HAL
archives-ouvertes.fr

Differential Attacks Against SPN: A Thorough Analysis

Anne Canteaut, Joëlle Roué

▶ **To cite this version:**

Anne Canteaut, Joëlle Roué. Differential Attacks Against SPN: A Thorough Analysis. Codes, Cryptology, and Information Security - C2SI 2015, May 2015, Rabat, Morocco. Springer, 9084, pp.45-62, 2015, Lecture Notes in Computer Science. <10.1007/978-3-319-18681-8_4>. <hal-01237293>

HAL Id: hal-01237293
https://hal.inria.fr/hal-01237293

Submitted on 3 Dec 2015

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

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

Differential attacks against SPN: a thorough analysis*

Anne Canteaut and Joëlle Roué

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

Abstract. This work aims at determining when the two-round maximum expected differential probability in an SPN with an MDS diffusion layer is achieved by a differential having the fewest possible active Sboxes. This question arises from the fact that minimum-weight differentials include the best differentials for the AES and several variants. However, we exhibit some SPN for which the two-round MEDP is achieved by some differentials involving a number of active Sboxes which exceeds the branch number of the linear layer. On the other hand, we also prove that, for some particular families of Sboxes, the two-round MEDP is always achieved for minimum-weight differentials.

Keywords. Differential cryptanalysis, linear layer, MDS codes, AES.

1 Introduction

Since the design of the AES and the seminal related work [12], it is known that the mixing layer which aims at providing diffusion within a block cipher must have a high differential branch number [10]. This quantity corresponds to the smallest number of active Sboxes within a two-round differential characteristic. Indeed, for a given choice of the Sbox, the maximal probability for an r-round differential characteristic decreases when the number of active Sboxes within r rounds increases. For this reason, many security analyses focus on the minimal number of active Sboxes within r consecutive rounds when r varies, not only for AES-like designs but for some other types of ciphers, including Present [5] or Feistel ciphers [23]. This approach is rather natural since, in differential attacks, cryptanalysts usually start by searching for a differential characteristic with the fewest possible active Sboxes. Therefore, the construction of MDS diffusion layers with an efficient implementation has been investigated by several authors, e.g., [22, 3, 1].

However, the complexity of a differential attack depends on the probability of *a differential*, *i.e.*, on the sum of the probabilities of all characteristics starting by a given input difference and ending by a given output difference. And, within two consecutive rounds of an SPN (Substitution-Permutation Networks), the

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

number of constituent characteristics increases with the Hamming weight of the differential. Then, the maximum expected probability (MEDP) for a two-round differential may result from a differential which contains a huge number of characteristics each with a low but nonzero probability, rather than from a differential which contains a few characteristics having a high probability. In other words, for two rounds of an SPN, there is *a priori* no reason to believe that the best differential corresponds to a differential with the lowest number of active Sboxes. However, it appears to be the case for most known examples, including the AES [17]. This aim of this paper is then to determine whether this phenomenon is more general and whether there are some general situations where it can be proved that the two-round MEDP is achieved by a differential with the smallest number of active Sboxes.

Our contributions. After recalling the main definitions in Section 2, we show in Section 3 that the choice of the MDS diffusion layer may affect the two-round MEDP even if the Sbox is fixed. In particular, we show that the form of the minimum-weight codewords plays an important role. Also, we provide some upper bound on the number of characteristics with nonzero probability within a given differential for an MDS linear layer. Section 4 focuses on the case where the Sbox is APN: in this case, it appears that the two-round MEDP is usually achieved by minimum-weight differentials. We prove this result for any APN Sbox over $\mathbf{F}_8$ and any $\mathbf{F}_8$-linear MDS diffusion layer. Finally, Section 5 exploits the previous analysis and exhibits some MDS mixing layers for which the maximum EDP over two rounds is achieved by a differential in which the number of active Sboxes exceeds the branch number.

2 Differential attacks against 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 [10, 11], which consists of an alternation of key-independent (usually similar) permutations and of round-key additions. The round permutation is usually composed of a nonlinear substitution function Sub which provides confusion, and of a linear permutation which provides diffusion. 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*.*

For instance, up to a linear transformation, two rounds of the AES can be seen as the concatenation of four similar *superboxes* [13]. 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 $X^8 + X^4 + X^3 + X + 1$.

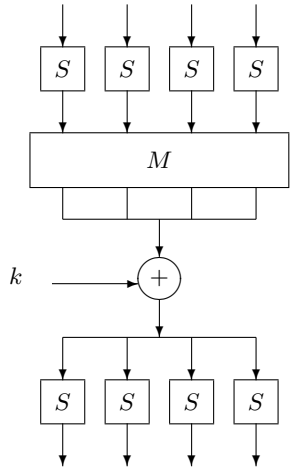

Fig. 1. The AES superbox.

2.2 Differential cryptanalysis

Differential [4] cryptanalysis is one of 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. Since computing the whole distribution of the probability of a differential is a very difficult task, 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.

2.3 Expected probability of a differential characteristic

Since computing the MEDP for most ciphers, even for a small number of rounds, is very difficult, most works focus on the expected probability of a differential characteristic.

Definition 3. *An r-round* differential characteristic Ω *is a collection of $(r+1)$ differences, $\Omega = (a_0, a_1, \ldots, a_r)$ where a_i corresponds to the difference obtained after the i-th round when encrypting two inputs which differ from a_0. The expected probability of the characteristic Ω is then defined as*

$$\mathrm{EDCP}_r(\Omega) = 2^{-\kappa} \sum_{k \in \mathbf{F}_2^\kappa} \Pr_{X_0}[X_1 + X_1' = a_1; \ldots; X_r + X_r' = a_r \,|\, X_0 + X_0' = a_0] \ ,$$

where X_i (resp. X_i') denotes the image of X_0 (resp. of X_0') after the i-th round of E_k.

We here use the specific notation EDCP for the expected probability of a characteristic in order to avoid confusion between differentials and characteristics.

A simple upper-bound on the expected probability of 2-round characteristics can be derived from the *differential branch number* of the linear layer and from the *differential uniformity* of the Sbox, in the sense of the following definition.

Definition 4 (Differential uniformity). *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.*

Then, for any two-round characteristic $\Omega = (a, M(b), M(c))$, the Markov assumption [19] implies that

$$\begin{aligned} \mathrm{EDCP}_2(\Omega) &= \mathrm{DP}_1^E(a, M(b)) \times \mathrm{DP}_1^E(M(b), M(c)) \\ &= \left(\prod_{i=1}^{t} \mathrm{DP}_1^S(a_i, b_i) \right) \left(\prod_{i=1}^{t} \mathrm{DP}_1^S(M(b)_i, c_i) \right) \ . \end{aligned} \tag{1}$$

Let $\mathsf{Supp}(x)$ and $wt(x)$ denote the support and the weight of a vector $x \in \mathbf{F}_2^{mt}$ seen as an element in $(\mathbf{F}_2^m)^t$. Then, the previous equation shows that

$\mathrm{EDCP}_2(\Omega) = 0$ unless $\mathsf{Supp}(a) = \mathsf{Supp}(b)$ and $\mathsf{Supp}(M(b)) = \mathsf{Supp}(c)$. Using this relation, we deduce that

$$\mathrm{EDCP}_2(\Omega) \leq \left(2^{-m}\Delta(S)\right)^{wt(b)+wt(M(b))} .$$

It then appears that the lowest possible value for the weight of a nonzero word of the form $(b, M(b))$ plays a major role in the resistance against differential attacks. This criterion on the diffusion layer of the cipher corresponds to the notion of *differential branch number*.

Definition 5 (Differential branch number [10]). *Let M be a permutation of $(\mathbf{F}_2^m)^t$. We associate to M the code $\mathcal{C}_M$ of length $2t$ and size 2^t over $\mathbf{F}_2^m$ defined by*

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

The differential branch number *of M is the minimum distance of the code $\mathcal{C}_M$.*

The following upper bound on the expected probability of any 2-round differential characteristic then follows:

$$\max_{\Omega} \mathrm{EDCP}_2(\Omega) \leq \left(2^{-m}\Delta(S)\right)^d , \tag{2}$$

where d is the differential branch number of the linear layer.

It is worth noticing that a similar notion is considered in the case of linear cryptanalysis. The *linear branch number* is then the minimum distance of the dual code $\mathcal{C}_M^{\perp}$ but this quantity is out of the scope of this paper. For this reason, in the following, *branch number* always refers to the differential branch number.

. From Singleton's bound, the highest possible value for the branch number of a permutation of $(\mathbf{F}_2^m)^t$ is $(t+1)$ and it corresponds to the case where the associated code $\mathcal{C}_M$ is an MDS (maximum distance separable) code.

3 From characteristics to differentials

The problem with the previous result is that differential cryptanalysis exploits *differentials* and not *characteristics* since the differences obtained after each intermediate round do not matter in the attack. The probability of a differential $(a, M(b))$ then corresponds to the sum of the probabilities of all characteristics with input difference a and output difference $M(b)$. Then, the relevant quantity for two rounds is the maximum of

$$\mathrm{EDP}_2(a, M(b)) = \sum_{x \in \mathbf{F}_2^{mt}} \mathrm{EDCP}_2^E(a, x, M(b)) .$$

Determining the expected probability of a differential, rather than focusing on a single characteristic, is difficult in general.

3.1 Expected probability of a 2-round differential

From Equation (1), any element x of $(\mathbf{F}_2^m)^t$ verifies

$$\mathrm{EDCP}_2(a, M(x), M(b)) = \left(\prod_{i=1}^{t} \mathrm{DP}_1^S(a_i, x_i)\right)\left(\prod_{i=1}^{t} \mathrm{DP}_1^S(M(x)_i, b_i)\right) .$$

If this probability is different from zero, we have that $\mathsf{Supp}(a) = \mathsf{Supp}(x)$ and $\mathsf{Supp}(M(x)) = \mathsf{Supp}(b)$, implying that $(x, M(x)) \in (\mathbf{F}_2^m)^{2t}$ is a word of $\mathcal{C}_M$ having the same support as (a, b). Moreover, by definition of the differential spectrum, $\mathrm{DP}_1^S(\alpha, \beta) = 2^{-m}\delta^S(\alpha, \beta)$. Thus, the two-round probability of a differential is

$$\mathrm{EDP}_2(a, M(b)) = 2^{-mwt(a,b)} \sum_{\substack{c \in \mathcal{C}_M : \\ \mathsf{Supp}(c) = \mathsf{Supp}(a,b)}} \left(\prod_{i \in \mathsf{Supp}(a)} \delta^S(a_i, c_i)\right)\left(\prod_{j \in \mathsf{Supp}(b)} \delta^S(c_{t+j}, b_j)\right) \tag{3}$$

A simple upper bound for the two-round MEDP can then be derived from the branch number of M and from the differential uniformity of the Sbox (see [15] and [12, Section B.2]):

$$\mathrm{MEDP}_2 \leq \left(2^{-m}\Delta(S)\right)^t .$$

This result has then been refined in [9, 21, 8]. The bounds in [15, 9, 21] are invariant under affine equivalence, *i.e.*, their values are the same for two Sboxes S and S' when there exist two affine permutations A_1 and A_2 such that $S' = A_1 \circ S \circ A_2$. However, the exact values of MEDP_2 may differ for Sboxes in the same equivalent class, and there can be a gap between these bounds and the exact value of MEDP_2. In [8], a new upper bound is introduced, that enhances the previously known bounds in the sense that it may vary when the Sbox is composed by an affine permutation. This new bound only applies when the diffusion layer M is linear over the field $\mathbf{F}_{2^m}$, where m is the size of the Sbox, exactly as in the AES. In this case, the linear layer and the Sbox can be represented as functions over the field $\mathbf{F}_{2^m}$ and the representation does not change the MEDP. In particular, the choice of the isomorphism that identifies the vector space $\mathbf{F}_2^m$ with the finite field $\mathbf{F}_{2^m}$ has no influence on the differential properties of the cipher. For this reason, we use the following alternative notation to define an SPN with this representation.

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 $\mathcal{M}$.*

When the Sbox is defined over $\mathbf{F}_{2^m}$, we equivalently define the differential spectrum as follows. 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 $(\alpha, \beta) \in \mathbf{F}_{2^m}$,

$$\delta_F^{\mathcal{S}}(\alpha, \beta) = \#\{x \in \mathbf{F}_{2^m}, \mathcal{S}(x+\alpha) + \mathcal{S}(x) = \beta\} = \delta^S(\varphi^{-1}(\alpha), \varphi^{-1}(\beta)) \ .$$

As the differential properties of any $\mathrm{SPN}(m, t, S, M)$ can be equivalently studied by considering the alternative representation $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$ [14], this paper focuses on the representation of an SPN_F in the field $\mathbf{F}_{2^m}$. 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.

The new bounds on MEDP_2 presented in [8] are derived from the particular structure of the set formed by all codewords in $\mathcal{C}_{\mathcal{M}}$ having a given support, when $\mathcal{M}$ is linear over $\mathbf{F}_{2^m}$. These bounds are expressed in terms of the following quantities. For any Sbox $\mathcal{S}$ over $\mathbf{F}_{2^m}$ with differential spectrum $(\delta_F(a,b))_{a,b \in \mathbf{F}_{2^m}}$ and any branch number d, we define for any $\mu \in \mathbf{F}_{2^m}$ and any integer $u > 0$,

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

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

In the rest of the paper, we will restrict ourselves to the case where *the diffusion layer is linear over* $\mathbf{F}_{2^m}$ *and MDS* (*i.e.*, with branch number $(t+1)$). We also assume that the well-known MDS conjecture [20] is valid, *i.e.*, in our context, that $t \le 2^{m-1}$ for $m > 3$ and $t \le 3$ for $m = 2$. For such MDS diffusion layers, Theorem 2 and Proposition 3 in [8] can be expressed as follows.

Theorem 1. *Let* $\mathcal{S}$ *be a permutation of* $\mathbf{F}_{2^m}$ *and* t *be any integer such that* $t \le 2^{m-1}$.

- *For any* $\mathbf{F}_{2^m}$*-linear diffusion layer* $\mathcal{M}$ *over* $\mathbf{F}_{2^m}^t$ *with maximal branch number, the block cipher* E *of the form* $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$ *satisfies*

$$\mathrm{MEDP}_2^E \le 2^{-m(t+1)} \max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}(\mu) \ .$$

- *There exists an* $\mathbf{F}_{2^m}$*-linear diffusion layer* $\mathcal{M}$ *over* $\mathbf{F}_{2^m}^t$ *with maximal branch number such that*

$$\mathrm{MEDP}_2^E \ge 2^{-m(t+1)} \mathcal{B}(0) \ .$$

In most cases, the values of the two-round MEDP for two ciphers of the form $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M}_1)$ and $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M}_2)$ where $\mathcal{M}_1$ and $\mathcal{M}_2$ are different MDS linear layers differ. The minimum-weight codewords of $\mathcal{C}_{\mathcal{M}}$ have a large influence on this value, as shown in the following example.

Example 1. Let us study the two-round MEDP of the SPN with the same building blocks as the Prøst permutation, which is the core function of several AEAD-schemes submitted to the CAESAR competition [16]. It is worth noticing that

the following results do not provide any direct information on the security of the Prøst permutation: indeed, we study the differential probabilities averaged over all keys while the key is fixed in the Prøst permutation. Two consecutive rounds of the Prøst permutation over $\mathbf{F}_2^{16d}$, $d \geq 1$, can be seen as the parallel application of d copies of a superbox defined over $\mathbf{F}_{16}$. This superbox is of the form $\mathrm{SPN}(4, d, S, M)$ where S is a 4-bit involution named `SubRows` and M corresponds the so-called `MixSlices` transformation. It has been shown in [8] that `MixSlices` is linear over $\mathbf{F}_{16}$ for some particular isomorphism between $\mathbf{F}_2^4$ and $\mathbf{F}_{16}$. Then, Theorem 1 applies and we get that, for any $\mathbf{F}_{16}$-linear layer $\mathcal{M}$, the block cipher E of the form $\mathrm{SPN}_F(4, d, \mathcal{S}, \mathcal{M})$ where $\mathcal{S}$ corresponds to the Prøst Sbox satisfies

$$\mathrm{MEDP}_2^E \leq 2^{-8} .$$

But, when the diffusion layer corresponds to `MixSlices`, we have computed the exact value of the MEDP_2 and obtained that $\mathrm{MEDP}_2 = 3 \times 2^{-11}$, which is smaller than the general upper bound.

However, since both lower and upper bounds in Theorem 1 are equal, we deduce that there exists another diffusion layer $\mathcal{M}$ such that

$$\mathrm{MEDP}_2^E = 2^{-8} .$$

An example of such a diffusion layer is

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

where α is a root of $X^4 + X^3 + 1$. Indeed, the set of codewords of the form $\{\lambda(0,0,0,1,1,1,1,1), \lambda \in \mathbf{F}_{16}^*\}$ belongs to the code associated with this diffusion layer. Then, the differences $a = (0,0,0,1)$ and $b = (1,1,1,1)$ satisfy $\mathrm{EDP}_2(a, \mathcal{M}(b)) = 2^{-8}$.

3.2 Influence of the weight of the differential

The previous example shows that, in some cases, the form of the minimum-weight codewords in $\mathcal{C}_{\mathcal{M}}$ plays an important role when determining the two-round MEDP. We observe from Equation (3) that these codewords are involved in the computation of $\mathrm{EDP}_2(a, \mathcal{M}(b))$ when the weight of the corresponding pair (a, b) is equal to the branch number of $\mathcal{M}$. We then call such a differential a minimum-weight differential. The role played by minimum-weight differentials appears in a more direct way when the Sbox $\mathcal{S}$ has the following additional property [8, Definition 7]. A mapping $\mathcal{S}$ of $\mathbf{F}_{2^m}$ 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

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

Power permutations, and more generally any function resulting from the composition on the right of a power permutation with an $\mathbf{F}_2$-linear permutation, has multiplicative-invariant derivatives. Another example of functions with multiplicative-invariant derivatives are the crooked permutations, which include all APN permutations of degree 2. When an Sbox has this property, the expression of $\mathcal{B}(\mu)$ (including $\mathcal{B}(0)$) simplifies but, more interestingly, we get some universal lower bound on MEDP_2, *i.e.*, which holds for *any* diffusion layer with maximal branch number. For instance, for all Sboxes $\mathcal{S}$ such that both $\mathcal{S}$ and $\mathcal{S}^{-1}$ have multiplicative-invariant derivatives, we obtain that, for *any* $\mathbf{F}_{2^m}$-linear diffusion layer $\mathcal{M}$ with maximal branch number, the corresponding block cipher satisfies

$$2^{-m(t+1)}\mathcal{B}(0) \leq \mathrm{MEDP}_2^E \leq 2^{-m(t+1)} \max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}(\mu) \ . \tag{4}$$

Moreover, $\mathrm{MEDP}_2^E = 2^{-m(t+1)}\mathcal{B}(0)$ if and only if the maximum expected differential probability is achieved by a minimum-weight differential.

Since the probability of a characteristic decreases when the weight of the underlying differential increases, a natural question is to determine in which situations the two-round MEDP is achieved by a minimum-weight differential. This is an important information: computing the two-round MEDP for a given cipher becomes obviously much easier once it is known that only the minimum-weight differentials need to be examined. Surprisingly enough, for all AES-like ciphers which have been investigated, the two-round MEDP is achieved by a minimum-weight differential. For instance, the bounds in [8] applied to the AES Sbox show that for any $\mathbf{F}_{2^8}$-linear layer $\mathcal{M}$, we have

$$53 \times 2^{-34} \leq \mathrm{MEDP}_2 \leq 55.5 \times 2^{-34} \ , \tag{5}$$

where the lower bound corresponds to some minimum-weight differentials. For the particular diffusion layer defined by `MixColumns` in the AES, the exact value of the two-round MEDP_2 computed by a pruning search algorithm [17], is $\mathrm{MEDP}_2 = 53 \times 2^{-34}$. It then corresponds to the lower bound of (5).

There also exist some SPN_F for which the exact value of the two-round MEDP can be directly deduced from the bounds in [8], for instance, when the Sbox $\mathcal{S}$ is an involution with multiplicative-invariant derivatives. In this case, the lower and upper bounds in (4) are equal and do not depend on the MDS diffusion layer. In other words, for any involution with multiplicative-invariant derivatives, the two-round MEDP is always achieved by a minimum-weight differential, for any choice of the MDS linear layer. This holds in particular for the so-called AES naive Sbox, *i.e.* the inversion in $\mathbf{F}_{2^m}$, which satisfies these conditions.

A natural question then arises from these examples: does there exist any cipher of the form SPN_F for which the two-round MEDP is not achieved by a minimum-weight differential? We now investigate this problem, and first exhibit some general families of ciphers for which this situation cannot occur.

3.3 Number of characteristics within a given 2-round differential

For the sake of simplicity, for any differential $(a, \mathcal{M}(b))$, we denote by $(a, c, \mathcal{M}(b))$ the corresponding characteristic where c is the codeword in $\mathcal{C}_{\mathcal{M}}$ defined by the

concatenation of the input and output differences of the first diffusion layer. With this notation, we have

$$\mathrm{EDP}_2(a, \mathcal{M}(b)) = \sum_{\substack{c \in \mathcal{C}_\mathcal{M}: \\ \mathsf{Supp}(c)=\mathsf{Supp}(a,b)}} \mathrm{EDCP}_2(a, c, \mathcal{M}(b)) \; .$$

In this differential, each characteristic having a nonzero probability is defined by a codeword in $\mathcal{C}_\mathcal{M}$ whose support is equal $\mathsf{Supp}(a, b)$. Therefore, we define the *weight of the differential* as the weight $w = wt(a) + wt(b)$. Then, the number of characteristics within a given differential $(a, \mathcal{M}(b))$ of weight w is defined by

$$\begin{aligned}\mathcal{A}_w(a, b) &= \#\{c \in \mathcal{C}_\mathcal{M} : \mathsf{Supp}(c) = \mathsf{Supp}(a, b) \text{ and } \mathrm{EDCP}_2(a, c, \mathcal{M}(b)) \neq 0\} \\ &= \#\{c \in \mathcal{C}_\mathcal{M} : \mathsf{Supp}(c) = \mathsf{Supp}(a, b), \delta_F^\mathcal{S}(a_i, c_i) \neq 0, \forall i \in \mathsf{Supp}(a) \\ &\qquad \text{and } \delta_F^\mathcal{S}(c_{t+j}, b_j) \neq 0, \forall j \in \mathsf{Supp}(b)\} \; .\end{aligned}$$

A first criterion to determine whether the two-round expected differential probability is maximized by a minimum-weight differential or not consists in estimating the number of characteristics involved in a differential having a given weight w. Since we only consider diffusion layers which are linear over $\mathbf{F}_{2^m}$, the codewords in $\mathcal{C}_\mathcal{M}$ having a given support can be gathered in *bundles* as pointed out in [13]: if c belongs to $\mathcal{C}_\mathcal{M}$, then the whole bundle $\mathcal{P}(c) = \{\gamma c, \gamma \in \mathbf{F}_{2^m}^*\}$ is also included in $\mathcal{C}_\mathcal{M}$. It follows that the number of codewords in $\mathcal{C}_\mathcal{M}$ having a given support is always divisible by $(2^m - 1)$. Moreover, for any pair $(\alpha, \beta) \in (\mathbf{F}_{2^m}^*)^2$, the values $\delta_F^\mathcal{S}(\alpha, \gamma\beta)$, when γ varies in $\mathbf{F}_{2^m}^*$, correspond to a row of the difference table of $\mathcal{S}$. Since these coefficients are all even and sum to 2^m, we deduce that, for any permutation $\mathcal{S}$, at least $2^{m-1} - 1$ coefficients among all $(\delta_F^\mathcal{S}(\alpha, \gamma\beta), \gamma \in \mathbf{F}_{2^m}^*)$ vanish, with equality if and only if $\mathcal{S}$ is APN. It then follows that, for any $c \in \mathcal{C}_\mathcal{M}$,

$$\#\{c' \in \mathcal{P}(c) : \mathrm{EDCP}_2(a, c', \mathcal{M}(b)) \neq 0\} \leq 2^{m-1} \; .$$

Differentials of weight $w = t + 1$. Recall that we focus on the case where the diffusion layer has maximal branch number, *i.e.*, where $\mathcal{C}_\mathcal{M}$ is MDS. It is well-known (e.g. [20, Page 319]) that if $\mathcal{C}_\mathcal{M}$ is an MDS code of length $2t$ and dimension t over $\mathbf{F}_{2^m}$, then for each support of size $(t + 1)$, there exist exactly $(2^m - 1)$ codewords (i.e., one bundle) having this support. From the previous discussion, we deduce that, for any minimum-weight differential (a, b)

$$\mathcal{A}_{t+1}(a, b) \leq 2^{m-1} \; .$$

Differentials of weight $w = t + 2$. We now provide a similar upper bound on the number of characteristics within a differential of weight $(t + 2)$.

Proposition 1. *Let $\mathcal{M}$ be an $\mathbf{F}_{2^m}$-linear MDS permutation of $\mathbf{F}_{2^m}^t$. Then, for any differential (a, b) of weight $(t + 2)$, we have*

$$\mathcal{A}_{t+2}(a, b) \leq 2^{m-1}(2^m - (t + 1)) \; .$$

Proof. From the previous discussion, we only have to prove that, for any support I of size $(t+2)$ there exist exactly $(2^m-(t+1))$ distinct bundles having I for support. Let $J=\{i_1,\ldots,i_{t-2}\}$ be the set formed by the $2t-(t+2)=t-2$ coordinates which do not belong to I. The codewords whose support is included in I then correspond to the codewords which vanish on J. Using that any t coordinates of $\mathcal{C}_{\mathcal{M}}$ is an information set [20, Page 321], we deduce that there are exactly $(2^{2m}-1)$ nonzero codewords whose support is included in I. Since we count the number of codewords whose support is equal to I, we need to remove the codewords of weight $(t+1)$ from the previous set. As previously mentioned, for any support of size $(t+1)$, there exists one bundle having this support. Since I contains $(t+2)$ subsets of size $(t+1)$, we need to remove $(t+2)(2^m-1)$ codewords from the previous set. It follows that the number of codewords having I for support is

$$2^{2m}-1-(t+2)(2^m-1)=(2^m-1)(2^m-(t+1))\,.$$

Therefore, $\mathcal{C}_{\mathcal{M}}$ contains exactly $(2^m-(t+1))$ bundles having I for support, implying that

$$\mathcal{A}_{t+2}(a,b)\leq 2^{m-1}(2^m-(t+1))\,.$$

□

Most notably, we deduce from this formula that, when $t=2^{m-1}$, $\mathcal{A}_{t+2}$ may be limited by the maximal value of $\mathcal{A}_{t+1}$. Some application of this result will be detailed in the next section.

4 SPN with an APN Sbox

In this section, we focus on the block ciphers SPN_F which use an APN Sbox. These ciphers are of particular interest in our context since the whole differential spectrum of the Sbox is known. It follows that, for any characteristic within a differential of weight w has probability either 0 or $2^{-w(m-1)}$. Then, we deduce that the expected probability of a differential of weight w only depends on the value of $\mathcal{A}_w(a,b)$:

$$\mathrm{EDP}_2(a,\mathcal{M}(b))=2^{-w(m-1)}\mathcal{A}_w(a,b)\,.$$

It follows that there exists a differential (a,b) of weight $(t+2)$ whose probability is higher than the probability of any minimum-weight differential if and only if, for any (α,β) of weight $(t+1)$,

$$2^{-(t+2)(m-1)}\mathcal{A}_{t+2}(a,b)\geq 2^{-(t+1)(m-1)}\mathcal{A}_{t+1}(\alpha,\beta)$$

or equivalently

$$\mathcal{A}_{t+2}(a,b)\geq 2^{m-1}\mathcal{A}_{t+1}(\alpha,\beta)\,.$$

From Proposition 1, we know that $\mathcal{A}_{t+2}(a,b)\leq 2^{m-1}(2^m-(t+1))$, implying that this situation can only occur if all minimum-weight differentials (α,β) satisfy

$$\mathcal{A}_{t+1}(\alpha,\beta)\leq(2^m-(t+1))\,. \tag{6}$$

For given parameters m and t, we can then directly deduce that, if the number of characteristics in a minimum-weight differential exceeds some bound, then the two-round MEDP cannot be achieved by a differential of weight $(t+2)$.

4.1 APN Sboxes over $\mathbf{F_8}$

We now show that, if $\mathcal{S}$ is an APN permutation over $\mathbf{F}_{2^3}$ (*i.e.*, $m=3$), then the maximum EDP is always achieved by a minimum-weight differential. This result is mainly due to the particular properties of 3-bit APN permutations.

Properties of APN Sboxes over $\mathbf{F_8}$. Since a permutation of $\mathbf{F}_{2^m}$ has degree at most $(m-1)$, all APN Sboxes over $\mathbf{F}_8$ are quadratic, and their inverses are also quadratic. Therefore, they are *crooked* [2, 18], *i.e.*, for any nonzero $a \in \mathbf{F}_{2^3}$, the set $\{b \in \mathbf{F}_{2^3} \ : \ \delta_F^{\mathcal{S}}(a,b) = 2\}$ is an affine hyperplane of $\mathbf{F}_{2^3}$. Furthermore, it is known that all these affine hyperplanes are distinct [7, Lemma 5]. Since the inverse $\mathcal{S}^{-1}$ is also a crooked permutation, the same property holds for the columns of the difference table of $\mathcal{S}$: for any nonzero b, the set $\{a \in \mathbf{F}_{2^3} \ : \ \delta_F^{\mathcal{S}}(a,b) = 2\}$ is an affine hyperplane and all these hyperplanes are distinct.

Minimum-weight differentials. From this algebraic structure, we deduce the maximal value of the expected differential probability of the minimum-weight differentials.

Proposition 2. *Let $\mathcal{S}$ be an APN permutation of $\mathbf{F}_{2^3}$. For any integer t and any $\mathbf{F}_{2^3}$-linear MDS diffusion layer $\mathcal{M}$ over $(\mathbf{F}_{2^3})^t$, the block cipher of the form $\mathrm{SPN}_F(3,t,\mathcal{S},\mathcal{M})$ satisfies*

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=t+1}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = 2^{-2t}.$$

Proof. Let $I = \{i_1, \ldots, i_{t+1}\}$ be a subset of $\{1, \ldots, 2t\}$ of size $(t+1)$. Our aim is to exhibit a pair (a,b) whose support equals I and such that $\mathcal{A}_{t+1}(a,b) = 4$. Such a differential leads to the result since $\mathcal{A}_{t+1}(a,b) = 4$ is the highest value we can have for a minimum-weight differential. Let c be a codeword in $\mathcal{C}_{\mathcal{M}}$ with $\mathsf{Supp}(c) = I$ since such a codeword always exists. Let us choose some nonzero element $a_{i_1} \in \mathbf{F}_{2^3}$. Then, we consider the set $H = \{\beta : \delta_F^{\mathcal{S}}(a_{i_1}, \beta) = 2\}$. Then H is an affine hyperplane. We now define

$$\Gamma = \{c_{i_1}^{-1}\lambda, \lambda \in H\} .$$

Obviously, Γ is also an affine hyperplane. Then, the four codewords in the bundle of c, $c' = \gamma c$ with $\gamma \in \Gamma$, satisfy

$$\delta_F^{\mathcal{S}}(a_{i_1}, c'_{i_1}) = \delta_F^{\mathcal{S}}(a_{i_1}, \lambda c_{i_1}^{-1} c_{i_1}) = 2.$$

Moreover, for any position i_j in I with $i_j \leq t$, the coordinates of these four codewords at position i_j vary in the set $c_{i_j}\Gamma$ which is an affine hyperplane. Therefore,

there exists some a_{i_j} such that this set corresponds to $\{\beta : \delta_F^{\mathcal{S}}(a_{i_j}, \beta) = 2\}$. Similarly, for any position $i_j \in I$ with $i_j > t$, there exists some b_{i_j} such that the affine hyperplane $c_{i_j}\Gamma$ corresponds to $\{\alpha : \delta(\alpha, b_{i_j}) = 2\}$. For this choice of (a, b), we get that, by construction,

$$\mathcal{A}_{t+1}(a, b) = 4 \ ,$$

implying that

$$\mathrm{EDP}_2(a, \mathcal{M}(b)) = 4 \times 2^{-2(t+1)} = 2^{-2t} \ .$$

□

It is worth noticing that we have proved a more general result: for any bundle, we can find a pair (a, b) such that the corresponding differential includes four characteristics from this bundle having a nonzero probability. However, this does not enable us to determine the maximum EDP for higher-weight differentials since the involved codewords correspond to several bundles, and we cannot control the different bundles together.

Higher-weight differentials. Since $\mathcal{C}_{\mathcal{M}}$ is an MDS code over $\mathbf{F}_8$, we have that t is at most 4. Moreover, we deduce from (6) that the maximum two-round EDP cannot be achieved by a differential of weight $(t + 2)$ when $t = 4$ since it would imply that all minimum-weight differentials would satisfy $\mathcal{A}_{t+1}(a, b) \leq 3$ while we have proved that $A_{t+1}(a, b)$ can be equal to 4.

Then, we need to examine all linear MDS codes of length $2t$ and dimension t over $\mathbf{F}_8$ for $t \in \{2, 3\}$. For each of these codes, we have computed the highest value of $\mathcal{A}_{t+2}(a, b)$ we can get for all (a, b) of weight $(t + 2)$. Since the difference tables of all crooked Sboxes over $\mathbf{F}_8$ have the same structure, the maximal value of $\mathcal{A}_{t+2}(a, b)$ over all (a, b) having a given support I corresponds to the largest set Γ of codewords c with support I such that, for each $i \in I$, c_i for all $c \in \Gamma$ belong to the same affine hyperplane.

For $t = 2$, the previous quantity has been computed for all $[4, 2, 3]$-codes over $\mathbf{F}_8$. For all of them, we get that the maximal value for $\mathcal{A}_4(a, b)$ is equal to 8. We then deduce that

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=3}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = 2^{-4} \text{ and } \max_{\substack{x\neq 0,\, y \\ wt(x,y)=4}} \mathrm{EDP}_2(x, \mathcal{M}(y)) = 2^{-8} \times 8 = 2^{-5}.$$

Then, the two-round MEDP is achieved by a minimum-weight differential only.

For $t = 3$, we have computed the highest possible value of $\mathcal{A}_5(a, b)$ for all $[6, 3, 4]$-codes over $\mathbf{F}_8$, and we have obtained that for all these codes, the maximal $\mathcal{A}_5(a, b)$ is 4, implying that

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=4}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = 2^{-6} \text{ and } \max_{\substack{x\neq 0,\, y \\ wt(x,y)=5}} \mathrm{EDP}_2(x, \mathcal{M}(y)) = 2^{-10} \times 4 = 2^{-8}.$$

Moreover, it can be checked that, for all these codes, the maximal $\mathcal{A}_6(a, b)$ is 32, implying that

$$\max_{\substack{x\neq 0,\, y \\ wt(x,y)=6}} \mathrm{EDP}_2(x, \mathcal{M}(y)) = 2^{-12} \times 32 = 2^{-7}.$$

We then deduce the following result.

Proposition 3. *Let $\mathcal{S}$ be an APN permutation of $\mathbf{F}_{2^3}$. For any integer t and any $\mathbf{F}_{2^3}$-linear MDS diffusion layer $\mathcal{M}$ over $(\mathbf{F}_{2^3})^t$, the block cipher of the form $\mathrm{SPN}_F(3, t, \mathcal{S}, \mathcal{M})$ satisfies*

$$\mathrm{MEDP}_2 = 2^{-2t},$$

and this value is achieved by some minimum-weight differentials only.

4.2 APN Sboxes over $\mathbf{F}_{32}$

APN permutations over $\mathbf{F}_{32}$ have been classified in [6] up to equivalence. But since APN permutations over $\mathbf{F}_{32}$ do not have the same algebraic structure as APN permutations over $\mathbf{F}_8$, each function from this classification has to be studied. Moreover, the number of MDS codes with these parameters is also much higher than in the previous case.

We have then computed the maximal value for $\mathcal{A}_{t+1}$ for several APN permutations and MDS permutations with $t = 2, 3$. For $t = 2$, we have always observed that the maximal $\mathcal{A}_{t+1}$ is at least 10. We should then find some differential of weight 4 with $\mathcal{A}_4 \geq 10 \times 2^{5-1} = 160$ to reach the same EDP than the best minimum-weight differential. However, the highest values we have observed for $\mathcal{A}_4$ are between 83 and 92. In other words, the maximum EDP for a differential of weight 4 is slightly higher than half of the maximum EDP for a minimum-weight differential.

For $t = 3$, we have observed that the maximal $\mathcal{A}_{t+1}$ is at least 9. We should then find some differential of weight 5 with $\mathcal{A}_5 \geq 9 \times 2^{5-1} = 144$, while the highest values we have observed for $\mathcal{A}_5$ lie between 54 and 60.

5 MEDP_2 can be tight for a differential of non-minimal weight

It seems that the number $\mathcal{A}_w$ of characteristics having a nonzero probability in a differential of weight $w > t + 1$ cannot be large enough to achieve a two-round EDP higher than the one which can be obtained with minimal-weight differentials. However, in the previously studied cases, the highest probability of a minimal-weight characteristic is always equal to the maximal value $(\Delta(\mathcal{S})/2^m)^{t+1}$. If the probability EDP_2 is minimized for any minimal-weight differential, that is, if the number $\mathcal{A}_{t+1}$ is small and the probabilities of the constituent characteristics are different from $(\Delta(\mathcal{S})/2^m)^{t+1}$, it should be possible to have a differential of weight $w > t + 1$ which has a higher probability than all minimal-weight differentials.

5.1 Examples where MEDP_2 is tight for a differential of weight $(t + 2)$

Sboxes such that only a few entries in the difference table are equal to $\Delta(\mathcal{S})$ are a good choice to avoid the existence of characteristics with probability

$(\Delta(\mathcal{S})/2^m)^{t+1}$ within any given minimum-weight differential. But for differentials of weight $t+2$, the probability of a characteristic also needs to be high. An Sbox with 4 to 6 entries in the difference table equal to $\Delta(\mathcal{S})$ seems to be a good tradeoff, as shown in the following examples. Note that the Sboxes are defined over the vectorial space $\mathbf{F}_2^m$ while the diffusion layer is defined over the field $\mathbf{F}_{2^m}$, as it is done in many concrete specifications (using the binary representation may be relevant to choose the Sbox, for instance in order to minimize the number of gates).

Let S be a permutation of $\mathbf{F}_2^3$ defined by

x	0	1	2	3	4	5	6	7
$S(x)$	0	1	2	3	4	6	7	5

Its differential uniformity is $\Delta(S) = 4$ and there are 6 coefficients equal to 4 in its difference table. Then there exist some $\mathbf{F}_8$-linear permutations with maximal branch number such that there are differentials of weight $(t+2)$ having a higher probability than all minimum-weight differentials. An example of such a diffusion layer with $t = 2$ is

$$\mathcal{M} = \begin{pmatrix} \alpha & \alpha+1 \\ \alpha^2 & \alpha^2+1 \end{pmatrix}$$

where α is a root of $X^3 + X + 1$. We compute the exact value of EDP_2 for all minimum-weight differentials first and then for differentials with weight $d+1 = 4$. We obtain:

$$\max_{\substack{a \neq 0,\, b \\ wt(a,b)=3}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = 2^{-4}$$

as there is only one characteristic of probability 2^{-4} in the differentials having the highest probability, and

$$\max_{\substack{x \neq 0,\, y \\ wt(x,y)=4}} \mathrm{EDP}_2(x, \mathcal{M}(y)) = 2^{-3}$$

as there are some differentials of weight 4 composed of two characteristics of probability 2^{-4}.

Let S be a permutation of $\mathbf{F}_2^4$ defined by

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

Its differential uniformity is $\Delta(S) = 6$ and there are 4 coefficients equal to 6 in its difference table. Then there exist some $\mathbf{F}_{16}$-linear permutations with maximal branch number such that there exist some differentials of weight $(t+2)$ having a higher probability than all minimum-weight differentials. An example of such a diffusion layer with $t = 4$ is

$$\mathcal{M} = \begin{pmatrix} 1 & 1 & \alpha^3 & \alpha^3 \\ \alpha^2+\alpha+1 & 1 & 1 & \alpha^2+\alpha \\ \alpha^2 & \alpha^3+1 & 1 & \alpha^3+\alpha^2+1 \\ \alpha^2+1 & \alpha^3+\alpha^2+\alpha & \alpha^3+\alpha & 1 \end{pmatrix}$$

where α is a root of $X^4 + X + 1$.

We compute the exact value of EDP_2 for all differentials of a given weight. We obtain

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=5}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = 1,2656 \times 2^{-8}, \qquad \max_{\substack{a\neq 0,\, b \\ wt(a,b)=6}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = 1,4238 \times 2^{-8},$$

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=7}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = 1,0942 \times 2^{-10} \text{ and } \max_{\substack{a\neq 0,\, b \\ wt(a,b)=8}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = 1,292 \times 2^{-12}.$$

5.2 Example where $\mathbf{MEDP_2}$ is tight for a differential of weight $(t+3)$

Similarly, we can exhibit an SPN whose two-round MEDP is achieved by some differentials of weight $(t+3)$ only.

Let S be a permutation of $\mathbf{F}_2^4$ defined by

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

It has differential uniformity $\Delta(S) = 8$ and has 4 coefficients equal to 8 in its difference table. An example of an MDS diffusion layer with $t = 3$ such that there are differentials of weight $t + 3 = 6$ having a higher probability than all differentials of weight $(t + 1)$ or $(t + 2)$ is

$$\mathcal{M} = \begin{pmatrix} 1 & \alpha & \alpha^3+\alpha^2+\alpha \\ \alpha^2 & \alpha+1 & \alpha^3+\alpha^2+\alpha+1 \\ \alpha^2+1 & \alpha^2+1 & \alpha^2+1 \end{pmatrix}$$

where α is a root of $X^4 + X + 1$.

By computing the exact value of EDP_2 for differentials with the same weight, we obtain:

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=4}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = \max_{\substack{a\neq 0,\, b \\ wt(a,b)=5}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = 2^{-6}$$

and

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=6}} \mathrm{EDP}_2(a, \mathcal{M}(b)) = 524288 \times 2^{-24} = 2^{-5}\ .$$

In these two examples, the Sboxes are such that there are only a few entries in their difference table which reach the maximum value $\Delta(S)$. Conversely, in the previous section, we have proved that the two-round MEDP is achieved by minimum-weight differentials when the Sbox is an APN permutation, that is, when all the nonzero coefficients of the difference table achieve the maximal value. Then we can wonder whether, when the number of entries in the difference table of the Sbox which are equal to the differential uniformity exceeds some bound, we can deduce that the two-round MEDP is tight for some minimum-weight differential only.

6 Conclusions

In this work, we have shown that the form of the minimum-weight codewords associated to the diffusion layer in an SPN_F affects the two-round MEDP. Moreover, we have exhibited for the first time some SPN such that the two-round MEDP is achieved by some differentials of weight higher than the branch number. On the other hand, we have also proved that this situation cannot occur in some cases, for instance when the Sbox is an APN permutation of $\mathbf{F}_8$. But, we give some concrete examples of round functions for which the highest differential probability is not achieved when the number of active Sboxes is minimized. This observation means that, while the branch number provides an upper bound on the two-round MEDP in any AES-like cipher [15, 12], an attacker searching for the best two-round differential has to consider all possible number of active Sboxes.

Acknowledgments. The authors want to thank Thierry Berger for many stimulating discussions, including the discussions around the relevance of the rank minimum distance of the diffusion layer, which have initiated our work.

References

1. Augot, D., Finiasz, M.: Direct Construction of Recursive MDS Diffusion Layers using Shortened BCH Codes. In: Fast Software Encryption - FSE 2014. LNCS, Springer (2014)
2. Bending, T.D., Fon-Der-Flaass, D.: Crooked Functions, Bent Functions, and Distance Regular Graphs. Electr. J. Comb. 5 (1998)
3. Berger, T.P.: Construction of recursive MDS diffusion layers from Gabidulin codes. In: Progress in Cryptology - INDOCRYPT 2013s. pp. 274–285. LNCS (2013)
4. Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. Journal of Cryptology pp. 3–72 (1991)
5. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J.B., Seurin, Y., Vikkelsoe, C.: PRESENT: an ultra-lightweight block cipher. In: Cryptographic Hardware and Embedded Systems - CHES 2007. LNCS, vol. 4727, pp. 450–466. Springer (2007)
6. Brinkmann, M., Leander, G.: On the classification of APN functions up to dimension five. Designs, Codes and Cryptography 49(1-3), 273–288 (2008)
7. Canteaut, A., Charpin, P.: Decomposing bent functions. IEEE Transactions on Information Theory 49(8), 2004–19 (2003)
8. Canteaut, A., Roué, J.: On the behaviors of affine equivalent sboxes regarding differential and linear attacks. In: Advances in Cryptology - EUROCRYPT 2015, part I. LNCS, vol. 9056. Springer (2015)
9. 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)
10. Daemen, J.: Cipher and hash function design strategies based on linear and differential cryptanalysis. Ph.D. thesis, K.U. Leuven (1995)
11. 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)

12. Daemen, J., Rijmen, V.: The Design of Rijndael: AES - The Advanced Encryption Standard. Information Security and Cryptography, Springer (2002)
13. 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)
14. 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)
15. 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)
16. 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`
17. 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)
18. Kyureghyan, G.M.: Crooked maps in $\mathbf{F}_{2^n}$. Finite Fields and Their Applications 13(3), 713–726 (2007)
19. 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)
20. MacWilliams, F., Sloane, N.: The Theory of Error-Correcting Codes, vol. 16. North-Holland (1977)
21. 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)
22. Sajadieh, M., Dakhilalian, M., Mala, H., Sepehrdad, P.: Efficient recursive diffusion layers for block ciphers and hash functions. J. Cryptology 28(2), 240–256 (2015)
23. Shibutani, K., Bogdanov, A.: Towards the optimality of feistel ciphers with substitution-permutation functions. Des. Codes Cryptography 73(2), 667–682 (2014), `http://dx.doi.org/10.1007/s10623-014-9970-4`