# Differential properties of power functions

Céline Blondeau, Anne Canteaut
and Pascale Charpin*

INRIA Paris-Rocquencourt,
Project-Team SECRET,
B.P. 105
Le Chesnay Cedex 78153, France
E-mail: celine.blondeau@inria.fr
E-mail: anne.canteaut@inria.fr
E-mail: pascale.charpin@inria.fr
*Corresponding author

**Abstract:** Some properties of power permutations, that is, monomials bijective mappings on $\mathbb{F}_{2^n}$, are investigated. In particular, the differential spectrum of these functions is shown to be of great interest for estimating their resistance to some variants of differential cryptanalysis. The relationships between the differential spectrum of a power permutation and the weight enumerator of a cyclic code with two zeroes are provided. The functions with a two-valued differential spectrum are also studied and the differential spectra of several infinite families of exponents are computed.

**Keywords:** differential uniformity; APN function; almost perfect non-linear function; boolean function; power function; permutation; cyclic codes; weight enumerator; differential cryptanalysis.

**Reference** to this paper should be made as follows: Blondeau, C., Canteaut, A. and Charpin, P. (2010) 'Differential properties of power functions', *Int. J. Information and Coding Theory*, Vol. 1, No. 2, pp.149–170.

**Biographical notes:** Céline Blondeau is a PhD Student at INRIA, the French National Institute for Research in Computer Science, within the SECRET Team. She is working on symmetric cryptography.

Anne Canteaut is a Director of Research at INRIA. Currently, she is the Scientific Head of the SECRET Research Team. Her current research interests include cryptography and coding theory.

Pascale Charpin is a Director of Research at INRIA, within the SECRET Team. Her research interests include finite algebra, error-correcting coding and cryptology.

## 1 Introduction

Differential cryptanalysis is the first statistical attack proposed for breaking iterated block ciphers. Its presentation (Biham and Shamir, 1991) then gave rise to numerous works which investigate the security offered by different types of functions with respect to differential attacks. This security is quantified by the so-called *differential uniformity* of the *Substitution box* used in the cipher. Most notably, finding appropriate S-boxes which guarantee that the cipher using them resist differential attacks is a major topic for the last 15 years. Power permutations, that is, monomial permutations, form a class of suitable candidates since they usually have a lower implementation cost in hardware. Moreover, their properties regarding differential attacks can be studied more easily since they are related to the weight enumerators of some cyclic codes with two zeroes (Carlet et al., 1998). However, using power permutations which are optimal for differential cryptanalysis might not be suitable in a cryptographic context.

One reason is that generally such permutations on $\mathbb{F}_{2^n}$ are not known for $n$ even (which is obviously the case in most applications). Actually, the non-existence of almost perfect nonlinear (APN) permutations for even $n$ was conjectured, until the recent announcement of such mappings for $n = 6$ by Dillon (2009). A second important point is that optimal functions usually correspond to extremal objects, which possess very strong algebraic structures. Then, optimal functions might introduce some unsuitable weaknesses within a cipher. Some examples of such weaknesses have been exploited in cryptanalysis, for instance in Jakobsen and Knudsen (1997); Courtois and Pieprzyk (2002) and Canteaut and Videau (2002). For all these reasons, it is important to find some functions which have an almost optimal low differential uniformity. Also, the security of the underlying cipher is affected by some other properties related to the behaviour of the function when the input difference is fixed, besides its differential uniformity.

In this context, this paper investigates the differential properties, namely the whole differential spectrum, of power permutations which have a low differential uniformity. Section 2 recalls some definitions and properties related to the resistance of a function to differential attacks. Section 3 then focuses on differentially four-uniform power permutations, and it points out that the whole differential spectrum of a power permutation may influence its security regarding some variants of differential cryptanalysis, especially truncated differential attacks. Section 4 then investigates the link between the differential spectrum of a power function and the weight enumerators of cyclic codes with two zeroes. Section 5 focuses on the special case of power permutation with a two-valued differential spectrum.

## 2 Definitions and basic properties

In the whole paper, $\#E$ denotes the cardinality of any set $E$. This paper investigates some properties of functions from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^m}$, $m \geq 1$. It mainly focuses on the case $m = n$, but Boolean functions, that is, with $m = 1$, are also involved. Thus, for the sake of clarity, capital letters (e.g. $F$) are used for denoting vectorial functions (i.e. for $m > 1$), and small letters (e.g. $f$) are dedicated to Boolean functions.

### *2.1 Differential characteristics of a function*

The resistance of a cipher to differential attacks and to its variants is quantified by some properties of the *derivatives* of its S(ubstitution)-box, in the sense of the following definition. It is worth noticing that this definition is general: it deals with mappings from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^m}$ for any $m \geq 1$.

Definition 1: *Let $F$ be a function from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^m}$. For any $a \in \mathbb{F}_{2^n}$, the derivative of $F$ with respect to $a$ is the function $D_aF$ from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^m}$ defined by*

$$D_aF(x) = F(x+a) + F(x), \quad \forall x \in \mathbb{F}_{2^n}$$

The resistance to differential cryptanalysis is related to the following quantities.

Definition 2: *Let $F$ be a function from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$. For any $a$ and $b$ in $\mathbb{F}_{2^n}$, we denote*

$$\delta(a,b) = \#\left\{x \in \mathbb{F}_{2^n},\ D_aF(x) = b\right\}$$

*Moreover, $\delta(1,b)$ will be often denoted $\delta(b)$. Then, the differential uniformity of $F$ is*

$$\delta(F) = \max_{a \neq 0,\ b \in \mathbb{F}_{2^n}} \delta(a,b)$$

*Those functions for which $\delta(F) = 2$ are said to be APN.*

The APN property can be equivalently defined as follows.

Proposition 1: *Let $F$ be any function on $\mathbb{F}_{2^n}$. Then, $F$ is APN if and only if, for any non-zero $a \in \mathbb{F}_{2^n}$, the set $\{D_aF(x), x \in \mathbb{F}_{2^n}\}$ has cardinality $2^{n-1}$, that is the functions $D_aF$ are 2-to-1.*

## *2.2 Walsh spectrum of a function*

We now recall some classical tools used for studying the functions from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^m}$.

Any function $F$ from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$ can be expressed as a univariate polynomial in $\mathbb{F}_{2^n}[X]$. The *degree* of $F$ is then the maximal Hamming weight of its exponents:

$$\deg\left(\sum_{i=0}^{2^n-1} \lambda_i X^i\right) = \max\ \left\{\mathrm{wt}(i) | \lambda_i \neq 0\right\}$$

where $\lambda_i \in \mathbb{F}_{2^n}$ and the weight is calculated on the 2-ary expansion of $i$. We denote by Tr the trace function on $\mathbb{F}_{2^n}$, that is, $\mathrm{Tr}(\beta) = \beta + \beta^2 + \cdots + \beta^{2^{n-1}}$.

The function $F$ can also be represented by $n$ Boolean functions of $n$ variables, its Boolean *coordinates*. Note that the coordinates are sometimes called the components of $F$, but it is more convenient for our purpose to use the following definition, like in Nyberg (1995).

Definition 3: *Let $F$ be a function from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$. The linear combinations of the coordinates of $F$ are the Boolean functions*

$$f_\lambda : x \in \mathbb{F}_{2^n} \mapsto Tr(\lambda F(x)), \quad \lambda \in \mathbb{F}_{2^n}$$

*where $f_0$ is the null function. The functions $f_\lambda$ are called the components of $F$.*

We denote by $\mathcal{B}_n$ the set of Boolean functions on $\mathbb{F}_{2^n}$. In our context, the linear functions in $\mathcal{B}_n$ are the functions $\varphi_a$, defined by

$$\varphi_a : x \in \mathbb{F}_{2^n} \mapsto \mathrm{Tr}(ax), \quad a \in \mathbb{F}_{2^n}^* \tag{1}$$

The following notation will be extensively used in this paper. For any $f \in \mathcal{B}_n$, we denote by $\mathcal{F}(f)$ the following value related to the Fourier (or Walsh) transform of $f$:

$$\mathcal{F}(f) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)} = 2^n - 2\mathrm{wt}(f) \tag{2}$$

where $\mathrm{wt}(f)$ is the Hamming weight of $f$, that is, the number of $x \in \mathbb{F}_{2^n}$ such that $f(x) = 1$. The function $f$ is said to be *balanced* if and only if $\mathcal{F}(f) = 0$ or, equivalently, $\mathrm{wt}(f) = 2^{n-1}$.

Definition 4: *The Walsh spectrum of $f \in \mathcal{B}_n$ is the multiset*

$$\left\{\mathcal{F}(f + \varphi_a), \quad a \in \mathbb{F}_{2^n}\right\}$$

*The non-linearity of $f$ is its Hamming distance to the set of all affine functions. It is given by*

$$2^{n-1} - \frac{1}{2}\mathcal{L}(f) \ \ where \ \ \mathcal{L}(f) = \max_{a\in\mathbb{F}_{2^n}} \left|\mathcal{F}\big(f + \varphi_a\big)\right|$$

The lowest possible value for $\mathcal{L}(f)$ is $2^{n/2}$ and this bound is achieved for *bent functions*. A special class of Boolean functions, which includes the bent functions, is the class of *plateaued functions*.

Definition 5 (Zhang and Zheng, 1999; Canteaut et al., 2000): *Let $f \in \mathcal{B}_n$. The function $f$ is said to be plateaued if its Walsh coefficients take at most three values, namely* $0, \pm\mathcal{L}(f)$. *Then, $\mathcal{L}(f) = 2^s$ with $s \geq n/2$.*

*If $s = n/2$ (and $n$ even) then $f$ is bent and its Walsh coefficients take two values only, namely $\pm 2^{n/2}$. Moreover, $f$ is said plateaued optimal if $s = (n + 1)/2$ for odd $n$ and $s = (n + 2)/2$ for even $n$.*

The fact that $s \geq n/2$ comes from the well-known *Parseval relation*:

$$\sum_{a\in\mathbb{F}_{2^n}} \mathcal{F}^2(f + \varphi_a) = 2^{2n}$$

An important remark is that the class of plateaued functions includes all quadratic functions.

The non-linearity of a function $F$ from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$ is now defined by means of the non–linearities of its components.

Definition 6: *Let $F$ be a function from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$ with components $f_\lambda$, $\lambda \in \mathbb{F}^*_{2^n}$. The non-linearity of $F$ is the minimal value of the non-linearities of the $f_\lambda$. It is equal to*

$$\mathcal{N}(F) = 2^{n-1} - \frac{\Lambda(F)}{2} \quad where \quad \Lambda(F) = \max_{\lambda\in\mathbb{F}^*_{2^n}} \mathcal{L}(f_\lambda)$$

The non-linearity of $F$ is a measure of its vulnerability to linear attacks. The functions that have maximal non-linearity are called almost bent (AB) functions. They exist for odd $n$ only.

Definition 7: *Let $F$ be a function from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$ with components $f_\lambda$, $\lambda \in \mathbb{F}_{2^n}$. Then,*

$$\Lambda(F) \geq 2^{(n+1)/2}$$

*The functions $F$ which satisfy*

$$\Lambda(F) = 2^{(n+1)/2}$$

*are said to be AB. They exist when $n$ is odd only. Moreover, if $F$ is AB, then for any $a \in \mathbb{F}_{2^n}$ and for any non-zero $\lambda$*

$$\left\{\mathcal{F}\big(f_\lambda + \varphi_a\big), \quad \lambda \in \mathbb{F}^*_{2^n}, \ a \in \mathbb{F}_{2^n}\right\} = \left\{0, \pm 2^{(n+1)/2}\right\} \tag{3}$$

*that is, all $f_\lambda$, $\lambda \neq 0$, are plateaued optimal.*

The Walsh spectrum of a Boolean function and its derivatives are related by the so-called *sum-of-square indicator* introduced in Zhang and Zheng (1995) and extensively studied in Canteaut et al. (2000, 2001) and Zhang and Zheng (1999). The proof of the following theorem can be found in Canteaut et al. (2000) and Zhang and Zheng (1999).

Definition 8: *The sum-of-square indicator of $f \in \mathcal{B}_n$ is defined by:*

$$\nu(f) = \sum_{a \in \mathbb{F}_{2^n}} \mathcal{F}^2\big(D_a f\big) = 2^{-n} \sum_{a \in \mathbb{F}_{2^n}} \mathcal{F}^4\big(f + \varphi_a\big)$$

Theorem 1: *Any $f \in \mathcal{B}_n$ satisfies $\nu(f) \leq 2^n \mathcal{L}^2(f)$. Equality occurs if and only if $f$ is plateaued, that is*

$$\mathcal{L}(f) = 2^s \text{ and } \nu(f) = 2^{n+2s}, \quad \frac{n}{2} \leq s \leq n \tag{4}$$

## 3 A new point of view on differential cryptanalysis

Most attacks on symmetric cryptographic algorithms are related to some properties of the functions describing the system. For iterated block ciphers, the efficiency of the main cryptanalytic techniques can be measured by some quantities related to the confusion part of the round function, usually named S(ubstitution)-box. In this paper, we focus on the S-boxes which guarantee a high resistance to differential cryptanalysis. This attack successfully applies when two plaintexts with fixed difference lead after the last but one round to outputs whose difference takes a certain value with a *high probability*.

More precisely, the attacker aims at exploiting this property for distinguishing the cipher from a random permutation. Then, the relevant quantity in the attack is the bias with respect to the uniform probability: for an S-box $F$ from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$, the attacker aims at finding a pair $(a, b)$ of input and output differences such that $\Pr[F(x+a)+F(x) = b]$ is significantly higher than $2^{-n}$.

We focus on the case where the S-box is a power function, that is, a monomial function on $\mathbb{F}_{2^n}$. In other words, $F(x) = x^d$ over $\mathbb{F}_{2^n}$. This power function will be denoted by $F_d$. Power functions are very popular S-boxes for symmetric ciphers since they have a relatively low implementation complexity in hardware environments. Studying their resistance to differential attacks is then of great interest. In the case of a power function, $F_d(x) = x^d$, the differential properties can be analysed more easily since, for any non-zero $a \in \mathbb{F}_{2^n}$, the equation $(x+a)^d + x^d = b$ can be written

$$a^d\left(\left(\frac{x}{a}+1\right)^d + \left(\frac{x}{a}\right)^d\right) = b$$

implying that $\delta(a, b) = \delta(1, b/a^d)$ for all $a \neq 0$. Then, if $F_d \colon x \mapsto x^d$ is a monomial function, the differential characteristics of $F_d$ are determined by the values $\delta(1, b)$, $b \in \mathbb{F}_{2^n}$. From now on, this quantity $\delta(1, b)$ is denoted by $\delta(b)$.

Since

$$\#\{b \in \mathbb{F}_{2^n} | \delta(a, b) = i\} = \#\{b \in \mathbb{F}_{2^n} | \delta(b) = i\} \quad \text{for all } a \neq 0,$$

the *differential spectrum* of $F_d$ can be defined as follows.

Definition 9: *Let $F_d(x) = x^d$ be a power function on $\mathbb{F}_{2^n}$. We denote by $\omega_i$ the number of output differences b that occur i times:*

$$\omega_i = \#\{b \in \mathbb{F}_{2^n} | \delta(b) = i\}$$

*The differential spectrum of $F_d$ is the set of $\omega_i$: $\mathbb{S} = \{\omega_0, \omega_2, \ldots, \omega_{\delta(F)}\}$.*

There are basic transformations which preserve $\mathbb{S}$.

Definition 10: *Let $F(x) = x^d$ and $G(x) = x^e$ be permutations of $\mathbb{F}_{2^n}$.*

- *We say that G is in the class of F if it exists k such that $e = 2^k d \bmod 2^n - 1$.*
- *We say that G is the inverse of F if $e = d^{-1} \bmod 2^n - 1$*

The following lemma is well-known.

Lemma 1: *Let $F(x) = x^d$ with $gcd(2^n-1, d) = 1$. Let G be another monomial permutation on $\mathbb{F}_{2^n}$. If G is in the class of F or if G is the inverse of F then G has the same differential spectrum as F.*

Now, we wish to point out that, other than the differential uniformity, the whole differential spectrum of the S-box affects the resistance of the cipher to differential attacks and to its variants.

### 3.1 Power permutations with $\delta(F) = 4$

We focus on S-boxes which are power permutations on $\mathbb{F}_{2^n}$. Some classes of APN power permutations are known when $n$ is odd but APN power permutations do not exist when $n$ is even. Hence, power permutations which are differentially 4-uniform are of great interest when $n$ is even. All APN permutations have the same differential spectrum, which is $\{2^{n-1}, 2^{n-1}\}$. But, when $\delta(F) = 4$, we have a number of distinct differential spectra as we show in the next example.

Example 1: *Let $F(x) = x^{2^n-2}$ over $\mathbb{F}_{2^n}$. Nyberg (1993), proved that $\delta(F) = 4$ with $\omega_4 = 1$ when n is even. So the differential spectrum of F is*

$$\{2^{n-1} + 1, 2^{n-1} - 2, 1\}$$

*On the other hand, the permutations obtained from quadratic and Kasami exponents which are differentially 4-uniform have a differential spectrum equal to $\{2^n - 2^{n-2}, 0, 2^{n-2}\}$ (see Section 5). Another differential spectrum is calculated in Example 4.*

If $\omega_4$ is large, the probability of having $\delta(a, b) = 4$ for a fixed input difference $a$ is not negligible. This affects the security of the corresponding cipher. Indeed, an obvious strategy for finding a good differential characteristic for the whole cipher consists in chaining several one-round differentials with $\delta(a, b) = 4$. This is usually much easier when there are some degrees of freedom in the choice of the output difference. So when $n$ is even, the power permutations which offer the best resistance to differential cryptanalysis are the differentially 4-uniform S-boxes with $\omega_4$ small. A fortiori $\omega_4 = 1$ is the best value. In this context, the inverse function has the best possible differential spectrum when $n$ is even.

### 3.1.1 Differentially 4-uniform power permutations for n between 6 and 25

Table 1 presents all power permutations that are differentially 4-uniform for $n$ between 6 and 25. According to Lemma 1, we calculate the differential spectrum of $x^d$ where the exponent $d$ is the smallest element of its cyclotomic coset modulo $2^n - 1$.

For $n = 12$ (resp. $n = 20$), $x^{73}$ (resp. $x^{1057}$) belongs to the class $x^d$ with $d = 2^{2r} + 2^r + 1$ and $n = 4r$ (see Example 4). For $n = 14$, note that 319 is in the class of the Kasami exponent $319 \times 2^6$. Similarly, for $n = 18, 1279$ is in the class of the Kasami exponent $1279 \times 2^8$. It is also worth noticing that there is no differentially 4-uniform power permutation for odd $n$ between 15 and 25.

## 3.2 Differential squares for truncated differential cryptanalysis

We know that the inverse function for $n$ even is the power permutation which offers the best resistance to differential cryptanalysis because it is differentially 4-uniform with $\omega_4 = 1$. But, other attacks may be mounted against the cipher, like algebraic cryptanalysis (Courtois and Pieprzyk, 2002). Algebraic attacks exploit the intrinsic algebraic structure of a block cipher. In its most common form, the attacker expresses the encryption transformation as a large set of multivariate polynomial equations, and she subsequently attempts to solve the system to recover information about the key. In a cipher, where the S-box is the inverse function, the attacker exploits the fact that $x \times x^{2^n-2} = 1$ for all non-zero $x$. Moreover, it has been pointed out in Canteaut and Videau (2002) that the use of AB functions introduce another weakness which may be exploited in a higher-order differential attack. This vulnerability comes from the fact that all values occurring in the Walsh spectrum of an AB function are divisible by a high power of 2. Now, we are going to present another type of vulnerability related to truncated differential cryptanalysis.

*Truncated differential cryptanalysis* form a class of attacks against block ciphers introduced by Knudsen (1995). Here, we propose a variant where we group several input and output differences together in a subset, named *a square*, of size $v$.

Definition 11: *Let $(a_1, \dots, a_v)$ be $v$ input differences and $(b_1, \dots, b_v)$ be $v$ output differences. A differential square of size $v$ and of parameter $\lambda$ are a set $(a_1, \dots, a_v, b_1, \dots, b_v)$ such that*

$$\forall i \in \{1 \dots v\} \text{ and } \forall j \in \{1 \dots v\} \quad \delta\big(a_i, b_j\big) \geq \lambda \tag{5}$$

*It is called a $(\lambda, v)$-differential square.*

Example 2: *To illustrate our purpose, we consider the APN power permutation $F(x) = x^3$ on $\mathbb{F}_{2^n}$ with $n = 3$. Some $(2, 2)$ differential squares are clearly identified in Table 2. For instance, as*

$$\delta(010, 010) = \delta(010, 100) = \delta(011, 010) = \delta(011, 100) = 2$$

$((010, 011), (010, 100))$ *is a $(2, 2)$-differential square.*

Our purpose, with this new variant, is to improve the complexity of the differential distinguisher for the S-boxes $F$ which have a small $\delta(F)$ but a high $\omega_{\delta(F)}$. We will show that such S-boxes may introduce some weaknesses regarding truncated differential cryptanalysis. Since we aim at presenting the context of our study only, this cryptanalytic aspect is not detailed: we only give an example in order to explain our approach.

**Table 1** Differentially 4-uniform power permutations for *n* between 6 and 25 and their differential spectra

| *n* | *Exponent/inverse* | $\omega_0$ | $\omega_2$ | $\omega_4$ | |
|---|---|---|---|---|---|
| 6 | 5/13 | 48 | 0 | 16 | Quadratic/Kasami |
| | 31/31 | 33 | 30 | 1 | Inverse |
| 7 | 19/47 | 85 | 22 | 21 | |
| 8 | 127/127 | 129 | 126 | 1 | Inverse |
| 9 | 45/125 | 292 | 184 | 36 | |
| 10 | 5/205 | 768 | 0 | 256 | Quadratic |
| | 13/79 | 768 | 0 | 256 | Kasami |
| | 17/181 | 768 | 0 | 256 | Quadratic |
| | 29/247 | 573 | 390 | 61 | |
| | 103/149 | 588 | 360 | 76 | |
| | 223/367 | 603 | 330 | 91 | |
| | 511/511 | 513 | 510 | 1 | Inverse |
| 11 | 79/183 | 1,156 | 760 | 132 | |
| | 109/695 | 1,189 | 694 | 165 | |
| | 251/367 | 1,255 | 562 | 231 | |
| | 463/703 | 1,222 | 628 | 198 | |
| 12 | 73/731 | 2,496 | 1,152 | 448 | Bracken and Leander |
| | 2,047/2,047 | 2,049 | 2,046 | 1 | Inverse |
| 13 | 303/947 | 4,603 | 3,082 | 507 | |
| 14 | 5/3,277 | 12,288 | 0 | 4,096 | Quadratic |
| | 13/1,339 | 12,288 | 0 | 4,096 | Kasami |
| | 17/2,893 | 12,288 | 0 | 4,096 | Quadratic |
| | 65/2,773 | 12,288 | 0 | 4,096 | Quadratic |
| | 205/241 | 12,288 | 0 | 4,096 | Kasami |
| | 319/979 | 12,288 | 0 | 4,046 | Kasami (4,033) |
| | 8,191/8,191 | 8,193 | 8,190 | 1 | Inverse |
| 16 | 32,767/32,767 | 32,769 | 32,766 | 1 | Inverse |
| 18 | 5/52,429 | 196,608 | 0 | 65,536 | Quadratic |
| | 13/20,165 | 196,608 | 0 | 65,536 | Kasami |
| | 17/46,261 | 196,608 | 0 | 65,536 | Quadratic |
| | 241/12,101 | 196,608 | 0 | 65,536 | Kasami |
| | 257/43,861 | 196,608 | 0 | 65,536 | Quadratic |
| | 1,279/12,605 | 196,608 | 0 | 65,536 | Kasami (65,281) |
| | 131,071/131,071 | 131,073 | 131,070 | 1 | Inverse |
| 20 | 1,057/306,539 | 651,264 | 270,336 | 126,976 | Bracken and Leander |
| | 524,287/524,287 | 524,289 | 524,286 | 1 | Inverse |
| 22 | 5/838,861 | 3,145,728 | 0 | 1,048,576 | Quadratic |
| | 13/322,639 | 3,145,728 | 0 | 1,048,576 | Kasami |
| | 17/740,173 | 3,145,728 | 0 | 1,048,576 | Quadratic |
| | 65/709,813 | 3,145,728 | 0 | 1,048,576 | Quadratic |
| | 241/87,019 | 3,145,728 | 0 | 1,048,576 | Kasami |
| | 257/734,419 | 3,145,728 | 0 | 1,048,576 | Quadratic |
| | 1,025/699,733 | 3,145,728 | 0 | 1,048,576 | Quadratic |
| | 3,277/16,639 | 3,145,728 | 0 | 1,048,576 | Kasami (65,281) |
| | 4,033/246,739 | 3,145,728 | 0 | 1,048,576 | Kasami |
| | 5,119/49,981 | 3,145,728 | 0 | 1,048,576 | Kasami (1,047,553) |
| | 2,097,151/2,097,151 | 2,097,153 | 2,097,150 | 1 | Inverse |
| 24 | 8,388,607/8,388,607 | 8,388,609 | 8,388,606 | 1 | Inverse |

**Table 2** $\delta(a, b)$ for the power permutation $F(x) = x^3$ for $n = 3$ ($a$ and $b$ are represented as elements in $\mathbb{F}_2^3$)

| $b$<br>$a$ | 000 | 001 | 010 | 011 | 100 | 101 | 110 | 111 |
|---|---|---|---|---|---|---|---|---|
| 000 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 001 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 |
| 010 | 0 | 0 | 2 | 2 | 2 | 2 | 0 | 0 |
| 011 | 0 | 2 | 2 | 0 | 2 | 0 | 0 | 2 |
| 100 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 |
| 101 | 0 | 2 | 0 | 2 | 2 | 0 | 2 | 0 |
| 110 | 0 | 0 | 2 | 2 | 0 | 0 | 2 | 2 |
| 111 | 0 | 2 | 2 | 0 | 0 | 2 | 2 | 0 |

Any APN permutation is such that $\omega_2 = 2^{n-1}$ (the maximal value). Fixing $v$, we study the $(2, v)$-differential squares for an APN S-box. Since the derivatives of an APN function are 2-to-1, the probability of any $(2, v)$-differential square is $q^* = v/2^{n-1}$. This must be compared to the uniform probability for a $(2, v)$-differential square, which is $q = v/2^n$. Then, the following example compares our truncated differential attack on an APN S-box with a classical differential cryptanalysis on a permutation which is differentially 8-uniform.

Example 3: *Let $n = 11$. We consider two S-boxes:*

- *$S_1$ is the inverse function, $x \mapsto x^{2^n-2}$, which is an APN permutation*
- *the second S-box $S_2$ is defined by a differentially 8-uniform function.*

Our computations lead to the following observations.

- For $S_1$, a $(2, 32)$-differential square can be obtained. Its probability is $q^* = (2 \times 32)/2^{11} = 2^{-5}$ while the uniform probability is $q = (32/2^{11}) = 2^{-6}$.
- For $S_2$, the maximal probability of a differential is $p^* = 8 \times 2^{-11} = 2^{-8}$ while the uniform probability of a differential is $p = 2^{-11}$.

To compare both attacks, we need to compare the minimal number of plaintexts required to distinguish both S-boxes from a random permutation. The number of samples needed for the attacks can be computed with the algorithm presented in Section 4 of Blondeau and Gérard (2009), with parameters $\alpha = 0.1$ and $\beta = 0.1$:

- For $S_1$, the number of samples is equal to 673. Then, the required number of plaintexts/ciphertexts is equal to $673 \times 33/32 = 694$.
- For $S_2$ the number of samples is equal to 955. Then, the required number of plaintexts/ciphertexts is equal to $955 \times 2 = 1,990$.

This example points out that the differential uniformity of an S-box does not completely determine the complexity of a differential attack which aims at distinguishing this S-box from a random permutation: the whole differential spectrum may affect the security of the cipher. Most notably, the existence of large differential squares may lead to a truncated differential attack which is more efficient than the differential attacks for another S-box with a higher differential uniformity.

Remark 1: *In the case where $v = 1$ and $\lambda = 2$, a $(2, 1)$-differential square corresponds to a classical differential. It is worth noticing that $(\lambda, v)$-differential squares do not exist* *if $\omega_\lambda < v$. For instance, for a differentially 4-uniform function with $\omega_4 = 1$, all $(\lambda, v)$-differential squares with $v > 1$ have parameter $\lambda = 2$. In this case, even if the function is differentially 4-uniform, its behaviour regarding differential squares is the same as the behaviour of APN functions. This occurs in particular when $F$ is the inverse function on $\mathbb{F}_{2^n}$, with $n$ even.*

## 4 Power functions and cyclic codes

Let $F$ be a function from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$. Studying the APN property and the AB property (for odd $n$) are equivalent to study the weights of an associated code $C_F$ and of its dual $C_F^\perp$. An extensive study of this link has been presented in Carlet et al. (1998) (see also Charpin, 1998).

In this section, we are mainly interested in *power functions*, $F_d : x \mapsto x^d$. The associated code $C_{F_d}$ will be then denoted by $C_d$. Since several results below do not need the condition $\gcd(2^n - 1, d) = 1$, we will specify when this condition is necessary.

Let $\alpha$ be a primitive root of $\mathbb{F}_{2^n}$. The code $C_d$ associated to $F_d : x \mapsto x^d$ is a binary code of length $N = 2^n - 1$, defined by the following $N \times 2n$-parity-check matrix

$$\mathcal{H}_d = \begin{pmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{N-1} \\ 1 & \alpha^d & \alpha^{2d} & \cdots & \alpha^{d(N-1)} \end{pmatrix} \tag{6}$$

where each entry is viewed as a binary vector. This code has dimension $2^n - 2n - 1$ except if $F_d$ has a linear component, that is, if $x \mapsto \mathrm{Tr}(\lambda x^d)$ has degree 0 or 1 for some non-zero $\lambda$.

Actually, $C_d$ is the cyclic code with two zeroes, $\alpha$ and $\alpha^d$. And $C_d^\perp$ can be defined by means of its so-called Mattson–Solomon polynomial. It is the space of binary codewords:

$$\left\{ \left( \mathrm{Tr}\left(a\alpha^{id} + b\alpha^i\right) \middle| i = 0, \ldots, N \right), \quad a \in \mathbb{F}_{2^n} \text{ and } b \in \mathbb{F}_{2^n} \right\}$$

When $\gcd(d, 2^n - 1) = 1$, the weight distribution of $C_d^\perp$ is determined by the weights of the above codewords for $a = 1$ and $b \in \mathbb{F}_{2^n}$ only.

It is well-known that the code $C_d$ has a minimum distance $\delta$ such that $3 \le \delta \le 5$ and that $F_d$ is APN if and only if $d = 5$ (Carlet et al., 1998, Theorem 5). We are going to look at the codewords of weight 3 and 4. Note that $\mathbf{c}$ is a codeword of $C_d$ if and only if $\mathcal{H}_d \mathbf{c}^t = 0$ where $\mathbf{c}^t$ is the transposed of the binary vector $\mathbf{c}$ of length $N$. The following result is currently known (see Charpin et al. (1997) for more details).

Proposition 2: *Let $d$ be an integer which is not a power of 2. The code $C_d$ has minimum distance 3 if and only if the polynomial*

$$U_d(x) = 1 + x^d + (1 + x)^d$$

*has at least one root in $\mathbb{F}_{2^n} \backslash \{0, 1\}$. Moreover, the number of codewords of weight 3 in $C_d$ is*

$$B_3 = \frac{(2^n - 1)}{6} \#\left\{ x \in \mathbb{F}_{2^n} \backslash \{0, 1\} \middle| U_d(x) = 0 \right\}$$

*Sketch of proof:* Any codeword of weight 3 is a pair $(a, b)$ of elements of $\mathbb{F}_{2^n}^*$ which satisfies

$$a^d + b^d + (a + b)^d = 0 \text{ with } a \neq b$$

Shifting by $a^{-1}$ we get $1 + (b/a)^d + (1 + (b/a))^d = 0$.

Let $W_3 = \{x \in \mathbb{F}_{2^n} \backslash \{0, 1\} | U_d(x) = 0\}$. For any $x \in W_3$, we have $x + 1 \in W_3$ and then get, by shift, $\#W_3(2^n - 1)/2$ distinct codewords. Moreover, each codeword is taken three times. □

On the other hand, any codeword of weight 4 in $C_d$ is a triple $(x, y, z)$ of three non-zero distinct elements of $\mathbb{F}_{2^n}$ satisfying

$$x^d + y^d + z^d + (x + y + z)^d = 0$$

Since $y = x + a$ for some $a$, we get an equivalent equation

$$x^d + (x + a)^d + z^d + (a + z)^d = 0$$

which is equivalent to

$$a^d \left( \left(\frac{x}{a}\right)^d + \left(\frac{x}{a} + 1\right)^d + \left(\frac{z}{a}\right)^d + \left(\frac{z}{a} + 1\right)^d \right) = 0$$

Thus, we have proved:

Lemma 2: *The numbers $B_3$ and $B_4$ of codewords of weight* 3 *and* 4 *in $C_d$ satisfy*

$$B_3 + B_4 = \frac{(2^n - 1)}{24} \#W_4$$

*where $W_4$ is the following set:*

$$W_4 = \left\{(x, y) \in \mathbb{F}_{2^n} \times \mathbb{F}_{2^n} | x \neq y \neq y + 1 \text{ and } U_d(x) = U_d(y)\right\} \tag{7}$$

According to Definition 8, it is clear that the sum-of-square indicator is related to the quantity $\#W_4$. A precise relationship is obtained when $x \mapsto x^d$ is a permutation.

Proposition 3: *Let $W_4$ be the set defined by (7). Let $f(x) = Tr(x^d)$ where $d$ is coprime to $2^n - 1$. Then*

$$\nu(f) = 2^n \#\left\{(x, y) \in \mathbb{F}_{2^n} \times \mathbb{F}_{2^n} | U_d(x) = U_d(y)\right\} = 2^{2n+1} + 2^n \#W_4$$

*Proof*: We simply compute $\nu(f)$. For clarity, we use the notation $e(P(x)) = (-1)^{\mathrm{Tr}(P(x))}$ where $P$ is any function on $\mathbb{F}_{2^n}$.

$$\begin{aligned}
\nu(f) &= \sum_{a \in \mathbb{F}_{2^n}} \left( \sum_{z \in \mathbb{F}_{2^n}} e\big(z^d + (z + a)^d\big) \right)^2 \\
&= \sum_{a,u,v} e\big(u^d + (u + a)^d + v^d + (v + a)^d\big) \\
&= \sum_{x,y} \sum_{a} e\big(a^d \big(x^d + (x + 1)^d + y^d + (y + 1)^d\big)\big) \\
&= 2^n \# \{(x, y) \in \mathbb{F}_{2^n} \times \mathbb{F}_{2^n} | U_d(x) = U_d(y)\}
\end{aligned}$$

since $a \mapsto a^d$ is a permutation. It remains to see that there are $2^{n+1}$ pairs $(x, y)$ such that $x = y$ or $x = y + 1$. □

The previous proposition implies that, as soon as $\nu(f)$ is known, the sum of the numbers of codewords of weight 3 and 4 in $C_d$ is known. But, in this case, we also have more information on the differential spectrum of $F_d$. Recall that $\delta(F_d)$ and $\delta(b) = \delta(1, b)$ (for any $b \in \mathbb{F}_{2^n}$) are introduced in Definition 2.

Proposition 4: *Let $F_d(x) = x^d$ be a power permutation of $\mathbb{F}_{2^n}$, that is, with $gcd(2^n - 1, d) = 1$. Let $f(x) = Tr(x^d)$. Then*

$$2^{-n}\nu(f) = \sum_{x\in\mathbb{F}_{2^n}} \delta\big(x^d + (x+1)^d\big)$$

*Consequently, $\delta(F_d) \geq 2^{-2n}\nu(f)$.*

*Proof*: Set $\Gamma = 2^{-n}\nu(f)$. From Proposition 3, we get

$$\Gamma = \#\big\{(x, y) \in \mathbb{F}_{2^n} \times \mathbb{F}_{2^n} \big| U_d(x) = U_d(y)\big\}$$

This implies

$$\Gamma = \sum_{x\in\mathbb{F}_{2^n}} \#\big\{y \in \mathbb{F}_{2^n} | y^d + (y+1)^d = b,\ b = x^d + (x+1)^d\big\}$$

$$= \sum_{x\in\mathbb{F}_{2^n}} \delta\big(x^d + (x+1)^d\big) \leq 2^n\delta\big(F_d\big)$$

since $\delta(b) \leq \delta(F_d)$ for any $b$. □

The previous relation between the sum-of-square indicator of a power permutation and its differential spectrum leads to the following expression of the numbers of words of weight 3 and 4 in the cyclic code $C_d$. Recall that $\omega_i = \#\{b \in \mathbb{F}_{2^n} | \delta(b) = i\}$ (see Definition 9).

Corollary 1: *Let $F_d(x) = x^d$ be a power permutation of $\mathbb{F}_{2^n}$, that is, with $gcd(2^n - 1, d) = 1$. The numbers $B_3$ and $B_4$ of codewords of weight 3 and 4 in $C_d$ are given by:*

$$B_3 = \frac{(2^n - 1)}{6}(\delta(1) - 2)$$

$$B_4 = \frac{(2^n - 1)}{24}\left(\sum_{b\in\mathbb{F}_{2^n}} \delta(b)^2 - 2^{n+1} - 4(\delta(1) - 2)\right)$$

$$= \frac{(2^n - 1)}{24}\left(\sum_{i=0}^{2^n}(i^2 - 2)\omega_i\right) - B_3$$

*Proof*: First, $B_3$ is deduced from Proposition 2 by noticing that the cardinality of $\{x \in \mathbb{F}_{2^n}, U_d(x) = 0\}$ equals $\delta(1)$. Now, Lemma 2 and Proposition 3 imply that

$$B_3 + B_4 = \frac{(2^n - 1)}{24}\big(2^{-n}\nu(f) - 2^{n+1}\big)$$

where $f : x \mapsto \mathrm{Tr}(x^d)$. Moreover, it is known from Proposition 4 that

$$2^{-n}\nu(f) = \sum_{b\in\mathbb{F}_{2^n}} \delta(b)^2 = \sum_{i=0}^{2^n} i^2\omega_i \tag{8}$$

which leads to the result. □

Another consequence of Proposition 4 is that the differential spectrum of a differentially 4-uniform power permutation is determined by its sum-of-square indicator $\nu(f)$.

Corollary 2: *Let $F_d(x) = x^d$ a power permutation of $\mathbb{F}_{2^n}$, that is, with* $\gcd(2^n - 1, d) = 1$. *Assume that $\delta(F_d) = 4$. Then, its differential spectrum is given by*

$$\omega_2 = 2^{n-1} - 2\omega_4 \text{ and } \omega_4 = \frac{\nu(f)}{2^{n+3}} - 2^{n-2} \tag{9}$$

*where $f(x) = Tr(x^d)$. Consequently, $\nu(f) = 2^{n+3}\kappa$ with $2^{n-2} < \kappa \leq 2^{n-1}$. In particular, if $\kappa = 2^{n-1}$ then $\omega_2 = 0$.*

*Proof*: Equation (8) can be written as

$$2^{-n}\nu(f) = 2^2\omega_2 + 2^4\omega_4 \text{ with } 2\omega_2 + 4\omega_4 = 2^n$$

By replacing $\omega_2$ by $(2^{n-1} - 2\omega_4)$, we get $\omega_4 = \nu(f)/2^{n+3} - 2^{n-2}$ and then prove (9).

We deduce that $\nu(f) = 2^{n+3}\kappa$ with $\kappa > 0$. Since $0 < \omega_4 \leq 2^{n-2}$, we must have $2^{n-2} < \kappa \leq 2^{n-1}$. In particular, $\omega_2 = 0$ if and only if $\kappa = 2^{n-1}$. □

Example 4: *According to Definition 8, $\nu(f)$ is known as soon as the Walsh spectrum of $f$ is known. It is the case for the following function defined on $\mathbb{F}_{2^n}$, where $n = 4r$ with $r$ odd:*

$$f(x) = Tr(F_d(x)), \quad F_d(x) = x^d, \quad d = 2^{2r} + 2^r + 1$$

*It is known that $F_d$ is a permutation whose components are highly non-linear. More precisely, it was proved by Dobbertin (1998) that the Walsh spectrum of $f$ is:*

| $\mathcal{F}(f + \varphi_u)$ | *Number of u* |
|---|---|
| $-2^{2r+1}$ | $(2^{n-2} - 2^{3(r-1)})/3 - 2^{2r-2}$ |
| $-2^{2r}$ | $(2^{n-1} + 2^{3r-1})/3$ |
| $0$ | $2^{n-1} - 2^{3r-2}$ |
| $2^{2r}$ | $(2^{n-1} + 2^{3r-1})/3$ |
| $-2^{2r+1}$ | $(2^{n-2} - 2^{3(r-1)})/3 + 2^{2r-2}$ |

*Then*

$$\begin{aligned} 2^n\nu(f) &= 2^{4(2r+1)}\frac{(2^{n-1} - 2^{3(r-1)+1})}{3} + 2^{8r}\frac{(2^n + 2^{3r})}{3} \\ &= 2^{8r}\frac{(2^{n+3} - 2^{3r+2} + 2^n + 2^{3r})}{3} \\ &= 2^{8r}\frac{(9.2^n - 3.2^{3r})}{3} = 2^{11.r}(3.2^r - 1) \end{aligned}$$

*so that $\nu(f) = 2^{7r}(3.2^r - 1)$. Recently, Bracken and Leander (2009) proved that $\delta(F_d) = 4$. Thus, with notation of Corollary 2, we get $\kappa = 2^{3r-1}(3.2^r - 1)$, implying*

$$\omega_4 = 2^{3r-3}(2^{r+1} + 2^r - 1) - 2^{4r-2} = 2^{3r-3}(2^r - 1)$$

*and*

$$\omega_2 = 2^{4r-1} - 2^{3r-2}(2^r - 1) = 2^{4r-2} + 2^{3r-2} = 2^{3r-2}(2^r + 1)$$

## 5 Two-valued differential spectra

A function $F$ is APN if and only if $\delta(F) = 2$. This last condition means that for any pair $(a, b)$, $a \in \mathbb{F}_{2^n}^*$ and $b \in \mathbb{F}_{2^n}$, we have $\delta(a, b) \in \{0, 2\}$. In this section, we will examine the case where $\delta(a, b)$ takes two values only, that is, $\delta(a, b) \in \{0, \kappa\}$ for any $(a, b)$ and for $\kappa \geq 2$.

### *5.1 General properties*

First, we observe that in this case $\kappa$ must be a power of 2. This result holds for any function $F$ from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$, not only for power functions. In this general case, the differential spectrum of the function is composed of the values $\delta(a, b)$, for all non-zero $a \in \mathbb{F}_{2^n}$ and all $b \in \mathbb{F}_{2^n}$.

Lemma 3: *Let $F$ be a function from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$. Assume that $F$ has a two-valued differential spectrum. Then $\delta(F) = 2^s$ for some $s$, $1 \leq s \leq n$.*

*Proof*: It is simply because in this case the function $x \mapsto D_a F(x)$ is $\delta(F)$-to-1 for any fixed $a$. Then, $\delta(F)$ is a power of 2 since

$$\#\{b | \delta(a, b) \neq 0\} = \frac{2^n}{\delta(F)}$$

□

One major characteristic of power permutations $F_d$ with a two-valued differential spectrum is that the sum-of-square indicator and the number of codewords of weight 3 and 4 in the associated code $C_d$ are completely determined by the value of $\delta(F_d)$.

Proposition 5: *Let $F_d(x) = x^d$ be a power permutation of $\mathbb{F}_{2^n}$ with a two-valued differential spectrum where $\delta(F_d) = 2^s$ for some $1 \leq s \leq n$. Let $f(x) = Tr(x^d)$. Then, $\nu(f) = 2^{2n+s}$. Moreover, the associated code $C_d$ has minimum distance 3 and the numbers $B_3$ and $B_4$ of codewords of weight 3 and 4 in $C_d$ are given by:*

$$B_3 = \frac{(2^n - 1)}{3}\left(2^{s-1} - 1\right)$$

$$B_4 = \frac{(2^n - 1)}{3}\left(2^{n-2} - 1\right)\left(2^{s-1} - 1\right)$$

*Proof*: The value of $\nu(f)$ is deduced from Proposition 4. The numbers of codewords of weight 3 and 4 are obtained by Corollary 1, using that $\delta(1) \neq 0$ since $x = 0$ and $x = 1$ both satisfy $(x + 1)^d + x^d = 1$. □

We are going to examine specific functions which may have a two-valued differential spectrum. The first family that we investigate is the family of plateaued functions.

Proposition 6: *Let $d$ be an integer such that* $\gcd(d, 2^n - 1) = 1$. *Let $F_d(x) = x^d$ and $f(x) = Tr(x^d)$. Assume that $f$ is a plateaued Boolean function with Walsh spectrum $\{0, \pm 2^{(n+k)/2}\}$.*

*Then $\delta(F_d) \geq 2^k$ with equality if and only if $\delta(b) \in \{0, 2^k\}$ for any $b$. Moreover, if any non-zero $\delta(b)$ is greater than or equal to $2^k$ then $\delta(b) \in \{0, 2^k\}$ for any $b$.*

*Proof*: Since $f$ is plateaued then $\nu(f) = 2^{2n+k}$ (see Definition 5 and Theorem 1). Now, from Proposition 4, we get

$$2^{n+k} = \sum_{x \in \mathbb{F}_{2^n}} \delta\left(x^d + (x+1)^d\right)$$

which implies $2^k \leq \delta(F_d)$. Thus, $\delta(F_d) \geq 2^k$ and equality holds if and only if any $\delta\left(x^d + (x+1)^d\right)$ above is equal to $2^k$. On the other hand, it is clearly impossible to have $\delta(b) \geq 2^k$, for any non-zero $\delta(b)$, unless $\delta(b) \in \{0, 2^k\}$ for any $b$. □

Note that, in the case of plateaued functions with Walsh spectrum $\{0, \pm 2^{(n+k)/2}\}$, the sum of the numbers of codewords of weight 3 and 4 in the associated code is fixed and given by

$$\begin{aligned} B_3 + B_4 &= \frac{2^n - 1}{24}\left(2^{-n}\nu(f) - 2^{n+1}\right) \\ &= \frac{2^n - 1}{24}\left(2^{n+k} - 2^{n+1}\right) \end{aligned}$$

In the remainder of this section, we will examine some examples.

## 5.2 *The quadratic exponents*

In this section, we consider functions $Q_t$ on $\mathbb{F}_{2^n}$ defined by $Q_t(x) = x^{2^t+1}$. Such a power function is said to have a *quadratic exponent*.

Recall that (cf. McEliece, 1987, Lemma 11.1)

$$\gcd(2^t + 1, 2^n - 1) = \begin{cases} 1, & \text{if } \gcd(t, n) = \gcd(2t, n) \\ 2^{\gcd(t,n)} + 1, & \text{if } 2\gcd(t, n) = \gcd(2t, n) \end{cases} \tag{10}$$

Also, the Walsh spectrum of $f(x) = \mathrm{Tr}(x^{2^t+1})$ is known (see McEliece, 1987, Chapter 11). It is, with notation $k = \gcd(2t, n)$:

| $\mathcal{F}(f + \varphi_u)$ | Number of $u$ |
|---|---|
| $0$ | $2^n - 2^{n-k}$ |
| $2^{(n+k)/2}$ | $2^{n-k-1} + 2^{(n-k-2)/2}$ |
| $-2^{(n+k)/2}$ | $2^{n-k-1} - 2^{(n-k-2)/2}$ |

(11)

For functions with quadratic exponents it is very easy to compute $\delta(Q_t)$. Indeed, we have

$$x^{2^t+1} + (x+1)^{2^t+1} = x^{2^t} + x + 1$$

And for any $b$, the equation $x^{2^t} + x + 1 + b = 0$ has either 0 solution or $2^s$ solutions, where $s = \gcd(t, n)$. Thus, for any $t$, we have clearly $\delta(Q_t) = 2^s$ with $\delta(b) \in \{0, 2^s\}$. Note that we find again a well-known result: the function $Q_t$ is APN if and only if $\gcd(t, n) = 1$. For further purposes, we will need a more precise result on the structure of the set of solutions of $Q_t(x) + Q_t(x + a) = b$ in this case.

Lemma 4: *Let $Q_t$ be the function on $\mathbb{F}_{2^n}$ defined by $Q_t(x) = x^{2^t+1}$ where $\gcd(t, n) = s$ with $s \geq 1$. Let us consider the equation*

$$Q_t(x) + Q_t(x + a) = b \tag{12}$$

*for any $a, b$ in $\mathbb{F}_{2^n}$. If Equation (12) has at least one solution $x$, then the set of its solutions is $x + a\mathbb{F}_{2^s}$.*

*Proof*: Suppose that the pair $(a, b)$ is such that (12) holds for at least one element, namely $y$. Then, there is at least one solution of

$$x^{2^t+1} + (x+a)^{2^t+1} = x^{2^t}a + a^{2^t}x + a^{2^t+1} = b \tag{13}$$

Since we have here an equation of the form $A(x) = 0$ where $A$ is an affine function on $\mathbb{F}_{2^n}$, the number of the solutions of (13) is either 0 or the same as the number of solutions of the linear part of $A(x)$. That is

$$x^{2^t}a + a^{2^t}x = 0 \tag{14}$$

But, $x^{2^t}a + a^{2^t}x = ax(x^{2^t-1} + a^{2^t-1})$. Thus, the set of solutions of Equation (14) is clearly $a\mathbb{F}_{2^s}$. We conclude that (13) has $2^s$ solutions; more precisely, by linearity the set of its solutions is $y + a\mathbb{F}_{2^s}$. □

### 5.3 Kasami exponents

In this section, we study a subclass of the power functions defined as follows.

Definition 12: *Let $t$ be an integer such that $2 \le t \le n/2$. Let us define the functions on $\mathbb{F}_{2^n}$:*

$$K_t : x \longmapsto x^{2^{2t}-2^t+1}$$

*Any such exponent is called a Kasami exponent.*

Remark 2: *Recall the following identity that we will use in the proof of the next theorem:*

$$2^{3t} + 1 = (2^t + 1)(2^{2t} - 2^t + 1) \tag{15}$$

*We can have $3t = n + k$ with $k \ge 0$ and in this case $2^{3t} + 1 \equiv 2^k + 1$ modulo $2^n - 1$. If $3t = n$ then the inverse function of $K_t$ is $x \mapsto x^{2^{3t-1}(2^t+1)}$. Indeed,*

$$\left(x^{2^{2t}-2^t+1}\right)^{2^t+1} = x^{2^{3t}+1} = x^2$$

*So, when $3t = n$, the differential spectrum of $K_t$ is the same as the differential spectrum of the quadratic function $Q_t$, its inverse function.*

It is well-known that $K_t$ is APN if and only if $\gcd(t, n) = 1$. So we will focus here on those $t$ satisfying $s = \gcd(t, n) > 1$. Moreover, we will suppose that $n/\gcd(t, n)$ is odd, which implies that $\gcd(2^{rt} + 1, 2^n - 1) = 1$ for any odd $r$. In this case, we deduce from (15) that $d = 2^{2t} - 2^t + 1$ is coprime to $2^n - 1$. Also, the Walsh spectrum of $f_t(x) = \mathrm{Tr}(K_t(x))$ is known from Kasami (1971): it consists precisely of $\{0, \pm 2^{(n+s)/2}\}$ if $s = \gcd(n, t) = \gcd(n, 2t)$.

We begin by recalling a result which is proved in a more general context in Charpin et al. (1997, Proposition 5). Note that, in the next lemma, the number $B_3$ of codewords of weight 3 of $C_d$ is given by Proposition 5.

Lemma 5: *Let $d = 2^{2t} - 2^t + 1$ with $t > 1$. Let $s = \gcd(t, n)$. Then*

$$U_d(x) = (x^{2^t} + x)\left(\frac{x^{2^t} + x}{x^2 + x}\right)^{2^t}$$

*Moreover, the associated code $C_d$ has minimal distance 3 for any $s > 1$.*

The first part of the next theorem was given in Hertel and Pott (2008). Recall that $\delta(b)$ is the number of $x$ such that $x^d + (x+1)^d = b$.

Theorem 2: *Let $K_t : x \mapsto x^{2^{2t}-2^t+1}$ over $\mathbb{F}_{2^n}$. We assume that $n \neq 3t$ and $s = \gcd(n,t)$ with $n/s$ odd. Then $\delta(b) \in \{0, 2^s\}$ for any $b$ and, consequently, $\delta(K_t) = 2^s$. More precisely, whenever the equation $x^d + (x+1)^d = b$ has a solution $x$ then the set of its solutions is*

$$(y + a\mathbb{F}_{2^s})^{2^t+1} \text{ where } x = y^{2^t+1},\ x+1 = (y+a)^{2^t+1}$$

*Proof*: We assume that $b$ is such that the following equation

$$x^{2^{2t}-2^t+1} + (x+1)^{2^{2t}-2^t+1} = b \tag{16}$$

has at least one solution. According to (18), $t$ satisfies $\gcd(2^t+1, 2^n-1) = 1$. Hence, there are $y$ and $z$ such that $x = y^{2^t+1}$ and $x+1 = z^{2^t+1}$. Moreover, there is $a$ such that $z = y + a$. Then (16) becomes

$$y^{2^{3t}+1} + (y+a)^{2^{3t}+1} = b \tag{17}$$

where $2^{3t}+1$ is computed modulo $(2^n-1)$ (if $3t = n + \ell$ with $\ell \geq 1$ then $2^{3t}+1$ is equivalent to $2^\ell + 1$). Let $\gcd(3t, n) = k$. Note that either $k = s$ or $k = 3s$. From Lemma 4, we know that if Equation (17) has a solution $y$ then the set of its solutions is $y + a\mathbb{F}_{2^k}$. Now we want to prove that, for any $\beta \in \mathbb{F}_{2^s}$, the element $(y+\beta a)^{2^t+1}$ is a solution of (16). Set $u = (y+\beta a)^{2^t+1}$ and $v = (y + (\beta+1)a)^{2^t+1}$. We have

$$\begin{aligned} u + v &= (y+\beta a)^{2^t} a + (y+\beta a)a^{2^t} + a^{2^t+1} \\ &= ya^{2^t} + y^{2^t}a + a^{2^t+1} \\ &= y^{2^t+1} + (y+a)^{2^t+1} \\ &= x + (x+1) = 1 \end{aligned}$$

Thus,

$$\begin{aligned} u^{2^{2t}-2^t+1} + (u+1)^{2^{2t}-2^t+1} &= u^{2^{2t}-2^t+1} + v^{2^{2t}-2^t+1} \\ &= (y+\beta a)^{2^{3t}+1} + (y+(\beta+1)a)^{2^{3t}+1} = b \end{aligned}$$

proving that (16) has at least $2^s$ solutions. Hence, we have $\delta(b) \geq 2^s$ for any non-zero $\delta(b)$. We deduce from Proposition 6 that $\delta(b) \in \{0, 2^s\}$ for any $b$, since the function $x \mapsto \mathrm{Tr}(x^d)$ is plateaued with spectrum $\{0, \pm 2^{(n+s)/2}\}$. Further, $\delta(F_t) = 2^s$ and the set of solutions are as expected. □

## *5.4 Other exponents and a conjecture*

Taking into account a number of numerical results (see Table 1), we propose the following conjecture.

Conjecture 1: *Any power permutation $x^d$ with a two-valued differential spectrum is such that $d$ is either a quadratic exponent or a Kasami exponent, up to any equivalence which preserves the differential spectrum.*

In particular, this conjecture implies that, for power functions, the same Walsh spectrum may correspond to several differential spectra. For instance, for any $n \equiv 2 \bmod 4$, there exist some plateaued optimal functions with different differential spectra: the quadratic function $x \mapsto x^5$ is differentially 4-uniform and has a two-valued differential spectrum. On the other hand, some other power functions $F_d$ whose components are plateaued optimal have been exhibited by Cusick and Dobbertin (1996), as pointed out in the following theorem. Some of these functions do not have a two-valued differential spectrum.

Theorem 3 (Cusick and Dobbertin, 1996): *Let $n = 2m$ with $m$ odd. Let $f_d(x) = Tr(x^d)$ denotes a Boolean function on $\mathbb{F}_{2^n}$. Then $f_d$ is plateaued with spectrum $\{0, \pm 2^{(n+2)/2}\}$ for the following values of $d$:*

*(i)* $d = 2^m + 2^{(m+1)/2} + 1$
*(ii)* $d = 2^{m+1} + 3$.

Note that in both cases, we have $\gcd(d, 2^n - 1) = 1$ so that the function $F_d : x \mapsto x^d$ is a permutation. Since any bijective power function on $\mathbb{F}_{2^n}$ (with $n$ even) cannot be APN, we know that $F_d$ cannot be APN. Proposition 6 implies that either these functions $F_d$ are differentially 4-uniform with a two-valued differential spectrum, or $\delta(F_d) \geq 6$. Our simulations show, that, for any $n \equiv 2 \bmod 4$, $10 \leq n \leq 18$, both these power permutations are differentially 8-uniform. Moreover, both of them have the same differential spectrum which takes all five values, 0, 2, 4, 6 and 8 (Table 3).

Conjecture 2: *Let $n = 2m$ with $m$ odd. Let $F_d : x \mapsto x^d$ be the power permutations defined by the following values of $d$:*

*(i)* $d = 2^m + 2^{(m+1)/2} + 1$
*(ii)* $d = 2^{m+1} + 3$.

*Then, for these values of $d$, $F_d$ is differentially* 8*-uniform and all values 0, 2, 4, 6 and 8 appear in its differential spectrum.*

Our first conjecture on the non-existence of power permutations with a two-valued differential spectrum, except the quadratic exponents and the Kasami exponents, is corroborated by the following results on the scarcity of such functions. Actually, it can be proved that power permutations over $\mathbb{F}_{2^n}$ with a two-valued differential spectrum do not exist for many sets of parameters.

**Table 3** Differential spectra of the power permutations studied by Cusick and Dobbertin (1996): $F_d : x \mapsto x^d$ over $\mathbb{F}_{2^n}$ with $d = 2^m + 2^{(m+1)/2} + 1$ and $d = 2^{m+1} + 3$, for $n \equiv 2 \bmod 4$, $10 \leq n \leq 18$

| $n$ | *Exponent/inverse* | $\delta(F_d)$ | $\omega_0$ | $\omega_2$ | $\omega_4$ | $\omega_6$ | $\omega_8$ |
|---|---|---|---|---|---|---|---|
| 10 | 41/25 | 8 | 698 | 200 | 76 | 40 | 10 |
| | 67/107 | 8 | 698 | 200 | 76 | 40 | 10 |
| 14 | 145/113 | 8 | 11,504 | 2,240 | 2,080 | 448 | 112 |
| | 259/1,613 | 8 | 11,504 | 2,240 | 2,080 | 448 | 112 |
| 18 | 545/481 | 8 | 182,496 | 40,320 | 29,248 | 8,064 | 2,016 |
| | 1,027/2,629 | 8 | 182,496 | 40,320 | 29,248 | 8,064 | 2,016 |

Proposition 7: *Let $p$ be a prime and $n = p^m$ for some $m \geq 1$. Let $F_d : x \mapsto x^d$ be a non-linear power permutation over $\mathbb{F}_{2^n}$ with a two-valued differential spectrum where $\delta(F_d) = 2^s$. Then, $p > 2$ and $p$ divides $(2^{s-1} - 1)$. Most notably,*

*(i) if $p = 2$, then there is no power permutation with a two-valued differential spectrum*

*(ii) for any $p$, $\delta(F_d) \neq 4$*

*(iii) for any $p \neq 3$, $\delta(F_d) \neq 8$*

*(iv) for any $p \neq 7$, $\delta(F_d) \neq 16$.*

*Proof*: Let us consider $\mathcal{E} = \{b \in \mathbb{F}_{2^n},\ \delta(b) \neq 0\}$. Since

$$\sum_{b \in \mathbb{F}_{2^n}} \delta(b) = 2^n$$

we have that, if $F_d$ has a two-valued differential spectrum with $\delta(F_d) = 2^s$, then $\#\mathcal{E} = 2^{n-s}$. But, for any $b$, $\delta(b) = \delta(b^2)$. Therefore, the set $\mathcal{E}$ consists of the union of some cyclotomic cosets modulo $(2^n - 1)$. Moreover, $\mathcal{E}$ includes $\{1\}$ since $\delta(1) \geq 2$. When $n = p^m$, the sizes of all cyclotomic cosets, except $\{1\}$, are divisible by $p$.

It follows that $\#\mathcal{E} = 1 + p\lambda$ for some integer $\lambda$, leading to $p\lambda = 2^{n-k} - 1$.

Note that $\lambda \geq 1$ since the case $\lambda = 0$ corresponds to the case where $\delta(F_d) = 2^n$, that is, $F_d(x) = x^{2^i}$. This situation does not occur since $F_d$ is assumed to be non-linear. Then, we have

$$2^{p^m - s} - 1 \equiv 0 \bmod p \tag{18}$$

Property (i) immediately follows since this cannot occur for $p = 2$ because $s < p^m$.

Let us now suppose that $p > 2$. Euler's totient theorem can be applied to $2^{p^{i-1}}$, since $p$ is an odd prime, for any $i \geq 1$. It leads to

$$\left(2^{p^{i-1}}\right)^{p-1} \equiv 1 \bmod p$$

implying $2^{p^i} \equiv 2^{p^{i-1}} \bmod p$. It then follows that

$$2^{p^m} \equiv 2 \bmod p.$$

Then, we deduce from (18) that $2^{p^m} \equiv 2^s \equiv 2 \bmod p$, that is, $p$ divides $2^{s-1} - 1$. □

With a very similar technique, we can prove the following result.

Proposition 8: *Let $p > 2$ be a prime and $n = 2p^m$ for some $m \geq 1$. Let $F_d : x \mapsto x^d$ be a non-linear power permutation over $\mathbb{F}_{2^n}$ with a two-valued differential spectrum. Then, $\delta(F_d) = 2^s$ and $p$ divides either $(2^{s-2} - 1)$ or $(3 \times 2^{s-2} - 1)$. Most notably,*

*(i) for any $p \neq 5$, $\delta(F_d) \neq 8$*

*(ii) for any $p \notin \{3, 11\}$, $\delta(F_d) \neq 16$*

*(iii) for any $p \notin \{7, 23\}$, $\delta(F_d) \neq 32$*

*(iv) for any $p \notin \{3, 5, 47\}$, $\delta(F_d) \neq 64$.*

If $n = p^m$, we know from Lemma 3 and Proposition 7 that the only power permutations with $\delta(F_d) \leq 6$ which have a two-valued differential spectrum are the APN power permutations. Now, we get some additional information on the differential spectrum of the power permutations with $\delta(F_d) \leq 6$ since we show that they all satisfy $\delta(1) = 2$.

Proposition 9: *Let $p > 2$ be a prime and $n = p^m$ for some $m \geq 1$. Let $F_d : x \mapsto x^d$ be a nonlinear power permutation over $\mathbb{F}_{2^n}$. Then, $p$ divides $(((\delta(1))/2) - 1)$. Most notably, if $\delta(F_d) = 4$ or $\delta(F_d) = 6$, then $\delta(1) = 2$, implying that $\delta(b) = \delta(F_d)$ for at least $p$ values of $b$.*

*Proof*: Let

$$\mathcal{E} = \{b \in \mathbb{F}_{2^n},\ \delta(b) \neq 0\}$$

The set $\mathcal{E}$ consists of the union of some cyclotomic cosets modulo $(2^n - 1)$. For $n = p^m$, the sizes of all cyclotomic cosets modulo $(2^n - 1)$ are divisible by $p$ except the coset $\{1\}$. It follows that

$$\sum_{b \in \mathbb{F}_{2^n}} \delta(b) = \delta(1) + p \sum_{b \in \mathcal{I}, b \neq 1} \delta(b) = 2^n$$

where $\mathcal{I}$ is a set of representatives of the cyclotomic cosets. Because all $\delta(b)$ are even, we deduce that

$$\frac{\delta(1)}{2} + p\lambda = 2^{n-1}$$

for some $\lambda$. Moreover, $\lambda \neq 0$ because $F_d$ is not linear. It follows that $p$ divides $\left(2^{p^m-1} - (\delta(1)/2)\right)$. By Euler's totient theorem, we have

$$2^{p^m} \equiv 2 \bmod p$$

Thus, $p$ divides $\left(2^{p^m-1} - (\delta(1)/2)\right)$ if and only if $p$ divides $((\delta(1)/2) - 1)$. It is worth noticing that this condition cannot hold for $\delta(1) \in \{4, 6\}$, implying that $\delta(1) = 2$ when $\delta(F_d) \leq 6$. □

Remark 3: *Let $p > 2$ be a prime and $n = p^m$ for some $m \geq 1$. Then, there is no differentially 4-uniform power permutation over $\mathbb{F}_{2^n}$ having the following differential spectrum: $\omega_4 = 1$, $\omega_2 = 2^{n-1} - 2$, $\omega_0 = 2^{n-1} + 1$.*

## 6 Conclusions

Differentially 4-uniform permutations are of great interest for the design of symmetric cryptographic primitives: in the lack of known APN permutations of an even number of variables (except for six variables), they are those which guarantee the best resistance to differential attacks in most practical cases. But, besides the differential uniformity, the whole differential spectrum of its S-box affects the security of a cipher as shown in Section 3. For power permutations, this differential spectrum is highly related to the number of low-weight codewords in some cyclic codes with two zeroes. In this context, we have studied the differential spectra of several infinite families of exponents and we have also investigated the case of power permutations with a two-valued differential spectrum. Most notably, several conjectures on such functions have been given.

## References

Biham, E. and Shamir, A. (1991) ‘Differential cryptanalysis of DES-like cryptosystems’, *Journal of Cryptology*, Vol. 4, No. 1, pp.3–72.

Blondeau, C. and Gérard, B. (2009) ‘On the data complexity of statistical attacks against block ciphers’, *Workshop on Coding and Cryptography - WCC 2009*.

Bracken, C. and Leander, G. (2009) ‘A highly nonlinear differentially 4-uniform power mapping that permutes fields of even degree’, *CoRR abs/0901.1824*, Available at: http://arxiv.org/abs/0901.1824.

Canteaut, A., Carlet, C., Charpin P. and Fontaine C. (2000) ‘Propagation characteristics and correlation-immunity of highly nonlinear boolean functions’, *Advances in Cryptology - EUROCRYPT’2000*, Vol. 1807 of *Lecture Notes in Computer Science*, Springer-Verlag, pp.507–522.

Canteaut, A., Carlet, C., Charpin, P. and Fontaine C. (2001) ‘On cryptographic properties of the cosets of $R(1, m)$’, *IEEE Transactions on Information Theory*, Vol. 47, No. 4, pp.1494–1513.

Canteaut, A. and Videau, M. (2002) ‘Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis’, *Advances in Cryptology - EUROCRYPT 2002*, Vol. 2332 of *Lecture Notes in Computer Science*, Springer-Verlag, pp.518–533.

Carlet, C., Charpin, P. and Zinoviev, V. (1998) ‘Codes, bent functions and permutations suitable for DES-like cryptosystems’, *Designs, Codes and Cryptography*, Vol. 15, No. 2, pp.125–156.

Charpin, P. (1998) ‘Chapter 11 – open problems on cyclic codes’, In V.S. Pless and W.C. Huffman, (Eds.), R.A. Brualdi. (Ass. Ed.), *Handbook of Coding Theory*, Amsterdam, The Netherlands: Elsevier, Vol. I, pp.963–1063.

Charpin, P., Tietäväinen, A. and Zinoviev, V. (1997) ‘On binary cyclic codes with minimum distance $d = 3$’, *Problems Information Transmission*, Vol. 33, No. 4, pp.287–296.

Courtois, N. and Pieprzyk, J. (2002) ‘Cryptanalysis of block ciphers with overdefined systems of equations’, *Advances in Cryptology – ASIACRYPT’02*, Vol. 2501 of *Lecture Notes in Computer Science*, Springer-Verlag, pp.267–287.

Cusick, T. and Dobbertin, H. (1996) ‘Some new 3-valued crosscorrelation functions of binary $m$-sequences’, *IEEE Transactions on Information Theory*, Vol. 42, No. 4, pp.1238–1240.

Dillon, J. (2009) ‘APN polynomials: an update’, *Fq9, The 9th International Conference on Finite Fields and Applications*, Dublin, Ireland, July 2009.

Dobbertin, H. (1998) ‘One-to-one highly nonlinear power functions on $GF(2^n)$’, *Applicable Algebra in Engineering, Communication and Computing*, Vol. 9, No. 2, pp.139–152.

Hertel, D. and Pott, A. (2008) ‘Two results on maximum nonlinear functions’, *Designs, Codes and Cryptography*, Vol. 47, No. 1–3, pp.225–235.

Jakobsen, T. and Knudsen, L.R. (1997) ‘The interpolation attack on block ciphers’, *Fast Software Encryption - FSE’97*, Vol. 1267 of *Lecture Notes in Computer Science*, Springer-Verlag, pp.28–40.

Kasami, T. (1971) ‘The weight enumerators for several classes of subcodes of the second order binary Reed-Muller codes’, *Information and Control*, Vol. 18, pp.369–394.

Knudsen, L.R. (1995) ‘Truncated and higher order differentials’, *Fast Software Encryption - FSE’94*, Vol. 1008 of *Lecture Notes in Computer Science*, Springer-Verlag, pp.196–211.

McEliece, R.J. (1987) *Finite Fields for Computer Scientists and Engineers*. Boston: Kluwer.

Nyberg, K. (1993) ‘Differentially uniform mappings for cryptography’, *Advances in Cryptology – EUROCRYPT’93*, Vol. 765 of *Lecture Notes in Computer Science*, Springer-Verlag, pp.55–64.

Nyberg, K. (1995) 'S-boxes and round functions with controllable linearity and differential uniformity', *Fast Software Encryption – FSE'94*, Vol. 1008 of *Lecture Notes in Computer Science*, Springer-Verlag, pp.111–130.

Zhang, X-M. and Zheng, Y. (1995) 'GAC – the criterion for global avalanche characterics of cryptographic functions', *Journal of Universal Computer Science*, Vol. 1, No. 5, pp.320–337.

Zhang, X-M. and Zheng, Y. (1999) 'Plateaued functions', *Information and Communication Security, ICICS'99*, Vol. 1726 of *Lecture Notes in Computer Science*, Springer-Verlag, pp.224–300.