HAL
open science

A generalisation of Dillon's APN permutation with the best known differential and nonlinear properties for all fields of size 2^{4k+2}

Anne Canteaut, Sébastien Duval, Léo Perrin

To cite this version:

Anne Canteaut, Sébastien Duval, Léo Perrin. A generalisation of Dillon's APN permutation with the best known differential and nonlinear properties for all fields of size 2^{4k+2}. IEEE Transactions on Information Theory, 2017, 63 (11), pp.7575–7591. 10.1109/TIT.2017.2676807 . hal-01589131

HAL Id: hal-01589131

https://inria.hal.science/hal-01589131v1

Submitted on 18 Sep 2017

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.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

A Generalisation of Dillon's APN Permutation with the Best Known Differential and Nonlinear Properties for all Fields of Size 2^{4k+2}

Anne Canteaut, Sébastien Duval and Léo Perrin

***Abstract*—The existence of Almost Perfect Nonlinear (APN) permutations operating on an even number of variables was a long-standing open problem, until an example with six variables was exhibited by Dillon et al. in 2009. However it is still unknown whether this example can be generalised to any even number of inputs. In a recent work, Perrin et al. described an infinite family of permutations, named butterflies, operating on $(4k+2)$ variables and with differential uniformity at most 4, which contains the Dillon APN permutation. In this paper, we generalise this family, and we completely solve the two open problems raised by Perrin et al. Indeed we prove that all functions in this larger family have the best known nonlinearity. We also show that this family does not contain any APN permutation besides the Dillon permutation, implying that all other functions have differential uniformity exactly four.**

***Index Terms*—Boolean function, Sbox, APN, differential uniformity, nonlinearity.**

I. Introduction

MODERN block ciphers are designed based on a methodology which guarantees that the cipher is resistant against all classical attacks. The differential cryptanalysis [1], [2] and the linear cryptanalysis [3], [4] are the two most prominent attacks against block ciphers, and a precise evaluation of their complexities has led to some design criteria on the so-called Sbox, *i.e.*, on the nonlinear mapping used in the cipher. The differential uniformity of the Sbox [5] and its nonlinearity respectively quantify its resistance to differential and linear attacks. These two design criteria are at the origin of a long line of work, including the search for infinite families of optimal permutations. Optimal permutations have been exhibited for more than twenty years when the number of variables is odd (e.g [6], [7]). In this case, the permutations with optimal nonlinearity are also APN, i.e., they have the lowest possible differential uniformity. But the situation is very different when the number of inputs is even, which is the case of practical interest. Indeed, in this case the values of the best differential uniformity and nonlinearity are unknown. For instance, it had been conjectured for many years that APN permutations of an even number of variables did not exist. This was disproved in 2009 by Dillon and his coauthors who exhibited an APN permutation of six variables [8]. However, the Dillon permutation is the only known example of an APN permutation depending on an even number of variables. It is unknown whether it can be generalised to any even number of inputs, or whether it is sporadic. In other words, the "(still) big APN problem" raised in [8] on the existence of an APN permutation of m variables for m even and greater than six is still unsolved.

Recently, Perrin *et al.* [9] exhibited a family of permutations acting on $2n$ variables, for any odd $n \geq 3$, with differential uniformity at most 4, and which includes the Dillon permutation for $n = 3$. The novel idea in their work is the representation of the permutation of $2n$ variables as a bivariate polynomial over the field $\mathbb{F}_{2^n}$. However, their simulations could not find any other APN permutation within this family than the one described by Dillon *et al.* The existence of an APN permutation of this form is left as an open problem at the end of [9]. A second open problem mentioned in [9] is the determination of the nonlinearity of these permutations, which is conjectured to be the same as the nonlinearity of the inverse mapping. Our work aims at solving these two open problems. More precisely, we define a new infinite family of permutations of $2n$ variables, n odd, which generalises the functions introduced in [9]. This new family includes all permutations defined in [9], but also some other mappings which are not CCZ-equivalent to the previous ones. We show that all permutations in this family have the best known differential uniformity and nonlinearity for any number of variables of the form $m = 4k+2$, $k \geq 1$. Most notably, the differential uniformity of all permutations in this family is equal to 4, except for a single (up to equivalence) permutation of 6 variables which is CCZ-equivalent to the Dillon APN permutation. It is worth noticing that the nonlinearity of the subclass of permutations introduced by Perrin *et al.* has also been determined in an independent work by Fu and Feng [10].

Therefore, our work solves all open questions raised in [9] and provides new permutations with very good cryptographic properties and a compact description, which may be appropriate for some implementation purposes. One specificity of this family is that it is the only known general family of mappings with such cryptographic properties which includes the Dillon permutation. However, it does not contain any new APN permutation of an even number of variables besides the one exhibited by Dillon *et al.*, therefore it does not solve the "big APN problem".

A. Canteaut is with Inria, SECRET project-team, Paris, France, e-mail: Anne.Canteaut@inria.fr.

S. Duval is with Sorbonne Universités/UPMC Univ. Paris 06/Inria, Paris, France, e-mail: Sebastien.Duval@inria.fr.

L. Perrin is with SnT, University of Luxembourg, Luxembourg, e-mail: leo.perrin@uni.lu. The work of Léo Perrin is supported by the CORE ACRYPT project (ID C12-15-4009992) funded by the *Fonds National de la Recherche* (Luxembourg).

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

Organization of the paper: The cryptographic properties of Boolean functions investigated in the paper are defined in Section II. Then, in Section III, the new infinite family of functions studied in this work is described and some equivalences between the different elements are identified. Following the terminology introduced in [9], the functions within this family are named *generalised butterflies*. The rest of the paper studies their cryptographic properties: their nonlinearity is computed in Section IV, their differential uniformity is determined in Section V while Section VI investigates their algebraic degree.

II. Boolean Functions

This paper mainly focuses on vectorial functions with the same number of inputs and outputs, i.e., on functions F from $\mathbb{F}_2^m$ into itself:

$$\begin{array}{rrcl} F: & \mathbb{F}_2^m & \to & \mathbb{F}_2^m \\ & (x_1,\ldots,x_m) & \mapsto & F(x_1,\ldots,x_m)=(y_1,\ldots,y_m)\ . \end{array}$$

The *components* of the vectorial function F are the Boolean functions corresponding to all linear combinations of its coordinates.

Alternatively, the vector space $\mathbb{F}_2^m$ can be identified with the finite field $\mathbb{F}_{2^m}$. In this case, F is seen as a univariate mapping from $\mathbb{F}_{2^m}$ into itself:

$$\begin{array}{rrcl} F: & \mathbb{F}_{2^m} & \to & \mathbb{F}_{2^m} \\ & x & \mapsto & F(x)=y\ . \end{array}$$

These two equivalent settings are related to two different notions of polynomials: a mapping from $\mathbb{F}_2^m$ into $\mathbb{F}_2^m$ can be uniquely represented by a collection of m *multivariate* polynomials with binary coefficients corresponding to the algebraic normal forms of its coordinates. A mapping from $\mathbb{F}_{2^m}$ into $\mathbb{F}_{2^m}$ is represented by a unique *univariate* polynomial with coefficients in $\mathbb{F}_{2^m}$. Therefore, we need to consider two different notions of *degree*.

Definition 1 (Univariate degree vs algebraic degree)**.** *Let F be a function from $\mathbb{F}_2^m$ into $\mathbb{F}_2^m$. The* algebraic degree *(aka multivariate degree) of F is the maximal degree of the algebraic normal forms of its coordinates. The* univariate degree *of F is the degree of the univariate polynomial in $\mathbb{F}_{2^m}[X]$ representing F when it is identified with a function from $\mathbb{F}_{2^m}$ into itself.*

Obviously, these two notions are different: for instance, the cube function x^3 over $\mathbb{F}_{2^m}$ has *univariate degree* 3 and *algebraic degree* 2. More generally, the algebraic degree of the univariate polynomial $x \mapsto x^e$ of $\mathbb{F}_{2^n}$ is the Hamming weight of the binary expansion of e [11, Lemma 1.1].

In the rest of the paper, we will use the following notation.

Definition 2 (Derivative of a function)**.** *Let F be a function from $\mathbb{F}_2^m$ into $\mathbb{F}_2^t$. The* derivative of F with respect to $a \in \mathbb{F}_2^m$ *is the function*

$$D_aF : x \in \mathbb{F}_2^m \mapsto F(x+a)+F(x)\ .$$

The resistance offered by a function to differential cryptanalysis is highly influenced by the following properties of its derivatives.

Definition 3 (Differential uniformity [5])**.** *Let F be a function from $\mathbb{F}_2^m$ into $\mathbb{F}_2^t$. For any $a \in \mathbb{F}_2^m$ and $b \in \mathbb{F}_2^t$, we define*

$$\delta(a,b) = |\{x \in \mathbb{F}_2^m, D_aF(x)=b\}|\ .$$

The multi-set $\{\delta(a,b),\ a \in \mathbb{F}_2^m \backslash \{0\}, b \in \mathbb{F}_2^t\}$ is the differential spectrum *of F, and its maximum*

$$\delta_F = \max_{a\neq 0,b} \delta(a,b)$$

is the differential uniformity *of F.*

Definition 4 (Linearity of a function)**.** *Let f be a Boolean function of m variables. The Walsh coefficients of f are the elements*

$$\widehat{f}(u) = \sum_{x\in\mathbb{F}_2^m} (-1)^{f(x)+u\cdot x}\ , u \in \mathbb{F}_2^m$$

where $u\cdot x$ denotes the scalar product between x and u. When F is a vectorial Boolean function from $\mathbb{F}_2^m$ into $\mathbb{F}_2^t$, the Walsh coefficients of F are the Walsh coefficients of its components, i.e.,

$$\widehat{F}(u,v) = \sum_{x\in\mathbb{F}_2^m} (-1)^{v\cdot F(x)+u\cdot x}\ , u \in \mathbb{F}_2^m, v \in \mathbb{F}_2^t$$

and the multi-set composed of all Walsh coefficients $\widehat{F}(u,v)$, $v \neq 0$, is called the Walsh spectrum *of F.*

The linearity *of F is the highest magnitude of its Walsh coefficients:*

$$\mathcal{L}(F) = \max_{v\in\mathbb{F}_2^t\backslash\{0\}} \max_{u\in\mathbb{F}_2^m} |\widehat{F}(u,v)|\ .$$

Note that we can alternatively define the nonlinearity of F as the lowest Hamming distance between a non-trivial component of F and the set of all affine functions. Then, the nonlinearity can be computed as

$$\mathcal{NL}(F) = 2^{m-1} - \frac{1}{2}\mathcal{L}(F)\ .$$

Definition 5 (CCZ-equivalence [12])**.** *Two mappings F and G from $\mathbb{F}_2^m$ into itself are said to be* CCZ-equivalent *if there exists a linear permutation L of $\mathbb{F}_2^{2m}$ such that*

$$\{(x,F(x)), \forall x \in \mathbb{F}_2^m\} = \{L(x,G(x)), \forall x \in \mathbb{F}_2^m\}\ .$$

CCZ-equivalence is the most relevant notion of equivalence with respect to the differential and linear properties of a mapping since it preserves both the differential and the Walsh spectra. It is worth noticing that neither the algebraic degree nor the fact that the mapping is a permutation is invariant under CCZ-equivalence.

III. Generalised Butterflies

A. Definition

Let $m = 2n$ be an even integer. In the paper, Boolean functions of m variables are identified with functions from $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$. Similarly, vectorial functions from $\mathbb{F}_2^m$ into $\mathbb{F}_2^m$ are identified with mappings from $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ to itself. It is worth noticing that the choice of the basis used for identifying $\mathbb{F}_{2^n}$ with $\mathbb{F}_2^n$ does not affect the cryptographic properties of

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

the functions we are studying since different bases lead to functions which are affine-equivalent.

In this setting, the scalar product between two elements (x_1, y_1) and (x_2, y_2) in $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ is defined as

$$\mathrm{Tr}(x_1 x_2) + \mathrm{Tr}(y_1 y_2)$$

where Tr is the trace function on $\mathbb{F}_{2^n}$, i.e., $\mathrm{Tr}(x) = x + x^2 + \ldots + x^{2^{n-1}}$.

We now define the family of vectorial functions that will be studied in the paper.

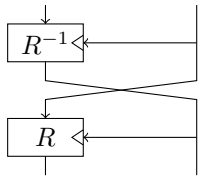
Definition 6 (Generalised Butterflies)**.** *Let R be a bivariate polynomial of $\mathbb{F}_{2^n}$ such that $R_y : x \mapsto R(x, y)$ is a permutation of $\mathbb{F}_{2^n}$ for all y in $\mathbb{F}_{2^n}$. The* closed butterfly V_R *is the function of $(\mathbb{F}_{2^n})^2$ defined by*

$$\mathsf{V}_R(x, y) \;=\; \big(R(x, y),\ R(y, x)\big)$$

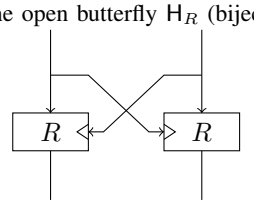
and the open butterfly H_R *is the permutation of $(\mathbb{F}_{2^n})^2$ defined by*

$$\mathsf{H}_R(x, y) \;=\; \big(R_{R_y^{-1}(x)}(y),\ R_y^{-1}(x)\big)$$

where $R_y(x) = R(x, y)$ and $R_y^{-1}(R_y(x)) = x$ for any y, x. A representation of H_R is given in Figure 1a and one of V_R is given in Figure 1b.



(a) The open butterfly H_R (bijective).



(b) The closed butterfly V_R.

Fig. 1: The butterfly constructions.

It can be easily checked that, for any choice of the keyed permutation R, the open butterfly H_R is an involution.

Lemma 1. *The permutation H_R and the function V_R of $(\mathbb{F}_{2^n})^2$ are CCZ-equivalent.*

Proof: The proof is identical to the proof of Lemma 2 in [9]. ■

In this paper, we restrict ourselves to the case where $R(x, y)$ has *univariate degree 3* as the butterflies first described in [9] have such structure. Therefore, the differential uniformity and the linearity of the generalised butterflies will be computed only for the corresponding closed butterfly V_R since it has algebraic degree 2 only.

The following lemma describes all polynomials R satisfying this degree condition which define a keyed permutation, as demanded by the butterfly definition. From a cryptographic point of view, the fact that R corresponds to a keyed permutation can be viewed as an integral property, as noted in [9].

Lemma 2 (Degree Restriction)**.** *Let R be a bivariate polynomial of $\mathbb{F}_{2^n}$ such that $R_y : x \mapsto R(x, y)$ is a permutation for any y and such that all terms in R are non-linear terms with degree at most 3. Then R can be described using two elements of $\mathbb{F}_{2^n}$ denoted α and β as*

$$R(x, y) = (x + \alpha y)^3 + \beta y^3 \ .$$

We denote butterflies based on such polynomials R as $\mathsf{V}_{\alpha,\beta}$ and $\mathsf{H}_{\alpha,\beta}$ for closed and open butterflies respectively.

The proof of this lemma relies on the following theorem.

Theorem 1 (Corollary 2.9 from [13])**.** *Let $\mathbb{F}_q$ have characteristic different from 3. Then $f : x \mapsto ax^3 + bx^2 + cx + d$ $(a \neq 0)$ permutes $\mathbb{F}_q$ if and only if $b^2 = 3ac$ and $q \equiv 2 \mod 3$.*

Proof of Lemma 2: Let

$$R(x, y) = Ax^3 + Bx^2y + Cxy^2 + Dy^3 + Exy.$$

Since $x \mapsto R(x, 0) = Ax^3$ must be a permutation, we deduce that $A \neq 0$. A multiplication by a non-zero constant changes neither the degree nor the integral property. Therefore, we consider a normalised case where $A = 1$. We need that $x \mapsto R(x, y)$ is a permutation for any y. We are in characteristic 2 and, using the notation of Theorem 1, we always have that $q \equiv 2 \mod 3$ because n is odd. Thus, in order to fulfill the integral condition, Theorem 1 imposes that $(By)^2 = Cy^2 + Ey$ for all y. This implies that $E = 0$ and $B^2 = C$. The polynomial can thus be written $R(x, y) = x^3 + Bx^2y + B^2xy^2 + Dy^3$ which we factor into $R(x, y) = (x + By)^3 + (B^3 + D)y^3$. Simply setting $\alpha = B$ and $\beta = B^3 + D$ gives us the lemma. ■

Remark 1. *Lemma 2 excludes terms in x^2, y^2, x and y from R. Since such terms have algebraic degree 1, they could be added without changing the linearity and the differential properties of V_R, which is why we ignore them.*

In this context, the results in [9] can be interpreted as handling the particular case $\beta = 1$. If $\alpha = 1$, the open butterflies and closed butterflies are functionally equivalent to the functions presented in Figure 2.

B. Equivalence Relations

As stated in Lemma 1, an open and a closed butterfly with identical parameters are CCZ-equivalent. There are other relations linking butterflies with each other.

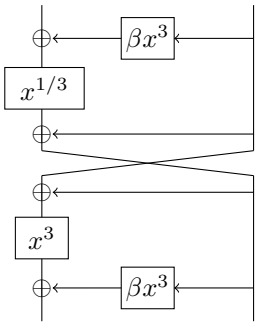
- If the exponent is equal to $e = 3 \times 2^t$, the corresponding closed butterfly is affine-equivalent to the closed butterfly with the same α, β. Therefore, all results presented in the paper also hold when

  $$R(x, y) = (x + \alpha y)^{3 \times 2^t} + \beta y^{3 \times 2^t}$$

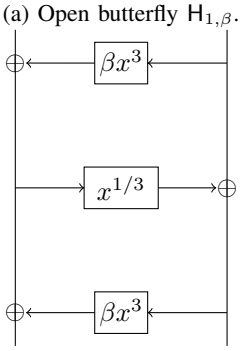
  for some t.
- The closed butterflies $\mathsf{V}_{\alpha,\beta}$ and $\mathsf{V}_{\alpha^2,\beta^2}$ are affine-equivalent as $\mathsf{V}_{\alpha^2,\beta^2}(x^2, y^2) = (\mathsf{V}_{\alpha,\beta}(x, y))^2$.

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

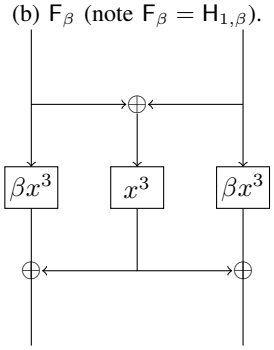
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory



(a) Open butterfly $\mathsf{H}_{1,\beta}$.



(b) F_β (note $\mathsf{F}_\beta = \mathsf{H}_{1,\beta}$).



(c) Closed butterfly $\mathsf{V}_{1,\beta}$.

Fig. 2: The equivalence between $\mathsf{H}_{1,\beta}$ and F_β.

- For any $\alpha \neq 1$, the closed butterflies $\mathsf{V}_{\alpha,\beta}$ and $\mathsf{V}_{\alpha,\beta'}$ with $\beta' = \beta^{-1}(1+\alpha)^6$ are affine-equivalent.
  This equivalence is obtained by composing $\mathsf{V}_{\alpha,\beta}$ with the inverse of the linear permutation

$$L : (x, y) \mapsto (z_1, z_2) = (\alpha x + y, x + \alpha y) \ .$$

  Indeed, $\mathsf{V}_{\alpha,\beta} \circ L^{-1}(x, y) = (A, B)$ with

$$\begin{aligned} A &= z_2^3 + \beta \left[(1+\alpha)^{-2}(z_1 + \alpha z_2)\right]^3 \\ &= (1+\alpha)^{-6}\left[(z_1 + \alpha z_2)^3 + \beta' z_2^3\right] \\ B &= z_1^3 + \beta \left[(1+\alpha)^{-2}(z_2 + \alpha z_1)\right]^3 \\ &= (1+\alpha)^{-6}\left[(z_2 + \alpha z_1)^3 + \beta' z_1^3\right] \ . \end{aligned}$$

- Let $\otimes$ be defined such that $(a, b) \otimes (c, d) = (ac, bd)$ for any pairs (a, b) and (c, d) of $(\mathbb{F}_{2^n})^2$. Then the generalised butterflies exhibit the same multiplicative stability as those described in [9], namely:

$$\mathsf{V}_{\alpha,\beta}\big((\lambda, \lambda) \otimes (x, y)\big) = (\lambda^3, \lambda^3) \otimes \mathsf{V}_{\alpha,\beta}(x, y),$$

  and

$$\mathsf{H}_{\alpha,\beta}\big((\lambda^3, \lambda) \otimes (x, y)\big) = (\lambda^3, \lambda) \otimes \mathsf{H}_{\alpha,\beta}(x, y).$$

  Note that these multiplicative properties correspond to the so-called *subspace property* introduced in [8] and investigated in [14].

C. Cryptographic Properties

Main Theorem. *The cryptographic properties of the generalised butterflies $\mathsf{V}_{\alpha,\beta}$ and $\mathsf{H}_{\alpha,\beta}$, which are based on functions $R : (x, y) \mapsto (x + \alpha y)^3 + \beta y^3$ with $\alpha, \beta \neq 0$ are as follows:*

- *the algebraic degree of $\mathsf{V}_{\alpha,\beta}$ is always equal to 2,*
- *if $n = 3$, $\alpha \neq 0$, $\mathrm{Tr}(\alpha) = 0$ and $\beta \in \{\alpha^3 + \alpha, \alpha^3 + 1/\alpha\}$ then the butterflies are APN, have a linearity equal to 2^{n+1} and the algebraic degree of $\mathsf{H}_{\alpha,\beta}$ is equal to $n+1$;*
- *if $\beta = (1+\alpha)^3$ then the differential uniformity is equal to 2^{n+1}, the linearity is equal to $2^{(3n+1)/2}$ and the algebraic degree of $\mathsf{H}_{\alpha,\beta}$ is equal to n;*
- *otherwise, the differential uniformity is equal to 4, the linearity is equal to 2^{n+1} and the algebraic degree of $\mathsf{H}_{\alpha,\beta}$ is either n or $n+1$. It is equal to n if and only if*

$$1 + \alpha\beta + \alpha^4 = (\beta + \alpha + \alpha^3)^2.$$

In particular, there are no APN butterflies operating on more than 6 bits.

Open generalised butterflies with $\beta \neq (1+\alpha)^3$ form a family of permutations operating on $2n$ bits with a linearity and a differential uniformity equal to the best known to be possible. Furthermore, the only known APN permutation on fields of even dimension is, up to affine-equivalence, a generalised butterfly as well.

The proof of this theorem is divided into several parts. Section IV treats the linearity, Section V the differential uniformity and Section VI the algebraic degree.

IV. On Linearity

In this section, we compute the linearity of generalised butterflies. We extensively use the fact that closed butterflies are quadratic, i.e., have algebraic degree 2 since the Walsh spectrum of quadratic Boolean functions can be easily computed from some properties of the derivatives. First, this general principle is detailed in Section IV-A. This method is then applied to the particular case of closed butterflies in Section IV-B.

A. General method for computing the linearity

Since $\mathsf{V}_{\alpha,\beta}$ has algebraic degree 2, its nonlinearity can be evaluated by computing the number of linear structures of its components, i.e., the number of constant derivatives of the components of the function. This relationship is described in the following proposition. Even if this result is well-known in the area of sequence design and Boolean functions (see e.g. [15, Chapter 15], [16, Proposition 15] or [17, Appendix A]), we give the proof for completeness.

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

Proposition 1. *Let f be a quadratic Boolean function of n variables. Let $\mathsf{LS}(f)$ denote the linear space of f, i.e.,*

$$\mathsf{LS}(f) = \{a \in \mathbb{F}_2^n : D_a f(x) = \varepsilon, \ \ \forall x \in \mathbb{F}_2^n\}$$

where $\varepsilon \in \{0,1\}$. Then, $s = \dim \mathsf{LS}(f)$ has the same parity as n and $\mathcal{L}(f) = 2^{\frac{n+s}{2}}$. Moreover, the Walsh coefficients of f take 2^{n-s} times the value $\pm 2^{\frac{n+s}{2}}$ and $(2^n - 2^{n-s})$ times the value 0.

Proof: First, it is clear that $\mathsf{LS}(f)$ is a linear subspace of $\mathbb{F}_2^n$. Moreover, since

$$D_a f(x) + D_b f(x) = D_{a+b} f(x+a),$$

only two situations can occur. Either $D_a f = 0$ for all $a \in \mathsf{LS}(f)$, or $D_a f = 0$ for exactly half of the elements in $\mathsf{LS}(f)$. In this second case, $\mathsf{LS}_1(f) = \{a : D_a f = 1\}$ is a coset of $\mathsf{LS}_0(f) = \{a : D_a f = 0\}$. Now, we use that the Walsh coefficients of f are related to the weight of its derivatives (see e.g. [18, Lemma 1]), namely

$$\left(\widehat{f}(u)\right)^2 = 2^n \sum_{a \in \mathbb{F}_2^n} (-1)^{u \cdot a} \widehat{D_a f}(0) \ .$$

Because f has algebraic degree 2, $D_a f$ has algebraic degree 1 or is constant. If it has degree 1, then it is balanced, i.e., $\widehat{D_a f}(0) = 0$. It follows that

$$\widehat{f}(u)^2 = 2^{2n} \left(\sum_{a \in \mathsf{LS}_0(f)} (-1)^{u \cdot a} - \sum_{a \in \mathsf{LS}_1(f)} (-1)^{u \cdot a} \right) \ .$$

If $\mathsf{LS}_1(f) = \emptyset$,

$$\widehat{f}(u)^2 \in \{0, 2^{2n + \dim \mathsf{LS}(f)}\} \ .$$

Otherwise, $\mathsf{LS}_1(f) = b + \mathsf{LS}_0(f)$ for some b, implying that

$$\widehat{f}(u)^2 = 2^{2n} \left(1 + (-1)^{u \cdot b}\right) \left(\sum_{a \in \mathsf{LS}_0(f)} (-1)^{u \cdot a} \right) \ .$$

Then,

$$\widehat{f}(u)^2 \in \{0, 2^{2n + \dim \mathsf{LS}(f)}\} \ .$$

The number of occurrences of the values 0 and $\mathcal{L}(f)$ in the Walsh spectrum is directly deduced from Parseval's relation. ■

Our proof also uses the following easy lemma, which is a special case of the more general result given in [19, Corollary 1].

Lemma 3. *Let n be an odd integer and a, b, c be three elements of $\mathbb{F}_{2^n}$ which are not all equal to zero. Then, the equation*

$$aX^{16} + bX^4 + cX = 0$$

has 1, 2 or 4 solutions in $\mathbb{F}_{2^n}$.

Proof: Let $P(X) = aX^{16} + bX^4 + cX$. Since P is a nonzero linearised polynomial with degree at most 16, its roots in $\mathbb{F}_{2^n}$ form a linear subspace of $\mathbb{F}_{2^n}$, seen as vector space over $\mathbb{F}_2$. Then, P has 2^r roots in $\mathbb{F}_{2^n}$, with $0 \le r \le 4$. Since $\mathbb{F}_{2^n}$ is a subfield of $\mathbb{F}_{2^{2n}}$, the roots of P in $\mathbb{F}_{2^n}$ are also included in the set of all roots of P in $\mathbb{F}_{2^{2n}}$. Let $\beta \in \mathbb{F}_{4^n}$ be such that $\beta^2 + \beta + 1 = 0$. Then, for any $x \in \mathbb{F}_{2^n}$ such that $P(x) = 0$, we have $P(\beta x) = 0$ and $P(\beta^2 x) = 0$, and none of these two other roots lie in $\mathbb{F}_{2^n}$. Moreover, if x and x' are two distinct roots of P in $\mathbb{F}_{2^n}$, then $\{x, \beta x, \beta^2 x, x', \beta x', \beta^2 x'\}$ are six distinct roots of P in $\mathbb{F}_{4^n}$. This implies that, if P has seven nonzero roots in $\mathbb{F}_{2^n}$, then it would have more than 16 roots in $\mathbb{F}_{4^n}$ which is impossible. Therefore, P has at most 4 roots in $\mathbb{F}_{2^n}$. ■

B. Linearity of generalised butterflies

Theorem 2. *Let $n > 1$ be an odd integer and (α, β) be a pair of nonzero elements in $\mathbb{F}_{2^n}$. Then the linearity of $\mathsf{V}_{\alpha,\beta}$ over $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ is 2^{n+1} if $\beta \neq (1+\alpha)^3$. Moreover, the Walsh coefficients of $\mathsf{V}_{\alpha,\beta}$ belong to $\{0, \pm 2^n, \pm 2^{n+1}\}$. If $\beta = (1+\alpha)^3$, the linearity of $\mathsf{V}_{\alpha,\beta}$ is equal to $2^{\frac{3n+1}{2}}$.*

Proof: The linearity of $\mathsf{V}_{\alpha,\beta}$ is the maximal linearity of its nontrivial components, i.e., of the Boolean functions

$$f_{\lambda,\mu} : (x,y) \mapsto \mathrm{Tr}(\lambda R(x,y)) + \mathrm{Tr}(\mu R(y,x))$$

for $\lambda, \mu \in \mathbb{F}_{2^n}$, where Tr denotes the trace function from $\mathbb{F}_{2^n}$ into $\mathbb{F}_2$.

Let us first determine the linearity of $f_{0,\mu}$. Since $x \mapsto R(x,y)$ is a permutation for every fixed y, we deduce that $f_{0,\mu}$ is balanced on any set $(a, \mathbb{F}_{2^n})$, implying that it is balanced over $\mathbb{F}_{2^n}^2$. The linearity of $f_{0,\mu}$ is determined by the number of pairs (a,b) such that $D_{(a,b)} f_{0,\mu}$ is constant. We have

$$\begin{aligned} R(y+b, x+a) + R(y,x) = {} & (\alpha^2 b + a(\alpha^3+\beta))x^2 \\ & + ((\alpha^3+\beta)a^2 + \alpha b^2)x + (b+\alpha a)y^2 \\ & + (b+\alpha a)^2 y + R(b,a) \ . \end{aligned}$$

Then, $D_{(a,b)} f_{0,\mu}$ is equal to

$$\begin{aligned} & \mathrm{Tr}\left(x^2 \left(\mu(\alpha^2 b + a(\alpha^3+\beta)) \ + \mu^2((\alpha^3+\beta)^2 a^4 + \alpha^2 b^4)\right)\right) \\ & \quad + \mathrm{Tr}\left(y^2 \left(\mu(b+\alpha a) + \mu^2 (b+\alpha a)^4\right)\right) + c, \ \text{with } c \in \mathbb{F}_2 \end{aligned}$$

implying that $D_{(a,b)} f_{0,\mu}$ is constant if and only if (a,b) satisfies

$$\begin{cases} \mu(\alpha^2 b + a(\alpha^3+\beta)) + \mu^2((\alpha^3+\beta)^2 a^4 + \alpha^2 b^4) & = & 0 \\ \mu(b+\alpha a) + \mu^2 (b+\alpha a)^4 & = & 0 \ . \end{cases}$$

Since $\mu \neq 0$, the second equation implies that

$$b = \alpha a + \delta \text{ with } \delta \in \{0, \mu^{-1/3}\} \ .$$

By replacing the value of b in the first equation, we get

$$\beta^2 a^4 + \beta a = \delta'$$

where δ' can take two values including 0. Since $\beta \neq 0$, we deduce that a takes two different values when $\delta' = 0$, and at most two values when $\delta' \neq 0$. Then we get that the number of pairs (a,b) satisfying these two equations is either 2 or 4. We know from Proposition 1 that the dimension of the linear space is even, implying that $D_{(a,b)} f_{0,\mu}$ is constant for four values (a,b). Thus $f_{0,\mu}$ has linearity exactly 2^{n+1}.

Now, we focus on the case when $\lambda \neq 0$. Then there exists $\lambda' \in \mathbb{F}_{2^n}$ such that $\lambda = \lambda'^3$ because $x \mapsto x^3$ is a permutation of $\mathbb{F}_{2^n}$ for any odd n. It follows that, for any nonzero λ,

$$\lambda R(x,y) = (x\lambda' + \alpha y \lambda')^3 + \beta (y\lambda')^3 = R(x\lambda', y\lambda') \ .$$

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

Therefore,

$$\begin{aligned} f_{\lambda,\mu}(x,y) &= \mathrm{Tr}\left(\lambda R(x,y)+\mu R(y,x)\right) \\ &= \mathrm{Tr}\left(R(x\lambda',y\lambda')+\mu\lambda^{-1}R(y\lambda',x\lambda')\right) \\ &= f_{1,\mu\lambda^{-1}}(x\lambda',y\lambda')\ . \end{aligned}$$

We deduce that, for $\lambda \neq 0$, $f_{\lambda,\mu}$ is linearly equivalent to $f_{1,\mu\lambda^{-1}}$, implying that these two functions have the same linearity. Therefore, we only need to compute the dimension of the linear space of the quadratic function $f_{1,\lambda}$ for all $\lambda \in \mathbb{F}_{2^n}^*$. Let $\gamma = \alpha^3+\beta$. Then, for any a and b in $\mathbb{F}_{2^n}$, we have

$$\begin{aligned} D_{(a,b)}f_{1,\lambda} &= \mathrm{Tr}\left[x^2\left(a+\alpha b+\lambda(\gamma a+\alpha^2 b)\right)\right] \\ &+ \mathrm{Tr}\left[x\left((a+\alpha b)^2+\lambda(\gamma a^2+\alpha b^2)\right)\right] \\ &+ \mathrm{Tr}\left[y^2\left(\alpha^2 a+\gamma b+\lambda(\alpha a+b)\right)\right] \\ &+ \mathrm{Tr}\left[y\left(\alpha a^2+\gamma b^2+\lambda(\alpha a+b)^2\right)\right]+f_{1,\lambda}(a,b)\ . \end{aligned}$$

Then, we write

$$D_{(a,b)}f_{1,\lambda} = \mathrm{Tr}(Ax^2)+\mathrm{Tr}(By^2)+f_{1,\lambda}(a,b)$$

implying that $D_{(a,b)}f_{1,\lambda}$ is constant if and only if

$$\left\{\begin{array}{lcl} A &=& (1+\lambda\gamma)^2a^4+(1+\lambda\gamma)a \\ && +(\lambda\alpha+\alpha^2)^2b^4+(\alpha+\lambda\alpha^2)b=0 \\ B &=& (\alpha+\lambda\alpha^2)^2a^4+(\alpha^2+\lambda\alpha)a \\ && +(\gamma+\lambda)^2b^4+(\gamma+\lambda)b=0 \end{array}\right. \quad (1)$$

We deduce that

$$\begin{aligned} E &= (\alpha+\lambda\alpha^2)^2A+(1+\lambda\gamma)^2B \\ &= a(1+\lambda\gamma)\left[(\alpha+\lambda\alpha^2)^2+(1+\lambda\gamma)(\alpha^2+\lambda\alpha)\right] \\ &\quad +b^4\left[(\alpha+\lambda\alpha^2)(\alpha^2+\lambda\alpha)+(\gamma+\lambda)(1+\lambda\gamma)\right]^2 \\ &\quad +b\left[(\alpha+\lambda\alpha^2)^3+(1+\lambda\gamma)^2(\gamma+\lambda)\right]=0\ . \end{aligned} \quad (2)$$

Let us assume that $\lambda \neq \gamma^{-1}$ (this case will be studied at the end of the proof). We first show that, if the coefficient of a in (2) vanishes, then the coefficient of b^4 does not vanish, unless $(\beta,\lambda) = ((1+\alpha)^3,1)$. Indeed, the coefficient of a vanishes if and only if

$$\lambda^2(\alpha^3+\gamma)+\lambda(\alpha\gamma+1) = \lambda\left(\lambda\beta+(\alpha^4+\alpha\beta+1)\right)=0\ . \quad (3)$$

There is then a single nonzero value of λ for which this coefficient vanishes, namely

$$\lambda = \beta^{-1}(\alpha^4+\alpha\beta+1). \quad (4)$$

The coefficient of b^4 vanishes if and only if

$$\begin{aligned} X &= (\alpha+\lambda\alpha^2)(\alpha^2+\lambda\alpha)+(\gamma+\lambda)(1+\lambda\gamma) \\ &= \lambda^2\beta+\lambda(\alpha^3+\alpha^2+\alpha+1+\beta)^2+\beta=0\ . \end{aligned}$$

Therefore, this coefficient vanishes for the value of λ given by (4) if and only if

$$\begin{aligned} &(\alpha^4+\alpha\beta+1)^2+(\alpha^4+\alpha\beta+1)(\alpha^3+\alpha^2+\alpha+1+\beta)^2+\beta^2 \\ &= \alpha\left(\beta+(1+\alpha)^3\right)^2\left(\beta+(1+\alpha)^2\alpha\right)=0\ . \end{aligned}$$

Furthermore, this equation has two roots: $\beta = (1+\alpha)^3$ (implying $\lambda=1$, from (4)) and $\beta=(1+\alpha)^2\alpha$. This second case cannot occur since (4) then leads to

$$\lambda = \alpha^{-1} \text{ and } \gamma=\alpha,$$

which contradicts the hypothesis $\lambda \neq \gamma^{-1}$. Therefore, when $(\beta,\lambda) \neq ((1+\alpha)^3,1)$, Equation (2) can be expressed either as

$$Ub^4+Vb=0 \text{ with } U\neq 0 \text{ or } a = Ub^4+Vb\ , \quad (5)$$

where the first situation occurs if and only if (3) holds.

In the first case, we deduce at most 2 solutions for b, including $b=0$. We now replace b by these two values in the first equation in (1). We get that

$$(1+\lambda\gamma)^2a^4+(1+\lambda\gamma)a=\delta$$

for two values of δ. Using that $\lambda\neq\gamma^{-1}$, we deduce that there are at most two solutions a for each value of δ, implying that there are at most four pairs (a,b) satisfying (1).

In the second case of (5), we replace a by its expression in the first equation in (1). This leads to

$$\begin{aligned} &b^{16}U^4(1+\lambda\gamma)^2+b^4\left[(\lambda\alpha+\alpha^2)^2+V^4(1+\lambda\gamma)^2+U(1+\lambda\gamma)\right] \\ &\quad +b\left[(\alpha+\lambda\alpha^2)+V(1+\lambda\gamma)\right]=0\ . \end{aligned}$$

From Lemma 3, this equation has at most 4 solutions b, unless all its coefficients vanish. Therefore, the system has at most 4 solutions except when $U=0$ and

$$\left\{\begin{array}{lcl} (\lambda\alpha+\alpha^2)+V^2(1+\lambda\gamma) &=& 0 \\ (\alpha+\lambda\alpha^2)+V(1+\lambda\gamma) &=& 0\ . \end{array}\right. \quad (6)$$

These two relations hold only either when $V=0$, which implies $\alpha=\lambda=1$, or when

$$\left\{\begin{array}{lcl} \alpha(1+\lambda\alpha)+V(1+\lambda\gamma) &=& 0 \\ (\alpha+\lambda)+V(1+\lambda\alpha) &=& 0\ . \end{array}\right.$$

When $V\neq 0$, by multiplying the first equation by $(1+\lambda\alpha)$ and the second one by $(1+\lambda\gamma)$, we deduce that

$$\alpha(1+\lambda\alpha)^2+(\alpha+\lambda)(1+\lambda\gamma)=\lambda^2\beta+\lambda(\alpha^4+\alpha\beta+1)=0$$

which contradicts the fact that the second situation of (5) occurs, because the second case of (5) means that (3) does not hold.

When $V=0$, we deduce that $\alpha=\lambda=1$. In this case, (1) corresponds to

$$\left\{\begin{array}{lcl} (1+\gamma)^2a^4+(1+\gamma)a &=& 0 \\ \beta^2b^4+\beta b &=& 0\ . \end{array}\right.$$

Using that $\beta=1+\gamma$ is nonzero, we deduce that this system of equations has at most 4 solutions (a,b), because each equation has at most 2 solutions.

We now need to handle the case $\lambda=\gamma^{-1}$. Then, the first equation in (1) equals

$$(\gamma^{-1}\alpha+\alpha^2)^2b^4+(\alpha+\gamma^{-1}\alpha^2)b=0\ .$$

The two coefficients of this equation cannot simultaneously vanish, unless

$$\gamma=\alpha^{-1}=\alpha$$

implying that $\alpha=\gamma=1$ which is impossible since $\beta\neq 0$. Therefore, at most two values of b (including $b=0$) satisfy this equation. Replacing b by these two values in the second equation in (1) leads to

$$(\alpha+\lambda\alpha^2)^2a^4+(\alpha^2+\lambda\alpha)a=\delta$$

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

for at most two values of δ (including $\delta = 0$). Again the coefficients of a^4 and of a cannot simultaneously vanish since $\alpha = \gamma = 1$ is impossible. This equation has at most two solutions for each value of b, leading to a total of at most 4 pairs (a, b) solutions of (1). We have thus proved that, unless $(\beta, \lambda) = ((1+\alpha)^3, 1)$, (1) has at most four solutions, i.e. $f_{1,\lambda}$ has at most four constant derivatives. We eventually deduce from Proposition 1 that $\mathcal{L}(f_{1,\lambda}) \in \{2^n, 2^{n+1}\}$. It follows that $\mathcal{L}(\mathsf{V}_{\alpha,\beta}) = 2^{n+1}$ since it corresponds to the linearity of $f_{0,\lambda}$.

The last case is when $(\beta, \lambda) = ((1+\alpha)^3, 1)$. Then, $(1+\gamma) = (\alpha^2 + \alpha)$. It follows that (1) is equivalent to

$$(\alpha^2 + \alpha)^2(a + b)^4 + (\alpha^2 + \alpha)(a + b) = 0 \ .$$

We deduce that $a + b$ takes exactly two values, implying that there are exactly 2^{n+1} pairs (a, b) solutions of (1). It follows that, for $\beta = (1 + \alpha)^3$, $\mathcal{L}(\mathsf{V}_{\alpha,\beta}) = 2^{\frac{3n+1}{2}}$. ■

V. On Differential Uniformity

In this section, we describe the differential properties of generalised butterflies. First, Section V-A is dedicated to Theorem 3, which shows that generalised butterflies $\mathsf{V}_{\alpha,\beta}$ and $\mathsf{H}_{\alpha,\beta}$ have differential uniformity at most 4, unless $\beta = (1 + \alpha)^3$. Then, Theorem 4 and its immediate consequence, Corollary 1, give a necessary and sufficient condition on α and β for a generalised butterfly to be APN. This corollary is then used to show Proposition 2 which states that there are no APN butterflies if $n > 3$. These results are presented and proved in Section V-B. Section V-C focuses on the special case $\alpha = \beta = 1$, for which the generalised butterflies are equivalent to a 3-round Feistel network. In this case, we recover, with a different proof, a result from [20], which states that the difference distribution tables of the corresponding butterflies do not contain the value 2. Finally, we show in Section V-D that the whole Walsh spectrum and difference distribution table of a generalised butterfly are determined by the number of bent components of the closed butterfly.

But first, we give a lemma playing a crucial role in these proofs. It allows to easily derive the maximum number of solutions of some particular degree-4 equations that appear several times in our proof. Since most proofs in this section rely on the number of solutions of univariate equations over $\mathbb{F}_{2^n}$, the notion of degree that will be used always refers to the univariate degree.

Lemma 4. *Let U, V be elements of $\mathbb{F}_{2^n}$ with n odd and let $Uz^4 + Vz^2 + (U + V)z = C$ be some linearised degree-4 equation in z. It has:*

- *0 or 2^n solutions if $U = V = 0$,*
- *0 or 4 solutions if $U \neq 0$, $U \neq V$ and $\mathrm{Tr}\,(V/U) = 1$,*
- *0 or 2 solutions otherwise, that is if one of the following is true:*
  - *$U = 0, V \neq 0$,*
  - *$U \neq 0$ and $V = U$,*
  - *$U \neq 0$ and $\mathrm{Tr}\,(V/U) = 0$,*

Proof: First of all, for any value of the constant C, the number of solutions of the equation is either zero or equal to the number of solutions of the linearised equation $Uz^4 + Vz^2 + (U + V)z = 0$. We then only need to study the case $C = 0$. Obviously, the number of solutions is always even as if z is a solution then $z + 1$ is too.

If $U = V = 0$ then the linearised equation does not involve z, meaning that all values of z satisfy it. We now suppose that either $U \neq 0$ or $V \neq 0$.

If $U = 0$ then the equation corresponds to $Vz(z + 1) = 0$, implying that it has 2 solutions.

Let us now suppose that $U \neq 0$. In this case, we can rewrite the equation as

$$Uz(z + 1)\big(1 + V/U + z(z + 1)\big) = 0.$$

Both $z = 0$ and $z = 1$ are obviously solutions. In fact, they are the only ones if $V = U$. Let us suppose that $V \neq U$. The term $(z^2 + z + 1 + V/U)$ can be equal to 0 if and only if $\mathrm{Tr}\,(V/U) = 1$, meaning that the linearised equation has 2 solutions if $\mathrm{Tr}\,(V/U) = 0$ and 4 otherwise. ■

A. The Non-APN Cases

Theorem 3 (Differential uniformity)**.** *Let $n > 1$ be an odd integer and (α, β) be a pair of nonzero elements in $\mathbb{F}_{2^n}$. If $\beta \neq (1+\alpha)^3$, the generalised butterfly with parameters α and β has differential uniformity at most 4. Moreover, it has differential uniformity* exactly 4 *unless $\beta \in \{(\alpha + \alpha^3), (\alpha^{-1} + \alpha^3)\}$.*

If $\beta = (1 + \alpha)^3$, the generalised butterfly with parameters α and β has differential uniformity 2^{n+1}.

Proof: In order to bound the differential uniformity of $\mathsf{V}_{\alpha,\beta}$, we must bound the number of solutions (x, y) of the following system:

$$\begin{cases} R(x, y) + R(x + a, y + b) = c \\ R(y, x) + R(y + b, x + a) = d \end{cases}$$

for any tuple (a, b, c, d) of $\mathbb{F}_{2^n}$ with $(a, b) \neq (0, 0)$. We have

$$\begin{aligned} R(x, y)+R(x + a, y + b) = &(ax^2 + a^2x) + \alpha(bx^2 + a^2y) \\ &+ \alpha^2(b^2x + ay^2) + (\alpha^3 + \beta)(by^2 + b^2y) \\ &+ R(a, b) \ . \end{aligned}$$

Let $u = a + \alpha b$. Then

$$\begin{aligned} R(x, y)+R(x + a, y + b) =&ux^2 + u^2x + (\alpha^2u + b\beta)y^2 \\ &+ (\alpha u^2 + b^2\beta)y + R(a, b) \end{aligned}$$

Similarly, for $v = \alpha a + b$, we have

$$\begin{aligned} R(y, x) + R(y + b, x + a) =&(\alpha^2v + a\beta)x^2 + (\alpha v^2 + a^2\beta)x \\ &+ vy^2 + v^2y + R(b, a) \ , \end{aligned}$$

implying that we search for the solutions of

$$\begin{cases} ux^2 + u^2x + (\alpha^2u + b\beta)y^2 + (\alpha u^2 + b^2\beta)y = c' \\ (\alpha^2v + a\beta)x^2 + (\alpha v^2 + a^2\beta)x + vy^2 + v^2y = d' \end{cases} \tag{7}$$

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

a) Special cases: We first focus on three special cases, namely $b = \alpha^{-1}a, \alpha a, 0$. The rest of the proof will be dedicated to the general case, when b differs from these three values. We also consider a fourth special case corresponding to $\beta = (1+\alpha)^3$ and $b = a$.

- $\boldsymbol{b = \alpha^{-1}a}$, or equivalently $u = 0$. Note that neither a nor b vanishes, since it would imply $a = b = 0$, which has been excluded. In this case, (7) can be written as

$$\begin{cases}(b\beta)y^2 + (b^2\beta)y = c' \\ (\alpha^2 v + a\beta)x^2 + (\alpha v^2 + a^2\beta)x + vy^2 + v^2 y = d' \ .\end{cases}$$

Since $\beta \neq 0$ and $b \neq 0$, we deduce that the first equation has at most two solutions y_0 and y_1. For each of these two solutions, the second equation has at most two solutions because the coefficients of x^2 and of x cannot simultaneously vanish. Indeed

$$(\alpha^2 v + a\beta) = (\alpha v^2 + a^2\beta) = 0 \tag{8}$$

implies that

$$a^2\beta = \alpha v^2 = \alpha^2 a v \ ,$$

leading to

$$\alpha v(v + a\alpha) = \alpha v b = 0 \ ,$$

which is impossible since $v = 0$ together with (8) would imply that $a = 0$. Therefore, (7) has at most four solutions when $u = 0$.

- $\boldsymbol{b = \alpha a}$, or equivalently $v = 0$. This case is similar to the previous one. Indeed, (7) now corresponds to

$$\begin{cases}ux^2 + u^2 x + (\alpha^2 u + b\beta)y^2 + (\alpha u^2 + b^2\beta)y = c' \\ a\beta x^2 + a^2\beta x = d'\end{cases}$$

Since $a\beta \neq 0$, the second equation has at most two solutions x_0 and x_1. For each of these solutions, the first equation has at most two solutions for y since the coefficients of y^2 and y cannot simultaneously vanish. Otherwise, we would have

$$b^2\beta = \alpha u^2 = \alpha^2 b u$$

implying $\alpha u a = 0$.

- $\boldsymbol{b = 0}$. Then, System (7) corresponds to

$$\begin{cases}ax^2 + a^2 x + \alpha^2 a y^2 + \alpha a^2 y = c' \\ (\alpha^3 a + a\beta)x^2 + (\alpha^3 a^2 + a^2\beta)x + \alpha a y^2 + \alpha^2 a^2 y = d' \ .\end{cases}$$

By summing the first equation and the second one multiplied by α, we get that

$$y\alpha a^2(1+\alpha^2) = (a + a\alpha^4 + a\alpha\beta)(x^2 + ax) + g$$

for some constant g. Let us first consider the case when $\alpha = 1$. Then, we get

$$a\beta(x^2 + ax) = g \ .$$

Since $a\beta \neq 0$, this equation has at most two solutions x_0 and x_1. Then, for each x_i, the first equation in the system provides at most two solutions for y, leading to at most four solutions (x, y) for the whole system.

Let us now assume that $\alpha \neq 1$. Then, we replace y by its value, i.e. $y = \mu(x^2 + ax) + g'$ in the first equation of the system, and we get

$$\alpha^2 a\mu^2 x^4 + \left[a + \alpha^2 a^3\mu^2 + \alpha a^2\mu\right] x^2 + \left[a^2 + \alpha a^3\mu\right] x = c' \ ,$$

where

$$\mu = \frac{(1 + \alpha^4 + \alpha\beta)}{\alpha a(1+\alpha^2)} \ .$$

By replacing $x = ax'$, we deduce that

$$Ux'^4 + Vx'^2 + (U+V)x' = c' \tag{9}$$

with

$$U = \alpha^2 a^5\mu^2 \text{ and } V = a^3 + \alpha^2 a^5\mu^2 + \alpha a^4\mu \ .$$

This equation has at most four solutions x_i, and each x_i leads to a single y, implying that the whole system has at most four solutions.

We now show that the whole system has at most two solutions for any $a \neq 0$ for two values of β only. It is worth noticing that, since $\mathsf{V}_{1,\beta}$ cannot be APN because any three-round Feistel network has differential uniformity at least 4 [20], $\alpha = 1$ can be excluded. If $\mathsf{V}_{\alpha,\beta}$ is APN, then the previous degree-4 equation (9) has at most two solutions for any $a \neq 0$ and any c'. We derive from Lemma 4 that this happens if and only if, for all $a \neq 0$,

$$U = 0 \text{ and } V \neq 0$$

or

$$U = V \text{ and } U \neq 0$$

or

$$U \neq 0 \text{ and } \mathrm{Tr}(V/U) = 0 \ .$$

We first observe that $V \neq 0$, otherwise

$$\alpha^2 a^2\mu^2 + \alpha a\mu + 1 = 0$$

which would mean that $(\alpha a\mu)$ is a root of $X^2 + X + 1$ while this polynomial is irreducible over $\mathbb{F}_{2^n}$, n odd. Then, the first condition means that

$$\mu = \frac{(1 + \alpha^4 + \alpha\beta)}{\alpha a(1+\alpha^2)} = 0$$

or equivalently

$$\beta = \alpha^{-1} + \alpha^3 \ .$$

The second condition corresponds to

$$\alpha a\mu = 1 \Leftrightarrow 1 + \alpha^4 + \alpha\beta = 1 + \alpha^2 \ ,$$

This condition is equivalent to

$$\beta = \alpha + \alpha^3 \ .$$

The last condition corresponds to

$$\begin{aligned}\mathrm{Tr}(V/U) &= \mathrm{Tr}(1) + \mathrm{Tr}\left(\frac{1 + \alpha a\mu}{a^2\alpha^2\mu^2}\right) \\ &= 1 + \mathrm{Tr}\left(\frac{1}{a^2\alpha^2\mu^2}\right) + \mathrm{Tr}\left(\frac{1}{a\alpha\mu}\right) = 1 = 0 \ ,\end{aligned}$$

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

which is impossible, Therefore, the only values of β for which $\mathsf{V}_{\alpha,\beta}$ can be APN are $\beta = \alpha^{-1} + \alpha^3$ and $\beta = \alpha + \alpha^3$.

- $\boldsymbol{b = a}$ and $\boldsymbol{\beta = (1+\alpha)^3}$. Note that $\beta = (1+\alpha)^3 \neq 0$ implies that $\alpha \neq 1$. In this case, (7) is equal to

$$\begin{cases} b(1+\alpha)x^2 + b^2(1+\alpha)^2x + b(1+\alpha)y^2 + b^2(1+\alpha)^2y = c' \\ b(1+\alpha)x^2 + b^2(1+\alpha)^2x + b(1+\alpha)y^2 + b^2(1+\alpha)^2y = d' \end{cases} .$$

Thus, it has no solution if $c' \neq d'$. If $c' = d'$, it is equivalent to the single equation

$$(x+y)^2 + b(1+\alpha)(x+y) = c'b^{-1}(1+\alpha)^{-1} ,$$

since $\alpha \neq 1$ and $b \neq 0$. Thus, either System (7) has no solution, or its solutions are of the form $x+y = \varepsilon$ for two values of ε, depending on (b, c'). In particular, System (7) has exactly 2^{n+1} solutions when $c' = d' = 0$.

b) General case: Let us now assume that u, v and b are all nonzero. We also suppose that $a = b$ and $\beta = (1+\alpha)^3$ do not hold simultaneously. Let ℓ_1 and ℓ_2 respectively denote the two equations in (7). Then the following expression must be constant:

$$\begin{aligned} v\ell_1 + u\ell_2 &= \left(uv(\alpha^2+1) + au\beta\right)x^2 + \left(uv(u+\alpha v) + a^2u\beta\right)x \\ &\quad + \left(uv(\alpha^2+1) + bv\beta\right)y^2 + \left(uv(\alpha u + v) + b^2v\beta\right)y \\ &= \left(uv(\alpha^2+1) + au\beta\right)(x^2+ax) \\ &\quad + \left(uv(\alpha^2+1) + bv\beta\right)(y^2+by) , \end{aligned}$$

where the last equality comes from the fact that $(u+\alpha v) = a(\alpha^2+1)$ and $(\alpha u + v) = b(\alpha^2+1)$. We have obtained a relation of the form

$$\lambda_0(x^2+ax) + \lambda_1(y^2+by) = \varepsilon \tag{10}$$

for some constant ε. We first prove that λ_0 and λ_1 cannot simultaneously vanish. Let us first consider the case $\alpha = 1$. Then

$$\lambda_0 = (a+b)a\beta \text{ and } \lambda_1 = (a+b)b\beta .$$

Since $u = a+b \neq 0$, $b \neq 0$ and $\beta \neq 0$, λ_1 does not vanish.

Let us now assume that $\alpha \neq 1$. Then we can write

$$\beta = (1+\alpha^2)(\alpha+\beta') .$$

It follows that

$$\begin{aligned} \lambda_0 &= uv(\alpha^2+1) + au\beta = u(\alpha^2+1)(b+a\beta') \\ \lambda_1 &= uv(\alpha^2+1) + bv\beta = v(\alpha^2+1)(a+b\beta') . \end{aligned}$$

Then (10) holds with

$$\lambda_0 = u(b+a\beta') \text{ and } \lambda_1 = v(a+b\beta') .$$

These two coefficients cannot simultaneously vanish: otherwise, it would lead to

$$a\beta' = b \text{ and } b\beta' = a$$

implying that

$$ab\beta' = a^2 = b^2 \text{ and } \beta' = 1 ,$$

which has been excluded since it implies that $\beta = (1+\alpha)^3$ and $a = b$.

We now combine (10) with one of the equations in (7). We need to consider two different cases:

- If $\lambda_0 = 0$, then (10), which can be written as

$$y^2 + by = \varepsilon' ,$$

has at most two solutions y_0 and y_1. By replacing y by these two values in the first equation in (7), we get at most two solutions for x for each y_i since $u \neq 0$.

- If $\lambda_0 \neq 0$, then (10) can be written as

$$x^2 = ax + \lambda_0^{-1}\lambda_1(y^2+by) + \varepsilon' .$$

We replace x^2 by this expression in the first equation in (7), and we get

$$\begin{aligned} &(ua+u^2)x + (u\lambda_0^{-1}\lambda_1 + \alpha^2u + b\beta)y^2 \\ &\qquad + (u\lambda_0^{-1}\lambda_1 b + \alpha u^2 + b^2\beta)y = c' . \end{aligned}$$

The coefficient of x does not vanish since $u = a$ is equivalent to $b = 0$. Then x can be written as a degree-2 polynomial in y, i.e.

$$x = \mu_2y^2 + \mu_1y + \mu_0 . \tag{11}$$

By replacing x by its value in (10), we derive that

$$\lambda_1(y^2+by) = \lambda_0\left(\mu_2^2y^4 + \mu_1^2y^2 + a\mu_2y^2 + a\mu_1y\right) + \varepsilon''$$

leading to

$$\lambda_0\mu_2^2y^4 + \left(\lambda_1 + \lambda_0\mu_1^2 + \lambda_0a\mu_2\right)y^2 + (\lambda_1b + \lambda_0a\mu_1)y + \varepsilon'' . \tag{12}$$

Let us first assume that the three coefficients of y^4, y^2 and y do not simultaneously vanish. Then (12) has at most four solutions, y_i, $0 \leq i < 4$. Moreover, we know from (11) that x is entirely determined by y. It follows that System (7) has at most four solutions (x, y).

Let us now suppose that all coefficients of (12) vanish. Then we have $\mu_2 = 0$ and

$$\lambda_1 + \lambda_0\mu_1^2 = 0 \text{ and } \lambda_1b + \lambda_0a\mu_1 = 0 .$$

This may occur in one of the following two situations:

- $\mu_1 = 0$ and $\lambda_1 = 0$. Using that λ_1 cannot vanish only when $\alpha \neq 1$, the definitions of λ_1 and μ_1 imply that

$$\alpha u^2 + b^2\beta = 0 \text{ and } u(\alpha^2+1) + b\beta = 0 ,$$

leading to

$$\alpha u^2 = u(\alpha^2+1)b$$

i.e.,

$$\alpha a + b = v = 0$$

which has been excluded.

- $b\mu_1 = a$ and $b^2\lambda_1 = a^2\lambda_0$. From the definition of μ_2, we deduce that $\mu_2 = 0$ together with this last relation implies that

$$\begin{aligned} 0 &= u\lambda_0^{-1}\lambda_1 + \alpha^2u + b\beta \\ &= b^{-2}\left(ua^2 + u\alpha^2b^2 + b^3\beta\right) \\ &= b^{-2}(u^3 + b^3\beta) \end{aligned}$$

i.e.,

$$\beta = (ub^{-1})^3 . \tag{13}$$

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

Since $\mu_2 = 0$ and $\mu_1 = ab^{-1}$, (11) can be written

$$ay = bx + \mu_0' \ .$$

By replacing ay by its value in the second equation of (7) multiplied by a^2 and with $A = vb^2 + \alpha^2 a^2 v + a^3\beta$ and $B = v^2 ab + \alpha a^2 v^2 + a^4\beta$, we get

$$\begin{aligned} Ax^2 + Bx &= d'' \\ \Leftrightarrow (v^3 + a^3\beta)(x^2 + ax) &= d'' \ . \end{aligned} \tag{14}$$

The coefficients of this equation do not vanish. Otherwise, combined with (13), it would yield

$$v^3 + a^3\beta = (\alpha a + b)^3 + (a^2 b^{-1} + \alpha a)^3 = 0$$

which implies

$$b = a^2 b^{-1}$$

i.e., $a = b$ and $\beta = (\alpha + 1)^3$ which has been excluded.

It follows that (14) has at most two solutions x_0 and x_1. Since y is entirely determined by x (or constant), it follows that System (7) has at most two solutions in this case.

■

B. On APN Butterflies

We first derive a necessary and sufficient condition for a generalised butterfly to be APN in Theorem 4. Then, we simplify these conditions in Corollary 1. Finally, we show in Proposition 2 that this condition can only be satisfied if $n = 3$.

Theorem 4 (APN Condition)**.** *Let $\alpha \neq 0, 1$. A generalised butterfly with parameters α and β is APN if and only if:*

$\beta \in \{\alpha + \alpha^3, \alpha^{-1} + \alpha^3\}$ *and* $\mathrm{Tr}\,(\mathcal{A}_\alpha(e)) = 1, \forall\, e \notin \{0, \alpha, 1/\alpha\}$,

where

$$\mathcal{A}_\alpha(e) = \frac{e\alpha(1+\alpha)^2}{(1+\alpha e)(\alpha + e)^2} \ .$$

Proof: Since we have proved in Section III-B that generalised butterflies with parameters (α, β_0) and (α, β_1) where $\beta_1 = \beta_0^{-1}(1+\alpha)^6$ are affine-equivalent, we only need to prove the result for $\beta = \alpha + \alpha^3$. As before, we need to count the number of solutions of

$$\begin{cases} R(x,y) + R(x+a, y+b) = c \\ R(y,x) + R(y+b, x+a) = d \end{cases} \tag{15}$$

for any tuple (a, b, c, d) of $\mathbb{F}_{2^n}$ with $(a, b) \neq (0, 0)$. Call

$$[\ell_1] : R(x,y) + R(x+a, y+b) = c$$

$$[\ell_2] : R(y,x) + R(y+b, x+a) = d$$

This system is equivalent to

$$\begin{aligned} [\ell_1] : ax^2 + a^2x + \alpha(bx^2 + a^2y) \\ + \alpha^2(b^2x + ay^2) \\ + (\alpha^3 + \beta)(by^2 + b^2y) = c_0 \ . \end{aligned}$$

and

$$\begin{aligned} [\ell_2] : by^2 + b^2y + \alpha(ay^2 + b^2x) \\ + \alpha^2(a^2y + bx^2) \\ + (\alpha^3 + \beta)(ax^2 + a^2x) = d_0 \ . \end{aligned}$$

As $\alpha \neq 1$, we can replace the lines ℓ_1 and ℓ_2 of this system by $\ell_1 + \alpha\ell_2$ and $\alpha\ell_1 + \ell_2$ to obtain a system with the exact same number of solutions. We obtain

$$\begin{aligned} [\ell_1] : (ax^2 + a^2x)(1 + \alpha\beta + \alpha^4) \\ + (\alpha + \alpha^3)(bx^2 + a^2y) \\ + (\alpha^3 + \alpha + \beta)(by^2 + b^2y) = c_0 \ . \end{aligned}$$

$$\begin{aligned} [\ell_2] : (by^2 + b^2y)(1 + \alpha\beta + \alpha^4) \\ + (\alpha + \alpha^3)(ay^2 + b^2x) \\ + (\alpha^3 + \alpha + \beta)(ax^2 + a^2x) = d_0 \ . \end{aligned}$$

For $\beta = \alpha + \alpha^3$, the system further simplifies using that $1 + \alpha\beta + \alpha^4 = (1 + \alpha^2)$ and $\alpha + \alpha^3 + \beta = 0$:

$$\begin{cases} (ax^2 + a^2x) + \alpha(bx^2 + a^2y) = c_1 \\ (by^2 + b^2y) + \alpha(ay^2 + b^2x) = d_1 \ . \end{cases} \tag{16}$$

We first consider the cases $a = 0$ and $b = 0$. Recall that $a = b = 0$ is excluded. If $a = 0$, then the first line of the system is equivalent to

$$x = \left(\frac{c_1}{\alpha b}\right)^{2^{n-1}} \ .$$

Replacing x by this value in the second line of System (16) yields a degree-2 equation in y with nonzero coefficients since $b \neq 0$, implying that (16) has at most two solutions (x, y). The case $b = 0$ is similar.

We now suppose $a \neq 0$ and $b \neq 0$, which allows us to set $x = ax'$ and $y = by'$. In this context, System (16) has as many solutions as

$$\begin{cases} a^3(x'^2 + x') + \alpha a^2 b(x'^2 + y') = c_1 \\ b^3(y'^2 + y') + \alpha a b^2(y'^2 + x') = d_1 \ , \end{cases}$$

which we rewrite using $e = a/b$ as

$$\begin{cases} e(x'^2 + x') + \alpha(x'^2 + y') = c_2 \\ e^{-1}(y'^2 + y') + \alpha(y'^2 + x') = d_2 \ . \end{cases} \tag{17}$$

Summing its lines yields

$$(x'^2 + x')(e + \alpha) + (y'^2 + y')(e^{-1} + \alpha) \ = \ c_2 + d_2 \ .$$

If $e = \alpha$, then y' is fixed to either y_0' or y_1' with $y_0' + y_1' = 1$. The first line of the system implies in this case that $x' = y_i' + c_2/\alpha$ as the terms in x^2 cancel each other, meaning that the system has at most two solutions. The case $e = \alpha^{-1}$ is similar. We now suppose $e \neq \alpha, \alpha^{-1}$.

The first line of System (17) allows us to express y' as a function of x':

$$y' = x'^2\left(\frac{e}{\alpha} + 1\right) + x'\frac{e}{\alpha} + \frac{c_2}{\alpha} \ .$$

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

We replace y' by this expression in the second line of (17) and obtain

$$\begin{aligned} Y &= y'^2\left(e^{-1}+\alpha\right) + y'e^{-1} + \alpha x' \\ &= \left(x'^2\left(\frac{e}{\alpha}+1\right) + x'\frac{e}{\alpha}\right)^2\left(e^{-1}+\alpha\right) \\ &\quad + \left(x'^2\left(\frac{e}{\alpha}+1\right) + x'\frac{e}{\alpha}\right)e^{-1} + \alpha x' \\ &= x'^4\left(1+\frac{e}{\alpha}\right)^2(\alpha+e^{-1}) \\ &\quad + x'^2\left(\frac{e^2}{\alpha^2}(e^{-1}+\alpha) + \left(\frac{e}{\alpha}+1\right)e^{-1}\right) + x'\left(\frac{1}{\alpha}+\alpha\right) \\ &= d_3 \end{aligned}$$

for some constant d_3. If we let $U = (1+e/\alpha)^2(\alpha+1/e)$ and $V = U + 1/\alpha + \alpha$, then the number of solutions of this equation can be computed using Lemma 4. First, $U \neq 0$ and $U + V \neq 0$ as $\alpha \neq 1$. Therefore, the possible number of solutions is at most equal to 4 and is given by the trace of V/U: if $\mathrm{Tr}\,(V/U) = 0$ then the equation has at most 2 solutions, otherwise it has 0 or 4 solutions. It holds that

$$\begin{aligned} \frac{V}{U} &= 1 + \frac{\alpha^{-1}+\alpha}{(e^{-1}+\alpha)(1+e\alpha^{-1})^2} \\ &= 1 + \frac{e\alpha(1+\alpha)^2}{(1+\alpha e)(\alpha+e)^2} \end{aligned}$$

so the function is APN if and only if

$$\begin{aligned} &\mathrm{Tr}\,(\mathcal{A}_\alpha(e)) = 1,\ \forall e \neq 0, \alpha, 1/\alpha, \\ &\text{with } \mathcal{A}_\alpha(e) = \frac{e\alpha(1+\alpha)^2}{(1+\alpha e)(\alpha+e)^2}\ . \end{aligned}$$

■

The condition provided by Theorem 4 is sufficient to describe all APN generalised butterflies but it can be greatly simplified. This is stated in the following corollary.

Corollary 1. *Let $\alpha \neq 1$, $\beta_0 = \alpha^3 + \alpha$ and $\beta_1 = \alpha^3 + 1/\alpha$. A generalised butterfly with parameters α and β is APN if and only if $\beta = \beta_0$ or β_1 and*

$$\mathrm{Tr}\,(\mathcal{C}_\alpha(v)) = 1,\ \forall\ u \notin \left\{0, 1, 1/(1+\alpha^{-2})\right\}\ .$$

with

$$\mathcal{C}_\alpha(v) = \left(\frac{1}{1+\alpha^{-1}}\right)^4 \frac{1}{u+u^3}.$$

Proof of Corollary 1: We know from Theorem 4 that a generalised butterfly with parameters α and β is APN if and only if $\beta \in \{\beta_0, \beta_1\}$ and $\mathrm{Tr}\,(\mathcal{A}_\alpha(e)) = 1$ for all e not in $\{0, \alpha, 1/\alpha\}$. Suppose that $\alpha \neq 1$ and let $\ell = (e+\alpha)(1+\alpha)^2$. Then we can rewrite some of the expressions involved in $\mathcal{A}_\alpha(e)$ as follows:

$$e(1+\alpha)^2 = \ell+\alpha+\alpha^3 \text{ and } (1+\alpha e)(1+\alpha)^2 = \alpha\left(\ell + \frac{(1+\alpha)^4}{\alpha}\right)$$

Recall that $\beta_0 = \alpha + \alpha^3$ and $\beta_1 = (\alpha+1)^4/\alpha$, so we can write:

$$\begin{aligned} \mathcal{A}_\alpha(e) &= \frac{e\alpha(1+\alpha)^2}{(1+\alpha e)(\alpha+e)^2} \\ &= \frac{\alpha\ e(1+\alpha^2)}{(1+\alpha e)(1+\alpha^2)\frac{((\alpha+e)(1+\alpha)^2)^2}{(1+\alpha)^6}} \\ &= (1+\alpha)^6\frac{\alpha(\ell+\beta_0)}{\alpha(\ell+\beta_1)\ell^2} \\ &= \frac{\beta_0\beta_1}{\ell^2}\frac{\ell+\beta_0}{\ell+\beta_1}\ . \end{aligned}$$

Let $\mathcal{B}_\alpha(v) = v^2(v+1)/(v+\beta_0/\beta_1)$. Then the following equality holds:

$$\mathcal{B}_\alpha\left(\frac{\beta_0}{\ell}\right) = \frac{\beta_0\beta_1}{\ell^2}\frac{\ell+\beta_0}{\ell+\beta_1} = \mathcal{A}_\alpha(e)\ .$$

It is therefore sufficient to study the trace of $\mathcal{B}_\alpha$. The condition $e \notin \{0, \alpha, \alpha^{-1}\}$ becomes $\ell \notin \{\beta_0, 0, \beta_1\}$ respectively and, equivalently, $\beta_0/\ell \notin \{0, 1, \beta_0/\beta_1\}$. As a consequence, the generalised butterfly with parameters α, β is APN if and only if $\beta = \beta_0$ or β_1 and

$$\mathrm{Tr}\,(\mathcal{B}_\alpha(v)) = 1,\ \forall v \notin \left\{0, 1, \frac{\alpha^2}{1+\alpha^2}\right\}$$

as $\beta_0/\beta_1 = \alpha^2/(1+\alpha^2)$. Finally, we note that the trace of $\mathcal{B}_\alpha(v)$ can be simplified:

$$\begin{aligned} \mathrm{Tr}\,(\mathcal{B}_\alpha(v)) &= \mathrm{Tr}\left(v^2\frac{1+v}{v+\frac{\alpha^2}{1+\alpha^2}}\right) \\ &= \mathrm{Tr}\left(v^2 + \frac{v^2}{(1+\alpha^2)v+\alpha^2}\right) \\ &= \mathrm{Tr}\left(v + \frac{v^2}{(1+\alpha^2)v+\alpha^2}\right) \\ &= \mathrm{Tr}\left(\frac{(1+\alpha^2)v^2+\alpha^2 v+v^2}{(1+\alpha^2)v+\alpha^2}\right) \\ &= \mathrm{Tr}\left(\frac{v^2+v}{\gamma v+1}\right)\ , \end{aligned}$$

where $\gamma = 1+\alpha^{-2}$. We deduce the following:

$$\begin{aligned} \mathrm{Tr}\left(\mathcal{B}_\alpha\left(u^{-1}\gamma^{-1}\right)\right) &= \mathrm{Tr}\left(\frac{\left(u^{-1}\gamma^{-1}\right)^2 + u^{-1}\gamma^{-1}}{u^{-1}+1}\right) \\ &= \mathrm{Tr}\left(\frac{\left(u^{-1}\gamma^{-1}\right)^2}{u^{-1}+1} + \frac{\left(u^{-1}\gamma^{-1}\right)^2}{u^{-2}+1}\right) \\ &= \mathrm{Tr}\left(\frac{(u^{-3}+u^{-2})\gamma^{-2}+u^{-2}\gamma^{-2}}{u^{-2}+1}\right) \\ &= \mathrm{Tr}\left(\frac{\gamma^{-2}}{u+u^3}\right). \end{aligned}$$

The condition $u^{-1}/\gamma \notin \{0, 1, \gamma^{-1}\}$ is equivalent to $u \notin \{0, \gamma^{-1}, 1\}$, the same set as before. This proves the corollary. ■

We now show that the last condition in Corollary 1 can hold if $n = 3$ only. In other words, APN generalised butterflies exist for $n = 3$ only. The proof relies on the following lemma.

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

Lemma 5. *[21] The cubic equation $x^3+ax+b=0$, where $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^n}^*$ has a unique solution in $\mathbb{F}_{2^n}$ if and only if $\mathrm{Tr}(a^3/b^2) \neq \mathrm{Tr}(1)$.*

Proposition 2. *Let $n>1$ be an odd integer, and $\lambda \in \mathbb{F}_{2^n}^*$. If*

$$\mathrm{Tr}\left(\frac{\lambda^2}{x+x^3}\right)=1,\ \forall\ x \notin \{0,1,\lambda\}\ , \tag{18}$$

then $n=3$.

Proof: Let z in $\mathbb{F}_{2^n}^*$ with $\mathrm{Tr}(z)=0$. Then, there exists a unique $x \in \mathbb{F}_{2^n} \setminus \mathbb{F}_2$ such that

$$\frac{1}{x^3+x}=z\ .$$

Indeed, since $z \neq 0$, this equivalently means that

$$x^3+x+\frac{1}{z}=0\ .$$

We know from Lemma 5 that this equation has a unique solution when $\mathrm{Tr}(z^2)=\mathrm{Tr}(z)=0$. Let us define z_λ as

$$z_\lambda=\frac{1}{\lambda^3+\lambda}\ ,$$

and

$$\mathcal{Z}=\{z \in \mathbb{F}_{2^n}^* \setminus \{z_\lambda\} : \mathrm{Tr}(z)=0\}\ .$$

Obviously, $\mathcal{Z}$ is either a hyperplane without 0 or a hyperplane without 0 and z_λ (depending on the value of $\mathrm{Tr}(z_\lambda)$). Then, Condition (18) implies that, for any $z \in \mathcal{Z}$,

$$\mathrm{Tr}(\lambda^2 z)=1\ .$$

Suppose that $n \geq 5$. Then, $\mathcal{Z}$ contains at least $(2^{n-1}-2) \geq 14$ elements and there exists at least two distinct elements z_0 and z_1 in $\mathcal{Z}$ such that $z_0+z_1 \in \mathcal{Z}$. Therefore, these two elements must satisfy

$$\mathrm{Tr}(\lambda^2 z_0)=\mathrm{Tr}(\lambda^2 z_1)=\mathrm{Tr}(\lambda^2(z_0+z_1))=1$$

which is impossible since

$$\mathrm{Tr}(\lambda^2(z_0+z_1))=\mathrm{Tr}(\lambda^2 z_0)+\mathrm{Tr}(\lambda^2 z_1)$$

When $n=3$, the situation is different since the condition may be satisfied when $\mathcal{Z}$ contains 2 elements only, i.e. when $\mathrm{Tr}(z_\lambda)=0$. ■

C. The Feistel case $\alpha=\beta=1$

In the case when $\alpha=\beta=1$, the generalised butterfly is equivalent to a 3-round Feistel network with round functions $x \mapsto x^3$, $x \mapsto x^{1/3}$ and $x \mapsto x^3$. Theorem 4 in [20] shows that, in this special case, the difference distribution table of the corresponding butterflies does not contain any 2. In other words, the number of solutions (x,y) of

$$\begin{cases} R(x,y)+R(x+a,y+b)=c \\ R(y,x)+R(y+b,x+a)=d \end{cases}$$

for any tuple (a,b,c,d) of $\mathbb{F}_{2^n}$ with $(a,b) \neq (0,0)$ is either 0 or 4. We now give an alternative proof of this result.

Proposition 3. *[20, Theorem 4] For $\alpha=\beta=1$, the difference distribution tables of the butterflies $\mathsf{V}_{1,1}$ and $\mathsf{H}_{1,1}$ contain the values 0 and 4 only.*

Proof: As in the proof of Theorem 3, we have to count the number of solutions of System (7), which simplifies to

$$\begin{cases} (a+b)x^2+(a+b)^2x+ay^2+a^2y=c' \\ bx^2+b^2x+(a+b)y^2+(a+b)^2y=d'\ . \end{cases} \tag{19}$$

- If $a=0$, the first line of the system equals $b(x^2+bx)=c'$ which has either 0 or 2 solutions, x_0 and x_1 (recall that a and b cannot simultaneously vanish). The second line of the system can be rewritten as

  $$b\left((x+y)^2+b(x+y)\right)=d'$$

  which has either 0 or 2 solutions, implying $y \in \{x+z_0, x+z_1\}$. Therefore, if the first line has two solutions, the second one has either 0 or 4 solutions. The case $b=0$ is similar.
- If $a=b$, the system is composed of two independent degree-2 equations, one in x and the second one in y. If one of these equations has no solution, then the whole system does not have any solution. Otherwise, each equation has two solutions, and the system has 4 solutions.
- If $ab(a+b) \neq 0$. Then, the first line ℓ_1 of (19) can be replaced by $b\ell_1+(a+b)\ell_2$, leading to

  $$\begin{cases} ab(a+b)x+(ab+a^2+b^2)y^2+(a^3+b^3+ab^2)y=\varepsilon \\ bx^2+b^2x+(a+b)y^2+(a+b)^2y=d'\ . \end{cases} \tag{20}$$

  We now multiply the second line by $a^2b(a+b)^2$ and replace $ab(a+b)x$ by the value given by the first line and get

  $$\begin{aligned}\varepsilon' &= y^4(ab+a^2+b^2)^2+y^2\left(a^6+b^6+ab^5+a^5b+a^3b^3\right) \\ &\quad + y\left(a^6b+a^5b^2+a^4b^3+a^3b^4+a^2b^5+ab^6\right)\ .\end{aligned}$$

  Replacing $y'=by$, we equivalently obtain

  $$Uy'^4+Vy'^2+Wy'=b^{-8}\varepsilon' \tag{21}$$

  where the coefficients U, V and W depend on $e=a/b$:

  $$\begin{aligned} U &= e^4+e^2+1=(e^2+e+1)^2 \\ V &= e^6+e^5+e^3+e+1=(e^2+e+1)^3 \\ W &= e^6+e^5+e^4+e^3+e^2+e=U+V\ . \end{aligned}$$

  Lemma 4 then applies. Clearly $U \neq 0$ since the polynomial X^2+X+1 has no root in $\mathbb{F}_{2^n}$ when n is odd. Also, $U \neq V$, otherwise $e^2+e+1=1$ which is not possible since the cases $e \in \{0,1\}$ (i.e., $a=0$ or $a=b$) have been excluded. Then, Equation (21) has two solutions only if $\mathrm{Tr}(V/U)=0$. But,

  $$\mathrm{Tr}\left(\frac{V}{U}\right)=\mathrm{Tr}(e^2+e+1)=\mathrm{Tr}(1)=1\ ,$$

  implying that Equation (21) has 0 or 4 solutions y_i, and each y_i leads to a unique value of x. Therefore, the whole system has either 0 or 4 solutions.

■

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

D. Walsh spectrum and difference distribution tables of generalised butterflies

Theorem 2 points out that, for $\beta \neq (1+\alpha)^3$, the nonzero components of $\mathsf{V}_{\alpha,\beta}$ are either bent, or their Walsh coefficients belong to $\{0, \pm 2^{n+1}\}$. Then, the whole multiset

$$\{|\widehat{\mathsf{H}_{\alpha,\beta}}(u,v)|, u \in \mathbb{F}_2^{2n}, v \in \mathbb{F}_2^{2n} \setminus \{0\}\} \\ = \\ \{|\widehat{\mathsf{V}_{\alpha,\beta}}(u,v)|, u \in \mathbb{F}_2^{2n}, v \in \mathbb{F}_2^{2n} \setminus \{0\}\}$$

is entirely determined by the number B of bent components of $\mathsf{V}_{\alpha,\beta}$. Moreover, it is well-known that there is a one-to-one correspondence between the squared Walsh transform of a vectorial function and its difference distribution table (see e.g. [22], [23]). Therefore, different values of B correspond to different difference distribution tables. Since all generalised butterflies have differential uniformity at most 4, the value of B equivalently determines the number of occurrences of 4 in the difference distribution table of the mapping. This correspondence is detailed in the following proposition, which is a variant of Corollary 3 in [24].

Proposition 4. *Let m be an even integer, and F be a differentially 4-uniform mapping from $\mathbb{F}_2^m$ into itself such that its nonzero components are bent or have Walsh spectrum $\{0, \pm 2^{\frac{m}{2}+1}\}$. Then, the number of 4 in the difference distribution table of F is equal to*

$$2^{m-2}(2^m-1) - 3 \times 2^{m-1} B$$

where B is the number of bent components of F.

Most notably,

- *F is APN if and only if $B = \frac{2(2^m-1)}{3}$;*
- *The difference distribution table of F does not contain any 2 if and only if $B = 0$.*

Proof: It is well-known [22], [23] that the squared Walsh transform of F is the Fourier transform of the mapping

$$(a,b) \in \mathbb{F}_2^m \times \mathbb{F}_2^m \mapsto \delta(a,b) = \#\{x \in \mathbb{F}_2^m : F(x+a)+F(x) = b\}.$$

A direct consequence is then the following result mentioned in [25, Theorem 2] for instance:

$$\sum_{\lambda \in \mathbb{F}_2^m} \sum_{\mu \in \mathbb{F}_2^m} \widehat{F}(\lambda,\mu)^4 = 2^{2m} \sum_{a,b \in \mathbb{F}_2^m} \delta(a,b)^2 .$$

Moreover, we have

$$\sum_{\lambda \in \mathbb{F}_2^m} \widehat{F}(\lambda,\mu)^4 = \begin{cases} 2^{4m} \text{ if } \mu = 0 \\ 2^{3m} \text{ if } F_\mu \text{ is bent} \\ 2^{3m+2} \text{ otherwise.} \end{cases}$$

It follows that

$$\sum_{\lambda \in \mathbb{F}_2^m} \sum_{\mu \in \mathbb{F}_2^m} \widehat{F}(\lambda,\mu)^4 = 2^{4m} + 2^{3m}\left(2^{m+2} - 4 - 3B\right) .$$

Let A_2 and A_4 respectively denote the number of occurrences of 2 and 4 in the difference distribution table of F. We know that

$$\sum_{a \in \mathbb{F}_2^m \setminus \{0\}} \sum_{b \in \mathbb{F}_{2^m}} \delta(a,b) = 2A_2 + 4A_4 = (2^m-1)2^m .$$

We derive that

$$\begin{aligned} \sum_{a,b \in \mathbb{F}_2^m} \delta(a,b)^2 &= 2^{2m} + 4A_2 + 16A_4 \\ &= 2^{2m} + 2^{m+1}(2^m-1) + 8A_4 . \end{aligned}$$

Therefore

$$\begin{aligned} 8A_4 &= 2^m\left(2^{m+2} - 4 - 3B - 2^{m+1} - 2\right) \\ &= 2^m\left(2^{m+1} - 2 - 3B\right) , \end{aligned}$$

leading to the result. Most notably, F is APN, i.e. $A_4 = 0$ if and only if

$$2^{m+1} - 2 - 3B = 0 ,$$

and $A_2 = 0$ if and only if $A_4 = 2^{m-2}(2^m-1)$, i.e., $B = 0$. ■

We have seen in Proposition 3 that, when $\alpha = \beta = 1$, the difference distribution tables of the corresponding butterflies do not contain any 2. The last item of the previous proposition shows that this situation corresponds to the case where $\mathsf{V}_{\alpha,\beta}$ has no bent component. In other words, the Walsh coefficients of $\mathsf{V}_{1,1}$ and of $\mathsf{H}_{1,1}$ over $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ take the values 0 and $\pm 2^{n+1}$ only. This result, corresponding to Theorem 5 in [20], is then a direct consequence of Proposition 3.

More generally, the whole Walsh spectrum and difference distribution table of generalised butterflies is derived by applying the previous proposition to $\mathsf{V}_{\alpha,\beta}$.

Corollary 2 (Walsh and differential spectra of generalised butterflies)**.** *Let α and β be two nonzero elements in $\mathbb{F}_{2^n}$ such that $\beta \neq (1+\alpha)^3$. The Walsh spectrum of $\mathsf{H}_{\alpha,\beta}$ and $\mathsf{V}_{\alpha,\beta}$, i.e., the multiset*

$$\{|\widehat{\mathsf{H}_{\alpha,\beta}}(u,v)|, u \in \mathbb{F}_2^{2n}, v \in \mathbb{F}_2^{2n} \setminus \{0\}\} \\ = \\ \{|\widehat{\mathsf{V}_{\alpha,\beta}}(u,v)|, u \in \mathbb{F}_2^{2n}, v \in \mathbb{F}_2^{2n} \setminus \{0\}\}$$

is given by

$$\left|\widehat{\mathsf{H}_{\alpha,\beta}}(u,v)\right| = \begin{cases} 0, & 3 \times 2^{2n-2}(2^n-1)(2^n+1-C) \text{ times} \\ 2^n, & 2^{2n}(2^n-1)C \text{ times} \\ 2^{n+1}, & 2^{2n-2}(2^n-1)(2^n+1-C) \text{ times.} \end{cases}$$

where $(2^n-1)C$ is the number of bent components of $\mathsf{V}_{\alpha,\beta}$.

The difference distribution table of both $\mathsf{H}_{\alpha,\beta}$ and $\mathsf{V}_{\alpha,\beta}$ contains the values 0, 2 and 4 with the following number of occurrences

$$\delta(a,b) = \begin{cases} 2, & 2^{2n-2}(2^n-1) \times 3C \text{ times} \\ 4, & 2^{2n-3}(2^n-1)(2^{n+2}+4-3C) \text{ times} \end{cases}$$

Proof: The result is directly deduced from Proposition 4, using that the number of bent components of $\mathsf{V}_{\alpha,\beta}$ is of the form $B = (2^n-1)C$. Indeed, let $f_{\lambda,\mu}$, for λ, μ in $\mathbb{F}_{2^n}$, denote the components of $\mathsf{V}_{\alpha,\beta}$, i.e.,

$$f_{\lambda,\mu} : (x,y) \mapsto \mathrm{Tr}(\lambda R(x,y)) + \mathrm{Tr}(\mu R(y,x)) .$$

We have seen in the proof of Theorem 2 that, for any $\mu \in \mathbb{F}_{2^n}^*$, $f_{0,\mu}$ is not bent, and that, for any nonzero $\lambda \in \mathbb{F}_{2^n}$, $f_{\lambda,\mu}$ is bent if and only if $f_{1,\lambda^{-1}\mu}$ is bent. We deduce that $B = (2^n-1)C$

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

where C is the number of $\mu \in \mathbb{F}_{2^n}$ such that $f_{1,\mu}$ is bent. The Walsh spectrum of the generalised butterfly is then derived by using that all Walsh coefficients of a bent component are equal to $\pm 2^n$ and, for a component with linearity 2^{n+1}, the Walsh transform takes 2^{2n-2} times the value $\pm 2^{n+1}$ and $(2^{2n} - 2^{2n-2})$ times the value 0 (see Proposition 1).

The differential spectrum is deduced from Proposition 4: the number of 4 in the difference distribution table is

$$\begin{aligned} A_4 &= 2^{m-2}(2^m - 1) - 3 \times 2^{m-1} B \\ &= 2^{2n-3}(2^n - 1)\left(2^{n+1} + 2 - 3C\right) , \end{aligned}$$

and the number of 2 in the difference distribution table is

$$A_2 = (2^{2n} - 1)2^{2n-1} - 2A_4 = 2^{2n-2}(2^n - 1) \times 3C .$$

■

The values of C for all $\mathsf{H}_{\alpha,\beta}$ of $2n$ variables, with $n \in \{3, 5\}$ are given in Tables I and II. We have checked by computer for $n \in \{3, 5, 7, 9\}$ that, when α and β vary in $\mathbb{F}_{2^n}^*$, C takes the value 0 and all even values between $(\frac{2^n+4}{3} - 2^{(n-1)/2})$ and $(\frac{2^n+4}{3} + 2^{(n-1)/2})$. This implies that, for these values of n, the family of generalised butterflies contains mappings with $(2^{(n-1)/2} + 2)$ different Walsh spectra (and therefore the same number of difference distribution tables).

It is also worth noticing that the family of generalised butterflies includes some mappings which are not CCZ-equivalent to the mappings in the family studied in [9]. Indeed, the case $\beta = 1$ does not include all possible values for C.

We can also check, from the values of C, that most of the generalised butterflies are not CCZ-equivalent to previously known differentially 4-uniform permutations. For instance, Table I shows that the generalised butterflies of 6 variables have a difference distribution table with a number of 4 equal to one of the following four values:

$$A_4 \in \{0, 336, 672, 1008\} .$$

While the last case corresponds to the same differential spectrum as the Gold and Kasami power permutations, the functions with the two intermediate values of A_4 cannot be CCZ-equivalent to a power function. Indeed, the only differentially 4-uniform power functions satisfy $A_4 = 1008$ (Gold and Kasami) or $A_4 = 63$ (inverse mapping). More generally, for any number of variables, none of the generalised butterflies is CCZ-equivalent to the inverse mapping since the formula of A_4 in the proof of Corollary 2 shows that A_4 is even, while the inverse mapping has an odd number of 4 in its difference distribution table. Similarly, it can be checked that the only permutation of $\mathbb{F}_2^6$ with optimal nonlinearity constructed in [26] has another differential spectrum (see [26, Table II]). It is also worth noticing that the generalised butterflies of six variables have a higher nonlinearity than the functions constructed in [27] and in [28].

VI. On Algebraic Degree

It is well-known that, if all Walsh coefficients of a Boolean function f of m variables are divisible by 2^ℓ, then the algebraic degree of f is at most $(m + 1 - \ell)$ (see e.g. [29, Prop. 1.5]). Thus, we deduce from Theorem 2 that the algebraic degree

TABLE I: Value of C, i.e., number of bent components divided by (2^3-1), of all $\mathsf{H}_{\alpha,\beta}$ for α and β in $\mathbb{F}_{2^3}^*$ where $\mathbb{F}_{2^3}$ is defined by the primitive element a such that $a^3 + a + 1 = 0$.

$\alpha \backslash \beta$	1	a	a^2	a^3	a^4	a^5	a^6
1	0	4	4	4	4	4	4
a	**6**	2	0	2	**6**	0	0
a^3	2	4	2	0	2	4	2

of any generalised butterfly of $2n$ variables does not exceed $(n + 1)$. We now show that this upper bound is actually tight for almost all values of (α, β).

Theorem 5. *Let α and β be two nonzero elements in $\mathbb{F}_{2^n}$. The generalised open butterfly $\mathsf{H}_{\alpha,\beta}$ has an algebraic degree equal to n or $n + 1$. It is equal to n if and only if*

$$(1 + \alpha\beta + \alpha^4)^3 \ = \ \beta(\beta + \alpha + \alpha^3)^3.$$

The closed butterfly $\mathsf{V}_{\alpha,\beta}$ has algebraic degree 2.

Remark 2. *The condition $(1+\alpha\beta+\alpha^4)^3 = \beta(\beta+\alpha+\alpha^3)^3$ can alternatively be written $Z(\alpha, \beta) = 0$, where:*

$$Z(\alpha, \beta) \ = \ \beta^4 + \alpha\beta^3 + \alpha(\alpha + 1)^6\beta + (1 + \alpha)^{12}.$$

Furthermore, Z can be factorised as follows:

$$\begin{aligned} Z(\alpha, \beta) &= \beta^4 + \alpha\beta^3 + \alpha(\alpha + 1)^6\beta + (1 + \alpha)^{12} \\ &= \left(\beta^2 + (1 + \alpha)^6\right)\left(\beta^2 + \alpha\beta + (1 + \alpha)^6\right) \\ &= \left(\beta^2 + (1 + \alpha)^6\right)\left(1 + \alpha\beta + \alpha^4 + (\beta + \alpha + \alpha^3)^2\right). \end{aligned}$$

Hence, if $\beta \neq (1 + \alpha)^3$ then $Z(\alpha, \beta) = 0$ if and only $1 + \alpha\beta + \alpha^4 = (\beta + \alpha + \alpha^3)^2$. It follows that $Z(\alpha, \beta)$ is equal to 0 when $\beta = (1 + \alpha)^3$ and, if $\mathrm{Tr}(\alpha^{-1}) = 1$, for two additional values of β. This includes the Feistel case, when $\alpha = \beta = 1$.

Proof: Obviously, $\mathsf{V}_\alpha(x, y)$ has algebraic degree 2. We then focus on the generalised open butterfly $\mathsf{H}_{\alpha,\beta}$. The right side of the output of such an open butterfly is equal to $(x + \beta y^3)^{1/3} + \alpha y$, where (x, y) is the input. We deduce from Theorem 1 of [30] (or equivalently from Proposition 5 of [5]) that the inverse of 3 modulo $(2^n - 1)$ for odd n is

$$1/3 \ \equiv \ \sum_{i=0}^{(n-1)/2} 2^{2i} \mod (2^n - 1),$$

which implies in particular that the algebraic degree of $x \mapsto x^{1/3}$ is equal to $(n+1)/2$. We deduce from this expression that the function $t(x, y) = (x + \beta y^3)^{1/3}$ is equal to $\prod_{i=0}^{(n-1)/2}(x + \beta y^3)^{2^{2k}}$. This sum can be developed as follows:

$$\begin{aligned} t(x, y) &= (x + \beta y^3)^{1/3} \\ &= \sum_{J \subseteq [0,(n-1)/2]} \underbrace{\prod_{j \in J} \beta^{2^{2j}} y^{3 \times 2^{2j}}}_{\deg < 2|J|} \ \underbrace{\prod_{j \in \overline{J}} x^{2^{2j}}}_{\deg = (n+1)/2 - |J|} , \end{aligned}$$

where $\overline{J}$ is the complement of J in $[0, (n-1)/2]$, i.e. $J \cap \overline{J} = \emptyset$ and $J \cup \overline{J} = [0, (n-1)/2]$. The algebraic degree of each term in this sum is at most equal to $|J| + (n + 1)/2$. Thus, if $|J| < (n - 1)/2$, then the degree of the corresponding term is smaller than n. If $\overline{J} = \emptyset$ then the corresponding term is equal

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

TABLE II: Value of C, i.e., number of bent components divided by (2^5-1), of all $\mathsf{H}_{\alpha,\beta}$ for α and β in $\mathbb{F}_{2^5}^*$ where $\mathbb{F}_{2^5}$ is defined by the primitive element a such that $a^5+a^2+1=0$.

$\alpha\backslash\beta$	1	a	a^2	a^3	a^4	a^5	a^6	a^7	a^8	a^9	a^{10}	a^{11}	a^{12}	a^{13}	a^{14}	a^{14}	a^{16}	a^{17}	a^{18}	a^{19}	a^{20}	a^{21}	a^{22}	a^{23}	a^{24}	a^{25}	a^{26}	a^{27}	a^{28}	a^{29}	a^{30}
0	0	16	16	12	16	12	12	8	16	12	12	12	12	12	8	12	16	12	12	8	12	12	12	12	12	8	12	12	8	12	12
1	12	10	12	12	8	10	12	10	10	12	10	8	12	12	10	12	10	14	12	10	12	10	14	0	14	10	12	10	12	14	10
3	14	8	10	10	14	12	8	10	12	12	12	12	10	8	12	14	10	10	8	14	14	8	10	12	14	0	14	12	10	8	14
5	14	0	10	10	12	12	0	12	12	10	10	0	14	10	14	12	12	10	14	14	12	12	12	12	14	14	10	12	12	14	10
7	12	12	14	16	0	16	14	12	12	12	8	10	10	14	14	10	0	10	14	12	12	14	10	0	10	14	14	10	10	8	12
11	8	16	14	16	8	14	12	14	10	10	10	10	10	10	14	12	14	8	16	14	16	8	0	14	12	10	0	10	12	14	0
15	12	14	10	16	12	8	12	10	10	10	0	10	10	10	12	8	12	16	10	14	12	14	10	12	8	10	10	8	12	10	14

to $\beta^{1/3}y$ and has degree 1. If $\overline{J}=\{j\}$ for some j, then the term is equal to

$$\begin{aligned} T &= x^{2^{2j}} \times \beta^{1/3}y \times (\beta y^3)^{2^n-1-2^{2j}} \\ &= \beta^{1/3-2^{2j}} \times x^{2^{2j}} \times y^{(2^n-1)-(2^{2j+1}+2^{2j}-1)} \ . \end{aligned}$$

If $j \neq (n-1)/2$, then its algebraic degree is

$$1+n-wt(2^{2j+1}+2^{2j}-1) = n-2j \ .$$

If $j=(n-1)/2$, then the term (omitting the constant factor) is equal to

$$x^{2^{n-1}} \times y \times y^{2^n-1-(2^n-2^{n-1})} = x^{2^{n-1}} y^{2^{n-1}-1} \ .$$

and has degree n. Therefore, $t(x,y)$ has two terms of degree n, corresponding to $j=0$ and $j=(n-1)/2$ namely

$$m_0(x,y) = \beta^{-2/3} x y^{2^n-3}$$

and

$$m_1(x,y) = \beta^{(2^{n-1}-1)/3} x^{2^{n-1}} y^{2^{n-1}-1}$$

Thus, the right side of the output has an algebraic degree equal to n. The left side is equal to

$$L(x,y) = \Big(y+\alpha\big((x+\beta y^3)^{1/3}+\alpha y\big)\Big)^3 + \beta\big((x+\beta y^3)^{1/3}+\alpha y\big)^3,$$

which we can re-write using the function $t(x,y) = (x+\beta y^3)^{1/3}$ as

$$L(x,y) \ = \ \big((\alpha^2+1)y+\alpha t(x,y)\big)^3 + \beta\big(t(x,y)+\alpha y\big)^3,$$

which we expand into

$$\begin{aligned} L(x,y) \ = \ & t(x,y)^3(\alpha^3+\beta) \ + \ y^3\big((\alpha^2+1)^3+\alpha^3\big) \\ & + \ y t(x,y)^2\big((\alpha^2+1)\alpha^2+\beta\alpha\big) \\ & + \ y^2 t(x,y)\big((\alpha^2+1)^2\alpha+\beta\alpha^2\big). \end{aligned}$$

The terms on the first line have degree at most 3. Let us focus on those of the second line and denote their sum $L'(x,y)$. First, we can simplify this expression as follows:

$$\frac{L'(x,y)}{\alpha} \ = \ C_0 y t(x,y)^2 \ + \ C_1 y^2 t(x,y)$$

where $C_0=(\beta+\alpha+\alpha^3)$ and $C_1=(1+\alpha\beta+\alpha^4)$.

Since $t(x,y)$ has algebraic degree n, we deduce that $L'(x,y)$ (and the left side of the output of $\mathsf{H}_{\alpha,\beta}$) has algebraic degree at most $(n+1)$, while the whole function has degree at least n because of the right side. Moreover, this upper bound is reached if and only if the terms of degree $(n+1)$ in $L'(x,y)$ do not cancel each other. The only terms in $L'(x,y)$ which may have degree $(n+1)$ correspond to terms of degree n in $t(x,y)$, namely (omitting the constant factors):

$$\begin{aligned} y^2 m_0(x,y) &= x y^{2^n-1}, & y^2 m_1(x,y) &= x^{2^{n-1}} y^{2^{n-1}+1}, \\ y m_0(x,y)^2 &= x^2 y^{(2^n-1)-3}, & y m_1(x,y)^2 &= x y^{2^n-1} \ . \end{aligned}$$

Only the first and the last terms have degree $(n+1)$. Therefore, the term of degree $(n+1)$ in $L'(x,y)$ is:

$$\begin{aligned} C_0 y m_1(x,y)^2 + C_1 y^2 m_0(x,y) &= x y^{2^n-1}\left(C_0\beta^{-1/3}+C_1\beta^{-2/3}\right) \\ &= x y^{2^n-1}\beta^{-1/3}\left(C_0+C_1\beta^{-1/3}\right). \end{aligned}$$

It follows that $\mathsf{H}_{\alpha,\beta}$ has algebraic degree $(n+1)$ if and only if

$$\beta C_0^3 \neq C_1^3 \ .$$

■

VII. Conclusion

We have solved the two open questions raised in [9] on the properties of butterflies. Moreover, a larger family of permutations of $\mathbb{F}_2^{2n}$, for any odd $n \geq 3$, with differential uniformity 4 and linearity 2^{n+1}, has been described. Several functions with similar cryptographic properties have already been exhibited, e.g. in [26]–[28], [31]–[35]. However, the family of open butterflies presents two interesting characteristics. First, it has a reasonably simple representation which may be suitable for cryptographic applications (some specific cases can even be implemented as 3-round Feistel networks) which is also easier to handle for determining its cryptographic properties. A second specificity is that it is the only known such infinite family which includes the APN permutation exhibited by Dillon *et al.* Unfortunately, it does not contain any other APN permutation, implying that the "big APN problem" raised by Dillon remains open.

Acknowledgment

The work of Léo Perrin is supported by the CORE ACRYPT project (ID C12-15-4009992) funded by the *Fonds National de la Recherche* (Luxembourg).

References

[1] E. Biham and A. Shamir, "Differential cryptanalysis of DES-like cryptosystems," in *CRYPTO'90*, ser. LNCS, A. J. Menezes and S. A. Vanstone, Eds., vol. 537. Springer, Berlin, Germany, 1991, pp. 2–21.

[2] ——, "Differential cryptanalysis of DES-like cryptosystems," *Journal of Cryptology*, vol. 4, no. 1, pp. 3–72, 1991.

[3] M. Matsui, "Linear cryptoanalysis method for DES cipher," in *EUROCRYPT'93*, ser. LNCS, T. Helleseth, Ed., vol. 765. Springer, Berlin, Germany, 1994, pp. 386–397.

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2676807, IEEE Transactions on Information Theory

[4] ——, "The first experimental cryptanalysis of the data encryption standard," in *CRYPTO'94*, ser. LNCS, Y. Desmedt, Ed., vol. 839. Springer, Berlin, Germany, 1994, pp. 1–11.

[5] K. Nyberg, "Differentially uniform mappings for cryptography," in *EUROCRYPT'93*, ser. LNCS, T. Helleseth, Ed., vol. 765. Springer, Berlin, Germany, 1994, pp. 55–64.

[6] K. Nyberg and L. R. Knudsen, "Provable security against differential cryptanalysis," in *CRYPTO'92*, ser. LNCS, E. F. Brickell, Ed., vol. 740. Springer, Berlin, Germany, 1993, pp. 566–574.

[7] R. Gold, "Maximal recursive sequences with 3-valued recursive cross-correlation functions," *IEEE Trans. Information Theory*, vol. 14, pp. 154–156, 1968.

[8] K. Browning, J. Dillon, M. McQuistan, and A. Wolfe, "An APN permutation in dimension six," in *Finite Fields: Theory and Applications - FQ9*, ser. Contemporary Mathematics, vol. 518. AMS, 2010, pp. 33–42.

[9] L. Perrin, A. Udovenko, and A. Biryukov, "Cryptanalysis of a theorem: Decomposing the only known solution to the big APN problem," in *Advances in Cryptology - CRYPTO 2016, Part II*, ser. LNCS, M. Robshaw and J. Katz, Eds., vol. 9815. Springer, 2016, pp. 93–122.

[10] S. Fu and X. Feng, "Further results of the cryptographic properties on the butterfly structure," Preprint arXiv:1607.08455, 2016.

[11] G. M. Kyureghyan, *Finite fields and their Applications - Character sums and polynomials*. De Gruyter, 2013, ch. Special mappings of finite fields, pp. 117–144.

[12] C. Carlet, P. Charpin, and V. Zinoviev, "Codes, Bent Functions and Permutations Suitable For DES-like Cryptosystems," *Des. Codes Cryptography*, vol. 15, no. 2, pp. 125–156, 1998.

[13] R. A. Mollin and C. Small, "On permutation polynomials over finite fields," *International Journal of Mathematics and Mathematical Sciences*, vol. 10, no. 3, pp. 535–543, 1987.

[14] F. Göloglu, "Almost perfect nonlinear trinomials and hexanomials," *Finite Fields and Their Applications*, vol. 33, pp. 258–282, 2015.

[15] F. J. MacWilliams and N. J. Sloane, *The theory of error-correcting codes*. North-Holland, 1977.

[16] C. Carlet, *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*. Cambridge University Press, 2010, ch. Boolean functions for cryptography and error correcting codes, pp. 257–397.

[17] A. Canteaut, C. Carlet, P. Charpin, and C. Fontaine, "On cryptographic properties of the cosets of R(1, m)," *IEEE Trans. Information Theory*, vol. 47, no. 4, pp. 1494–1513, 2001.

[18] ——, "Propagation characteristics and correlation-immunity of highly nonlinear boolean functions," in *EUROCRYPT 2000*, ser. LNCS, B. Preneel, Ed., vol. 1807. Springer, Berlin, Germany, 2000, pp. 507–522.

[19] C. Bracken, E. Byrne, N. Markin, and G. McGuire, "Determining the nonlinearity of a new family of APN functions," in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes - AAECC-17*, ser. LNCS, vol. 4851. Springer, 2007, pp. 72–79.

[20] Y. Li and M. Wang, "Constructing S-boxes for lightweight cryptography with Feistel structure," in *CHES 2014*, ser. LNCS, L. Batina and M. Robshaw, Eds., vol. 8731. Springer, Berlin, Germany, 2014, pp. 127–146.

[21] E. Berlekamp, H. Rumsey, and G. Solomon, "On the solution of algebraic equations over finite fields," *Inform. Contr.*, vol. 12, no. 5, pp. 553–564, October 1967.

[22] F. Chabaud and S. Vaudenay, "Links between differential and linear cryptoanalysis," in *EUROCRYPT'94*, ser. LNCS, A. D. Santis, Ed., vol. 950. Springer, Berlin, Germany, 1995, pp. 356–365.

[23] C. Blondeau and K. Nyberg, "Links between truncated differential and multidimensional linear properties of block ciphers and underlying attack complexities," in *EUROCRYPT 2014*, ser. LNCS, P. Q. Nguyen and E. Oswald, Eds., vol. 8441. Springer, Berlin, Germany, 2014, pp. 165–182.

[24] T. P. Berger, A. Canteaut, P. Charpin, and Y. Laigle-Chapuy, "On almost perfect nonlinear functions over $\mathbf{F}_2^n$," *IEEE Trans. Information Theory*, vol. 52, no. 9, pp. 4160–4170, 2006.

[25] Y. Tan, G. Gong, and B. Zhu, "Enhanced criteria on differential uniformity and nonlinearity of cryptographically significant functions," *Cryptography and Communications*, vol. 8, no. 2, pp. 291–311, 2016.

[26] L. Qu, Y. Tan, C. H. Tan, and C. Li, "Constructing differentially 4-uniform permutations over $\mathbf{F}_{2^{2k}}$ via the switching method," *IEEE Trans. Information Theory*, vol. 59, no. 7, pp. 4675–4686, 2013.

[27] D. Tang, C. Carlet, and X. Tang, "Differentially 4-uniform bijections by permuting the inverse function," *Des. Codes Cryptography*, vol. 77, no. 1, pp. 117–141, 2015.

[28] Z. Zha, L. Hu, and S. Sun, "Constructing new differentially 4-uniform permutations from the inverse function," *Finite Fields and Their Applications*, vol. 25, pp. 64–78, 2014.

[29] P. Langevin, "Covering radius of RM (1, 9) in RM (3, 9)," in *International Symposium on Coding Theory and Applications - EUROCODE '90*, ser. LNCS, G. D. Cohen and P. Charpin, Eds., vol. 514. Springer, 1990, pp. 51–59.

[30] G. M. Kyureghyan and V. Suder, "On inverses of APN exponents," in *IEEE International Symposium on Information Theory - ISIT 2012*. IEEE, 2012, pp. 1207–1211.

[31] C. Bracken and G. Leander, "A highly nonlinear differentially 4 uniform power mapping that permutes fields of even degree," *Finite Fields and Their Applications*, vol. 16, no. 4, pp. 231–242, 2010.

[32] P. Charpin, G. M. Kyureghyan, and V. Suder, "Sparse permutations with low differential uniformity," *Finite Fields and Their Applications*, vol. 28, pp. 214–243, 2014.

[33] Y. Li and M. Wang, "Constructing differentially 4-uniform permutations over $GF(2^{2m})$ from quadratic APN permutations over $GF(2^{2m+1})$," *Des. Codes Cryptography*, vol. 72, no. 2, pp. 249–264, 2014.

[34] L. Qu, Y. Tan, C. Li, and G. Gong, "More constructions of differentially 4-uniform permutations on $\mathbf{F}_{2^{2k}}$," *Des. Codes Cryptography*, vol. 78, no. 2, pp. 391–408, 2016.

[35] J. Peng and C. H. Tan, "New explicit constructions of differentially 4-uniform permutations via special partitions of $\mathbf{F}_{2^{2k}}$," *Finite Fields and Their Applications*, vol. 40, pp. 73–89, 2016.

Anne Canteaut received the French engineer's degree from the École Nationale Supérieure de Techniques Avancées in 1993 and the Ph.D. degree in computer science from the University of Paris VI, France, in 1996. Since 1997, she has been a researcher with the French National Research Institute in Computer Science (Inria), Paris. She is currently Director of Research and the scientific head of the SECRET research team at Inria. Her research interests include symmetric cryptography and coding theory.

Dr. Canteaut has served on program committees for more than 50 international conferences such as Eurocrypt, Crypto, Asiacrypt and FSE. Most notably, she served as program chair for Indocrypt 2004 (Chennai, India), for WCC 2011 (Workshop on Coding and Cryptography, Paris) and for FSE 2012 (Fast Software Encryption, Washington DC, USA). She currently serves on, or has served on the Editorial Boards of the following journals: IEEE Transactions on Information Theory (2005 to 2008); Finite Fields and their Applications; Applicable Algebra in Engineering, Communication and Computing; IACR Transactions on Symmetric Cryptology. Since 2017, she is the vice-head of science of the Inria Paris research center.

Sébastien Duval received the French engineer's degree from Télécom ParisTech in 2015 and is currently a Ph.D. student at the French National Research Institute in Computer Science (Inria), Paris, within the SECRET team. His work focuses on symmetric cryptography.

Léo Perrin received the French engineer's degree from Centrale Lyon and the Swedish engineering's degree from the Royal Institute of Technology in 2013. He is currently a Ph.D. student at the Univsersity of Luxembourg in the CryptoLux team. Since 2017, he serves on the editorial board of the IACR Transactions on Symmetric Cryptology. His research interests include symmetric cryptography and Boolean functions.

0018-9448 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.