{F}_2^s$, $y \in \mathbb{F}_2^k$, ϕ is any mapping from $\mathbb{F}_2^s$ to $\mathbb{F}_2^k$, and g is an arbitrary Boolean function with k variables (taking $s = k$ and ϕ as an arbitrary permutation on $\mathbb{F}_2^k$ results in the rather large class of bent functions discovered independently by Maiorana and McFarland [16], [32]).

Construction 1. *Let $n \geq 3$ be odd. For every mapping*

$$\phi : \mathbb{F}_2^{\frac{n-1}{2}} \mapsto \mathbb{F}_2^{\frac{n+1}{2}}$$

injective whose image set is the complement of a linear hyperplane, and every $\frac{n-1}{2}$-variable Boolean function g, we define the n-variable Boolean function f (in the M-M class) as $f(x, y) = x \cdot \phi(y) \oplus g(y)$, where $x \in \mathbb{F}_2^{\frac{n+1}{2}}$, $y \in \mathbb{F}_2^{\frac{n-1}{2}}$.

Proposition 1. *Let f be any function obtained by construction 1. Assume that $E = \{\mathbf{0}_{\frac{n+1}{2}}, \omega\}^\perp$ is the linear hyperplane equal to the complement of the image set of ϕ. Then f is balanced, near-bent and is such that:*

$$r_f(d) = \begin{cases} 2^n, & \text{if } d = \mathbf{0}_n \\ -2^n, & \text{if } d = (\omega, \mathbf{0}_{\frac{n-1}{2}}) \\ 0, & \text{if } d = \mathbb{F}_2^{n*} \setminus (\omega, \mathbf{0}_{\frac{n-1}{2}}) \end{cases}.$$

Hence, $\Gamma_f = 0$.

Proof. We have $r_f(\mathbf{0}_n) = 2^n$ as for any n-variable function. From the definition of the Walsh-Hadamard transform, we have for every $u \in \mathbb{F}_2^{\frac{n+1}{2}}$ and every $v \in \mathbb{F}_2^{\frac{n-1}{2}}$:

$$W_f(u, v) = 2^{\frac{n+1}{2}} \sum_{y \in \phi^{-1}(u)} (-1)^{g(y) \oplus v \cdot y} \in \{0, \pm 2^{\frac{n+1}{2}}\}. \quad (10)$$

Hence f is near-bent. Since $\phi^{-1}(\mathbf{0}_{\frac{n+1}{2}})$ is empty, we have $W_f(\mathbf{0}_{\frac{n+1}{2}}, \mathbf{0}_{\frac{n-1}{2}}) = 0$ and f is balanced.

For every $u \in \mathbb{F}_2^{\frac{n+1}{2}}$ and $v \in \mathbb{F}_2^{\frac{n-1}{2}}$, we have

$$f(x,y) \oplus f(x \oplus u, y \oplus v) = \\ x \cdot (\phi(y) \oplus \phi(y \oplus v)) \oplus u \cdot \phi(y \oplus v) \oplus g(y) \oplus g(y \oplus v).$$

If $v \neq \mathbf{0}$, we have then that $f(x,y) \oplus f(x \oplus u, y \oplus v)$ is balanced since ϕ being injective, we have $\phi(y) \oplus \phi(y \oplus v) \neq \mathbf{0}$ for every y, and hence $r_f(u,v) = 0$.

If $v = \mathbf{0}$, we have then:

$$f(x,y) \oplus f(x \oplus u, y) = u \cdot \phi(y) \text{ and} \\ r_f(u, \mathbf{0}) = 2^{\frac{n+1}{2}} \sum_{y \in \mathbb{F}_2^{\frac{n-1}{2}}} (-1)^{u \cdot \phi(y)}.$$

The value set of $\phi(y)$ being the complement of the linear hyperplane $\{\mathbf{0}_{\frac{n+1}{2}}, \omega\}^{\perp}$, we have $\sum_{y \in \mathbb{F}_2^{\frac{n-1}{2}}} (-1)^{u \cdot \phi(y)} = 0$ if $u \neq \omega$ and $\sum_{y \in \mathbb{F}_2^{\frac{n-1}{2}}} (-1)^{u \cdot \phi(y)} = -2^{\frac{n-1}{2}}$ if $u = \omega$, since we have

$$\sum_{z \in E} (-1)^{u \cdot \phi(y)} = 2^{\frac{n-1}{2}} \text{ and } \sum_{z \in \mathbb{F}_2^{\frac{n+1}{2}}} (-1)^{u \cdot z} = -2^{\frac{n-1}{2}}.$$

This completes the proof. □

There are $2^{\frac{n+1}{2}} - 1$ distinct linear hyperplanes in $\mathbb{F}_2^{\frac{n+1}{2}}$. For each of them there are $2^{\frac{n-1}{2}}!$ distinct mappings ϕ. The number of functions g is $2^{2^{\frac{n-1}{2}}}$. Let us prove that the different choices of a hyperplane, a mapping ϕ and a function g provide distinct functions; this will show that the number of functions generated by Construction 1 is $(2^{\frac{n+1}{2}} - 1)(2^{\frac{n-1}{2}}!)2^{2^{\frac{n-1}{2}}}$. Suppose that

$$\begin{array}{rcl} f(x,y) & = & x \cdot \phi(y) \oplus g(y) \\ f'(x,y) & = & x \cdot \phi'(y) \oplus g'(y) \end{array} \quad \text{and}$$

are the same function, that is, that we have

$$x \cdot (\phi(y) \oplus \phi'(y)) = g(y) \oplus g'(y)$$

for every x, y. Suppose first that $g(y) \neq g'(y)$ for some $y \in \mathbb{F}_2^{\frac{n-1}{2}}$, then $x \cdot (\phi(y) \oplus \phi'(y)) = 1$ for all $x \in \mathbb{F}_2^{\frac{n+1}{2}}$ and this is impossible. We deduce that $g(y) = g'(y)$ for all $y \in \mathbb{F}_2^{\frac{n-1}{2}}$, and $x \cdot (\phi(y) \oplus \phi'(y)) = 0$ for all $x \in \mathbb{F}_2^{\frac{n+1}{2}}$, that is, $\phi(y) = \phi'(y)$. Hence, $\phi = \phi'$ and $g = g'$. This completes the proof.

We have computed that for $n = 7$, Construction 1 generates 154828800 functions among which 2580480 are of degree 2 and 152248320 are of degree 3.

For any function f generated by Construction 1, we have $deg(f) = \max\{deg(\phi) + 1, deg(g)\} \leq \frac{n+1}{2}$, $\Delta_f = 2^n$ (since there exists d such that $r_f(d) = -2^n$), and a nonlinearity $NL_f = 2^{n-1} - 2^{\frac{n-1}{2}}$ (see e.g. [8]).

Remark 2. *The attacker does not have the choice of the parity of n. So the case n even should be also considered. The same Maiorana-McFarland construction can be used to build functions in even numbers of variables, but it does not seem to allow reaching null, nor even small, value for Γ_f.*

1. For example, let $\phi : \mathbb{F}_2^{\frac{n-2}{2}} \mapsto \mathbb{F}_2^{\frac{n+2}{2}}$ be injective and have for image set a coset of a linear $\frac{n-2}{2}$-dimensional subspace of $\mathbb{F}_2^{\frac{n+2}{2}}$, different from this linear subspace (i.e. not containing $\mathbf{0}$), and let g be an $\frac{n-2}{2}$-variable Boolean function; we can define again the n-variable Boolean function $f(x,y) = x \cdot \phi(y) \oplus g(y)$, where $x \in \mathbb{F}_2^{\frac{n+2}{2}}$, $y \in \mathbb{F}_2^{\frac{n-2}{2}}$. The same calculations show that f is balanced and semi-bent, that $f(x,y) \oplus f(x \oplus u, y \oplus v)$ is balanced for $v \neq \mathbf{0}$, and that $r_f(u, \mathbf{0}) = 2^{\frac{n+2}{2}} \sum_{y \in \mathbb{F}_2^{\frac{n-2}{2}}} (-1)^{u \cdot \phi(y)}$; but we have now $r_f(u, \mathbf{0}) \neq 0$ for three values of $u \neq \mathbf{0}$: two giving $r_f(u, \mathbf{0}) = -2^n$ and one giving $r_f(u, \mathbf{0}) = 2^n$. We have then $\Gamma_f = 2^n$ and this choice of f is then optimal for the defender. We leave open the determination whether optimal solutions for the attacker (i.e. such that $\Gamma_f = 0$) can be constructed for n even.

2. It is possible to reach values of Γ_f significantly smaller than 2^n with the Maiorana-McFarland construction, but these values cannot be considered as small: relaxing the condition that the image set of ϕ is a coset of a linear $\frac{n-2}{2}$-dimensional subspace of $\mathbb{F}_2^{\frac{n+2}{2}}$, the value of Γ_f being equal to $2^{\frac{n+2}{2}}$ times the maximal value of $\sum_{z \in S} (-1)^{u \cdot z} = \widehat{1_S}(u) = -\frac{1}{2} W_{1_S}(u)$ for $u \neq \mathbf{0}$, where S is the image set of ϕ and can be any set of size $2^{\frac{n-2}{2}}$ not containing $\mathbf{0}$ and 1_S is its indicator function, we can take for 1_S a Boolean function over $\mathbb{F}_2^{\frac{n+2}{2}}$ of best known nonlinearity among functions of weight $2^{\frac{n-2}{2}}$.

3. In fact a better option is not to use the Maiorana-McFarland construction, but rather to modify a bent function in $2^{\frac{n}{2}-1}$ positions (that is, to add to a bent function a Boolean function g of Hamming weight $2^{\frac{n}{2}-1}$) so as to make it balanced (see Subsection IV-D1 for a proper way called Dobbertin's iterative construction to do so while keeping good nonlinearity). Since $r_{f \oplus g}(d) \leq r_f(d) + 2w_H(g)$, this gives $\Gamma_{f \oplus g} \leq 2^{\frac{n}{2}}$.

Construction 2. *For any $(n-1)$-variable bent function g, the concatenation $f = g || (g \oplus 1)$ is a function with $r_f(\mathbf{0}_{n-1}, 1) = -2^n$ and $r_f(d) = 0$ for all $d \in \mathbb{F}_2^{n*} \setminus (\mathbf{0}_{n-1}, 1)$.*

Proposition 2. *Any function resulting from Construction 2 is balanced near-bent. It satisfies $\Gamma_f = 0$.*

Proof. It is clear that f is balanced and that $W_f(\omega) \in \{0, \pm 2^{\frac{n+1}{2}}\}$. This proves the first part. For every $f = g || h$ and every nonzero $a \in \mathbb{F}_2^{n-1}$, we have $r_f(a, 0) = \sum_{x \in \mathbb{F}_2^{n-1}} (-1)^{D_a g(x)} + \sum_{x \in \mathbb{F}_2^{n-1}} (-1)^{D_a h(x)} = 0$ and $r_f(a, 1) = 2 \sum_{x \in \mathbb{F}_2^{n-1}} (-1)^{g(x) \oplus h(x \oplus a)}$ and here we deduce $r_f(a, 1) = 0$ if $a \neq \mathbf{0}_{n-1}$ and $r_f(\mathbf{0}_{n-1}, 1) = -2^n$. This proves the second part. □

Note that $deg(f) = deg(g)$ (if $deg(g) \geq 1$). We shall see that Construction 2 is cryptographically relevant for the design of S-boxes as well.

We still have that f is balanced near-bent and r_f takes only one nonzero value equal to -2^n if we concatenate a shift $g(x \oplus u)$ and $g(x) \oplus 1$. It should be noted that if g is an M-M bent function then $f = g || (g \oplus 1)$ can be obtained by Construction 1 but not, in general, if we concatenate a shift $g(x \oplus u)$ and $g \oplus 1$. And there are many constructions of bent functions outside the M-M class.

Note that, for two $(n-1)$-variable bent functions g, h, the only possibility for the function $f = g||h$ to have the property $r_f(d) \leq 0$ for any $d \in \mathbb{F}_2^{n*}$ is $h(x) = g(x \oplus u) \oplus 1$ for some u. Indeed, we have already seen that

$$r_f(a,0) = \sum_{x \in \mathbb{F}_2^{n-1}} (-1)^{D_a g(x)} + \sum_{x \in \mathbb{F}_2^{n-1}} (-1)^{D_a h(x)} = 0$$

and we have

$$\begin{aligned} r_f(a,1) &= 2 \sum_{x \in \mathbb{F}_2^{n-1}} (-1)^{g(x) \oplus h(x \oplus a)} \\ &= 2^{-2(n-1)} \sum_{x,u,v \in \mathbb{F}_2^{n-1}} W_g(u) W_h(v) (-1)^{u \cdot x \oplus v \cdot (x \oplus a)} \\ &= 2^{-(n-1)} \sum_{u \in \mathbb{F}_2^{n-1}} W_g(u) W_h(u) (-1)^{u \cdot a} \\ &= \sum_{u \in \mathbb{F}_2^{n-1}} (-1)^{\widetilde{g}(u) \oplus \widetilde{h}(u) \oplus u \cdot a}, \end{aligned}$$

where $\widetilde{g}$ and $\widetilde{h}$ are respectively the dual bent functions of g and h, and it is well-known that a function (here $\widetilde{g} \oplus \widetilde{h}$) can not have all its nonzero Walsh transform values of the same sign, except if it is affine, see e.g. [11]; this completes the observation since $\widetilde{g} \oplus \widetilde{h}$ is affine if and only if $h(x) = g(x \oplus u) \oplus \epsilon$ for some $u \in \mathbb{F}_2^{n-1}$ and $\epsilon \in \mathbb{F}_2$, and $\epsilon = 0$ does not provide the correct sign.

2) Vectorial functions: The condition seen at Subsection III-B1 being on each coordinate function, any general construction of Boolean functions gives a general construction of vectorial functions satisfying the condition. The condition of Subsection III-B2 is more demanding.

Construction 3. *Let $n = 2k+1$ be an odd integer no less than 5. We construct an (n,k)-function F whose coordinate functions f_i's $(1 \leq i \leq k)$ are defined as follows:*

$$f_i(x,y) = x \cdot \phi_i(y) \oplus g_i(y)$$

where

(1) $x \in \mathbb{F}_2^{k+1}$ and $y \in \mathbb{F}_2^k$,
(2) ϕ_i's are mappings from $\mathbb{F}_2^k$ to $\mathbb{F}_2^{k+1}$ such that for any $(l_1, l_2, \cdots, l_k) \in \mathbb{F}_2^{k*}$ the linear combination

$$l_1\phi_1 \oplus l_2\phi_2 \oplus \cdots \oplus l_k\phi_k$$

is an injective mapping from $\mathbb{F}_2^k$ to $\mathbb{F}_2^{k+1*}$,
(3) g_i's are arbitrary Boolean functions on $\mathbb{F}_2^k$.

The following result is a consequence of the proof of Proposition 1:

Proposition 3. *Let $n = 2k+1 \geq 5$ be an odd integer and F be an (n,k)-function generated by Construction 3. Then any component function f of F is a balanced near-bent function with $\Gamma_f = 0$.*

An (n,m)-function is called bent vectorial if and only if all of its component functions are bent. It is well-known that the bent vectorial functions exist only for even n and $m \leq n/2$. Bent vectorial functions are characterized by the fact that all their derivatives $D_aF(x) = F(x) + F(x+a)$, with $a \in (\mathbb{F}_2^n)^*$, are balanced (i.e. take each value of $\mathbb{F}_2^m$ the same number of times 2^{n-m}) and are then also called perfect nonlinear (PN). By Construction 2, we present the following construction.

Construction 4. *Let $n = 2k+1$ be an odd integer no less than 5 and $G = (g_1, g_2, \cdots, g_k)$ be an (n,k)-bent vectorial function. We construct an (n,k)-function F whose coordinate functions f_i's $(1 \leq i \leq k)$ are defined as follows:*

$$f_i = g_i || g_i \oplus 1.$$

It follows from Proposition 2 that:

Proposition 4. *Let $n = 2k+1 \geq 5$ be an odd integer and F be an (n,k)-function generated by Construction 4. Then any component function f of F is a balanced near-bent function with $\Gamma_f = 0$.*

Remark 3. *It is known that, for any n-variable Boolean function, we have $\sum_{u \in \mathbb{F}_2^n} {W_f}^2(u) = 2^{2n}$ and $2^n \sigma_f = \sum_{a \in \mathbb{F}_2^n} W_f^4(a)$. This implies that, for any n odd and any n-variable near-bent function f, we have $\sigma_f = 2^{2n+1}$, which implies that $2^{2n} = \sum_{d \in \mathbb{F}_2^{n*}} r_f^2(d)$. Consider the case $\Delta_f = 2^n$. Then, there exists a value $d' \in \mathbb{F}_2^{n*}$ such that $|r_f(d')| = 2^n$ and $r_f(d) = \mathbf{0}$ for every $d \in \mathbb{F}_2^{n*} \setminus \{d'\}$. Notice that if f is balanced, then, since $\sum_{d \in \mathbb{F}_2^n} r_f(d) = r_f(d') + r_f(\mathbf{0}) = 0$, we have $r_f(d') = -2^n$. Hence, $\Gamma_f = 0$. Note that $r_f(d) \in \{0, \pm 2^n\}$ for any quadratic Boolean function $f \in \mathcal{B}_n$. So the non-absolute indicator of any balanced quadratic near-bent function is equal to 0. It should be noted that, for odd n, there exist (n,n)-functions such that any of their component functions f is balanced near-bent function with $\Gamma_f = 0$. For example, one can mention the quadratic functions $F(x) = x^{2^i+1}$ over $\mathbb{F}_{2^n}$, where $1 \leq i \leq (n-1)/2$ is co-prime with n. Such power function is called a Gold functions. From a mathematically (theoretical) point of view, balanced near-bent function $f \in \mathcal{B}_n$ with $\Gamma_f = 0$ have optimal non-absolute indicator. But we should point out that in the field of Boolean functions and S-boxes for cryptographic use, several design criteria co-exist, related to known attacks on the cryptosystems in which they are involved. Constructions have been found to ensure the best possible tradeoffs between the parameters that quantify the levels at which the functions satisfy these criteria. However, when a new criterion appears, because of the invention of a new attack, tradeoffs need to be redefined. One noticeable recent historical example is, in the domain of single-output Boolean functions, upon the introduction of the algebraic attacks: the new constraint of having a good algebraic immunity has been added to the global tradeoff, and the requirement on the values of the previously existing parameters has been slightly lowered. In the framework of this paper, we consider mitigation of side-attacks in addition to traditional cryptanalytic attacks. The attacks are very unbalanced in terms of risk, as side-channels are much more powerful and practical than classical cryptanalyses: they can recover the key within a few thousands of traces, whereas cryptanalyses require more than 2^{80} pairs of plaintext/ciphertext to succeed. Hence, the tradeoff is clearly in favor of side-channel attacks mitigation. Therefore, we include in our exploration S-box constructions which may be suboptimal according to those*

standards considered when only classical attacks are taken into account. This does not make block cipher designs less strong concretely, nor presents regression with respect to the state-of-the-art. We simply place ourselves in the situation of embedded cryptography which is subject to side-channel attacks on their implementations

D. Constructions of Boolean and vectorial functions for the defender

1) Boolean functions: In the following, we construct Boolean functions that are optimum from a defender's point of view, that is, such that Γ_f is large. We have seen with Remark 2 a first example with an even number of variables, using the Maiorana-McFarland construction. For obtaining another example with an even number of variables as well, we need to recall Dillon's direct sum of functions [16].

Lemma 2. *Let three positive integers n, r and e be such that $n = r + e$. Let $f(x_1, \ldots, x_n) = g(x_1, \ldots, x_r) + h(x_{r+1}, \ldots, x_n)$, where $g \in \mathcal{B}_r$ and $h \in \mathcal{B}_e$. For any $\beta \in \mathbb{F}_2^n$, we have:*

1) $W_f(\beta) = W_g(\beta') \cdot W_h(\beta'')$,
2) $r_f(\beta) = r_g(\beta') \cdot r_h(\beta'')$,

where $\beta = (\beta', \beta'') \in \mathbb{F}_2^r \times \mathbb{F}_2^e$ with $\beta' = (\beta_1, \ldots, \beta_r)$ and $\beta'' = (\beta_{r+1}, \ldots, \beta_k)$.

Construction 5. *Let $n \geq 6$ be an even number such that $n = r + e$ for odd numbers $r, e \geq 3$. Let $g \in \mathcal{B}_r$ and $h \in \mathcal{B}_e$ be Boolean functions constructed by Construction 1 (resp. Construction 2). We consider the function $f(x_1, \ldots, x_n) = g(x_1, \ldots, x_r) + h(x_{r+1}, \ldots, x_n)$.*

Proposition 5. *The function f defined in Construction 5 satisfies $\Gamma_f = 2^n$.*

Proof. This result directly follows from Lemma 2, since we have seen that r_g takes value -2^r at some (nonzero) input and r_h takes value -2^e at some (nonzero) input. □

The construction 5 is suitable for cryptographic applications, as:

Proposition 6. *The function f defined in Construction 5 has nonlinearity $NL_f = 2^{n-1} - 2^{\frac{n}{2}}$, is balanced and has algebraic degree $deg(f) \leq \max\left\{\frac{r-1}{2}, \frac{e-1}{2}\right\}$, where this latter bound is tight in both cases.*

Proof. This nonlinearity directly follows from Lemma 2 and Relation (1). Besides, f is balanced and we have $deg(f) \leq \max\left\{\frac{r-1}{2}, \frac{e-1}{2}\right\}$ because in Construction 1, ϕ is injective and has image a hyperplane, then it can have algebraic degree at most $\frac{r-1}{2}$ (resp. $\frac{e-1}{2}$), and in Construction 2, the function is an affine extension of a bent function in $r-1$ (resp $e-1$) variables, which can have algebraic degree at most $\frac{r-1}{2}$ (resp. $\frac{e-1}{2}$). □

In Proposition 6, all three values hold also if we let $n = r + 1$, and if instead of taking h from a construction, we take a well-chosen single variable Boolean function. The following example illustrates this situation.

Example 1. Let $h(0) = 0$ and $h(1) = 1$. Let the truth table of g be the following:

```
00FFA956CC33659AF00F59A63CC3956A .
```

Then the truth table of f is obtained as follows:

```
5555AAAA99966669A5A55A5A69669699 \
AA5555AA669699695AA5A55A96666999
```
(11)

for which $NL_f = 112$, $\Gamma_f = 256$, and $deg(f) = 3$.

Remark 4. *Suppose there exists a balanced Boolean function g on even number r of variables for which $\Gamma_g = 0$. Let h be a bent function with e variables and $n = r + e$. Then $\Gamma_f = 0$ for the function $f(x_1, \ldots, x_n) = g(x_1, \ldots, x_r) + h(x_{r+1}, \ldots, x_n)$.*

Hence, if one can find any balanced Boolean function g on even number of variables r with $\Gamma_g = 0$, then we can construct a balanced Boolean function f on even number n of variables with $n > r$ such that $\Gamma_f = 0$. However, it seems that such balanced Boolean function g are difficult to find and their existence is an open question after several computer investigations. We have completed an exhaustive search for the class of 8-variable rotation-symmetric Boolean functions (RSBFs) for which the search space is 2^{36} (the number of balanced ones is $\approx 2^{30.2}$) and found that the minimum value of Γ_f is 16. We also performed several heuristic searches in the whole space of 8-variable Boolean functions, which did not yield a better result.

Now let us recall Dobbertin's iterative construction based on normal bent functions for constructing balanced Boolean function with very high nonlinearity, which will be employed in our next construction.

Dobbertin's iterative construction [17]: Let n be an even integer no less than 4. Write $n = 2^t m$ such that $t \geq 1$ and m is an odd integer. Then a balanced Boolean function $f(x, y) \in \mathcal{B}_n$ over $\mathbb{F}_2^n$ is defined by

$$f(x, y) = \begin{cases} f_0(x, y), & \text{if } x \neq \mathbf{0}_{\frac{n}{2}} \\ g_1(y), & \text{if } x = \mathbf{0}_{\frac{n}{2}} \end{cases}, \tag{12}$$

where $f_0(x, y)$ is an arbitrary n-variable bent function with $f_0(\mathbf{0}_{\frac{n}{2}}, y) = cst$ and g_1 is generated by an iterative procedure as

$$g_i(x, y) = \begin{cases} f_i(x, y), & \text{if } x \neq \mathbf{0}_{\frac{n}{2^{i+1}}} \\ g_{i+1}(y), & \text{if } x = \mathbf{0}_{\frac{n}{2^{i+1}}}, \end{cases} \quad i = 1, 2, \cdots \tag{13}$$

where $x, y \in \mathbb{F}_2^{\frac{n}{2^{i+1}}}$ and in each step f_i is an arbitrary $\frac{n}{2^i}$-variable bent function with $f_i(\mathbf{0}_{\frac{n}{2^i}}, y) = cst$. The iterative process will continue until $i = t - 1$ with $g_t = s \in \mathcal{B}_m$ being a balanced m-variable Boolean function with the best known nonlinearity and $s(0) = 0$.

Theorem 3 [17]**.** *Let f be a balanced Boolean function given by* (12)*. Then*

$$NL_f \geq 2^{n-1} - 2^{\frac{n}{2}} + NL_{g_1}.$$

The following theorem can be easily checked.

Theorem 4. *Let f be a balanced Boolean function given by* (12)*. Then*

$$deg(f) = \frac{n}{2} + deg(g_1).$$

Corollary 3. *Let n be a power of 2 and f be a balanced Boolean function given by* (12)*. Then*

$$deg(f) = n - 1.$$

By using the class of M-M bent functions and employing the balanced Boolean functions generated by Dobbertin's iterative construction, we propose the following construction.

Construction 6. *For $n = 2k$, we define a balanced Boolean function $f(x, y) \in \mathcal{B}_n$ over $\mathbb{F}_2^n$ as follows:*

$$f(x,y) = \begin{cases} \phi(x) \cdot y, & \text{if } x \neq \mathbf{0}_k \\ g(y), & \text{if } x = \mathbf{0}_k \end{cases}, \tag{14}$$

where $x, y \in \mathbb{F}_2^k$, ϕ is a permutation on $\mathbb{F}_2^k$ such that $\phi(\mathbf{0}_k) = 0$, and g is a balanced Boolean function on $\mathbb{F}_2^k$ generated by (12) *and* (13)*.*

Theorem 5. *Let f be an $n = 2k$-variable Boolean function generated by Construction 6, then for any $(a, b) \in \mathbb{F}_2^k \times \mathbb{F}_2^k$ we have*

$$r_f(a,b) = \begin{cases} 2^n & \text{if } a = b = \mathbf{0}_k \\ -2^n + r_g(b), & \text{if } a = \mathbf{0}_k, b \in \mathbb{F}_2^k \\ 2(-1)^{\phi(a)\cdot b} W_g(\phi(a)), & \text{if } a \in \mathbb{F}_2^{k*}, b \in \mathbb{F}_2^k \end{cases}.$$

Proof. It follows from the definition of autocorrelation function that

$$r_f(a,b) = \sum_{(x,y)\in\mathbb{F}_2^k\times\mathbb{F}_2^k} (-1)^{f(x,y)\oplus f(x\oplus a, y\oplus b)} \tag{15}$$

for any $(a, b) \in \mathbb{F}_2^k \times \mathbb{F}_2^k$. Clearly, we have $r_f(\mathbf{0}_k, \mathbf{0}_k) = 2^n$. We now consider the values of $r_f(a, b)$ for all $(a, b) \in \mathbb{F}_2^k \times \mathbb{F}_2^k \setminus \{(\mathbf{0}_k, \mathbf{0}_k)\}$. Basically, our discussion is built on the facts that $\sum_{x\in\mathbb{F}_2^{k*}} (-1)^{c\cdot x}$ equals -1 if $c \in \mathbb{F}_2^{k*}$ and equals $2^k - 1$ if $c = \mathbf{0}_k$. We consider the following two cases:

[**Case 1**.] $(a, b) \in \{\mathbf{0}_k\} \times \mathbb{F}_2^{k*}$. It can be easily seen that in this case Eq. (15) becomes

$$\begin{aligned} r_f(a,b) &= \sum_{(x,y)\in\{\mathbf{0}_k\}\times\mathbb{F}_2^k} (-1)^{f(\mathbf{0}_k,y)\oplus f(\mathbf{0}_k,y\oplus b)} + \\ &\quad \sum_{(x,y)\in\mathbb{F}_2^{k*}\times\mathbb{F}_2^k} (-1)^{f(x,y)\oplus f(x,y\oplus b)} \\ &= \sum_{y\in\mathbb{F}_2^k} (-1)^{g(y)\oplus g(y\oplus b)} + \\ &\quad \sum_{(x,y)\in\mathbb{F}_2^{k*}\times\mathbb{F}_2^k} (-1)^{\phi(x)\cdot y\oplus\phi(x)\cdot(y\oplus b)} \\ &= r_g(b) + \sum_{(x,y)\in\mathbb{F}_2^{k*}\times\mathbb{F}_2^k} (-1)^{\phi(x)\cdot b} \\ &= r_g(b) + 2^k \sum_{x\in\mathbb{F}_2^{k*}} (-1)^{\phi(x)\cdot b} \\ &= -2^k + r_g(b). \end{aligned}$$

[**Case 2**.] $(a, b) \in \mathbb{F}_2^{k*} \times \mathbb{F}_2^k$. In this case, the value of $r_f(a, b)$ of Eq. (15) becomes

$$\begin{aligned} &\sum_{\substack{x\in\{\mathbf{0}_k,a\}\\ y\in\mathbb{F}_2^k}} (-1)^{f(x,y)\oplus f(x\oplus a,y\oplus b)} + \\ &\sum_{\substack{x\in\mathbb{F}_2^k\setminus\{\mathbf{0}_k,a\}\\ y\in\mathbb{F}_2^k}} (-1)^{f(x,y)\oplus f(x\oplus a,y\oplus b)} \\ &= \sum_{y\in\mathbb{F}_2^k} \left((-1)^{f(\mathbf{0}_k,y)\oplus f(a,y\oplus b)} + (-1)^{f(a,y)\oplus f(\mathbf{0}_k,y\oplus b)}\right) \\ &\quad + \sum_{\substack{x\in\mathbb{F}_2^k\setminus\{\mathbf{0}_k,a\}\\ y\in\mathbb{F}_2^k}} (-1)^{f(x,y)\oplus f(x\oplus a,y\oplus b)} \\ &= 2\sum_{y\in\mathbb{F}_2^k} (-1)^{g(y)\oplus\phi(a)\cdot(y\oplus b)} + \\ &\quad \sum_{\substack{x\in\mathbb{F}_2^k\setminus\{\mathbf{0}_k,a\}\\ y\in\mathbb{F}_2^k}} (-1)^{\phi(x)\cdot y\oplus\phi(x\oplus a)\cdot(y\oplus b)} \\ &= 2(-1)^{\phi(a)\cdot b} W_g(\phi(a)) + \\ &\quad \sum_{x\in\mathbb{F}_2^k\setminus\{\mathbf{0}_k,a\}} (-1)^{\phi(x\oplus a)\cdot b} \sum_{y\in\mathbb{F}_2^k} (-1)^{z\cdot y} \\ &= 2(-1)^{\phi(a)\cdot b} W_g(\phi(a)), \end{aligned}$$

where $z = \phi(a) \oplus \phi(x \oplus a)$ is nonzero for any $a \in \mathbb{F}_2^{k*}$ and $x \in \mathbb{F}_2^k \setminus \{\mathbf{0}_k, a\}$. □

2) Vectorial functions: Here again, any general construction of Boolean functions satisfying the condition seen at Subsection III-B1 gives a general construction of vectorial functions satisfying the same condition. The condition of Subsection III-B2 is more demanding.

For $n = 2k$, we present a construction of (n, k)-functions obtained by modifying (n, k)-bent vectorial functions.

Construction 7. *Let $n = 2k$ be an even integer no less than 4 and $G = (g_1, g_2, \cdots, g_k)$ be an (n, k)-bent vectorial function such that $g_i(\mathbf{0}_k, y) = cst$ for all $1 \leq i \leq k$, where $y \in \mathbb{F}_2^k$. Let $G' = (g'_1, g'_2, \cdots, g'_k)$ be a balanced (k, k)-function with nonlinearity no less than $2^{k-1} - 2^{\lfloor k/2 \rfloor}$. We construct an (n, k)-function F whose coordinate functions f_i's $(1 \leq i \leq k)$ are defined as follows:*

$$f_i(x,y) = \begin{cases} g_i(x,y), & \text{if } x \neq \mathbf{0}_k \\ g'_i(y), & \text{if } x = \mathbf{0}_k \end{cases},$$

where $x, y \in \mathbb{F}_2^k$.

It can be easily checked that every (n, k)-function generated by Construction 7 is balanced. Further, according to the proof of Theorem 5 we have the following result.

Proposition 7. *Let $n = 2k \geq 4$ be an even integer and F be an (n, k)-function generated by Construction 7 with g_i's in M-M class. Then any component function f of F has nonlinearity no less than $2^{n-1} - 2^{k-1} - 2^{\lfloor k/2 \rfloor}$ and $\Gamma_f \leq 2^{\lceil (k+1)/2 \rceil + 1}$.*

Proof. For every $c = (c_1, c_2, \ldots, c_k) \in \mathbb{F}_2^{k*}$, we define $f = \bigoplus_{i=1}^k c_i f_i$, $g = \bigoplus_{i=1}^k c_i g_i$, and $g' = \bigoplus_{i=1}^k c_i g'_i$. where

f_i's, g's and g'_i's are defined in Construction 7. On the one hand, it follows from [17] that any component function f has nonlinearity $2^{n-1} - 2^{k-1} - \frac{1}{2}\max_{a\in\mathbb{F}_2^k}|W_{g'}(a)|$ which is no less than $2^{n-1}-2^{k-1}-2^{\lfloor k/2\rfloor}$, since $G' = (g'_1, g'_2, \cdots, g'_k)$ is a balanced (k,k)-function with nonlinearity no less than $2^{k-1} - 2^{\lfloor k/2\rfloor}$ which implies that $\max_{a\in\mathbb{F}_2^k}|W_{g'}(a)| \leq 2^{\lceil (k+1)/2\rceil}$. On the other hand, similar to the proof of Theorem 5 we can immediately get that $r_f(a,b) = -2^n + r_{g'}(b)$ if $a = \mathbf{0}_k, b \in \mathbb{F}_2^k$ and $r_f(a,b) \in \{\pm 2W_{g'}(a)\}$ if $a \in \mathbb{F}_2^{k*}, b \in \mathbb{F}_2^k$. This gives $\Gamma_f \leq 2^{\lceil (k+1)/2\rceil+1}$. This completes the proof. □

V. Some Specific S-box Constructions

A. In dimension 6

We consider the rotation-symmetric S-boxes [50] (RSSBs) that are bijective from $\mathbb{F}_2^6 \to \mathbb{F}_2^6$. In [25], using a sieving strategy, the search space of all bijective RSSBs in dimension 6 is reduced from $2^{47.9}$ to 2^{40}. The sieving strategy in [25] is mainly based on the facts that some of the component functions of an RSSB are the generalized k-rotation symmetric Boolean functions [27], where $k = 1$, 2, and 3, and there are some affine equivalence relations among these functions yielding that the nonlinearity of an RSSB can be found by computing the nonlinearities of only 13 (instead of 63) component functions. Thanks to this, all possible candidates for some of the 13 component functions (with nonlinearity greater than or equal to 24) are obtained. After that, using those component functions, all the 2^{40} RSSBs containing them are generated and the RSSBs with nonlinearity 24 are found efficiently (as it is enough to find the nonlinearities of the remaining component functions to find the nonlinearity of an RSSB). It is found in [25] that there are 23102464×12 ($\approx 2^{28}$) RSSBs with nonlinearity 24 (known maximal nonlinearity) and, among them, the number of those with differential uniformity 4 (there is only one example [5] with differential uniformity 2 in the literature) is 2332288×12 ($\approx 2^{24.7}$).

We have checked that among them there are only four (up to the affine equivalence) that are optimum in terms of the defender's objective (i.e., $\min_{c\in\mathbb{F}_2^{6*}}\Gamma_{c\cdot F} = 64$ for each S-box F). However, these S-boxes (given in Appendix A) are quadratic.

Let $Q_k : \mathbb{F}_{2^6} \to \mathbb{F}_{2^6}$ denote the polynomial representation S-box # k, where $k = 1, 2, 3,$ and 4. Using a normal basis, the representations are obtained as follows by Lagrange interpolation:

$$Q_1(\alpha) = \alpha^{24}+\alpha^{40}+\alpha^{44}+\alpha^{48}+\alpha^{52}+\alpha^{55}+\alpha^{56}+\alpha^{61}+\alpha^{63}$$
$$Q_2(\alpha) = \alpha^{44}+\alpha^{47}+\alpha^{52}+\alpha^{54}+\alpha^{55}+\alpha^{59}+\alpha^{62}$$
$$Q_3(\alpha) = \alpha^{30}+\alpha^{31}+\alpha^{40}+\alpha^{48}+\alpha^{55}+\alpha^{56}+\alpha^{59}+\alpha^{61}+\alpha^{63}$$
$$Q_4(\alpha) = \alpha^{28}+\alpha^{30}+\alpha^{32}+\alpha^{47}+\alpha^{60}+\alpha^{61}+\alpha^{63}$$

Considering the APN S-box in dimension 6 [5], we find that it is very weak from defender's point of view. Among the 63 component functions, 28 of them have non-absolute indicator value 8 (which is the minimum possible) and the rest non-absolute indicator value 16. Notice that among these component functions, the minimum algebraic degree is 3 and the maximum algebraic degree is 4.

B. In dimension 8

Some well-known constructions generating permutations in dimension 8 with differential uniformity 4 and nonlinearity $2^{n-1} - 2^{\frac{n}{2}}$ are given in III, where $tr_1^k(\alpha)$ for $\alpha \in \mathbb{F}_{2^k}$ is the trace function from $\mathbb{F}_{2^k}$ to $\mathbb{F}_2$ defined by $tr_1^k(\alpha) = \sum_{i=0}^{k-1}\alpha^{2^i}$. In Table IV, the value of $\min_i \Gamma_{f_i}$ is computed along with the other cryptographic properties. It is seen from Table IV that the worst value is 32 among these constructions. Specifically, for the inverse function we find that $\Gamma_{f_i}(d) = 32$ for each coordinate function f_i. We have performed some heuristics in the class of RSSBs, in order to find the bijective S-boxes having $\max_i \Gamma_{f_i} < 32$. Note that RSSBs are affine equivalent to the S-boxes obtained from the (sum of) power maps and most of the known constructions correspond to some power maps, e.g. the inverse function. The search yielded bijective RSSBs for which $\max_i \Gamma_{f_i}(d) = 24 < 32$, hence worse than the constructions in Table III in terms of the defender's objective.

Remark 5. *It is shown that the nonlinearity of an RSSB is bounded above by the nonlinearity of f, where f is the coordinate function defining the RSSB, see [36] for more details. It can be easily checked that this property is also true for Γ_f. This could maybe ease the search of RSSB with good Γ_f in higher dimensions. We leave this as a topic for future research.*

VI. Discussion

A. Practical Evaluation

The strength of a side-channel attack can be measured by its probability of success [53, §3.1]. It is defined as the probability of the event $\hat{k} = k^*$, where:

- k^* is the actual secret key, and
- $\hat{k}$ is the key guessed (recall (2)) by the optimal side-channel distinguisher, which consists in maximizing the probability of the observations knowing the assumed model for the given keys (recall (4)).

This probability can be estimated as a success rate by repeating several independent attacks. In addition, the standard deviation of the success probability can also be estimated, as that of a Bernoulli distribution [33].

We compare in this section the result of the optimal distinguisher (which, by definition, maximizes the success probability) for various S-boxes, namely:

- The worst (from an attacker standpoint) Boolean function (5555-6999) displayed in Eqn. (11).
- The optimal (from an attacker standpoint) Boolean function (18CA-9029) displayed in Eqn. (9).
- The LSB of $x \mapsto x^{101}$ in $\mathbb{F}_{256}$, seen as $\mathbb{F}_2[X]/\langle X^8+X^4+X^3+X+1\rangle$ (S-box also used as an example in [23, §4.3]); this S-box has average properties regarding side-channel attacks, hence should be regarded as a representative average case.
- The LSB of the AES [44] S-box (nicknamed `SubBytes`), which is $x \mapsto \mathtt{63} + \mathtt{8f}x^{127} + \mathtt{b5}x^{191} + \mathtt{01}x^{223} + \mathtt{f4}x^{239} + \mathtt{25}x^{247} + \mathtt{f9}x^{251} + \mathtt{09}x^{253} + \mathtt{05}x^{254}$. This S-box is renown as very relevant from a

TABLE III. Differentially-4 uniform permutations with the best known nonlinearity $2^{n-1} - 2^{\frac{n}{2}}$.

Function	Condition	Reference
α^{2^n-2}	n even	[45]
$\left(F(\alpha) \oplus f(\alpha)^{2^i}\right)\Big\vert_H$	$H{=}\{\alpha \in \mathbb{F}_{2^{n+1}} \mid tr_1^{n+1}(\alpha) = 0\}$, $F(\alpha){=}\alpha^{\frac{1}{2^i+1}} + tr_{m/3}^{n+1}(\alpha + \alpha^{2^{2s}})$, $s \equiv i \bmod 3$, $\gcd(i, n+1){=}1$, $n \equiv 2 \bmod 6$	[29]
$\left(\beta\alpha^{\frac{2^i}{2^i+1}} + \beta^{2^i}\alpha^{\frac{1}{2^i+1}}\right)\Big\vert_{H_\beta}$	$\beta \in \mathbb{F}_{2^{n+1}}^*$, $H_\beta{=}\{\beta\alpha^{2^i} + \beta^{2^i}\alpha \mid \alpha \in \mathbb{F}_{2^{n+1}}\}$, $\gcd(i, n+1){=}1$, $n \geq 4$ and even	
$\left(\beta\alpha^{\frac{2^i}{2^i+1}} + \beta^{2^i}\alpha^{\frac{1}{2^i+1}} + \alpha\right)\Big\vert_{H_\beta}$		

TABLE IV. Cryptographic properties of the S-boxes in dimension 8 generated from the constructions in Table III.

Function*	Absolute Indicator	Degree	$\min_i \Gamma_{f_i}$	Nonlinearity
α^{254}	32	7	32	112
$\left(F(\alpha) \oplus f(\alpha)^{2^i}\right)\Big\vert_H$	72, 80	5	48	112
$\left[\left(F(\alpha) \oplus f(\alpha)^{2^i}\right)\Big\vert_H\right]^{-1}$	256	5	48	112
$\left(\beta\alpha^{\frac{2^i}{2^i+1}} + \beta^{2^i}\alpha^{\frac{1}{2^i+1}}\right)\Big\vert_{H_\beta}$	72, 80	5	$32, 40, 48, 56$	112
$\left[\left(\beta\alpha^{\frac{2^i}{2^i+1}} + \beta^{2^i}\alpha^{\frac{1}{2^i+1}}\right)\Big\vert_{H_\beta}\right]^{-1}$	256	3	32	112
$\left(\beta\alpha^{\frac{2^i}{2^i+1}} + \beta^{2^i}\alpha^{\frac{1}{2^i+1}} + \alpha\right)\Big\vert_{H_\beta}$	72, 80	5	$40, 48, 56$	112
$\left[\left(\beta\alpha^{\frac{2^i}{2^i+1}} + \beta^{2^i}\alpha^{\frac{1}{2^i+1}} + \alpha\right)\Big\vert_{H_\beta}\right]^{-1}$	72, 80	5	$40, 48, 56$	112

* $[F(\alpha)]^{-1}$ denotes the compositional inverse of $F(\alpha)$, $H{=}\{\alpha \in \mathbb{F}_{2^9} \mid tr_9(\alpha) = 0\}$, and $H_\beta{=}\{\beta\alpha^{2^i} + \beta^{2^i}\alpha \mid \alpha \in \mathbb{F}_{2^9}\}$, where $\beta \in \mathbb{F}_{2^9}^*$ and $i = 1, 2, 4, 5, 7, 8$.

cryptographic point of view, hence it should be considered as a representative for the best case.

- The 2nd-LSB of SAFER [37] Exponentiating S-box, namely $x \mapsto 45^x \bmod 257$ (except that the output is taken to be 0 if the modular result is 256, which occurs for input $x = 128$). It has $\Gamma_f = 256$.

The properties of these S-boxes are gathered in Table V.

The superiority of the function (9) over the LSB (and also other bits—not shown) of AES can be seen in Fig. 3. In this figure, the success rate has been computed based on 1 million acquisition campaigns. Each campaign is independent and has been carried our on 20 side-channel measurements (or 20 traces).

Without surprise, all curves start, for no side-channel trace ($Q = 0$), at success probability $2^{-n} = 1/256$. The success probability is then increasing with the number of traces $Q > 0$. We observe that:

- The attack on the optimal Boolean function (18CA-9029) is even more favorable to an attacker than that of the LSB of AES S-box;
- The attack on the worst Boolean function (5555-6999) results in the smallest success rate (for a given number of traces Q), and furthermore feature ties, hence the asymptotic success rate of $1/2$ when $Q \to +\infty$. The same behavior is observed for f equal to SAFER exponentiating S-box 2nd-LSB, which also has $\Gamma_f = 256$.

Notice that the curves in Fig. 3 are equipped with error bars. Those are intentionally narrow, owing to the large number of attacks to validate the success rate with great accuracy.

Notice that the attack success rate for Boolean function (18CA-9029) is *unambiguously* greater (though not a lot) than that corresponding to the attack on the AES SubBytes LSB. The greater success rate for the new S-box is illustrated in the inset at top left corner of Fig. 3, where it is clear that the error bars at $\pm\sigma$ do not interpenetrate, hence the new 8×8 S-box is *strictly* more favorable to the attacker than that of AES. This analysis reveals that the AES S-Box is **not** the worst with respect to side-channel analysis. However, it is not that far from being the worst, thus we recommend that designers resort to protection schemes, such as leakage balancing (hiding) or S-box randomization (masking) [34, Chap. 7 & 9].

Same conclusions hold in the presence of measurement noise, as shown in Fig. 4. Information-theoretical analysis [15, Theorem 1, Eqn. (4) & (5)] allows to derive two bounds on the minimum number of traces for succeeding a key extraction irrespective of the attack method. These bounds are

TABLE V. Properties of studied S-boxes (differential uniformity compares between resp. cases 1, 2, 5 and cases 3, 4)

S-box F name	Construction	Balanced?	Nonlinearity	Diff. uniformity	Alg. degree	Γ_f (our paper)
5555-6999	Eqn. (11), from construction 5	Yes	112	256	3	256
18CA-9029	Eqn. (9), copied from [26]	Yes	116	136	7	16
$x \mapsto x^{101}$ in finite field $\mathbb{F}_{256}$	A power S-box, as proposed in [23, §4.3]	Yes	80	30	4	All 255 component functions f have $\Gamma_f = 112$
AES `SubBytes`	Linear transformation in $\mathbb{F}_2^8$ of $x \mapsto x^{-1} \in \mathbb{F}_{256}$	Yes	112	4	7	All 255 component functions f have $\Gamma_f = 32$
SAFER exponentiating S-Box (2nd LSB)	$\forall x \in \{0, \ldots, 255\} \backslash \{128\}$, $x \mapsto 45^x \bmod 257$, and $128 \mapsto 0$	Yes	82	128	7	48, 56, 64, 72, 80 88, 96, 104, 112, 120 128, 136, 144, 152, 160 168, 176, 184, 192, 200 208, 216, 224, 232, 240 248, 256

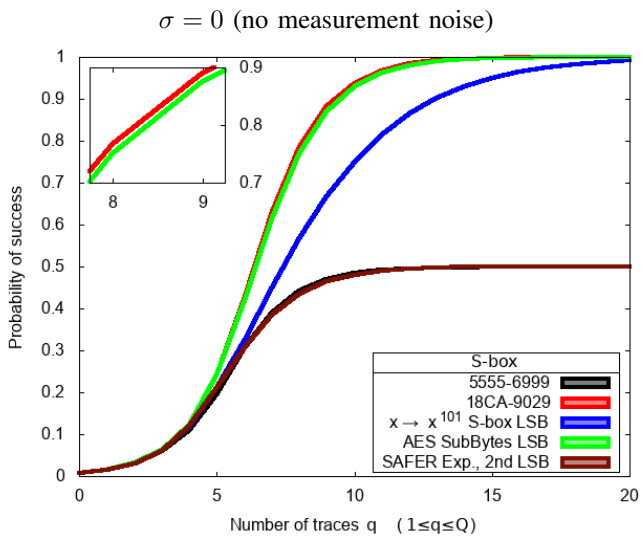

Fig. 3. Success rate of attack without noise for various S-boxes, and inset magnification on a portion where the new S-box is successfully attacked faster than AES S-box.

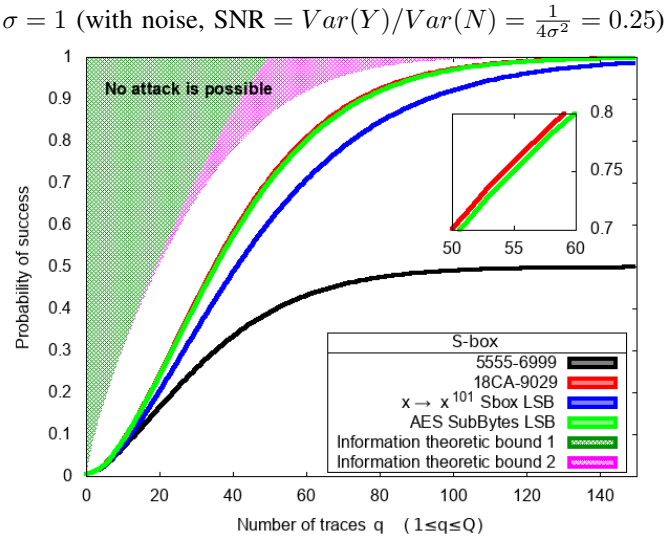

Fig. 4. Success rate of attack with noise for various S-boxes and attack bounds, with similar inset magnification as in Fig. 3.

superimposed on the graph (for $\sigma = 1$, since the bound does not work for $\sigma = 0$). The bounds are clearly close to the actual success rates on attacks instantiated on S-boxes (and all the closer as Γ_f is small), which confirms that S-boxes are relevant and critical architectural blocks as targets to side-channel attacks within block ciphers.

B. Critical Analysis

Let us mention some remaining open problems.

First, the existence of balanced 8×8 S-boxes with $\Gamma = 8$ is an open problem.

Second, we have suggested some constructions of S-boxes which, in addition to the customarily criteria, need to optimize the component's auto-correlation. Our constructions 1-4 all yield semi-bent functions, which feature linear structures, that should be avoided in block ciphers. This is the case also for Construction 5: since it is based on the direct sum, it generates decomposable functions, which also have linear structures and are vulnerable to divide-and-conquer attacks. There is therefore room for further better tradeoffs.

Third, our examples range over some well-known construction methods, but there is no guarantee that our analysis is covering all S-boxes. Therefore, new designs can be imagined.

Fourth, Remark 3 provides balanced Boolean functions f with $\Gamma_f = 0$, but they are of low algebraic degree. A more general question would be that of finding the maximum achievable degree of an n-variable balanced Boolean function with $\Gamma_f = 0$.

Eventually, S-boxes are usually protected against side-channel attacks by their random masking. Provable computation of masked S-boxes resort to the interpolation of their truth table by a Lagrange polynomial. In order to minimize the computational overhead, S-boxes which can be interpolated by small polynomials are preferred. The complexity is quantified by the number of multiplications [1] when evaluating the polynomial. Therefore, another challenge is to find suitable S-boxes which in addition also meet this criterion.

VII. Conclusion

It is well-known that S-boxes are appealing targets for side-channel attacks on cryptographic algorithms. Indeed, they

allow to distinguish clearly between the correct key and erroneous key guesses. Considering the most powerful attacker, we derive that exploited property of the S-box is the (signed) auto-correlation of its components. We therefore consider the question to know which S-box is the easiest (harder) to attack with side-channel analysis. Interestingly, we show that some S-boxes exist which maximize (resp. minimize) the side-channel attack efficiency. Leveraging known S-box constructions, we put forward S-boxes with optimal values in terms of components auto-correlation. Namely, we provide concrete instantiations for 6- and 8-bit S-Boxes. Our S-boxes have two applications: first, they are interesting per se as objects which shall be considered in the context of facing both cryptanalysis and side-channel attacks. Second, they allow to validate in which respect some security bounds on side-channel attacks are close to the "easiest attackable" S-boxes (in termes of side-channel analysis).

Acknowledgement

We thank Stjepan Picek who made intensive experiments and (also) obtained 16 as the smallest value of Γ_f for balanced functions in 8 variables. The authors would like to thank the anonymous reviewers and the Associate Editor, Prof. Georg Sigl, for their valuable comments that improved the editorial as well as technical quality of this paper.

References

[1] Sonia Belaïd, Fabrice Benhamouda, Alain Passelègue, Emmanuel Prouff, Adrian Thillard, and Damien Vergnaud. Private multiplication over finite fields. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology - CRYPTO 2017, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part III*, volume 10403 of *LNCS*, pages 397–426. Springer, 2017.

[2] Roderick Bloem, Hannes Groß, Rinat Iusupov, Bettina Könighofer, Stefan Mangard, and Johannes Winter. Formal Verification of Masked Hardware Implementations in the Presence of Glitches. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2018, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part II*, volume 10821 of *LNCS*, pages 321–353. Springer, 2018.

[3] Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, Yannick Seurin, and Charlotte Vikkelsoe. PRESENT: An Ultra-Lightweight Block Cipher. In *CHES*, volume 4727 of *LNCS*, pages 450–466. Springer, September 10-13 2007. Vienna, Austria.

[4] Éric Brier, Christophe Clavier, and Francis Olivier. Correlation power analysis with a leakage model. In Marc Joye and Jean-Jacques Quisquater, editors, *Cryptographic Hardware and Embedded Systems - CHES 2004: 6th International Workshop Cambridge, MA, USA, August 11-13, 2004. Proceedings*, volume 3156 of *Lecture Notes in Computer Science*, pages 16–29. Springer, 2004.

[5] Browning, K. A., Dillon, J. F., McQuistan, M. T., Wolfe, A. J. An APN permutation in dimension six. The 9th Conference on Finite Fields and Applications - Fq9, Contemporary Mathematics, Vol. 518, 33-42, 2010.

[6] Paul Camion, Claude Carlet, Pascale Charpin, and Nicolas Sendrier. On correlation-immune functions, *Proceedings of CRYPTO'91, Lecture Notes in Computer Science*, vol. 576, pp. 86-100, 1991.

[7] Claude Carlet. On Highly Nonlinear S-Boxes and Their Inability to Thwart DPA Attacks. In Subhamoy Maitra, C. E. Veni Madhavan, and Ramarathnam Venkatesan, editors, *Progress in Cryptology - INDOCRYPT 2005, 6th International Conference on Cryptology in India, Bangalore, India, December 10-12, 2005, Proceedings*, volume 3797 of *Lecture Notes in Computer Science*, pages 49–62. Springer, 2005.

[8] Claude Carlet. Boolean functions for cryptography and error correcting codes. *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, 2:257–397, 2010.

[9] Claude Carlet. Vectorial Boolean Functions for Cryptography. Chapter of the monography *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, Y. Crama and P. Hammer eds, Cambridge University Press, pp. 398-469 (2010).

[10] Claude Carlet, Annelie Heuser, and Stjepan Picek. Trade-Offs for S-Boxes: Cryptographic Properties and Side-Channel Resilience. In Dieter Gollmann, Atsuko Miyaji, and Hiroaki Kikuchi, editors, *Applied Cryptography and Network Security - 15th International Conference, ACNS 2017, Kanazawa, Japan, July 10-12, 2017, Proceedings*, volume 10355 of *Lecture Notes in Computer Science*, pages 393–414. Springer, 2017.

[11] Claude Carlet and Sihem Mesnager. On Semi-bent Boolean Functions. *IEEE Transactions on Information Theory*, Vol. 58, no. 5, pp. 3287-3292, 2012.

[12] Kaushik Chakraborty, Sumanta Sarkar, Subhamoy Maitra, Bodhisatwa Mazumdar, Debdeep Mukhopadhyay, and Emmanuel Prouff. Redefining the transparency order. *Des. Codes Cryptography*, 82(1-2):95–115, 2017.

[13] Pascale Charpin and Enes Pasalic. On Propagation Characteristics of Resilient Functions. *Selected Areas in Cryptography 2002, Lecture Notes in Computer Science 2595*, pp. 175-195, 2002.

[14] Eloi de Chérisey, Sylvain Guilley, and Olivier Rioul. Confused yet successful: - theoretical comparison of distinguishers for monobit leakages in terms of confusion coefficient and SNR. In Fuchun Guo, Xinyi Huang, and Moti Yung, editors, *Information Security and Cryptology - 14th International Conference, Inscrypt 2018, Fuzhou, China, December 14-17, 2018, Revised Selected Papers*, volume 11449 of *Lecture Notes in Computer Science*, pages 533–553. Springer, 2018.

[15] Eloi de Chérisey, Sylvain Guilley, Olivier Rioul, and Pablo Piantanida. Best Information is Most Successful Mutual Information and Success Rate in Side-Channel Analysis. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2019(2):49–79, 2019.

[16] John F. Dillon, Elementary Hadamard difference sets. PhD thesis, Univ. of Maryland, 1974.

[17] Hans Dobbertin, Construction of bent functions and balanced Boolean functions with high nonlinearity. FSE 1994, LNCS Vol. 1008, 61-74, Springer, Berlin, Heidelberg, 1995.

[18] M. Abdelaziz Elaabid and Sylvain Guilley. Portability of templates. *Journal of Cryptographic Engineering*, 2(1):63–74, 2012. 10.1007/s13389-012-0030-6.

[19] Yunsi Fei, Qiasi Luo, and A. Adam Ding. A Statistical Model for DPA with Novel Algorithmic Confusion Analysis. In Emmanuel Prouff and Patrick Schaumont, editors, *CHES*, volume 7428 of *LNCS*, pages 233–250. Springer, 2012.

[20] Wieland Fischer and Berndt M. Gammel. Masking at Gate Level in the Presence of Glitches. In *CHES*, volume 3659 of *Lecture Notes in Computer Science*, pages 187–200. Springer, August 29 – September 1 2005. Edinburgh, UK.

[21] Sylvain Guilley, Annelie Heuser, and Olivier Rioul. A Key to Success - Success Exponents for Side-Channel Distinguishers. In Alex Biryukov and Vipul Goyal, editors, *Progress in Cryptology - INDOCRYPT 2015, Bangalore, India, December 6-9, 2015, Proceedings*, volume 9462 of *Lecture Notes in Computer Science*, pages 270–290. Springer, 2015.

[22] Sylvain Guilley, Philippe Hoogvorst, and Renaud Pacalet. Differential Power Analysis Model and Some Results. In Jean-Jacques Quisquater, Pierre Paradinas, Yves Deswarte, and Anas Abou El Kalam, editors, *Smart Card Research and Advanced Applications VI, IFIP 18th World Computer Congress, TC8/WG8.8 & TC11/WG11.2 Sixth International Conference on Smart Card Research and Advanced Applications (CARDIS), 22-27 August 2004, Toulouse, France*, volume 153 of *IFIP*, pages 127–142. Kluwer/Springer, 2004.

[23] Annelie Heuser, Olivier Rioul, and Sylvain Guilley. A Theoretical Study of Kolmogorov-Smirnov Distinguishers — Side-Channel Analysis vs. Differential Cryptanalysis. In Emmanuel Prouff, editor, *Constructive Side-Channel Analysis and Secure Design - 5th International Workshop, COSADE 2014, Paris, France, April 13-15, 2014. Revised Selected Papers*, volume 8622 of *Lecture Notes in Computer Science*, pages 9–28. Springer, 2014.

[24] Annelie Heuser, Olivier Rioul, and Sylvain Guilley. Good Is Not Good Enough - Deriving Optimal Distinguishers from Communication Theory. In Lejla Batina and Matthew Robshaw, editors, *Cryptographic Hardware and Embedded Systems - CHES 2014 - 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings*, volume 8731 of *Lecture Notes in Computer Science*, pages 55–74. Springer, 2014.

[25] Selçuk Kavut, Results on rotation-symmetric S-boxes. Information Sciences, 201, 93-113, 2012.

[26] Selçuk Kavut and Melek Diker Yücel, A new algorithm for the design of strong Boolean functions (in Turkish). Proceedings of the First National Cryptology Symposium, pp. 95-105, METU, 2005.

[27] Selçuk Kavut and Melek Diker Yücel. Generalized Rotation Symmetric and Dihedral Symmetric Boolean Functions - 9 Variable Boolean Functions with Nonlinearity 242. In Serdar Boztaş and Hsiao-Feng (Francis) Lu, editors, *17th International Applied Algebra, Algebraic Algorithms, and Error Correcting Codes Symposium, AAECC 2007, Bangalore, India, December 16-20, 2007, Proceedings*, volume 4851 of *Lecture Notes in Computer Science*, pages 321–329. Springer, 2007.

[28] Paul C. Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In Michael J. Wiener, editor, *CRYPTO*, volume 1666 of *Lecture Notes in Computer Science*, pages 388–397. Springer, 1999.

[29] Yongqiang Li and Mingsheng Wang, Constructing differentially 4-uniform permutations over $GF(2^{2m})$ from quadratic APN permutations over $GF(2^{2m+1})$. Des. Codes Cryptogr., 72(2), 249-264, 2014.

[30] Rudolf Lidl and Harald Niederreiter. Finite fields. 1983.

[31] Hongying Liu, Guoyu Qian, Satoshi Goto, and Yukiyasu Tsunoo. Correlation Power Analysis Based on Switching Glitch Model. In Yongwha Chung and Moti Yung, editors, *WISA*, volume 6513 of *Lecture Notes in Computer Science*, pages 191–205. Springer, 2010.

[32] Robert L. McFarland, A family of difference sets in non-cyclic groups. Journal of Combinatorial Theory, Series A, 15(1), 1-10, 1973.

[33] Houssem Maghrebi, Olivier Rioul, Sylvain Guilley, and Jean-Luc Danger. Comparison between Side-Channel Analysis Distinguishers. In Tat Wing Chim and Tsz Hon Yuen, editors, *ICICS*, volume 7618 of *LNCS*, pages 331–340. Springer, 2012.

[34] Stefan Mangard, Elisabeth Oswald, and Thomas Popp. *Power Analysis Attacks: Revealing the Secrets of Smart Cards*. Springer, December 2006. ISBN 0-387-30857-1, http://www.dpabook.org/.

[35] Stefan Mangard and Kai Schramm. Pinpointing the side-channel leakage of masked AES hardware implementations. In Louis Goubin and Mitsuru Matsui, editors, *Cryptographic Hardware and Embedded Systems - CHES 2006, 8th International Workshop, Yokohama, Japan, October 10-13, 2006, Proceedings*, volume 4249 of *Lecture Notes in Computer Science*, pages 76–90. Springer, 2006.

[36] Luca Mariot, Stjepan Picek, Alberto Leporati, and Domagoj Jakobovic. Cellular automata based S-boxes. *Cryptography and Communications* 11(1): 41-62 (2019)

[37] James L. Massey. SAFER K-64: A byte-oriented block-ciphering algorithm. In Ross J. Anderson, editor, *Fast Software Encryption, Cambridge Security Workshop, Cambridge, UK, December 9-11, 1993, Proceedings*, volume 809 of *Lecture Notes in Computer Science*, pages 1–17. Springer, 1993.

[38] Sihem Mesnager. *Bent functions: fundamentals and results*. Springer, ISBN 978-3-319-32593-4, pp. 1-544, August 2016.

[39] Amir Moradi, Sylvain Guilley, and Annelie Heuser. Detecting Hidden Leakages. In Ioana Boureanu, Philippe Owesarski, and Serge Vaudenay, editors, *Applied Cryptography and Network Security - 12th International Conference, ACNS 2014, Lausanne, Switzerland, June 10-13, 2014. Proceedings*, volume 8479 of *Lecture Notes in Computer Science*, pages 324–342. Springer, 2014.

[40] David P. Montminy, Rusty O. Baldwin, Michael A. Temple, and Eric D. Laspe. Improving cross-device attacks using zero-mean unit-variance normalization. *J. Cryptographic Engineering*, 3(2):99–110, 2013.

[41] Maxime Nassar, Youssef Souissi, Sylvain Guilley, and Jean-Luc Danger. RSM: A small and fast countermeasure for AES, secure against 1st and 2nd-order zero-offset SCAs. In Wolfgang Rosenstiel and Lothar Thiele, editors, *2012 Design, Automation & Test in Europe Conference & Exhibition, DATE 2012, Dresden, Germany, March 12-16, 2012*, pages 1173–1178. IEEE, 2012.

[42] Svetla Nikova, Christian Rechberger, and Vincent Rijmen. Threshold Implementations Against Side-Channel Attacks and Glitches. In *ICICS*, volume 4307 of *LNCS*, pages 529–545. Springer, December 4-7 2006. Raleigh, NC, USA.

[43] Svetla Nikova, Vincent Rijmen, and Martin Schläffer. Secure Hardware Implementation of Non-linear Functions in the Presence of Glitches. In *ICISC*, volume 5461 of *Lecture Notes in Computer Science*, pages 218–234. Springer, 2008. Seoul, Korea.

[44] NIST. AES Proposal: Rijndael (now FIPS PUB 197), April 2003. http://csrc.nist.gov/archive/aes/rijndael/Rijndael-ammended.pdf.

[45] Kaisa Nyberg, Differentially uniform mappings for cryptography. EUROCRYPT'93, LNCS Vol. 765, 55-64, Springer, Berlin, Heidelberg, 1994.

[46] Stjepan Picek, Bodhisatwa Mazumdar, Debdeep Mukhopadhyay, and Lejla Batina. Modified Transparency Order Property: Solution or Just Another Attempt. In Rajat Subhra Chakraborty, Peter Schwabe, and Jon A. Solworth, editors, *Security, Privacy, and Applied Cryptography Engineering - 5th International Conference, SPACE 2015, Jaipur, India, October 3-7, 2015, Proceedings*, volume 9354 of *Lecture Notes in Computer Science*, pages 210–227. Springer, 2015.

[47] Emmanuel Prouff. DPA Attacks and S-Boxes. In Henri Gilbert and Helena Handschuh, editors, *FSE*, volume 3557 of *Lecture Notes in Computer Science*, pages 424–441. Springer, 2005.

[48] Emmanuel Prouff, editor. *Constructive Side-Channel Analysis and Secure Design - 5th International Workshop, COSADE 2014, Paris, France, April 13-15, 2014. Revised Selected Papers*, volume 8622 of *Lecture Notes in Computer Science*. Springer, 2014.

[49] Oscar Reparaz, Benedikt Gierlichs, and Ingrid Verbauwhede. A note on the use of margins to compare distinguishers. In Prouff [48], pages 1–8.

[50] Rijmen, Vincent, Barreto, Paulo S. L. M., and Gazzoni Filhob, Dcio L. Rotation symmetry in algebraically generated cryptographic substitution tables. Inf. Process. Lett., 106(6), 246-250, 2008.

[51] Claude E. Shannon. Communication theory of secrecy systems. *Bell system technical journal*, 28, pp. 656-715, 1949.

[52] Werner Schindler, Kerstin Lemke, and Christof Paar. A Stochastic Model for Differential Side Channel Cryptanalysis. In LNCS, editor, *CHES*, volume 3659 of *LNCS*, pages 30–46. Springer, Sept 2005. Edinburgh, Scotland, UK.

[53] François-Xavier Standaert, Tal Malkin, and Moti Yung. A Unified Framework for the Analysis of Side-Channel Key Recovery Attacks. In *EUROCRYPT*, volume 5479 of *LNCS*, pages 443–461. Springer, April 26-30 2009. Cologne, Germany.

[54] Nicolas Veyrat-Charvillon and François-Xavier Standaert. Mutual Information Analysis: How, When and Why? In Christophe Clavier and Kris Gaj, editors, *Cryptographic Hardware and Embedded Systems - CHES 2009, 11th International Workshop, Lausanne, Switzerland, September 6-9, 2009, Proceedings*, volume 5747 of *Lecture Notes in Computer Science*, pages 429–443. Springer, 2009.

[55] Carolyn Whitnall and Elisabeth Oswald. A Comprehensive Evaluation of Mutual Information Analysis Using a Fair Evaluation Framework. In Phillip Rogaway, editor, *CRYPTO*, volume 6841 of *Lecture Notes in Computer Science*, pages 316–334. Springer, 2011.

[56] Carolyn Whitnall and Elisabeth Oswald. A Fair Evaluation Framework for Comparing Side-Channel Distinguishers. *J. Cryptographic Engineering*, 1(2):145–160, 2011.

[57] Xian-Mo Zhang and Yulian Zheng, GAC − the criterion for global avalanche characteristics of cryptographic functions. Journal of Universal Computer Science 1(5), 320-337, 1995.

[58] Hui Zhao, Yongbin Zhou, François-Xavier Standaert, and Hailong Zhang. Systematic Construction and Comprehensive Evaluation of Kolmogorov-Smirnov Test Based Side-Channel Distinguishers. In Robert H. Deng and Tao Feng, editors, *ISPEC*, volume 7863 of *Lecture Notes in Computer Science*, pages 336–352. Springer, 2013.

APPENDIX

A. Truth Table for the four Found RSSB S-boxes $\mathbb{F}_2^6 \to \mathbb{F}_2^6$

S-box # 1:

0, 3, 6, 29, 12, 8, 58, 38, 24, 27, 16, 11, 53, 49, 13, 17, 48, 2, 54, 28, 32, 21, 22, 59, 43, 25, 35, 9, 26, 47, 34, 15, 33, 46, 4, 19, 45, 37, 56, 40, 1, 14, 42, 61, 44, 36, 55, 39, 23, 41, 50, 20, 7, 62, 18, 51, 52, 10, 31, 57, 5, 60, 30, 63

S-box # 2:

0, 10, 20, 48, 40, 26, 33, 61, 17, 27, 52, 16, 3, 49, 59, 39, 34, 38, 54, 28, 41, 21, 32, 50, 6, 2, 35, 9, 55, 11, 15, 29, 5, 24, 13, 62, 45, 8, 56, 51, 19, 14, 42, 25, 1, 36, 37, 46, 12, 31, 4, 57, 7, 44, 18, 23, 47, 60, 22, 43, 30, 53, 58, 63

S-box # 3:

0, 10, 20, 29, 40, 26, 58, 11, 17, 27, 52, 61, 53, 7, 22, 39, 34, 38, 54, 49, 41, 21, 59, 4, 43, 47, 14, 9, 44, 16, 15, 48, 5, 46, 13, 37, 45, 62, 35, 51, 19, 56, 42, 2, 55, 36, 8, 24, 23, 50, 31, 57, 28, 1, 18, 12, 25, 60, 32, 6, 30, 3, 33, 63

S-box # 4:

0, 10, 20, 58, 40, 31, 53, 38, 17, 27, 62, 16, 43, 28, 13, 30, 34, 55, 54, 7, 61, 21, 32, 44, 23, 2, 56, 9, 26, 50, 60, 48, 5, 29, 47, 19, 45, 8, 14, 15, 59, 35, 42, 22, 1, 36, 25, 24, 46, 41, 4, 39, 49, 11, 18, 12, 52, 51, 37, 6, 57, 3, 33, 63