On the behaviors of affine equivalent Sboxes regarding differential and linear attacks*

Anne Canteaut and Joëlle Roué

Inria, project-team SECRET, France
Anne.Canteaut@inria.fr, Joelle.Roue@inria.fr

Abstract. This paper investigates the effect of affine transformations of the Sbox on the maximal expected differential probability MEDP and linear potential MELP over two rounds of a substitution-permutation network, when the diffusion layer is linear over the finite field defined by the Sbox alphabet. It is mainly motivated by the fact that the 2-round MEDP and MELP of the AES both increase when the AES Sbox is replaced by the inversion in $\mathbf{F}_{2^8}$. Most notably, we give new upper bounds on these two quantities which are not invariant under affine equivalence. Moreover, within a given equivalence class, these new bounds are maximal when the considered Sbox is an involution. These results point out that different Sboxes within the same affine equivalence class may lead to different two-round MEDP and MELP. In particular, we exhibit some examples where the basis chosen for defining the isomorphism between $\mathbf{F}_2^m$ and $\mathbf{F}_{2^m}$ affects these values. For Sboxes with some particular properties, including all Sboxes of the form $A(x^s)$ as in the AES, we also derive some lower and upper bounds for the 2-round MEDP and MELP which hold for any MDS linear layer.
Keywords. Sboxes, affine equivalence, differential cryptanalysis, linear cryptanalysis, AES.

1 Introduction

Cryptographic functions, including the so-called Sboxes, are usually classified up to affine equivalence (see *e.g.* [6, 34, 11]) since many of the relevant cryptographic properties are invariant under affine transformations. Indeed, both Sboxes S and $A_2 \circ S \circ A_1$, where A_1 and A_2 are two affine permutations, have the same algebraic degree, the same non-linearity (even the same square Walsh spectrum) and the same differential uniformity (even the same differential spectrum), which are the usual criteria measuring the resistance of an Sbox against higher-order differential attacks [29, 31], linear cryptanalysis [44, 37] and differential cryptanalysis [5] respectively. However, it is well-known that equivalent Sboxes may have different implementation costs and may also provide different security levels. For instance, the number of terms in their polynomial representations may highly vary within an equivalent class. This has motivated the choice of the AES

* Partially supported by the French Agence Nationale de la Recherche through the BLOC project under Contract ANR-11-INS-011.

Sbox: it corresponds to the inversion in $\mathbf{F}_{2^8}$, which is a power permutation with the best known resistance against the previously mentioned attacks; but this power permutation is then composed with an $\mathbf{F}_2$-affine permutation of $\mathbf{F}_2^8$ which makes its polynomial representation much more complex. Composing the inverse function with an affine permutation then thwarts potential attacks exploiting a simple algebraic representation of the Sbox, like an extremely sparse polynomial. Some other relevant properties (usually of minor importance) are also affected by composition with affine transformations, like the number of fixed points and the bitwise branch number [43].

But, when focusing on statistical attacks, especially on differential and linear cryptanalyses, the Sboxes within the same equivalence class are often considered to have similar behaviors. The main reason is that all known upper bounds on the maximal expected differential probability, and on the maximal expected square correlation (aka maximum expected linear potential) [41] are invariant under the affine transformations of the Sbox. However, the exact values of these two quantities for two rounds of the AES have been computed by Keliher and Sui with a sophisticated pruning algorithm [28], and it appears that the values obtained for the multiplicative inverse in $\mathbf{F}_{2^8}$ and for the original AES Sbox are different, while these two Sboxes belong to the same equivalence class. Going further in the analysis, Daemen and Rijmen have then determined the expected probabilities of all two-round differentials with 5 or 6 active Sboxes in the AES for both Sboxes [20]. After this analysis, they have even conjectured that, for any number of rounds, the maximal expected differential probability of the AES is always higher with the inversion in $\mathbf{F}_{2^8}$ than with the AES Sbox [17]. The aim of this paper is then to have a better understanding of this phenomenon. For instance, we would like to determine whether these different behaviors originate from the Sboxes only, independently of the choice of the diffusion layer, or not. One of our main motivations is to help the designers choose an Sbox within a given equivalence class. Indeed, in most situations, some appropriate equivalence classes are known (e.g. 4-bit permutations are classified up to affine equivalence [34]) and the search is often restricted to these classes.

Our contribution. In this paper, we investigate the maximal expected differential probability MEDP and linear potential MELP over two rounds of an SPN. We focus on diffusion layers which are linear over the field of size 2^m, where m is the number of bits of the Sbox, exactly as in the AES and several other ciphers like LED [25], KLEIN [24], mCrypton [35], Prøst [27]... We give a new upper bound on the two-round MEDP and MELP which supersedes the best previous result [41], and which is not invariant under affine equivalence. This result is combined with the lower bounds corresponding to some minimum-weight differentials (or linear masks). We are then able to exhibit different behaviors regarding differential and linear attacks on two rounds depending on the choice of the Sbox within a given equivalence class. This includes some unexpected differences since we point out that, for a given m-bit Sbox, the choice of the basis used for defining the finite field in the description of the linear layer may also affect

the value of the two-round MEDP or MELP. This is due to the $\mathbf{F}_{2^m}$-linearity of the mixing layer.

More interestingly from the designers' viewpoint, for some classical families of Sboxes including all functions of the form $A(x^s)$ or $(A(x))^s$ where A is an affine function, like the AES Sbox, our results yield some lower and upper bounds on the two-round MEDP and MELP which are independent of the choice of the MDS linear layer. In other words, we show that, for these families of Sboxes, the two-round MEDP and MELP are two quantities which essentially depend on the Sbox only. Therefore, the designer can choose an Sbox and get a precise estimation of the corresponding two-round MEDP and MELP, while all previous methods [28] involved the specifications of both the Sbox and the diffusion layer together. As an illustration, we prove that the previously known upper bounds on MEDP_2 and MELP_2 due to Park et al. [41] are always tight for the multiplicative inverse over $\mathbf{F}_{2^m}$ and for any MDS linear layer. In other words, the inversion is the mapping within its equivalence class which has the highest two-round MEDP and MELP, independently of the choice of the MDS linear layer. This situation mainly originates from the fact that this Sbox is an involution.

2 Maximum Expected Differential Probability and Linear Potential for Substitution-Permutation Networks

2.1 Substitution-Permutation Networks

One of the most widely-used constructions for iterated block ciphers is the so-called key-alternating construction [15, 18] (aka iterated Even-Mansour construction), which consists of an alternation of key-independent (usually similar) permutations and of round-key additions. The round permutation usually follows the principles introduced by Shannon. It is decomposed into a nonlinear substitution function Sub which provides confusion, and a linear permutation which provides diffusion[1]. In order to reduce the implementation cost of the substitution layer, which is usually the most expensive part of the cipher in terms of circuit complexity, a usual choice for Sub consists in concatenating several copies of a permutation S which operates on a much smaller alphabet. In the whole paper, we will concentrate on such block ciphers, and use the following notation to describe the corresponding round permutation.

Definition 1. *Let m and t be two positive integers. Let S be a permutation of $\mathbf{F}_2^m$ and M be a linear permutation of $\mathbf{F}_2^{mt}$. Then, $\mathrm{SPN}(m, t, S, M)$ denotes any substitution-permutation network defined over $\mathbf{F}_2^{mt}$ whose substitution function consists of the concatenation of t copies of S and whose diffusion function corresponds to M.*

[1] Here, the terminology *substitution-permutation* has to be understood in a broad sense without any restriction on the linear permutation, while in some other papers, it is limited to the class of bit permutations.

For instance, up to a linear transformation, two rounds of the AES can be seen as the concatenation of four similar *superboxes* [20]. The superbox, depicted on Fig. 1, is linearly equivalent to a two-round permutation of the form $\mathrm{SPN}(8, 4, S, M)$ where the AES Sbox S corresponds to the composition of the inversion in $\mathbf{F}_{2^8}$ with an affine permutation A. More precisely, $S(x) = A \circ \varphi^{-1}\left(\varphi(x)^{254}\right)$ where φ is the isomorphism from $\mathbf{F}_2^8$ into $\mathbf{F}_{2^8}$ defined by the basis $\{1, \alpha, \alpha^2, \ldots, \alpha^7\}$ with α a root of $X^8 + X^4 + X^3 + X + 1$.

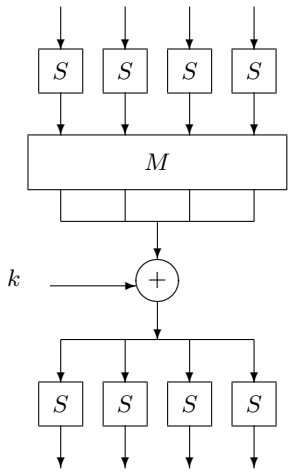

Fig. 1. The AES superbox.

Differential [5] and linear [44, 37] cryptanalyses are the most prominent statistical attacks. The complexity of differential attacks depends critically on the distribution over the keys k of the probability of the differentials (a, b), *i.e.*,

$$\mathrm{DP}(a, b) = \Pr_X[E_k(X) + E_k(X + a) = b]$$

where E_k corresponds to the (possibly round-reduced) encryption function under key k. This probability may highly vary with the key especially when a small number of rounds is considered (see *e.g.* [32], [19, Section 8.7.2], [21], [22] and [7]). But computing the whole distribution of the probability of a differential is a very difficult task, and cryptanalysts usually focus on its expectation.

Definition 2. *Let $(E_k)_{k \in \mathbf{F}_2^\kappa}$ be an r-round iterated cipher with key-size κ. Then, the* expected probability of an r-round differential (a, b) *is*

$$\mathrm{EDP}_r^E(a, b) = 2^{-\kappa} \sum_{k \in \mathbf{F}_2^\kappa} \Pr_X[E_k(X) + E_k(X + a) = b] \ .$$

The maximum expected differential probability for r rounds *is*

$$\mathrm{MEDP}_r^E = \max_{a \neq 0, b} \mathrm{EDP}_r^E(a, b) \ .$$

The index in MEDP_r^E will be omitted when the number of rounds is not specified. It is worth noticing that the MEDP is relevant for estimating the resistance against classical differential cryptanalysis, not against its variants like truncated differential attacks (which provide better attacks on the AES since the AES resists differential cryptanalysis by design).

Similarly, the resistance of a cipher against linear cryptanalysis can be evaluated by determining the distribution over the keys of the correlation of each r-round mask (u, v):

$$\mathrm{C}(u, v) = 2^{-n} \sum_{x \in \mathbf{F}_2^n} (-1)^{u \cdot x + v \cdot E_k(x)},$$

where n is the block-size. For a key-alternating cipher with independent round keys, the average over all keys of the correlation $C(u, v)$ is zero for any nonzero mask (u, v) (see e.g. [19, Section 7.9] or [1, Prop. 1]). Then, the major parameter investigated in this paper is the variance of the distribution of the correlation, which corresponds to the average square correlation. The way it affects the complexity of linear cryptanalysis is discussed for instance in [40, 19, 38, 33, 1].

Definition 3. *Let $(E_k)_{k \in \mathbf{F}_2^\kappa}$ be an r-round iterated cipher with block-size n and key-size κ. Then, the* expected square correlation (aka linear potential [40]) of an r-round mask (u, v) *is*

$$\mathrm{ELP}_r^E(u, v) = 2^{-2n-\kappa} \sum_{k \in \mathbf{F}_2^\kappa} \left(\sum_{x \in \mathbf{F}_2^n} (-1)^{u \cdot x + v \cdot E_k(x)} \right)^2 .$$

The maximum expected square correlation for r rounds *is*

$$\mathrm{MELP}_r^E = \max_{u, v \neq 0} \mathrm{ELP}_r^E(u, v).$$

2.2 Known results on two-round MEDP and MELP

Computing the MEDP and MELP for an SPN, even for a small number of rounds, is usually non-trivial. An easier task consists in computing the expected probability of an r-round differential characteristic (*i.e.*, a collection of $(r+1)$ differences), or the expected square correlation of a linear trail (*i.e.*, a collection of $(r + 1)$ linear masks). In particular, a simple upper bound on this quantity can be derived from the differential uniformity [39] (resp. the nonlinearity) of the Sbox, and from the differential (resp. linear) branch number of the linear layer. We will then extensively use the following notation for these quantities.

Definition 4. *Let S be a function from $\mathbf{F}_2^m$ into $\mathbf{F}_2^m$.*

- *For any a and b in $\mathbf{F}_2^m$, we define*

$$\delta^S(a, b) = \#\{x \in \mathbf{F}_2^m, S(x + a) + S(x) = b\} .$$

The multi-set $\{\delta^S(a, b),\ a, b \in \mathbf{F}_2^m\}$ is the differential spectrum *of S and its maximum $\Delta(S) = \max_{a \neq 0, b} \delta^S(a, b)$ is the* differential uniformity *of S.*

– *For any u and v in $\mathbf{F}_2^m$, we define*

$$\mathcal{W}^S(u,v) = \sum_{x \in \mathbf{F}_2^m} (-1)^{u \cdot x + v \cdot S(x)} ,$$

where $\cdot$ is the usual scalar product in $\mathbf{F}_2^m$. The multi-set $\{\mathcal{W}^S(u,v),\ u \in \mathbf{F}_2^m, v \in \mathbf{F}_2^m\}$ is the Walsh spectrum *of S, and its highest magnitude $\mathcal{L}(S) = \max_{u,v \neq 0} \left|\mathcal{W}^S(u,v)\right|$ is the* linearity *of S.*

The branch number of the diffusion layer then determines the minimum number of active Sboxes within a differential or linear trail.

Definition 5. *[15] Let M be an $\mathbf{F}_2$-linear permutation of $(\mathbf{F}_2^m)^t$. We associate with M the codes $\mathcal{C}_M$ and $\mathcal{C}_M^\perp$ of length $2t$ and size 2^t over $\mathbf{F}_2^m$ defined by*

$$\mathcal{C}_M = \{(c, M(c)),\ c \in (\mathbf{F}_2^m)^t\} \text{ and } \mathcal{C}_M^\perp = \{(M^*(c), c),\ c \in (\mathbf{F}_2^m)^t\} ,$$

where M^* is the adjoint of M, i.e.*, the linear map such that $x \cdot M(y) = M^*(x) \cdot y$ for any (x,y). The* differential branch number *(resp.* linear branch number*) of M is the minimum distance of the code $\mathcal{C}_M$ (resp. of $\mathcal{C}_M^\perp$).*

From Singleton's bound, the maximum branch number of M is $(t+1)$ and is achieved when $\mathcal{C}_M$ is MDS. Since the codes $\mathcal{C}_M$ and $\mathcal{C}_M^\perp$ are dual to each other, M has optimal differential branch number if and only if it has optimal linear branch number. A simple upper bound for both the two-round MEDP and MELP can then be derived from the branch numbers of M, and from the differential uniformity and the linearity of the Sbox (see [26] and [19, Section B.2]):

$$\mathrm{MEDP}_2 \leq \left(2^{-m} \Delta(S)\right)^t \text{ and } \mathrm{MELP}_2 \leq \left(2^{-m} \mathcal{L}(S)\right)^{2t} . \tag{1}$$

This result has then be refined in [14, 41].

Theorem 1 (FSE 2003 bounds). *[41, 14] Let E be a block cipher of the form $\mathrm{SPN}(m,t,S,M)$ where M is a linear permutation with differential (resp. linear) branch number d (resp. $d^\perp$). Then, we have*

$$\mathrm{MEDP}_2^E \leq 2^{-md} \max \left(\max_{a \in (\mathbf{F}_2^m)^*} \sum_{\gamma \in (\mathbf{F}_2^m)^*} \delta^S(a,\gamma)^d, \max_{b \in (\mathbf{F}_2^m)^*} \sum_{\gamma \in (\mathbf{F}_2^m)^*} \delta^S(\gamma,b)^d \right) ,$$

$$\mathrm{MELP}_2^E \leq 2^{-2md^\perp} \max \left(\max_{u \in (\mathbf{F}_2^m)^*} \sum_{\gamma \in (\mathbf{F}_2^m)^*} \mathcal{W}^S(u,\gamma)^{2d^\perp}, \max_{v \in (\mathbf{F}_2^m)^*} \sum_{\gamma \in (\mathbf{F}_2^m)^*} \mathcal{W}^S(\gamma,v)^{2d^\perp} \right)$$

It is worth noticing that the FSE 2003 bounds always supersede (1).

The main question is now to determine the gap between the FSE 2003 bounds and the exact values of MEDP_2 and MELP_2 for a given cipher. An interesting property is that the FSE 2003 bound is invariant under affine equivalence, *i.e.*, under left or right composition of the Sbox with an affine permutation. Actually, the following well-known property holds.

Lemma 1. *Let S be permutation of $\mathbf{F}_2^m$ and A_1 and A_2 be two affine permutations of $\mathbf{F}_2^m$. Then $S' = A_2 \circ S \circ A_1$ satisfies*

$$\delta^{S'}(a,b) = \delta^S(L_1(a), L_2^{-1}(b)) \text{ and } \mathcal{W}^{S'}(u,v)^2 = \mathcal{W}^S((L_1^{-1})^*(u), L_2^*(v))^2$$

where L_1 and L_2 correspond to the linear parts of A_1 and A_2, and L^* denotes the adjoint of L.

However, while the previous bounds are invariant under affine equivalence, it appears that the exact values of MEDP_2 and MELP_2 may vary when the Sbox is composed with an affine permutation. For instance, for the AES with its original Sbox, the exact values of the two-round MEDP_2 and MELP_2 have been computed by a pruning search algorithm [28]: $\mathrm{MEDP}_2 = 53 \times 2^{-34}$ and $\mathrm{MELP}_2 = 109,953,193 \times 2^{-54} \approx 1.638 \times 2^{-28}$. But, if the AES Sbox is replaced by the so-called *naive Sbox* [17], obtained by removing the affine permutation from the AES Sbox, $\mathrm{MEDP}_2 = 79 \times 2^{-34}$ [20] which corresponds to the FSE 2003 bound. To our best knowledge, the exact value of MELP_2 for the naive Sbox has not been computed, but we will deduce from our results in Section 4.3 that, for the multiplicative inverse over $\mathbf{F}_{2^m}$ and any MDS $\mathbf{F}_{2^m}$-linear layer, the FSE 2003 bound is always tight. In particular, for $m = 8$, $\mathrm{MELP}_2 = 192,773,764 \times 2^{-54} \approx 2.873 \times 2^{-28}$. Then, the AES Sbox provides a better resistance against differential and linear cryptanalyses for two rounds of the AES than the naive Sbox. More generally, it has been conjectured in [17, Conjecture 1] that, for any number of rounds r, MEDP_r is smaller for the AES Sbox than for the naive Sbox.

2.3 SPNs over $\mathbf{F}_{2^m}$

A special case of affine equivalent Sboxes corresponds to the mappings over $\mathbf{F}_2^m$ which are derived from the same function over the finite field $\mathbf{F}_{2^m}$, but from different correspondences between $\mathbf{F}_{2^m}$ and the vector space $\mathbf{F}_2^m$. Such equivalent Sboxes appear in several situations. Indeed, a simple construction for an optimal linear layer consists in choosing for M a permutation of $\mathbf{F}_2^{mt}$ associated with a code $\mathcal{C}_M$ which is linear over the field $\mathbf{F}_{2^m}$, where m is the size of the Sbox. Then, this diffusion layer has to be defined over $\mathbf{F}_{2^m}^t$, instead of $\mathbf{F}_2^{mt}$. To this end, we need to identify the vector space $\mathbf{F}_2^m$ with the finite field $\mathbf{F}_{2^m}$ by the means of an isomorphism φ associated to a basis $(\alpha_0, \ldots, \alpha_{m-1})$, namely:

$$\begin{array}{rccc} \varphi : & \mathbf{F}_2^m & \to & \mathbf{F}_{2^m} \\ & (x_0, \ldots, x_{m-1}) & \mapsto & \sum_{i=0}^{m-1} x_i \alpha_i \ . \end{array}$$

Then, both the Sbox and the diffusion layer can be represented as functions over the field $\mathbf{F}_{2^m}$ by

$$\mathcal{S} = \varphi \circ S \circ \varphi^{-1} \text{ and } \mathcal{M} = \widetilde{\varphi} \circ M \circ \widetilde{\varphi}^{-1} \ ,$$

where $\widetilde{\varphi}$ is the concatenation of t copies of φ. In this case, as noticed in [19, Section A.5], any r rounds of $\mathrm{SPN}(m,t,S,M)$ can be written as $\widetilde{\varphi}^{-1} \circ \mathrm{Add}_{k_r} \circ \ldots \circ$

$\mathcal{R} \circ \mathrm{Add}_{k_1} \circ \mathcal{R} \circ \mathrm{Add}_{k_0} \circ \widetilde{\varphi}$ where the round function $\mathcal{R} = \mathcal{M} \circ (\mathcal{S}, \ldots, \mathcal{S})$ is a permutation of $(\mathbf{F}_{2^m})^t$ and Add_x denotes the addition of x in $(\mathbf{F}_{2^m})^t$. Obviously, composing by $\widetilde{\varphi}$ at the beginning and by $\widetilde{\varphi}^{-1}$ at the end changes neither the MEDP nor the MELP. This implies that MEDP_r^E and MELP_r^E depend on $\mathcal{M}$ and $\mathcal{S}$ only, *i.e.*, on the representations of the Sbox and of the diffusion layer over $\mathbf{F}_{2^m}$. In particular, *the choice of the basis* $(\alpha_0, \ldots, \alpha_{m-1})$ *has no influence on the differential and linear properties of the cipher*. For this reason, we use the following alternative notation for defining an SPN from these representations.

Definition 6. *Let m and t be two positive integers. Let $\mathcal{S}$ be a permutation of $\mathbf{F}_{2^m}$ and $\mathcal{M}$ be a permutation of $(\mathbf{F}_{2^m})^t$ which is linear over $\mathbf{F}_{2^m}$. Then, we denote by* $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$ *a substitution-permutation network defined over $(\mathbf{F}_{2^m})^t$ whose substitution function consists of the concatenation of t copies of $\mathcal{S}$ and whose diffusion function corresponds to $\mathcal{M}$.*

For the sake of clarity, all quantities related to the representation in the field $\mathbf{F}_{2^m}$ will be indexed by F, and all functions defined over $\mathbf{F}_{2^m}$ will be denoted by calligraphic letters. As pointed out in [23], the differential and linear properties of any $\mathrm{SPN}(m, t, S, M)$ can be equivalently studied by considering the alternative representation $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$. This alternative analysis then involves the differential spectrum and the Walsh spectrum of the Sbox $\mathcal{S}$ over $\mathbf{F}_{2^m}$, which are related to the spectra of the corresponding function S over $\mathbf{F}_2^m$ as follows.

Proposition 1. *(see* e.g. *[23]) Let $(\alpha_0, \ldots, \alpha_{m-1})$ be a basis of $\mathbf{F}_{2^m}$, and φ the corresponding isomorphism from $\mathbf{F}_2^m$ into $\mathbf{F}_{2^m}$. Let S be a mapping over $\mathbf{F}_2^m$, and $\mathcal{S} = \varphi \circ S \circ \varphi^{-1}$. Then, for any $(\alpha, \beta) \in \mathbf{F}_{2^m}$,*

$$\delta_F^{\mathcal{S}}(\alpha, \beta) = \#\{x \in \mathbf{F}_{2^m}, \mathcal{S}(x + \alpha) + \mathcal{S}(x) = \beta\} = \delta^S(\varphi^{-1}(\alpha), \varphi^{-1}(\beta))$$
$$\mathcal{W}_F^{\mathcal{S}}(\alpha, \beta) = \sum_{x \in \mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(\alpha x + \beta \mathcal{S}(x))} = \mathcal{W}^S(\psi^{-1}(\alpha), \psi^{-1}(\beta))$$

where ψ is the isomorphism from $\mathbf{F}_2^m$ into $\mathbf{F}_{2^m}$ defined by the dual basis, i.e., *the basis $(\beta_0, \ldots, \beta_{m-1})$ such that $\mathrm{Tr}(\alpha_i \beta_j) = 0$ if $i \neq j$ and $\mathrm{Tr}(\alpha_i \beta_i) = 1$.*

3 New upper bounds on the 2-round MEDP and MELP

Now, we study the exact values of the two-round MEDP and MELP for any cipher of the form $\mathrm{SPN}(m, t, S, M)$ *where the diffusion layer M is linear over $\mathbf{F}_{2^m}$*, like in the AES. We aim at obtaining a better approximation of the MEDP_2 and MELP_2 by finding some improved lower and upper bounds. In particular, we would like to be able to differentiate affine equivalent Sboxes.

3.1 The new upper bounds

From now on, when considering a cipher of the form $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$, $\delta_F(\alpha, \beta)$ and $\mathcal{W}_F(\alpha, \beta)$ will denote the differential and Walsh spectra of the Sbox $\mathcal{S}$. The considered Sbox will be mentioned in the notation in case of ambiguity only.

Theorem 2. *Let E be a block cipher of the form $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$ where $\mathcal{M}$ is linear over $\mathbf{F}_{2^m}$ and has differential (resp. linear) branch number d (resp. $d^\perp$). For $\mu \in \mathbf{F}_{2^m}$ and $u > 0$, we define*

$$\mathcal{B}_u(\mu) = \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \delta_F(\alpha,\gamma)^u \delta_F(\gamma\lambda+\mu,\beta)^{(d-u)} , \tag{2}$$

$$\mathcal{B}_u^\perp(\mu) = \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \mathcal{W}_F(\alpha,\gamma)^{2u} \mathcal{W}_F(\gamma\lambda+\mu,\beta)^{2(d^\perp-u)} , \tag{3}$$

$$\mathcal{B}(\mu) = \max_{1 \le u < d} \mathcal{B}_u(\mu) \text{ and } \mathcal{B}^\perp(\mu) = \max_{1 \le u < d^\perp} \mathcal{B}_u^\perp(\mu) .$$

Then,

$$\mathrm{MEDP}_2^E \le 2^{-md} \max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}(\mu) \text{ and } \mathrm{MELP}_2^E \le 2^{-2md^\perp} \max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}^\perp(\mu) .$$

The proof is given in Appendix A. It mainly exploits the special form of the codewords in an $\mathbf{F}_{2^m}$-linear code (see Lemma 2 in Appendix A). In the whole paper, the proofs in the context of differential attacks and of linear attacks are similar. Actually, all results can be written in a more generic way, by replacing the $2^m \times 2^m$ matrix with coefficients $2^{-m}\delta_F(\alpha,\beta)$, $\alpha,\beta \in \mathbf{F}_{2^m}$, or with coefficients $2^{-2m}\mathcal{W}_F(\alpha,\beta)^2$, by any matrix $(\Lambda(\alpha,\beta))_{\alpha,\beta \in \mathbf{F}_{2^m}}$, such that the coefficients $\Lambda(\alpha,\beta)$ lie between 0 and 1 and satisfy $\Lambda(\alpha,0) = \Lambda(0,\alpha) = 0$ for any nonzero α and $\sum_{\beta \in \mathbf{F}_{2^m}} \Lambda(\alpha,\beta) = \sum_{\beta \in \mathbf{F}_{2^m}} \Lambda(\beta,\alpha) = 1$ for all $\alpha \in \mathbf{F}_{2^m}$. Clearly, both normalized differential and Walsh spectra satisfy these conditions.

Computing this new bound is obviously more expensive than computing the FSE 2003 bound since we have to take the maximum of a similar quantity over all λ and μ. We will see in Section 4 that this bound simplifies in some cases, for instance for all Sboxes corresponding to the composition of a power permutation with an affine mapping, like the AES Sbox. Also, we will show that this refined bound may enable us to deduce the exact values of the MEDP_2 and MELP_2 in a much more efficient way than the ad hoc search algorithm presented in [28].

In the case of the AES Sbox over $\mathbf{F}_{2^8}$ and $d = d^\perp = 5$, these new bounds lead to $\mathrm{MEDP}_2 \le 55.5 \times 2^{-34}$ instead of $\mathrm{MEDP}_2 \le 79 \times 2^{-34}$ for the FSE 2003 bound, and $\mathrm{MELP}_2 \le 31,231,767 \times 2^{-52}$ instead of $\mathrm{MELP}_2 \le 48,193,441 \times 2^{-52}$. This seems to be a minor improvement since there is only a factor $\rho \simeq 0.7$ between the two bounds. However, in AES-like constructions, the 2-round MEDP and MELP correspond to the average differential uniformity and linearity of the average superbox. Upper-bounds on the 4-round MEDP and MELP can then be derived from these values using (1). Then, we get a factor ρ^{d-1} (resp. $\rho^{d^\perp-1}$) between the bounds on MEDP_4 and MELP_4.

While the FSE 2003 bound corresponds to the highest d-th power moment of a row or a column in the difference table of the Sbox (or in the square correlation table), our new bound involves together a row and a column in the table. In other words, this new bound depends on the link between some quantity (*e.g.* a derivative or the squared Walsh transform of a component) for $\mathcal{S}$, and the same

quantity for the inverse permutation $\mathcal{S}^{-1}$. This clearly appears when $\mathcal{S}$ has a two-valued differential spectrum, since the expression of $\mathcal{B}(\mu)$ simplifies to

$$\mathcal{B}(\mu) = \Delta(\mathcal{S})^d \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \#\left(\mathsf{Im}(D_\alpha \mathcal{S}) \cap \left[\lambda \mathsf{Im}(D_\beta \mathcal{S}^{-1}) + \mu\right]\right) ,$$

where $D_\alpha \mathcal{S}$ denotes the derivative of $\mathcal{S}$ at point α, *i.e.*, the function $x \mapsto \mathcal{S}(x+\alpha) + \mathcal{S}(x)$. Similarly, if $\mathcal{S}$ is a plateaued function [45], *i.e.*, a function whose Walsh spectrum contains the values 0 and $\pm\mathcal{L}(S)$ only, we get

$$\mathcal{B}^\perp(\mu) = \mathcal{L}(\mathcal{S})^{2d^\perp} \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \#\left(\mathsf{Supp}(\widehat{\mathcal{S}_\alpha^{-1}}) \cap \left[\lambda \mathsf{Supp}(\widehat{\mathcal{S}_\beta}) + \mu\right]\right) ,$$

where $\mathcal{S}_\alpha$ denotes the Boolean function $x \mapsto \mathrm{Tr}(\alpha \mathcal{S}(x))$, and $\widehat{f}$ denotes the Walsh transform of f, *i.e.*, $\alpha \mapsto \sum_{x \in \mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(f(x)+\alpha x)}$. It appears from these formulas that the cardinality of the intersection of such sets cannot exceed the cardinality of each set (equal to $2^m/\Delta(\mathcal{S})$ and $2^{2m}/\mathcal{L}(\mathcal{S})^2$ respectively), and that this maximum is obviously tight when $\mathcal{S}$ is an involution, *i.e.*, $\mathcal{S}^{-1} = \mathcal{S}$. But when the Sbox is composed with a randomly chosen affine permutation, the two sets can be considered as independent. Then, the expected cardinality of their intersection is about $2^m \pi_\Delta^2 = 2^m/\Delta(\mathcal{S})^2$ (resp. $2^m \pi_\mathcal{L}^2 = 2^{3m}/\mathcal{L}(\mathcal{S})^4$) where $\pi_\Delta = 1/\Delta(\mathcal{S})$ is the proportion of nonzero elements within a row or a column of the difference table, and $\pi_\mathcal{L} = 2^m/\mathcal{L}(\mathcal{S})^2$ is the proportion of nonzero elements within a row or a column of the square correlation table. For instance, for an almost bent Sbox, i.e., with m odd, $\Delta(\mathcal{S}) = 2$ and $\mathcal{L}(\mathcal{S}) = 2^{(m+1)/2}$, the expected cardinality of the two sets involved in the previous formulas is 2^{m-2}, while it is equal to 2^{m-1} when $\mathcal{S}$ is an involution. More generally, our new upper bound is always smaller than or equal to the corresponding FSE 2003 bound, with equality when $\mathcal{S}$ is an involution, as stated in the following proposition (see the proof in Appendix A).

Proposition 2. *Let $\mathcal{S}$ be a permutation of $\mathbf{F}_{2^m}$ and d be some positive integer. Then, each of the two upper bounds defined in Theorem 2 is less than or equal to the corresponding FSE 2003 bound. Moreover, equality holds if $\mathcal{S}$ is an involution, since for any integer $u < d$,*

$$\max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}_u(\mu) = \mathcal{B}_u(0) = \max_{a \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \delta_F(a,\gamma)^d = \max_{b \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \delta_F(\gamma,b)^d$$

and $$\max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}_u^\perp(\mu) = \mathcal{B}_u^\perp(0) = \max_{a \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \mathcal{W}_F(a,\gamma)^{2d} = \max_{b \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \mathcal{W}_F(\gamma,b)^{2d}.$$

3.2 Some lower bounds

An interesting question is to determine whether these new bounds are optimal, in the sense that, for a given Sbox, there exists a linear layer such that the bounds are tight. Here, we exhibit some functions $\mathcal{M}$ with optimal branch number such

that the EDP (resp. ELP) of some minimum-weight differential (resp. linear mask) over two rounds is related to the upper bound of Theorem 2. This lower bound results from the fact that the minimum-weight codewords with a given support in an $\mathbf{F}_{2^m}$-linear MDS code form a set of the form $\{\lambda c, \lambda \in \mathbf{F}_{2^m}^*\}$ for some fixed codeword c. Such a set is named *a bundle* in [20]. It is not difficult to observe that $\mathcal{B}(0)$ (resp. $\mathcal{B}^\perp(0)$) corresponds to the maximum EDP (resp. to the maximum ELP) of some particular minimum-weight bundles. Moreover, for any such particular bundle, a function $\mathcal{M}$ such that $\mathcal{C}_\mathcal{M}$ contains this bundle can be constructed from some Generalized Reed-Solomon code.

Proposition 3. *Let $\mathcal{S}$ be a permutation of $\mathbf{F}_{2^m}$ and t be any integer such that $t \leq 2^{m-1}$. Then, there exist two $\mathbf{F}_{2^m}$-linear diffusion layers $\mathcal{M}_1$ and $\mathcal{M}_2$ over $\mathbf{F}_{2^m}^t$ with maximal branch number $d = t+1$ such that any block cipher E_1 of the form $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M}_1)$ and E_2 of the form $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M}_2)$ satisfy*

$$\mathrm{MEDP}_2^{E_1} \geq 2^{-m(t+1)}\mathcal{B}(0) \text{ and } \mathrm{MELP}_2^{E_2} \geq 2^{-2m(t+1)}\mathcal{B}^\perp(0)$$

where $\mathcal{B}(0)$ and $\mathcal{B}^\perp(0)$ are defined as in Theorem 2.

Proof. We give here the proof for MEDP_2 only, but it is similar for MELP_2. Actually, the proposition can be formulated in a generic way as the results in Appendix A. Let $\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda} \in \mathbf{F}_{2^m}^*$ and $1 \leq \widehat{u} \leq t$ be some values such that $\sum_{\gamma \in \mathbf{F}_{2^m}^*} \delta_F(\widehat{\alpha}, \gamma)^{\widehat{u}} \delta_F(\gamma\widehat{\lambda}, \widehat{\beta})^{t+1-\widehat{u}} = \mathcal{B}(0)$. Let $a \in \mathbf{F}_{2^m}^t$ be the input difference whose first $\widehat{u}$ coordinates equal $\widehat{\alpha}$ and whose last $(t-\widehat{u})$ coordinates equal 0. Similarly, $b \in \mathbf{F}_{2^m}^t$ denotes the output difference whose first $(t+1-\widehat{u})$ coordinates equal $\widehat{\beta}$ and whose last $(\widehat{u}-1)$ coordinates equal 0. Since

$$\mathrm{EDP}_2(a, M(b)) = \sum_{c \in \mathcal{C}_\mathcal{M}} \left(\prod_{i=1}^{t} \delta_F(a_i, c_i) \right) \left(\prod_{j=1}^{t} \delta_F(c_{t+j}, b_j) \right) ,$$

it is equal to $\mathcal{B}(0)$ if the words of the form

$$\gamma(\underbrace{1, \ldots, 1}_{\widehat{u}}, \underbrace{0, \ldots, 0}_{t-\widehat{u}}, \underbrace{\widehat{\lambda}, \ldots, \widehat{\lambda}}_{t+1-\widehat{u}}, \underbrace{0, \ldots, 0}_{\widehat{u}-1}) \tag{4}$$

are the codewords in $\mathcal{C}_\mathcal{M}$ having the same support as (a, b). Therefore, we aim at finding a linear MDS diffusion layer $\mathcal{M}$ such that $\mathcal{C}_\mathcal{M}$ contains these codewords. Since $t \leq 2^{m-1}$, we can choose $2t$ distinct elements $x_1, \ldots, x_{2t}$ in $\mathbf{F}_{2^m}$. For any choice of $2t$ elements $v_1, \ldots, v_{2t}$, we define the $t \times t$ matrix

$$R = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ x_1v_1 & x_2v_2 & \ldots & x_tv_t \\ x_1^2v_1 & x_2^2v_2 & \ldots & x_t^2v_t \\ & & \ldots & \\ x_1^{t-1}v_1 & x_2^{t-1}v_2 & \ldots & x_t^{t-1}v_t \end{bmatrix}^{-1} \times \begin{bmatrix} 1 & 1 & \ldots & 1 \\ x_{t+1}v_{t+1} & x_{t+2}v_{t+2} & \ldots & x_{2t}v_{2t} \\ x_{t+1}^2v_{t+1} & x_{t+2}^2v_{t+2} & \ldots & x_{2t}^2v_{2t} \\ & & \ldots & \\ x_{t+1}^{t-1}v_{t+1} & x_{t+2}^{t-1}v_{t+2} & \ldots & x_{2t}^{t-1}v_{2t} \end{bmatrix} .$$

Then, the code $\mathcal{C}_\mathcal{M} = \{(x, xR), x \in \mathbf{F}_{2^m}^t\}$ is the *generalized Reed-Solomon code* $\mathsf{GRS}_t(x_1, \ldots, x_{2t}; v)$. It is well-known [36, Page 303] that this code is MDS and is

composed of all words of the form $(v_1F(x_1), \ldots, v_{2t}F(x_{2t}))$ where F ranges over all polynomials in $\mathbf{F}_{2^m}[X]$ of degree strictly less than t. Then, the codewords in $\mathcal{C}_{\mathcal{M}}$ having the same support as (a, b) correspond to the polynomials of degree at most $(t-1)$ which vanish at all $(t-1)$ points x_i for $i \notin \mathsf{Supp}((a, b))$. Then, these polynomials can be written as $\gamma\widehat{F}(x)$, $\gamma \in \mathbf{F}_{2^m}^*$, and $\widehat{F}(x_i) \neq 0$ for $i \in \mathsf{Supp}((a, b))$ since $\widehat{F}$ cannot have more than $(t-1)$ roots. Therefore, we can choose for v a vector such that $v_i = 1/\widehat{F}(x_i)$ for $1 \leq i \leq \widehat{u}$ and $v_i = \widehat{\lambda}/\widehat{F}(x_i)$ for $t+1 \leq i \leq 2t+1-\widehat{u}$. This guarantees that the words in $\mathcal{C}_{\mathcal{M}}$ having the same support as (a, b) are the words of the form (4). It follows that

$$\begin{aligned}\mathrm{EDP}_2(a, M(b)) &= \sum_{\gamma \in \mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{\widehat{u}} \delta_F(\widehat{\alpha}, \gamma v_i \widehat{F}(x_i)) \right) \left(\prod_{j=1}^{t+1-\widehat{u}} \delta_F(\gamma v_{t+j} \widehat{F}(x_{t+j}), \widehat{\beta}) \right) \\ &= \sum_{\gamma \in \mathbf{F}_{2^m}^*} \delta_F(\widehat{\alpha}, \gamma)^{\widehat{u}} \delta_F(\gamma\widehat{\lambda}, \widehat{\beta})^{t+1-\widehat{u}} \, .\end{aligned}$$

□

Remark 1. In some particular cases, we can find a generalized Reed-Solomon code corresponding to a linear layer $\mathcal{M}$ for which the two bounds hold together. Indeed, we want to construct a code $\mathcal{C}_{\mathcal{M}}$ which contains the words (4) and whose dual $\mathcal{C}_{\mathcal{M}}^{\perp}$ contains the words

$$\gamma(\underbrace{0, \ldots, 0}_{t-\bar{u}}, \underbrace{\bar{\lambda}, \ldots, \bar{\lambda}}_{\bar{u}}, \underbrace{0, \ldots, 0}_{\bar{u}-1}, \underbrace{1, \ldots, 1}_{t+1-\bar{u}})$$

for some given $\bar{\lambda}$ and $\bar{u}$. But the dual of $\mathsf{GRS}_t(x_1, \ldots, x_{2t}; v)$ is another generalized Reed-Solomon code, $\mathsf{GRS}_t(x_1, \ldots, x_{2t}; w)$ with $w_i^{-1} = v_i \prod_{j \neq i}(x_i + x_j)$. In particular, if $\bar{u} + \widehat{u} = t$, we can find a vector $(v_1, \ldots, v_{2t})$ such that both conditions hold together. This situation occurs for instance when $\mathcal{S}$ is an involution since $\mathcal{B}(0)$ (resp. $\mathcal{B}^{\perp}(0)$) is attained for all $\widehat{u} < d$ (resp. for all $\bar{u} < d^{\perp}$).

An interesting situation is the case where the maximum over all $\mu \in \mathbf{F}_{2^m}$ of $\mathcal{B}(\mu)$ (resp. of $\mathcal{B}^{\perp}(\mu)$) is attained for $\mu = 0$. Then there exists some $\mathcal{M}$ for which the upper bound from Theorem 2 is tight for $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$, implying that it is impossible to find a general better bound which depends on $\mathcal{S}$ and t only. This situation occurs in particular for any involutional Sbox. Indeed, by combining Prop. 2 and 3, we deduce that, for any involutional Sbox and any $t \leq 2^{m-1}$, there exists a linear layer over $\mathbf{F}_{2^m}^t$ such that the exact values of MEDP_2 and of MELP_2 are equal to the FSE 2003 bounds.

Corollary 1. *Let $\mathcal{S}$ be an involution of $\mathbf{F}_{2^m}$ and t be any integer with $t \leq 2^{m-1}$. Then, there exist an $\mathbf{F}_{2^m}$-linear diffusion layer $\mathcal{M}$ over $\mathbf{F}_{2^m}^t$ with maximal branch number such that $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$ satisfies*

$$\mathrm{MEDP}_2 = \max_{a \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \left(\frac{\delta_F(a, \gamma)}{2^m} \right)^{(t+1)} = \max_{b \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \left(\frac{\delta_F(\gamma, b)}{2^m} \right)^{(t+1)}$$

$$\textit{and } \mathrm{MELP}_2 = \max_{a \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \left(\frac{\mathcal{W}_F(a,\gamma)}{2^m} \right)^{2(t+1)} = \max_{b \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \left(\frac{\mathcal{W}_F(\gamma,b)}{2^m} \right)^{2(t+1)}.$$

Proof. We know from Prop. 2 that for any u, $\max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}_u(\mu) = \mathcal{B}_u(0) = \mathcal{B}(0)$, and $\max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}_u(\mu)^\perp = \mathcal{B}_u(0)^\perp = \mathcal{B}(0)$. Moreover, all these values are equal to the FSE 2003 bounds. By combining Th. 2 and Prop. 3, we deduce the existence of some linear layers for which MEDP_2 (resp. MELP_2) are lower- and upper-bounded by $\mathcal{B}(0)$ (resp. $\mathcal{B}^\perp(0)$). Moreover, we have proved in Prop. 2 that $\mathcal{B}(0)$ (resp. $\mathcal{B}^\perp(0)$) is attained for all values of u. This is a case where we can construct a GRS code satisfying the conditions for MEDP_2 and MELP_2 together. □

Example 1. The Prøst permutation over $\mathbf{F}_2^{16d}$, $d \geq 1$, is the core function of several AEAD-schemes submitted to the CAESAR competition [27]. This permutation is of the form $\mathrm{SPN}(4, 4d, S, M)$ where S is a 4-bit involution named `SubRows` and M corresponds to the composition of two linear permutations, `MixSlices` and `ShiftPlanes`. The round-constant addition is omited here since it does not have any impact in our context. Similarly to the AES, two consecutive rounds of the Prøst permutation can be seen as the parallel application of d copies of a superbox defined over $\mathbf{F}_{16}$. This superbox corresponds to two `SubRows` layers separated by a `MixSlices` transformation. Moreover, even if it is not mentioned in the design rationale, it can be checked that `MixSlices` is linear over $\mathbf{F}_{16}$ if $\mathbf{F}_{16}$ is identified with $\mathbf{F}_2^4$ by the following isomorphism:

$$\varphi : (x_0, \ldots, x_3) \mapsto x_1 + \alpha x_2 + \alpha^2 x_3 + \alpha^3 x_0$$

where α is a root of $X^4 + X^3 + 1$. Indeed, the function defined over $\mathbf{F}_{16}^4$ by $\mathcal{M} = \tilde{\varphi} \circ \texttt{MixSlices} \circ \tilde{\varphi}^{-1}$ corresponds to the multiplication by

$$\begin{pmatrix} 1 & \alpha & \alpha+\alpha^2 & \alpha^2 \\ \alpha & 1 & \alpha^2 & \alpha+\alpha^2 \\ \alpha+\alpha^2 & \alpha^2 & 1 & \alpha \\ \alpha^2 & \alpha+\alpha^2 & \alpha & 1 \end{pmatrix}.$$

The previous framework then directly applies[2]. In particular, since the Sbox is an involution, our bounds are equal to the FSE 2003 bounds (Prop. 2): the Prøst permutation with any $\mathbf{F}_2$-linear MDS `MixSlices` transformation satisfies

$$\mathrm{MEDP}_2 \leq 2^{-8} \text{ and } \mathrm{MELP}_2 \leq 2^{-8}.$$

These bounds are tight as stated in Corollary 1: using the construction described in the proof of Prop. 3, we obtain that the following matrix over $\mathbf{F}_{16}$ (with the previously described representation) leads to a variant of the Prøst permutation

[2] We here focus on the MEDP and MELP of the SPN with the same building blocks as the Prøst permutation, but these expectations do not provide any direct information on the security of the Prøst permutation in which the key is fixed.

with $\mathrm{MEDP}_2 = \mathrm{MELP}_2 = 2^{-8}$:

$$\begin{pmatrix} \alpha^2+\alpha+1 & \alpha^3+\alpha & \alpha^3+\alpha+1 & 1 \\ \alpha+1 & \alpha^3+\alpha^2+\alpha & \alpha^2+\alpha+1 & 1 \\ \alpha^2+1 & \alpha^3+\alpha^2+1 & \alpha^3 & 1 \\ \alpha^2 & \alpha^3+\alpha^2 & \alpha^3+1 & 1 \end{pmatrix} .$$

This implies that, for this particular Sbox, the `MixSlices` transformation must be chosen with care to guarantee small MEDP_2 and MELP_2. Instead, for some Sboxes within the same equivalence class as `SubRows`, we can guarantee that, for any $\mathbf{F}_{16}$-linear MDS `MixSlices`, $\mathrm{MEDP}_2 \leq 3 \times 2^{-10}$. This does not make a big difference in the case of Prøst since the alphabet is small and the exact MEDP_2 and MELP_2 can be easily computed. For instance, for the `MixSlices` transformation chosen by the designers, we have $\mathrm{MEDP}_2 = 3 \times 2^{-11}$ and $\mathrm{MELP}_2 = 81 \times 2^{-16}$. However, for Sboxes over $\mathbf{F}_{2^8}$, computing the exact MEDP_2 and MELP_2 is rather expensive and obtaining a better upper bound is very helpful.

3.3 Influence of the field representation

Clearly, there is no reason why the two-round MEDP or MELP should be the same for affine equivalent Sboxes in general. Then, it makes sense that our new bounds are not invariant under affine equivalence. More surprisingly, by combining the upper bound from Theorem 2 with the lower bound provided by Prop. 3, we can exhibit some examples showing that the choice of the field $\mathbf{F}_{2^m}$, *i.e.*, the choice of the isomorphism φ between $\mathbf{F}_2^m$ and $\mathbf{F}_{2^m}$, may influence the value of the MEDP and MELP.

Example 2. Let us consider 2 rounds of a cipher of the form $\mathrm{SPN}(4, 4, S, M)$, where S is one of the permutations of $\mathbf{F}_2^4$ used in the PRINCE-family [9], namely permutation S_6 in [10, Table 3]:

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S(x)$	0	1	2	13	4	7	15	6	8	14	11	10	9	3	12	5

where each element in $\mathbf{F}_2^4$ is here represented as an integer between 0 and 15. This Sbox is differentially 4-uniform and has linearity $\mathcal{L}(\mathcal{S}) = 8$. For this Sbox, the FSE 2003 bound gives $\mathrm{MEDP}_2^E \leq 34 \times 2^{-14}$, for any $\mathbf{F}_2$-linear diffusion layer M over $\mathbf{F}_2^{16}$ with branch number 5. If we now consider a diffusion layer $\mathcal{M}$ with branch number 5 which is linear over $\mathbf{F}_{2^4}$ where $\mathbf{F}_{2^4}$ is identified with $\mathbf{F}_2^4$ by the basis $\{1, \alpha, \alpha^2, \alpha^3\}$ where α is a root of the irreducible polynomial $X^4 + X^3 + X^2 + X + 1$, we get from Theorem 2 that

$$\mathrm{MEDP}_2^E \leq 33 \times 2^{-14} ,$$

and this inequality holds for any such function $\mathcal{M}$. However, we can now consider a permutation $\mathcal{M}'$ which is linear over $\mathbf{F}_{2^4}$, but where the isomorphism between $\mathbf{F}_2^4$ and $\mathbf{F}_{2^4}$ is defined by a different basis, namely $\{1, \beta, \beta^2, \beta^3\}$ where β is a

root of the primitive polynomial $X^4 + X + 1$. Then, the value $\mathcal{B}(0)$ involved in Prop. 3 equals 17×2^7, implying that there exists an $\mathbf{F}_{2^4}$-linear function $\mathcal{M}'$ with branch number 5 such that

$$\mathrm{MEDP}_2^E = 34 \times 2^{-14} ,$$

which is strictly higher that the upper bound we have for any MDS diffusion layer $\mathbf{F}_{2^4}$-linear where $\mathbf{F}_{2^4}$ is defined with the basis $\{1, \alpha, \alpha^2, \alpha^3\}$.

There is no contradiction here since the two theorems apply to the representations of the Sbox and of the diffusion function over $\mathbf{F}_{2^4}$ only. Here, we have proved that there is a particular $\mathcal{M}'$ such that $\mathrm{SPN}(4, 4, S, \widetilde{\psi}^{-1} \circ \mathcal{M}' \circ \widetilde{\psi})$ has $\mathrm{MEDP}_2 = 34 \times 2^{-14}$, where $\widetilde{\psi}$ is the concatenation of 4 copies of the isomorphism ψ from $\mathbf{F}_2^4$ into $\mathbf{F}_{2^4}$ defined by the basis $\{1, \beta, \beta^2, \beta^3\}$. But if we consider the basis $\{1, \alpha, \alpha^2, \alpha^3\}$ and the corresponding isomorphism φ, Th. 2 provides a bound for all $\mathrm{SPN}_F(4, 4, \varphi \circ S \circ \varphi^{-1}, \mathcal{M})$ which does not include the previous case because the permutation defined by $\mathcal{M} = \widetilde{\varphi} \circ \widetilde{\psi}^{-1} \circ \mathcal{M}' \circ \widetilde{\psi} \circ \widetilde{\varphi}^{-1}$ is not linear over $\mathbf{F}_{2^4}$, since $(\psi \circ \varphi^{-1})$ is not a ring isomorphism. This unexpected result comes from the fact that the definitions of the Sbox and of the linear layer do not use the same representation: the Sbox is defined over $\mathbf{F}_2^4$ while the linear layer is defined over $\mathbf{F}_{2^4}$. This is why the choice of the basis affects the MEDP while this is obviously not the case when the two functions are defined over the same alphabet. But, even if this does not correspond to a natural mathematical description, it may be relevant to use the binary representation for the Sbox (chosen to minimize the number of gates for instance), while the field representation is used for the mixing layer since it is $\mathbf{F}_{2^m}$-linear (see e.g. [25]).

It is worth noticing that the previous situation is not related to the fact that one of the field representations is defined by a non-primitive polynomial. Indeed, the following example shows that even changing the primitive polynomial used for constructing $\mathbf{F}_{2^m}$ may affect the two-round MEDP and MELP.

Example 3. We now consider two rounds of a cipher $\mathrm{SPN}(5, 4, S, M)$ where S is the following permutation of $\mathbf{F}_2^5$:

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S(x)$	0	1	18	20	25	16	6	27	17	3	22	15	31	7	30	26
x	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$S(x)$	4	23	29	21	9	10	24	2	14	5	13	8	28	19	12	11

When $\mathbf{F}_{2^5}$ is identified with $\mathbf{F}_2^5$ by the basis $\{1, \alpha, \alpha^2, \alpha^3, \alpha^4\}$ where α is a root of the primitive polynomial $X^5 + X^2 + 1$, Theorem 2 shows that any MDS diffusion layer linear over $\mathbf{F}_{2^5}$ with this representation satisfies

$$\mathrm{MEDP}_2^E \leq 13 \times 2^{-20} \text{ and } \mathrm{MELP}_2^E \leq 8407 \times 2^{-27} .$$

When $\mathbf{F}_{2^5}$ is constructed from the primitive polynomial $X^5 + X^3 + 1$, the lower and upper bounds from Th. 2 and Prop. 3 are equal, and show that there exists some MDS layer linear over $\mathbf{F}_{2^5}$ with this alternative representation such that

$$\mathrm{MEDP}_2^E = 14 \times 2^{-20} \text{ and } \mathrm{MELP}_2^E = 8663 \times 2^{-27} .$$

The first primitive polynomial then guarantees lower two-round MEDP and MELP than the second one.

Another example is the LED block cipher. In [25, Section 3.2], the designers provide an upper bound on the four-round MEDP and MELP of a cipher of the form $\mathrm{SPN}_F(4,4,\mathcal{S},\mathcal{M})$[3] where $\mathcal{M}$ is an $\mathbf{F}_{2^4}$-linear function with branch number 5, $\mathcal{S}$ corresponds to the Present Sbox and $\mathbf{F}_{2^4}$ is defined by the basis $(1,\alpha,\alpha^2,\alpha^3)$, with α a root of X^4+X+1. To this end, they use the FSE 2003 bound, which leads to $\mathrm{MEDP}_2 \leq 2^{-8}$ and $\mathrm{MELP}_2 \leq 2^{-8}$, implying that $\mathrm{MEDP}_4 \leq 2^{-32}$ and $\mathrm{MELP}_4 \leq 2^{-32}$. For this cipher, our new upper bound is equal to the FSE 2003 bound and then does not improve the result. However, if we consider the same Sbox, but modify the representation of $\mathbf{F}_{2^4}$ and choose the basis defined by X^4+X^3+1, Theorem 2 leads to $\mathrm{MEDP}_2 \leq 3\times 2^{-10}$. Then, with this minor modification, the upper bound on MEDP_4 is improved by a factor $(3/4)^4 = 0.3164$ (while the bound on MELP_4 is unchanged).

4 Multiplicative invariance for Sboxes

Power permutations are often considered as suitable Sboxes since determining their differential and Walsh spectra is easier and also because they usually have a lower implementation cost. This family of Sboxes is also of great interest in our context since our bounds provide a very good approximation of the exact two-round MEDP and MELP which depends on the Sbox and on the branch number only. Indeed, for power permutations, we get a universal lower bound in the sense that the bound provided in Prop. 3 holds for any $\mathbf{F}_{2^m}$-linear permutation $\mathcal{M}$. This comes from the fact that all rows in the difference table (resp. in the correlation table) of a power permutation can be deduced from a single one. This is because any power function $\mathcal{S}$ is an endomorphism over the multiplicative group $\mathbf{F}_{2^m}^*$, *i.e.*, $\mathcal{S}(xy) = \mathcal{S}(x)\mathcal{S}(y)$ for any pair of nonzero elements (x,y). Unfortunately, there is no hope to capture a larger family of Sboxes with a straightforward generalization of this property since it can be easily shown that any function $\mathcal{S}$ satisfying $\mathcal{S}(xy) = \mathcal{S}(x)\mathcal{S}'(y)$ for some $\mathcal{S}'$ is of the form $\mathcal{S}(x) = cx^s$. However, we can define this suitable multiplicative property on the difference table (resp. on the Walsh transform) of $\mathcal{S}$, and not on the function itself.

4.1 Generalizing the multiplicative property

Definition 7. *Let $\mathcal{S}$ be a mapping of $\mathbf{F}_{2^m}$.*

- *$\mathcal{S}$ is said to have* multiplicative-invariant derivatives *if, for any $x \in \mathbf{F}_{2^m}^*$ there exists a permutation π_x of $\mathbf{F}_{2^m}^*$ such that*

$$\delta_F(\alpha, xy) = \delta_F(\pi_x(\alpha), y), \quad \forall y \in \mathbf{F}_{2^m}^*.$$

[3] As for Prøst, this result does not directly apply to LED since the round keys are inserted every four rounds only.

– $\mathcal{S}$ *is said to have a* multiplicative-invariant Walsh transform *if, for any* $x \in \mathbf{F}_{2^m}^*$ *there exists a permutation* ψ_x *of* $\mathbf{F}_{2^m}^*$ *such that*

$$\mathcal{W}_F(\alpha, xy)^2 = \mathcal{W}_F(\psi_x(\alpha), y)^2, \quad \forall y \in \mathbf{F}_{2^m}^*.$$

These definitions include all functions resulting from the composition on the right of a power permutation with an $\mathbf{F}_2$-linear permutation (cf. proof in Appendix B).

Proposition 4. *Let* $\mathcal{S} = \mathcal{S}' \circ L$ *where* L *is an* $\mathbf{F}_2$*-linear permutation of* $\mathbf{F}_{2^m}$ *and* $\mathcal{S}' : x \mapsto x^s$ *is a power permutation over* $\mathbf{F}_{2^m}$*. Then, both the derivatives of* $\mathcal{S}$ *and its Walsh transform are multiplicative-invariant.*

It is worth noticing that the fact that a permutation has multiplicative-invariant derivatives (resp. Walsh transform) does not imply that a similar property holds for its inverse. In other words, Prop. 4 does not apply to the composition on the left of a power permutation with a linear permutation. The following proposition shows that the permutations with multiplicative-invariant derivatives (resp. Walsh transform) are not all affine equivalent to a power permutation.

Proposition 5. *Let* m *be an odd integer and* $\mathcal{S}$ *be a quadratic permutation of* $\mathbf{F}_{2^m}$ *with* $\Delta(S) = 2$ *(aka APN permutation). Then,* $\mathcal{S}$ *has multiplicative-invariant derivatives and* $\mathcal{S}^{-1}$ *has a multiplicative-invariant Walsh transform.*

This result actually applies to a (possibly) more general class of permutations known as *crooked permutations*, which includes all quadratic APN permutations (see details in Appendix B). Prop. 5 applies for instance to the infinite family of APN permutations of degree 2

$$x \longmapsto x^{2^i+1} + ux^{2^{j\frac{m}{3}}+2^{(3-j)\frac{m}{3}+i}} \text{ with } \gcd(i, m) = 1 \text{ and } j = im/3 \bmod 3$$

over $\mathbf{F}_{2^m}$, m odd, divisible by 3 and not by 9, which is not affine equivalent to a power mapping [12].

4.2 A universal lower bound for Sboxes with some multiplicative invariance

We now show that for Sboxes with multiplicative-invariant derivatives (resp. Walsh transform), the previously established bounds simplify.

Proposition 6. *Let* $\mathcal{S}$ *be a permutation of* $\mathbf{F}_{2^m}$ *such that either* $\mathcal{S}$ *or* $\mathcal{S}^{-1}$ *has multiplication-invariant derivatives (resp. Walsh transform). For any integers* d *and* $d^\perp$*, we define*

$$\mathcal{B}'_u(\mu) = \max_{\alpha,\beta \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \delta_F(\alpha, \gamma)^u \delta_F(\gamma + \mu, \beta)^{(d-u)}, \text{ with } 1 \le u < d,$$

$$\mathcal{B}'^{\perp}_u(\mu) = \max_{\alpha,\beta \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \mathcal{W}_F(\alpha, \gamma)^{2u} \mathcal{W}_F(\gamma + \mu, \beta)^{2(d^\perp-u)}, \text{ with } 1 \le u < d^\perp .$$

Then, for any u, we have

$$\mathcal{B}_u(0) = \mathcal{B}'_u(0) \text{ and } \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}_u(\mu) = \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}'_u(\mu)$$

$$\text{resp. } \mathcal{B}_u^{\perp}(0) = \mathcal{B}_u'^{\perp}(0) \text{ and } \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}_u^{\perp}(\mu) = \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}_u'^{\perp}(\mu) \ ,$$

where $\mathcal{B}_u(\mu)$ and $\mathcal{B}_u^{\perp}(\mu)$ are defined as in Theorem 2.

It follows that the upper bounds defined by Theorem 2 simplify to

$$\mathrm{MEDP}_2^E \leq 2^{-md} \max_{1 \leq u < d} \max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}'_u(\mu) \text{ and } \mathrm{MELP}_2^E \leq 2^{-2md^{\perp}} \max_{1 \leq u < d^{\perp}} \max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}_u'^{\perp}(\mu).$$

More interestingly, we now get some universal lower bound on MEDP_2 and MELP_2, *i.e.*, which hold for *any* diffusion layer with maximal branch number.

Theorem 3. *Let $\mathcal{S}$ be a permutation of $\mathbf{F}_{2^m}$. Then, for* any *$\mathbf{F}_{2^m}$-linear diffusion layer $\mathcal{M}$ over $(\mathbf{F}_{2^m})^t$ with maximal branch number $d = t + 1$, the MEDP_2 and MELP_2 of any block cipher E of the form $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$ satisfy the following.*

- *If both $\mathcal{S}$ and $\mathcal{S}^{-1}$ have multiplicative-invariant derivatives, then*

$$\mathrm{MEDP}_2^E \geq 2^{-m(t+1)} \max_{1 \leq u < d} \mathcal{B}'_u(0);$$

- *if both $\mathcal{S}$ and $\mathcal{S}^{-1}$ have a multiplicative-invariant Walsh transform, then*

$$\mathrm{MELP}_2^E \geq 2^{-2m(t+1)} \max_{1 \leq u < d} \mathcal{B}_u'^{\perp}(0) \ ;$$

- *if $\mathcal{S}$ has multiplicative-invariant derivatives (resp. Walsh transform), then*

$$\mathrm{MEDP}_2^E \geq 2^{-m(t+1)} \mathcal{B}'_t(0), \text{ resp. } \mathrm{MELP}_2^E \geq 2^{-2m(t+1)} \mathcal{B}_t'^{\perp}(0).$$

- *if $\mathcal{S}^{-1}$ has multiplicative-invariant derivatives (resp. Walsh transform), then*

$$\mathrm{MEDP}_2^E \geq 2^{-m(t+1)} \mathcal{B}'_1(0), \text{ resp. } \mathrm{MELP}_2^E \geq 2^{-2m(t+1)} \mathcal{B}_1'^{\perp}(0).$$

Let us focus on all permutations of $\mathbf{F}_{2^8}$ of the same form as the AES Sbox: $\mathcal{S}(x) = A(x^{254})$, where A is an $\mathbf{F}_2$-affine permutation of $\mathbf{F}_{2^8}$. Since $\mathcal{S}^{-1}$ has multiplication-invariant derivatives and Walsh transform (cf. Prop. 4), we derive from Theorems 2 and 3, and Prop. 6 that, for $t = 4$,

$$2^{-40} \mathcal{B}'_1(0) \leq \mathrm{MEDP}_2 \leq 2^{-40} \max_{1 \leq u \leq 4} \max_{\mu \in \mathbf{F}_{2^8}^*} \mathcal{B}'_u(\mu)$$

$$\text{and } 2^{-80} \mathcal{B}_1'^{\perp}(0) \leq \mathrm{MELP}_2 \leq 2^{-80} \max_{1 \leq u \leq 4} \max_{\mu \in \mathbf{F}_{2^8}^*} \mathcal{B}_u'^{\perp}(\mu) \ .$$

These bounds do not depend on the isomorphism between $\mathbf{F}_2^8$ and $\mathbf{F}_{2^8}$ since their expressions do not involve any multiplication in $\mathbf{F}_{2^8}$, while this was not the case of the more general bound in Theorem 2. Then, we get the following results for different choices of the affine permutation A.

– For the affine function A used in the AES, $\mathrm{SPN}_F(8,4,\mathcal{S},\mathcal{M})$ satisfies

$$53\times 2^{-34} \leq \mathrm{MEDP}_2 \leq 55.5\times 2^{-34} \text{ and } 1.638\times 2^{-28} \leq \mathrm{MELP}_2 \leq 1.86\times 2^{-28}$$

for any $\mathbf{F}_{2^8}$-linear MDS diffusion layer $\mathcal{M}$ and any isomorphism between $\mathbf{F}_{2^8}$ and $\mathbf{F}_2^8$. The exact values for the diffusion layer used in the AES correspond to the lower bounds in both cases. But, we have exhibited in Prop. 3 an MDS diffusion layer for which $\mathrm{MELP}_2 \geq 1.66 \times 2^{-28}$. Then, the choice of the MDS linear layer affects the value of MELP_2 within this interval.

– For the affine function A' used in SHARK [42] and SQUARE [16] which are two predecessors of the AES, $\mathrm{SPN}_F(8,4,\mathcal{S},\mathcal{M})$ satisfies

$$53\times 2^{-34} \leq \mathrm{MEDP}_2 \leq 56\times 2^{-34} \text{ and } 1.7169\times 2^{-28} \leq \mathrm{MELP}_2 \leq 1.9847\times 2^{-28}$$

for any $\mathbf{F}_{2^8}$-linear MDS diffusion layer $\mathcal{M}$. Then, the affine function chosen in the AES Sbox offers a slightly better guarantee than the one chosen in SQUARE. Indeed, it is impossible with the SQUARE affine function to obtain a two-round MELP which is as small as the one of the AES. Note that the isomorphism between $\mathbf{F}_2^8$ and $\mathbf{F}_{2^8}$ is different in SQUARE and in the AES [2].

– We have exhibited a linear permutation A'' of $\mathbf{F}_2^8$ for which the corresponding Sbox is such that $\mathrm{SPN}_F(8,4,\mathcal{S},\mathcal{M})$ satisfies

$$\mathrm{MEDP}_2 = 56 \times 2^{-34} \text{ and } 1.8354 \times 2^{-28} \leq \mathrm{MELP}_2 \leq 1.8684 \times 2^{-28}$$

for any $\mathbf{F}_{2^8}$-linear MDS diffusion layer $\mathcal{M}$. Then, this Sbox always provides a higher two-round MEDP than the AES Sbox.

Even if we are not able to explicitly construct an affine permutation A which minimizes the values of MEDP_2 and MELP_2, our results clearly simplify the task of the designer. Indeed, the affine permutation A and the diffusion layer M can be chosen separately since a very good estimate of MEDP_2 and MELP_2 is obtained independently of the diffusion layer. This is more efficient than computing these values for many pairs (A, M).

4.3 Involutions with some multiplicative invariance

A particular case of interest is when $\mathcal{S}$ is an involution with multiplicative-invariant derivatives (or Walsh transform). Then, the lower bound in the previous theorem corresponds to the upper bound in Theorem 2, and both values are equal to the FSE 2003 bound.

Corollary 2. *Let $\mathcal{S}$ be an involution of $\mathbf{F}_{2^m}$ with multiplicative-invariant derivatives (resp. Walsh transform). Then, for any t and any $\mathbf{F}_{2^m}$-linear diffusion layer $\mathcal{M}$ over $\mathbf{F}_{2^m}^t$ with branch number $t+1$, any block cipher of the form $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$ satisfies*

$$\mathrm{MEDP}_2^E = 2^{m(t+1)} \max_{\alpha\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \delta_F(\alpha,\gamma)^{t+1},$$

$$\textit{resp. } \mathrm{MELP}_2^E = 2^{2m(t+1)} \max_{\alpha\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \mathcal{W}_F(\alpha,\gamma)^{2(t+1)}.$$

The naive Sbox, *i.e.* the inversion in $\mathbf{F}_{2^m}$, satisfies all hypotheses of the previous corollary. The exact values of MEDP_2 and MELP_2 for an SPN combining the naive Sbox over $\mathbf{F}_{2^m}$ and any $\mathbf{F}_{2^m}$-linear layer with maximal branch number are then always equal to the FSE 2003 bounds. For instance, for the two-round AES with the naive Sbox, we have $\mathrm{MEDP}_2 = 79 \times 2^{-34}$ and $\mathrm{MELP}_2 = 48,193,409 \times 2^{-52}$, and this is independent of the $\mathbf{F}_{2^8}$-linear MDS layer. In particular, the exact MEDP_2 and MELP_2 do not depend on the field representation since Coro. 2 provides the same value for any basis. Also, this explains why, among all Sboxes in the same equivalence class, the naive Sbox is the one which leads after two rounds both to the highest MEDP and to the highest MELP for any $\mathbf{F}_{2^m}$-linear diffusion layer with maximal branch number. And this situation is independent of the size of the Sbox, and of the choice of the $\mathbf{F}_{2^m}$-linear MDS layer.

5 Conclusions

We have improved the general upper bounds on the two-round MEDP and MELP for a given Sbox over $\mathbf{F}_2^m$ and any $\mathbf{F}_{2^m}$-linear diffusion layer with given branch numbers. One of the main properties of these new bounds is that they are not invariant under affine equivalence, and then they enable the designers to choose an appropriate Sbox within an equivalence class, independently of the diffusion layer. These bounds point out the importance of some interactions between the Sbox and its inverse. In particular, the involutions play a special role since there always exists some diffusion layer for which both MEDP_2 and MELP_2 achieve the highest possible value we can obtain for an Sbox in the same equivalence class. Also, we have shown that, for the Sboxes with multiplicative-invariant derivatives or Walsh transform, we can compute a lower bound on MEDP_2 and MELP_2 independently of the choice of the MDS diffusion layer. This result applies for instance to all Sboxes of the form $x \mapsto A(x^s)$, as in the AES. In particular, we have proved that, independently of the specifications of the MDS diffusion layer, the naive Sbox leads to the highest possible MEDP_2 and MELP_2. The exact MEDP_2 and MELP_2 may even vary with the basis used for defining $\mathbf{F}_{2^m}$. Our work then raises several open questions. We have shown that involutional power permutations are the weakest Sboxes in their equivalence class whatever MDS linear layer is chosen. For involutions which do not have any multiplicative-invariant property, this result holds but for some MDS layers only. Then, it would be interesting to determine whether this weakness is more general, and whether an involution is always the worst choice within an equivalent class. This issue is of practical interest since involutional Sboxes are a natural choice for minimizing the implementation overhead of decryption on top of encryption. Another open question is whether the use of an involutional Sbox, especially of the naive Sbox, introduces a similar weakness for a higher number of rounds, in the sense of the conjecture in [17]. The difficulty comes from the fact that, exactly as for the FSE 2003 bound, applying our upper bound twice successively requires the knowledge of the whole difference table of the superbox. Our new bound can then be combined with (1) only, to get a bound of the 4-round MEDP and MELP.

Acknowledgments. The authors would like to thank Daniel Augot, Matthieu Finiasz, María Naya Plasencia for valuable discussions, and the reviewers for their constructive comments.

References

1. Abdelraheem, M.A., Ågren, M., Beelen, P., Leander, G.: On the Distribution of Linear Biases: Three Instructive Examples. In: Advances in Cryptology - CRYPTO 2012. LNCS, vol. 7417, pp. 50–67. Springer (2012)
2. Barreto, P.S.: Implementation of the SQUARE block cipher. `http://www.larc.usp.br/~pbarreto/sqjava21.zip`
3. Bending, T.D., Fon-Der-Flaass, D.: Crooked Functions, Bent Functions, and Distance Regular Graphs. Electr. J. Comb. 5 (1998)
4. Bierbrauer, J., Kyureghyan, G.M.: Crooked binomials. Designs, Codes and Cryptography 46(3), 269–301 (2008)
5. Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. Journal of Cryptology pp. 3–72 (1991)
6. Biryukov, A., De Cannière, C., Braeken, A., Preneel, B.: A Toolbox for Cryptanalysis: Linear and Affine Equivalence Algorithms. In: Advances in Cryptology - EUROCRYPT 2003. LNCS, vol. 2656, pp. 33–50. Springer (2003)
7. Blondeau, C., Bogdanov, A., Leander, G.: Bounds in Shallows and in Miseries. In: Advances in Cryptology - CRYPTO 2013 (1). LNCS, vol. 8042, pp. 204–221. Springer (2013)
8. Blondeau, C., Nyberg, K.: New Links between Differential and Linear Cryptanalysis. In: Advances in Cryptology - EUROCRYPT 2013. LNCS, vol. 7881, pp. 388–404. Springer (2013)
9. Borghoff, J., Canteaut, A., Güneysu, T., Kavun, E.B., Knezevic, M., Knudsen, L.R., Leander, G., Nikov, V., Paar, C., Rechberger, C., Rombouts, P., Thomsen, S.S., Yalçin, T.: PRINCE - A Low-Latency Block Cipher for Pervasive Computing Applications - Extended Abstract. In: Advances in Cryptology - ASIACRYPT 2012. LNCS, vol. 7658, pp. 208–225. Springer (2012)
10. Borghoff, J., Canteaut, A., Güneysu, T., Kavun, E.B., Knezevic, M., Knudsen, L.R., Leander, G., Nikov, V., Paar, C., Rechberger, C., Rombouts, P., Thomsen, S.S., Yalçin, T.: PRINCE - A Low-latency Block Cipher for Pervasive Computing Applications (Full version). IACR Cryptology ePrint Archive 529 (2012)
11. Brinkmann, M., Leander, G.: On the classification of APN functions up to dimension five. Designs, Codes and Cryptography 49(1-3), 273–288 (2008)
12. Budaghyan, L., Carlet, C., Leander, G.: Two Classes of Quadratic APN Binomials Inequivalent to Power Functions. IEEE Transactions on Information Theory 54(9), 4218–4229 (2008)
13. Canteaut, A., Charpin, P.: Decomposing bent functions. IEEE Transactions on Information Theory 49(8), 2004–19 (2003)
14. Chun, K., Kim, S., Lee, S., Sung, S.H., Yoon, S.: Differential and linear cryptanalysis for 2-round SPNs. Inf. Process. Lett. 87(5), 277–282 (2003)
15. Daemen, J.: Cipher and hash function design strategies based on linear and differential cryptanalysis. Ph.D. thesis, K.U. Leuven (1995)
16. Daemen, J., Knudsen, L.R., Rijmen, V.: The Block Cipher Square. In: Fast Software Encryption - FSE'97. LNCS, vol. 1267, pp. 149–165. Springer (1997)

17. Daemen, J., Lamberger, M., Pramstaller, N., Rijmen, V., Vercauteren, F.: Computational aspects of the expected differential probability of 4-round AES and AES-like ciphers. Computing 85(1-2), 85–104 (2009)
18. Daemen, J., Rijmen, V.: The Wide Trail Design Strategy. In: IMA International Conference - Coding and Cryptography 2001. LNCS, vol. 2260, pp. 222–238. Springer (2001)
19. Daemen, J., Rijmen, V.: The Design of Rijndael: AES - The Advanced Encryption Standard. Information Security and Cryptography, Springer (2002)
20. Daemen, J., Rijmen, V.: Understanding Two-Round Differentials in AES. In: Security and Cryptography for Networks - SCN 2006. LNCS, vol. 4116, pp. 78–94. Springer (2006)
21. Daemen, J., Rijmen, V.: Probability distributions of correlation and differentials in block ciphers. Journal of Mathematical Cryptology 1(3), 221–242 (2007)
22. Daemen, J., Rijmen, V.: New criteria for linear maps in AES-like ciphers. Cryptography and Communications 1(1), 47–69 (2009)
23. Daemen, J., Rijmen, V.: Advanced Linear Cryptanalysis of Block and Stream Ciphers, chap. Correlation Analysis in $GF(2^n)$, pp. 115–131. Cryptology and information security, IOS Press (2011)
24. Gong, Z., Nikova, S., Law, Y.W.: KLEIN: A New Family of Lightweight Block Ciphers. In: RFID. Security and Privacy - RFIDSec 2011. LNCS, vol. 7055, pp. 1–18. Springer (2012)
25. Guo, J., Peyrin, T., Poschmann, A., Robshaw, M.: The LED Block Cipher. In: Cryptographic Hardware and Embedded Systems - CHES 2011. LNCS, vol. 6917, pp. 326–341. Springer (2011)
26. Hong, S., Lee, S., Lim, J., Sung, J., Cheon, D.H., Cho, I.: Provable Security against Differential and Linear Cryptanalysis for the SPN Structure. In: Fast Software Encryption - FSE 2000. LNCS, vol. 1978, pp. 273–283. Springer (2000)
27. Kavun, E.B., Lauridsen, M.M., Leander, G., Rechberger, C., Schwabe, P., Yalçın, T.: Prøst v1.1. Submission to the CAESAR competition (2014), `http://proest.compute.dtu.dk/proestv11.pdf`
28. Keliher, L., Sui, J.: Exact maximum expected differential and linear probability for two-round Advanced Encryption Standard. IET Information Security 1(2), 53–57 (2007)
29. Knudsen, L.R.: Truncated and higher order differentials. In: Fast Software Encryption - FSE'94. LNCS, vol. 1008, pp. 196–211. Springer-Verlag (1995)
30. Kyureghyan, G.M.: Crooked maps in $\mathbf{F}_{2^n}$. Finite Fields and Their Applications 13(3), 713–726 (2007)
31. Lai, X.: Higher order derivatives and differential cryptanalysis. In: Symposium on Communication, Coding and Cryptography. Kluwer Academic Publishers (1994)
32. Lai, X., Massey, J.L., Murphy, S.: Markov ciphers and differential cryptanalysis. In: Advances in Cryptology - EUROCRYPT'91. LNCS, vol. 547, pp. 17–38. Springer-Verlag (1991)
33. Leander, G.: On linear hulls, statistical saturation attacks, PRESENT and a cryptanalysis of PUFFIN. In: Advances in Cryptology - EUROCRYPT 2011. LNCS, vol. 6632, pp. 303–322. Springer (2011)
34. Leander, G., Poschmann, A.: On the Classification of 4 Bit S-Boxes. In: Arithmetic of Finite Fields - WAIFI 2007. LNCS, vol. 4547, pp. 159–176. Springer (2007)
35. Lim, C.H., Korkishko, T.: mCrypton - A Lightweight Block Cipher for Security of Low-Cost RFID Tags and Sensors. In: Information Security Applications - WISA 2005. LNCS, vol. 3786, pp. 243–258. Springer (2006)

36. MacWilliams, F.J., Sloane, N.J.: The theory of error-correcting codes. North-Holland (1977)
37. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Advances in Cryptology - EUROCRYPT'93. LNCS, vol. 765. Springer-Verlag (1994)
38. Murphy, S.: The effectiveness of the linear hull effect. J. Mathematical Cryptology 6(2), 137–147 (2012)
39. Nyberg, K.: Differentially uniform mappings for cryptography. In: Advances in Cryptology - EUROCRYPT'93. LNCS, vol. 765, pp. 55–64. Springer-Verlag (1993)
40. Nyberg, K.: Linear Approximation of Block Ciphers. In: Advances in Cryptology - EUROCRYPT'94. LNCS, vol. 950. Springer-Verlag (1995)
41. Park, S., Sung, S.H., Lee, S., Lim, J.: Improving the Upper Bound on the Maximum Differential and the Maximum Linear Hull Probability for SPN Structures and AES. In: Fast Software Encryption - FSE 2003. LNCS, vol. 2887, pp. 247–260. Springer (2003)
42. Rijmen, V., Daemen, J., Preneel, B., Bosselaers, A., Win, E.D.: The Cipher SHARK. In: Fast Software Encryption - FSE'96. LNCS, vol. 1039, pp. 99–111. Springer (1996)
43. Saarinen, M.J.O.: Cryptographic analysis of all 4×4-bit sboxes. In: Selected Areas in Cryptography - SAC 2011. LNCS, vol. 7118, pp. 118–133. Springer (2012)
44. Tardy-Corfdir, A., Gilbert, H.: A known plaintext attack of FEAL-4 and FEAL-6. In: Advances in Cryptology - CRYPTO'91. LNCS, vol. 576, pp. 172–182. Springer-Verlag (1991)
45. Zheng, Y., Zhang, X.M.: Plateaued functions. In: Information and Communication Security - ICICS'99. LNCS, vol. 1726, pp. 224–300. Springer-Verlag (1999)

A Proofs of Theorem 2, Prop. 2 and 6, and Theorem 3

The new upper bounds on MEDP_2 and MELP_2 exploit the particular structure of the codewords in an $\mathbf{F}_{2^m}$-linear code, which is related to the notion of bundle introduced in [20]. In particular, we use the structure of the subsets of the code of the following form.

Definition 8. *Consider a word c of length n and a subset $I \subseteq \{1, \ldots, n\}$. The decomposition of c with respect to I is denoted by $(x, y)_I$: x corresponds to the restriction of c to I, and y corresponds to the restriction of c to the complement subset $\bar{I}$. For the sake of simplicity, the $|I|$ coordinates of x (resp. the coordinates of y) will be indexed by the elements of I (resp. of $\bar{I}$),* i.e., *$x_i = c_i$ for all $i \in I$ and $y_j = c_j$ for all $j \in \bar{I}$.*

Lemma 2. *Let $\mathcal{C}$ be a linear code of length n, dimension k and minimum distance d over $\mathbf{F}_{2^m}$. For any subset $I \subset \{1, \ldots, n\}$ of size $(n-d)$, and any $x \in (\mathbf{F}_{2^m})^{n-d}$, we define*

$$Z(I, x) = \{y \ : (x, y)_I \in \mathcal{C}\} \ .$$

Then, for any I of size $(n-d)$,

- *either $Z(I, 0)$ is empty or there exists some $y_0 \in (\mathbf{F}_{2^m}^*)^d$ such that $Z(I, 0) = \{\gamma y_0, \ \gamma \in \mathbf{F}_{2^m}\}$;*

– *For any $x \neq 0$, either $Z(I,x)$ is empty or there exist some $y_0 \in (\mathbf{F}_{2^m}^*)^d$ and some $y_1 \in (\mathbf{F}_{2^m})^d$ such that $Z(I,x) \subseteq \{y_1 + \gamma y_0,\ \gamma \in \mathbf{F}_{2^m}\}$.*

Proof. – Assume that $Z(I,0)$ is not empty. Since $\mathcal{C}$ is $\mathbf{F}_{2^m}$-linear, for any $y_0 \in Z(I,0)$, $(0,y_0)_I$ belongs to $\mathcal{C}$, implying that all $\gamma(0,y_0)_I$ with $\gamma \in \mathbf{F}_{2^m}$ belong to $\mathcal{C}$ too.

– Let $x \neq 0$. Since the result obviously holds if $|Z(I,x)| \leq 1$, we suppose that $|Z(I,x)| \geq 2$. For any distinct y and y' in $Z(I,x)$, we get that both $c = (x,y)_I$ and $c' = (x,y')_I$ belong to $\mathcal{C}$, implying that $(y+y') \in Z(I,0)$. From the previous result, there exists some y_0 such that $y+y' = \gamma y_0$ for some $\gamma \in \mathbf{F}_{2^m}$. It follows that y' is of the form $y' = y + \gamma y_0$. Since $wt(c+c') = wt(y+y')$ cannot be less than d, all coordinates of y_0 should be nonzero. □

A.1 Proofs of Theorem 2 and Proposition 2

As in [41], we will use the following generalized version of Hölder inequality.

Lemma 3. *[41, Lemma 1] Let $\{x_i^{(j)}\}_{i=1}^n$, $1 \leq j \leq p$, be p sequences of n real numbers. Then*

$$\sum_{i=1}^{n} \left| \prod_{j=1}^{p} x_i^{(j)} \right| \leq \prod_{j=1}^{p} \left(\sum_{i=1}^{n} |x_i^{(j)}|^p \right)^{\frac{1}{p}}.$$

Now, we prove the following generic version of Theorem 2.

Theorem 4. *Let m and t be two positive integers. Let Λ be a $2^m \times 2^m$ matrix with coefficients $\Lambda(\alpha,\beta), (\alpha,\beta) \in (\mathbf{F}_{2^m})^2$ in $[0;1]$ such that $\Lambda(\alpha,0) = \Lambda(0,\alpha) = 0$ for any $\alpha \neq 0$, and*

$$\sum_{\beta \in \mathbf{F}_{2^m}} \Lambda(\alpha,\beta) = \sum_{\beta \in \mathbf{F}_{2^m}} \Lambda(\beta,\alpha) = 1, \text{ for all } \alpha \in \mathbf{F}_{2^m}.$$

Then, for any $\mathbf{F}_{2^m}$-linear code $\mathcal{C}$ of length $(2t)$ with minimum distance d and for any nonzero a and b in $\mathbf{F}_{2^m}^t$, we have:

$$\Lambda_{a,b} = \sum_{c \in \mathcal{C}} \left(\prod_{i=1}^{t} \Lambda(a_i, c_i) \right) \left(\prod_{j=1}^{t} \Lambda(c_{t+j}, b_j) \right) \leq \max_{1 \leq u < d} \max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}_u(\mu)$$

$$\text{where } \mathcal{B}_u(\mu) = \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha,\gamma)^u \Lambda(\gamma\lambda + \mu, \beta)^{d-u}.$$

Proof. Let a, b be nonzero elements of $(\mathbf{F}_{2^m})^t$. For any codeword c such that $\mathsf{Supp}(c) \neq \mathsf{Supp}(a,b)$, there exists $\ell \in \{1,\ldots,t\}$ such that $\Lambda(a_\ell, c_\ell) = 0$ or $\Lambda(c_{t+\ell}, b_\ell) = 0$. Then,

$$\Lambda_{a,b} = \sum_{c \in \mathcal{C}:\mathsf{Supp}(c) = \mathsf{Supp}(a,b)} \left(\prod_{i=1}^{t} \Lambda(a_i, c_i) \right) \left(\prod_{j=1}^{t} \Lambda(c_{t+j}, b_j) \right).$$

We assume that $wt(a) + wt(b) \geq d$, otherwise the value $\Lambda_{a,b}$ is equal to zero, as there is no $c \in \mathcal{C}$ such that $\mathsf{Supp}(c) = \mathsf{Supp}(a, b)$. Then we can choose a pair of subsets I_1 and I_2 of $\{1, \ldots, t\}$ such that $I_1 \subseteq \mathsf{Supp}(a)$, $I_2 \subseteq \mathsf{Supp}(b)$ and $|I_1| + |I_2| = d$. We decompose any codeword c whose support equals $\mathsf{Supp}((a, b))$ into two parts: $c = (y, x)_I$ where $I = (\{1, \ldots, t\} \setminus I_1) \cup \{t + j, j \notin I_2\}$. In other words, y corresponds to the restriction of c to the positions outside I_1 and I_2, while x corresponds to the other d positions. Recall that, following Definition 8, the coordinates of y (resp. of x) are indexed by the elements of I (resp. of $I_1 \cup \{t + j, j \in I_2\}$). Then, for $Z(I, y) = \{x \ : (y, x)_I \in \mathcal{C}\}$,

$$\Lambda_{a,b} = \sum_{y \in \mathbf{F}_{2^m}^{n-d}} \left(\prod_{i \notin I_1} \Lambda(a_i, y_i) \right) \left(\prod_{j \notin I_2} \Lambda(y_{t+j}, b_j) \right) \mathcal{Q}_{a,b}(I, y) \tag{5}$$

$$\text{where } \mathcal{Q}_{a,b}(I, y) = \sum_{x \in Z(I,y)} \left(\prod_{i \in I_1} \Lambda(a_i, x_i) \right) \left(\prod_{j \in I_2} \Lambda(x_{t+j}, b_j) \right) .$$

We aim at finding an upper bound on $\mathcal{Q}_{a,b}(I, y)$. Let $u = |I_1|$. From Lemma 3,

$$\begin{aligned} \mathcal{Q}_{a,b}(I, y) &= \sum_{x \in Z(I,y)} \prod_{i \in I_1} \left[\Lambda(a_i, x_i) \left(\prod_{j \in I_2} \Lambda(x_{t+j}, b_j) \right)^{\frac{1}{u}} \right] \\ &\leq \prod_{i \in I_1} \left[\sum_{x \in Z(I,y)} \Lambda(a_i, x_i)^u \left(\prod_{j \in I_2} \Lambda(x_{t+j}, b_j) \right) \right]^{\frac{1}{u}} . \end{aligned}$$

For any $i \in I_1$, we apply Lemma 3 again:

$$\begin{aligned} \sum_{x \in Z(I,y)} \Lambda(a_i, x_i)^u \left(\prod_{j \in I_2} \Lambda(x_{t+j}, b_j) \right) &= \sum_{x \in Z(I,y)} \prod_{j \in I_2} \left(\Lambda(a_i, x_i)^{\frac{u}{d-u}} \Lambda(x_{t+j}, b_j) \right) \\ &\leq \prod_{j \in I_2} \left(\sum_{x \in Z(I,y)} \Lambda(a_i, x_i)^u \Lambda(x_{t+j}, b_j)^{d-u} \right)^{\frac{1}{d-u}} \end{aligned}$$

Now, we know from Lemma 2 that, if $Z(I, y) \neq \emptyset$, there exist $\alpha \in (\mathbf{F}_{2^m}^*)^d$ and $\beta \in (\mathbf{F}_{2^m})^d$ such that $Z(I, y) \subseteq \{\gamma\alpha + \beta, \gamma \in \mathbf{F}_{2^m}\}$. Then, for any pair $(i, j) \in I_1 \times I_2$, we can write:

$$\begin{aligned} \sum_{x \in Z(I,y)} \Lambda(a_i, x_i)^u \Lambda(x_{t+j}, b_j)^{d-u} &\leq \sum_{\gamma \in \mathbf{F}_{2^m}} \Lambda(a_i, \gamma\alpha_i + \beta_i)^u \Lambda(\gamma\alpha_{t+j} + \beta_{t+j}, b_j)^{d-u} \\ &= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(a_i, \gamma')^u \Lambda(\gamma'\lambda + \mu, b_j)^{d-u} , \end{aligned}$$

where the last equality is obtained by replacing $\gamma\alpha_i + \beta_i$ by γ' since $\alpha_i \neq 0$, and by setting $\lambda = \alpha_{t+j}\alpha_i^{-1}$ and $\mu = \beta_{t+j} + \alpha_{t+j}\alpha_i^{-1}\beta_i$. Moreover the sum can be taken over all nonzero γ' since $\Lambda(a_i, \gamma') = 0$ for $\gamma' = 0$. Let

$$\mathcal{B}_u = \max_{a,b,\lambda \in \mathbf{F}_{2^m}^*} \max_{\mu \in \mathbf{F}_{2^m}} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(a,\gamma)^u \Lambda(\gamma\lambda + \mu, b)^{d-u} \ .$$

Then, we get

$$\mathcal{Q}_{a,b}(I,y) \leq \prod_{i \in I_1} \prod_{j \in I_2} \left(\sum_{x \in Z(I,y)} \Lambda(a_i, x_i)^u \Lambda(x_{t+j}, b_j)^{d-u} \right)^{\frac{1}{u(d-u)}} \leq \mathcal{B}_u^{\frac{u(d-u)}{u(d-u)}} = \mathcal{B}_u \ .$$

Using (5) and $\sum_{\beta \in \mathbf{F}_{2^m}} \Lambda(\alpha, \beta) = \sum_{\alpha \in \mathbf{F}_{2^m}} \Lambda(\alpha, \beta) = 1$, we eventually deduce

$$\begin{aligned} \Lambda_{a,b} &= \sum_{y \in \mathbf{F}_{2^m}^{n-d}} \left(\prod_{i \notin I_1} \Lambda(a_i, y_i) \right) \left(\prod_{j \notin I_2} \Lambda(y_{t+j}, b_j) \right) \mathcal{Q}_{a,b}(I,y) \\ &\leq \mathcal{B}_u \sum_{y \in \mathbf{F}_{2^m}^{n-d}} \left(\prod_{i \notin I_1} \Lambda(a_i, y_i) \right) \left(\prod_{j \notin I_2} \Lambda(y_{t+j}, b_j) \right) \leq \mathcal{B}_u \ . \end{aligned}$$

□

Theorem 2 is derived by observing that, up to a constant factor, $\Lambda_{a,b} = \mathrm{EDP}_2(a, \mathcal{M}(b))$ for $\Lambda(\alpha,\beta) = 2^{-m}\delta_F(\alpha,\beta)$ and $\mathcal{C}$ is the code $\mathcal{C}_{\mathcal{M}}$ defined in Definition 5, while $\Lambda_{a,b} = \mathrm{ELP}_2(a, (\mathcal{M}^*)^{-1}(b))$ for $\Lambda(\alpha,\beta) = 2^{-2m}\mathcal{W}_F(\alpha,\beta)^2$ when $\mathcal{C}$ is the code $\mathcal{C}_{\mathcal{M}}^{\perp}$ and $\mathcal{M}^*$ denotes the adjoint of $\mathcal{M}$.

In the same way, we now prove the following generic version of Prop. 2.

Proposition 7. *Let m and d be two positive integers and Λ be a $2^m \times 2^m$ matrix satisfying the same hypotheses as in Theorem 4. Then, for any $1 \leq u < d$ and any $\mu \in \mathbf{F}_{2^m}$, we have*

$$\mathcal{B}_u(\mu) \leq \max \left(\max_{a \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(a,\gamma)^d, \max_{b \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\gamma,b)^d \right) \ .$$

Moreover, if $\Lambda(\alpha,\beta) = \Lambda(\beta,\alpha)$ for any $(\alpha,\beta) \in (\mathbf{F}_{2^m})^2$, we have that, for any $1 \leq u < d$,

$$\max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}_u(\mu) = \mathcal{B}_u(0) = \max_{a \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(a,\gamma)^d = \max_{b \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\gamma,b)^d \ .$$

Proof. Lemma 2 implies that, for any set of p sequences $\{x_i^{(j)}\}_{i=1}^n$, $1 \leq j \leq p$,

$$\sum_{i=1}^{n} \left| \prod_{j=1}^{p} x_i^{(j)} \right| \leq \max_{1 \leq j \leq p} \sum_{i=1}^{n} |x_i^{(j)}|^p \ .$$

Using this inequality with $p = d$, we get that, for any $1 \leq u < d$, $\alpha, \beta, \lambda \in \mathbf{F}_{2^m}^*$ and $\mu \in \mathbf{F}_{2^m}$

$$\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma\lambda+\mu)^u \Lambda(\gamma, \beta)^{d-u} \leq \max \left(\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^d, \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\gamma\lambda + \mu, \beta)^d \right) .$$

Since $\lambda \neq 0$, we have $\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\gamma\lambda+\mu, \beta)^d = \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(\gamma', \beta)^d$. The inequality then follows.

Now, we assume that $\Lambda(a, b) = \Lambda(b, a)$ for any pair (a, b). Then,

$$\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\gamma\lambda + \mu, \beta)^{d-u} = \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\beta, \gamma\lambda + \mu)^{d-u} .$$

For $\mu = 0$, the maximum of this value over all nonzero α, β, λ is then greater than or equal to the value obtained for $\beta = \alpha$ and $\lambda = 1$, implying that

$$\mathcal{B}_u(0) \geq \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\alpha, \gamma)^{d-u} = \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^d .$$

Then, $\max_{a \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(a, \gamma)^d$ is a lower bound for $\mathcal{B}_u(0)$, and then for $\max_\mu \mathcal{B}_u(\mu)$. Since we have proved that it is also an upper bound, we conclude that both quantities are equal. □

A.2 Proofs of Proposition 6 and of Theorem 3

We now prove that for any Sbox $\mathcal{S}$ such that either $\mathcal{S}$ or $\mathcal{S}^{-1}$ has multiplicative-invariant derivatives (resp. Walsh transform), the bound defined in Theorem 2 simplifies as explained in Prop. 6. Again, we give a generic version of this proposition which captures both settings.

Proposition 8. *Let m and d be two positive integers and Λ be a $2^m \times 2^m$ matrix satisfying the same hypotheses as in Theorem 4. Let*

$$\mathcal{B}'_u(\mu) = \max_{\alpha, \beta \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\gamma + \mu, \beta)^{(d-u)}, \text{ with } 1 \leq u < d .$$

Assume that one of the following two conditions holds:

- **(i)** *for any $x \in \mathbf{F}_{2^m}^*$ there is a permutation π_x of $\mathbf{F}_{2^m}^*$ such that $\Lambda(\alpha, xy) = \Lambda(\pi_x(\alpha), y)$, $\forall y \in \mathbf{F}_{2^m}^*$;*
- **(ii)** *for any $x \in \mathbf{F}_{2^m}^*$ there is a permutation ψ_x of $\mathbf{F}_{2^m}^*$ such that $\Lambda(xy, \alpha) = \Lambda(y, \psi_x(\alpha))$, $\forall y \in \mathbf{F}_{2^m}^*$.*

Then, the quantities $\mathcal{B}_u(\mu)$ defined in Theorem 4 satisfy

$$\mathcal{B}_u(0) = \mathcal{B}'_u(0) \text{ and } \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}_u(\mu) = \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}'_u(\mu) .$$

Proof. If Condition (i) holds, we have for any $\alpha, \beta, \lambda \in \mathbf{F}_{2^m}^*$ and any $\mu \in \mathbf{F}_{2^m}$,

$$\begin{aligned}\mathcal{B}_u(\alpha,\beta,\lambda,\mu) &= \sum_{\gamma\in\mathbf{F}_{2^m}^*} \Lambda(\alpha,\gamma)^u \Lambda(\gamma\lambda+\mu,\beta)^{d-u} \\ &= \sum_{\gamma'\in\mathbf{F}_{2^m}^*} \Lambda(\alpha,\lambda^{-1}(\gamma'+\mu))^u \Lambda(\gamma',\beta)^{d-u} = \mathcal{B}_u(\pi_{\lambda^{-1}}(\alpha),\beta,1,\mu)\ .\end{aligned}$$

The result then follows. If (ii) holds, we get $\mathcal{B}_u(\alpha,\beta,\lambda,\mu) = \mathcal{B}_u(\alpha,\psi_\lambda(\beta),1,\mu\lambda^{-1})$ in a similar way. □

A generic version of Theorem 3 is then the following.

Theorem 5. *Let m and t be 2 positive integers and Λ a $2^m \times 2^m$ matrix satisfying the hypotheses of Th. 4. Assume that one of the following holds:*

(i) *for any $x \in \mathbf{F}_{2^m}^*$ there is a permutation π_x of $\mathbf{F}_{2^m}^*$ such that $\Lambda(\alpha, xy) = \Lambda(\pi_x(\alpha), y)$, $\forall y \in \mathbf{F}_{2^m}^*$;*

(ii) *for any $x \in \mathbf{F}_{2^m}^*$ there is a permutation ψ_x of $\mathbf{F}_{2^m}^*$ such that $\Lambda(xy, \alpha) = \Lambda(y, \psi_x(\alpha))$, $\forall y \in \mathbf{F}_{2^m}^*$.*

Let $M\Lambda$ be defined by

$$M\Lambda = \max_{a,b\neq 0} \sum_{c\in\mathcal{C}} \left(\prod_{i=1}^{t} \Lambda(a_i, c_i)\right)\left(\prod_{j=1}^{t} \Lambda(c_{t+j}, b_j)\right)\ .$$

with $\mathcal{C}$ any $\mathbf{F}_{2^m}$-linear code of length $2t$, dimension t and $d_{\min} = t+1$. Then,

- *If both (i) and (ii) hold, then $M\Lambda \geq \max_{1\leq u<d} \mathcal{B}'_u(0)$.*
- *If (i) holds, then $M\Lambda \geq \mathcal{B}'_t(0)$.*
- *If (ii) holds, then $M\Lambda \geq \mathcal{B}'_1(0)$.*

Proof. For any fixed u, $1 \leq u \leq t$, we consider $\widehat{\alpha}, \widehat{\beta} \in \mathbf{F}_{2^m}^*$ some values for which

$$\sum_{\gamma\in\mathbf{F}_{2^m}^*} \Lambda(\widehat{\alpha},\gamma)^u \Lambda(\gamma,\widehat{\beta})^{(d-u)} = \mathcal{B}'_u(0)\ .$$

Since $\mathcal{C}$ is MDS, any set of $(t+1)$ positions is the support of a minimum-weight codeword [36, Page 319]. Let then $c \in \mathcal{C}$ with support $I = \{1,\ldots,u\} \cup \{t+1,\ldots,2t+1-u\}$. From Lemma 2, we know that the codewords with support I are the elements $\gamma c, \gamma \in \mathbf{F}_{2^m}^*$. We now examine the 3 cases.

- If both (i) and (ii) hold, then for any pair (a,b), we have

$$\begin{aligned}\Lambda_{a,b} &= \sum_{c\in\mathcal{C}} \left(\prod_{i=1}^{t} \Lambda(a_i,c_i)\right)\left(\prod_{j=1}^{t} \Lambda(c_{t+j},b_j)\right) \\ &= \sum_{\gamma\in\mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(a_i,\gamma c_i)\right)\left(\prod_{j=1}^{t} \Lambda(\gamma c_{t+j},b_j)\right) \\ &= \sum_{\gamma\in\mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(\pi_{c_i}(a_i),\gamma)\right)\left(\prod_{j=1}^{t} \Lambda(\gamma,\psi_{c_{t+j}}(b_j))\right)\ .\end{aligned}$$

We choose a and b as $a_i = \pi_{c_i}^{-1}(\widehat{\alpha})$ for $1 \le i \le u$, $a_i = 0$ otherwise, and $b_j = \psi_{c_{t+j}}^{-1}(\widehat{\beta})$ for $1 \le j \le t+1-u$, $b_j = 0$ otherwise. Then, for these values,

$$\Lambda_{a,b} = \sum_{\gamma \in \mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(\widehat{\alpha}, \gamma) \right) \left(\prod_{j=1}^{t} \Lambda(\gamma, \widehat{\beta}) \right) = \mathcal{B}'_u(0) \ .$$

Since such a pair (a, b) can be defined for any $1 \le u < d$, we deduce that $M\Lambda \ge \max_{1 \le u < d} \mathcal{B}'_u(0)$.

– If only (i) holds, then we consider $u = t$ and we define a and b by $a_i = \pi_{c_i c_{t+1}^{-1}}^{-1}(\widehat{\alpha})$ for $1 \le i \le t$, $b_1 = \widehat{\beta}$ and $b_j = 0$ for $j > 1$. Then, we get

$$\begin{aligned}
\Lambda_{a,b} &= \sum_{\gamma \in \mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(a_i, \gamma c_i) \right) \Lambda(\gamma c_{t+1}, b_1) \\
&= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(a_i, \gamma' c_i c_{t+1}^{-1}) \right) \Lambda(\gamma', b_1) \\
&= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(\pi_{c_i c_{t+1}^{-1}}(a_i), \gamma') \right) \Lambda(\gamma', b_1) \\
&= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(\widehat{\alpha}, \gamma')^t \Lambda(\gamma', \widehat{\beta}) = \mathcal{B}'_t(0) \ .
\end{aligned}$$

– If only (ii) holds, then we choose $u = 1$ and define a and b by $a_1 = \widehat{\alpha}$, $a_i = 0$ for $i > 1$, and $b_j = \varphi_{c_{t+j} c_1^{-1}}^{-1}(\widehat{\beta})$ for $1 \le j \le t$. Then we get that $\Lambda_{a,b} = \mathcal{B}'_1(0)$ □

B Proofs of Propositions 4 and 5

We now prove that for any mapping $\mathcal{S} = \mathcal{S}' \circ A$ where A is an $\mathbf{F}_2$-affine permutation of $\mathbf{F}_{2^m}$ and $\mathcal{S}' : x \mapsto x^s$, both the derivatives of $\mathcal{S}$ and its Walsh transform are multiplicative-invariant.

Proof. From Lemma 1, it is known that

$$\delta_F^{\mathcal{S}}(a, b) = \delta_F^{\mathcal{S}'}(L(a), b) \text{ and } \mathcal{W}_F^{\mathcal{S}}(a, b)^2 = \mathcal{W}_F^{\mathcal{S}'}((L^{-1})^*(a), b)^2 \ ,$$

where $L : x \mapsto A(x) + A(0)$. Since $\mathcal{S}'(x) = x^s$, we have

$$\begin{aligned}
\delta_F^{\mathcal{S}'}(a, bc) &= \#\{x \in \mathbf{F}_{2^m}, (x+a)^s + x^s = bc\} \\
&= \#\{x \in \mathbf{F}_{2^m}, \left(c^{-e}x + c^{-e}a\right)^s + (c^{-e}x)^s = b\} = \delta_F^{\mathcal{S}'}(c^{-e}a, b)
\end{aligned}$$

where $x \mapsto x^e$ is the compositional inverse of $\mathcal{S}'$, *i.e.*, e is the inverse of s modulo $(2^m - 1)$, and

$$\mathcal{W}_F^{\mathcal{S}'}(a, bc) = \sum_{x \in \mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(bcx^s + ax)} = \sum_{x \in \mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(by^s + ac^{-e}y)} = \mathcal{W}_F^{\mathcal{S}'}(c^{-e}a, b) \ .$$

Therefore, it follows that

$$\delta_F^{\mathcal{S}}(a,bc) = \delta_F^{\mathcal{S}'}(c^{-e}L(a),b) = \delta_F^{\mathcal{S}}(\pi_c(a),b) \text{ with } \pi_c(a) = L^{-1}(c^{-e}L(a)) \ ,$$

$$\text{and } \mathcal{W}_F^{\mathcal{S}}(a,bc) = \mathcal{W}_F^{\mathcal{S}'}(c^{-e}(L^{-1})^*(a),b)^2 = \mathcal{W}_F^{\mathcal{S}}(\psi_c(a),b)$$

with $\psi_c(a) = L^*(c^{-e}(L^{-1})^*(a))$, since $(L^{-1})^* = (L^*)^{-1}$. Clearly, both π_c and ψ_c are permutations for any nonzero c. □

Now, we prove a generalized version of Prop. 5, which applies to a (possibly) larger family of mappings named crooked permutations.

Definition 9. *[3] A function $\mathcal{S}$ from $\mathbf{F}_{2^m}$ into $\mathbf{F}_{2^m}$ is said to be* crooked *if, for any nonzero $\alpha \in \mathbf{F}_{2^m}$, $\mathsf{Im}(D_\alpha\mathcal{S})$ is a linear or affine subspace of codimension 1, where $D_\alpha\mathcal{S} : x \mapsto \mathcal{S}(x+\alpha) + \mathcal{S}(x)$.*

It is known that all crooked permutations are APN and almost bent [3], and exist for m odd only. Clearly, any quadratic APN permutation is crooked. And it is highly conjectured that the crooked functions exactly correspond to the quadratic APN functions. This has been proved in [30] in the case of monomial functions and in [4] in the case of binomials. Now we can prove the following.

Proposition 9. *Let $\mathcal{S}$ be a crooked permutation. Then, $\mathcal{S}$ has multiplicative-invariant derivatives and $\mathcal{S}^{-1}$ has a multiplicative-invariant Walsh transform.*

Proof. Since $\mathcal{S}$ is a permutation, for any nonzero a, $D_a\mathcal{S}$ cannot vanish implying that $\mathsf{Im}(D_a\mathcal{S})$ is an affine hyperplane. Moreover, it is known that the $(2^m - 1)$ affine hyperplanes corresponding to $\mathsf{Im}(D_a\mathcal{S})$ for all $a \neq 0$ are distinct [13, Lemma 5]. Therefore, there exists a permutation φ of $\mathbf{F}_{2^m}$ with $\varphi(0) = 0$ such that $\mathsf{Im}(D_aS) = \mathbf{F}_{2^m} \setminus \langle\varphi(a)\rangle^\perp$ for any nonzero a. Moreover, it is known (see *e.g.* [8]) that, for $u, v \in \mathbf{F}_{2^m}^*$,

$$\mathcal{W}_F^2(u,v) = \sum_{a,b\in\mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(au+bv)}\delta_F(a,b) = 2^m + \sum_{a,b\in\mathbf{F}_{2^m},a\neq 0} (-1)^{\mathrm{Tr}(au+bv)}\delta_F(a,b).$$

The differential spectrum of $\mathcal{S}$ is determined by φ: for any $a \neq 0$, $\delta_F(a,b) = 1 - (-1)^{\mathrm{Tr}(\varphi(a)b)}$. Then, for any $v \neq 0$, we get

$$\begin{aligned}
\mathcal{W}_F^2(u,v) &= 2^m + \sum_{a,b\in\mathbf{F}_{2^m},a\neq 0} (-1)^{\mathrm{Tr}(au+bv)} - \sum_{a,b\in\mathbf{F}_{2^m},a\neq 0} (-1)^{\mathrm{Tr}(au+bv+\varphi(a)b)} \\
&= 2^m - \sum_{a\in\mathbf{F}_{2^m},a\neq 0} (-1)^{\mathrm{Tr}(au)} \left(\sum_{b\in\mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(b(v+\varphi(a)))} \right) \\
&= 2^m - 2^m(-1)^{\mathrm{Tr}(u\varphi^{-1}(v))}
\end{aligned}$$

where the last equality uses the fact that $\varphi^{-1}(v) \neq 0$ when $v \neq 0$. It follows that

$$\mathcal{W}_F^2(xy,v) = 2^m - 2^m(-1)^{\mathrm{Tr}(xy\varphi^{-1}(v))} = 2^m - 2^m(-1)^{\mathrm{Tr}(y\varphi^{-1}(\pi_x(v)))}$$

where $\pi_x(v) = \varphi(x\varphi^{-1}(v))$. Moreover, for any nonzero x, π_x is a permutation. □