On the distance between APN functions

Lilya Budaghyan[1], Claude Carlet[1,2], Tor Helleseth[1], and Nikolay Kaleyski[1]

[1]*Department of Informatics, University of Bergen, Norway*
[2]*Department of Mathematics, Universities of Paris VIII and XIII, France*

***Abstract*—We investigate the differential properties of a vectorial Boolean function G obtained by modifying an APN function F. This generalizes previous constructions where a function is modified at a few points. We characterize the APN-ness of G via the derivatives of F, and deduce an algorithm for searching for APN functions whose values differ from those of F only on a given $U \subseteq \mathbb{F}_{2^n}$.**

We introduce a value Π_F associated with any F, which is invariant under CCZ-equivalence. We express a lower bound on the distance between a given APN function F and the closest APN function in terms of Π_F. We show how Π_F can be computed efficiently for F quadratic. We compute Π_F for all known APN functions over $\mathbb{F}_{2^n}$ up to $n \leq 8$. This is the first new CCZ-invariant for APN functions to be introduced within the last ten years.

We derive a mathematical formula for this lower bound for the Gold function $F(x) = x^3$, and observe that it tends to infinity with n. Finally, we describe how to efficiently find all sets U such that taking $G(x) = F(x) + v$ for $x \in U$ and $G(x) = F(x)$ for $x \notin U$ is APN.

I. Introduction

A vectorial (n, m)-Boolean function is any mapping $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$, where $\mathbb{F}_{2^n}$ is the finite field with 2^n elements. Such a function can also be seen as mapping sequences of n bits (zeros and ones) to sequences of m bits, which more clearly reveals their practical importance. Vectorial Boolean functions are of central interest in cryptography since they can be used to represent virtually all components of a block cipher; in particular, its non-linear components (whose cryptographic properties directly influence the cipher's security) can be expressed as vectorial Boolean functions. For instance, the Advanced Encryption Standard (AES) and algorithms based on Feistel networks such as the Data Encryption Standard (DES), all utilize vectorial Boolean functions in the role of so-called "substitution boxes". The resistance of the encryption to various categories of cryptanalytic attacks then directly depends on the properties of the underlying Boolean functions (see e.g. [22] for basic background on cryptography and encryption schemes).

Almost Perfect Nonlinear (APN) functions were introduced by Nyberg [20] as the functions that provide optimal resistance to the so-called differential attack invented by Biham and Shamir [2]. More precisely, we say that a function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is APN if the equation $F(x) + F(x + a) = b$ in x has at most 2 solutions for any $a \in \mathbb{F}_{2^n}^*$ and any $b \in \mathbb{F}_{2^n}$. Despite the simplicity of this definition, finding and investigating the properties of APN functions, even in finite fields of relatively low dimension, is a challenging task. For this reason, various methods of constructing such functions have been considered by researchers.

In [6], a construction in which a function $G : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is obtained from a given function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ by modifying one of its values is introduced in an attempt to resolve the open problem of the existence of APN functions over $\mathbb{F}_{2^n}$ of algebraic degree n. A number of nonexistence results are obtained in the paper, which support the conjecture that this is impossible. The idea of the construction is interesting in its own right, however, and it can naturally be generalized to the modification of more than one point.

The particular case of swapping two points of a given function is studied [24] in the context of constructing differentially 4-uniform permutations, and the more general question of arbitrarily modifying the values of a given function at two points, as well as swapping two points in a more general context, is investigated in [17].

In this paper, we consider the general case of arbitrarily changing K points. To be more accurate, given a function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, some K distinct field elements $u_1, \dots, u_K \in \mathbb{F}_{2^n}$ and some K elements $v_1, v_2, \dots, v_K \in \mathbb{F}_{2^n}^*$, we define G as

$$G(x) = \begin{cases} F(u_i) + v_i & x = u_i \\ F(x) & x \notin \{u_1, u_2, \dots, u_K\} \end{cases}$$

and try to find some correlation between the properties of F and those of G. We derive sufficient and necessary conditions that the derivatives of F must satisfy in order for G to be APN, and obtain an efficient filtering procedure for finding all possible values of $v_1, v_2, \dots, v_K$ in the case that $u_1, u_2, \dots, u_K$ are known. In the case when F is itself APN, we define the values Π_F and m_F, which count the number of derivatives of F satisfying a certain condition, and express a lower bound on the distance between F and the closest APN function in terms of m_F. We further demonstrate that these values are invariant under CCZ-equivalence and that their computation is particularly efficient when F is quadratic. In addition, we show how an exact formula for m_F can be computed in the case of $F(x) = x^3$.

We experimentally compute Π_F and m_F for all known APN functions over $\mathbb{F}_{2^n}$ for $n \leq 8$. We notice that over fields of odd dimension, this new invariant tends to take the same value for all known APN functions except the inverse function, but for fields of even dimension, it can take a large number of distinct values which make it a useful tool for disproving CCZ-equivalence between a given pair of functions. These experimental results are summarized in Section IV and Table II, and a detailed table of the computational results can be found online at *https://boolean.h.uib.no/mediawiki/*.

In the case when $v_1 = v_2 = \cdots = v_K$, we show how all possible combinations of points $u_1, u_2, \dots, u_K$ can be found (for all values of K) by solving a system of linear equations. We note that constructions of the form $G(x) = F(x) + vf(x)$ for $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$ have been investigated in [7], [16].

II. Preliminaries

A. Basic Notation

Let n be a positive integer. We denote by $\mathbb{F}_{2^n}$ the finite field with 2^n elements; in particular, $\mathbb{F}_2$ is the field with two elements. For any positive integer m, $\mathbb{F}_2^m$ is the vector space of dimension m over $\mathbb{F}_2$. Given any set S, we denote by S^* the set $S \setminus \{0\}$; in particular, $\mathbb{F}_{2^n}^*$ is the multiplicative group of $\mathbb{F}_{2^n}$.

The characteristic function of the set S is denoted by $1_S(x)$ and is defined as

$$1_S(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S. \end{cases}$$

This paper was presented in part at the Third International Workshop on Boolean Functions and their Applications (BFA-2018) which took place in Loen, Norway on June 17-22, 2018 [5]

For a finite set $S = \{s_1, s_2, \dots, s_k\}$ we will use $1_{s_1,s_2,\dots,s_k}(x)$ as shorthand for $1_{\{s_1,s_2,\dots,s_k\}}(x)$.

B. Representation of Vectorial Functions

Given two positive integers n and m, a vectorial Boolean (n,m)-function, or simply (n,m)-function, is any function $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$. It can be uniquely expressed in the so-called *algebraic normal form* (ANF) as follows [10]:

$$F(x_1, x_2, \dots, x_n) = \sum_{I \subseteq \{1,2,\dots,n\}} a_I \left(\prod_{i \in I} x_i \right) = \sum_{I \subseteq \{1,2,\dots,n\}} a_I x^I, a_I \in \mathbb{F}_2^m.$$

The *algebraic degree* of $F(x_1, x_2, \dots, x_n)$ is defined as the degree of its ANF, namely

$$\deg(F) = \max\{|I| : a_I \neq (0,0,\dots,0), I \subseteq \{1,2,\dots,n\}\}.$$

Clearly, $\deg(F) \leq n$.

Vectorial Boolean $(n,1)$-functions, i.e. functions of the form $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$, are referred to as *Boolean functions*.

When $m = n$, one can identify the vector space $\mathbb{F}_2^n$ with the finite field $\mathbb{F}_{2^n}$. Note that any basis $\{e_1, e_2, \dots, e_n\}$ for $\mathbb{F}_{2^n}$, viewed as a vector space over $\mathbb{F}_2$, determines a correspondence between $\mathbb{F}_{2^n}$ and $\mathbb{F}_2^n$ via $x = \sum_{i=1}^n x_i e_i$. The algebraic degree does not depend on the choice of the basis since any change of basis corresponds to a linear permutation. Then any (n,n)-function has a unique representation as a univariate polynomial over $\mathbb{F}_{2^n}$ of the form

$$F(x) = \sum_{i=0}^{2^n-1} a_i x^i, a_i \in \mathbb{F}_{2^n}.$$

Let $x = \sum_{i=1}^n x_i e_i$ and $i = \sum_{s=0}^{n-1} i_s 2^s$ where $i_s \in \{0,1\}$. Then F can be rewritten as

$$F(x) = \sum_{i=0}^{2^n-1} a_i \left(\sum_{i=1}^n x_i e_i \right)^i = \sum_{i=0}^{2^n-1} a_i \prod_{s=0}^{n-1} \left(\sum_{i=1}^n x_i e_i^{2^s} \right)^{i_s}$$

which, after expansion, gives the ANF of F. Moreover, let $w_2(i) = \sum_{s=0}^{n-1} i_s$ denote the 2-weight of i, where $0 \leq i \leq 2^n - 1$ has binary expansion $i = \sum_{s=0}^{n-1} 2^s i_s$. Then the algebraic degree of F in univariate polynomial form is equal to

$$\deg(F) = \max\{w_2(i) : a_i \neq 0, 0 \leq i \leq 2^n - 1\}.$$

Given two functions $F, G : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, the *Hamming distance* $d(F,G)$ is defined as the number of points $x \in \mathbb{F}_{2^n}$ on which the values of F and G differ, i.e.

$$d(F,G) = |\{x \in \mathbb{F}_{2^n} : F(x) \neq G(x)\}|.$$

C. Almost Perfect Nonlinear Functions and Bent Functions

Let F be a function from $\mathbb{F}_{2^n}$ to itself. The *derivative of F in direction a* for any $a \in \mathbb{F}_{2^n}$ is the function $D_a F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ defined as

$$D_a F(x) = F(x) + F(a + x).$$

The *differential sets* $H_a F$ are the image sets of the derivatives of F, i.e. the sets

$$H_a F = \{D_a F(x) : x \in \mathbb{F}_{2^n}\} = \{F(x) + F(a + x) : x \in \mathbb{F}_{2^n}\}.$$

Alongside the derivatives $D_a F$, we define the *shifted derivative* $D_a^\beta F$ of F *in direction a with shift β*, which is a function over $\mathbb{F}_{2^n}$ defined as

$$D_a^\beta F(x) = D_a F(x) + F(a + \beta) = F(x) + F(a + x) + F(a + \beta)$$

for any fixed $a, \beta \in \mathbb{F}_{2^n}$. The *shifted differential sets* $H_a^\beta F$ are then the image sets of the shifted derivatives, i.e.

$$H_a^\beta F = \{D_a^\beta F(x) : x \in \mathbb{F}_{2^n}\} = \{F(x) + F(a + x) + F(a + \beta) : x \in \mathbb{F}_{2^n}\}.$$

For any $a, b \in \mathbb{F}_{2^n}$, define $\Delta_F(a,b) = |\{x \in \mathbb{F}_{2^n} : F(x + a) + F(x) = b\}|$; that is, $\Delta_F(a,b)$ is the number of solutions x of the equation $D_a F(x) = b$ for some given a and b. Then the *differential uniformity* of F is defined as

$$\Delta_F = \max\{\Delta_F(a,b) : a, b \in \mathbb{F}_{2^n}, a \neq 0\}.$$

A function F from $\mathbb{F}_{2^n}$ to itself is called *differentially δ-uniform* if $\Delta_F \leq \delta$. If $\delta = 2$, then F is called *almost perfect nonlinear* (APN). Note that this is optimal in the case of a finite field of characteristic two, since if some x solves $F(x) + F(a + x) = b$, then so does $(a + x)$, and thus $\Delta_F(a,b)$ is always even.

Note that the definition of differential uniformity can be extended to functions $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ between fields of different dimensions. A *perfect nonlinear* (PN) function is one whose differential uniformity is 2^{n-m}; as observed above, for $n = m$ such functions cannot exist. In fact, PN functions are the same as bent functions (briefly discussed below) and do not exist whenever $m > n/2$ [19].

A number of useful characterizations of APN functions can be given in terms of the so-called Walsh transform. The Walsh transform of a Boolean function $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$ is defined as

$$W_f(a) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x) + \mathrm{Tr}_1^n(ax)}, \quad a \in \mathbb{F}_2,$$

where $\mathrm{Tr}_k^n(x) = \sum_{i=0}^{n-1} x^{2^{ki}}$ is the trace function from $\mathbb{F}_{2^n}$ to its subfield $\mathbb{F}_{2^k}$, for $k \mid n$. We will also use the inverse Walsh transform formula, defined as

$$\sum_{a \in \mathbb{F}_{2^n}} W_f(a) = 2^n (-1)^{f(0)}.$$

The Walsh transform of an (n,m)-function is defined in terms of the Walsh transform of its *component functions* $\mathrm{Tr}_1^m(bF(x))$ for $b \in \mathbb{F}_{2^m}^*$ as

$$W_F(a,u) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}_1^m(uF(x)) + \mathrm{Tr}_1^n(ax)}.$$

If the Walsh transform of a Boolean function $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$ satisfies $W_f(a) \in \{0, \pm\mu\}$ for all $a \in \mathbb{F}_{2^n}$, then f is called a *plateaued* function with *amplitude* μ. An (n,n)-function F is called *plateaued* if all of its component functions are plateaued (possibly with different amplitudes). If all of the component functions of F are plateaued with the same amplitude, then F is called *plateaued with single amplitude*. Plateaued functions are an important class of vectorial Boolean functions since their additional structure makes them more tractable than the general case.

The following characterizations of APN functions by means of the power moments of their Walsh transform are often very useful in the investigation of APN functions.

Lemma 1 ([14])**.** Let F be an (n,n)-function. Then F is APN if and only if

$$\sum_{a \in \mathbb{F}_{2^n}} \sum_{u \in \mathbb{F}_{2^n}^*} W_F^4(a,u) = 2^{3n+1}(2^n - 1).$$

Lemma 2 ([10])**.** Let F be an APN function over $\mathbb{F}_{2^n}$ satisfying $F(0) = 0$. Then

$$\sum_{a,b \in \mathbb{F}_{2^n}} W_F^3(a,b) = 3 \cdot 2^{3n} - 2^{2n+1}.$$

Note that while Lemma 2 expresses only a necessary condition for F to be APN in the general case, in the case of a plateaued function F this condition becomes necessary and sufficient [11].

The following lemma provides an alternative characterization of the APN-ness of a vectorial Boolean function in terms of the second power moments of its derivatives.

Lemma 3 ([21], [1])**.** A function F over $\mathbb{F}_{2^n}$ is APN if and only if for all $a \in \mathbb{F}_{2^n}^*$ we have

$$\sum_{b \in \mathbb{F}_{2^n}} W_{D_a F}(0,b)^2 = 2^{2n+1}.$$

The nonlinearity $\mathcal{NL}_F$ of an (n,m)-function F is the minimum Hamming distance between its component functions and the affine functions. The nonlinearity of any (n,m)-function satisfies the so-called covering radius bound $\mathcal{NL}_F \leq 2^{n-1} - 2^{n/2-1}$. The nonlinearity can be expressed as

$$\mathcal{NL}_F = 2^{n-1} - \frac{1}{2} \max_{a \in \mathbb{F}_{2^m}, u \in \mathbb{F}_{2^n}^*} |W_F(a,u)|.$$

Functions meeting this bound are called *bent*. These coincide with the class of PN functions and exist only for $m \leq n/2$ [21]. In particular, for $m = n$, which is our case of interest, bent functions do not exist.

When n is odd, the optimal (n,n)-functions from the point of view of nonlinearity are the almost bent functions. An (n,n)-function F is called *almost bent* (AB) if it satisfies $W_F(a,u) \in \{0, \pm 2^{(n+1)/2}\}$ for all $a \in \mathbb{F}_{2^n}$ and nonzero $u \in \mathbb{F}_{2^n}^*$. Any AB function is APN, but not vice versa. However, for n odd, every quadratic APN function is also AB [12]. An (n,n)-function F is AB if and only if all the values $W_F(u,v)$ in its Walsh spectrum are divisible by $2^{\frac{n+1}{2}}$ [9].

D. Equivalence Relations of Functions

There are several equivalence relations of functions for which differential uniformity and nonlinearity are invariant. Due to these equivalence relations, having only one APN (respectively, AB) function, one can generate a huge class of APN (respectively, AB) functions.

Two functions F and F' from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^n}$ are called

- *affine equivalent* (*linear equivalent*) if $F' = A_1 \circ F \circ A_2$, where the mappings A_1 and A_2 are affine (linear) permutations of $\mathbb{F}_{2^n}$;
- *extended affine equivalent* (EA-equivalent) if $F' = A_1 \circ F \circ A_2 + A$, where the mappings $A, A_1, A_2 : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ are affine, and A_1, A_2 are permutations;
- *Carlet-Charpin-Zinoviev equivalent* (CCZ-equivalent) if for some affine permutation $\mathcal{L}$ of $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ the image of the graph of F is the graph of F', that is, $\mathcal{L}(G_F) = G_{F'}$ where $G_F = \{(x, F(x)) : x \in \mathbb{F}_{2^n}\}$ and $G_{F'} = \{(x, F'(x)) : x \in \mathbb{F}_{2^n}\}$.

Although different, these equivalence relations are related. It is obvious that linear equivalence is a particular case of affine equivalence, and that affine equivalence is a particular case of EA-equivalence. As shown in [12], EA-equivalence is a particular case of CCZ-equivalence and every permutation is CCZ-equivalent to its inverse. The algebraic degree of a function (if it is not affine) is invariant under EA-equivalence but, in general, it is not preserved by CCZ-equivalence. Let us recall why the structure of CCZ-equivalence implies this: for a function F from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^n}$ and an affine permutation $\mathcal{L}(x,y) = \big(L_1(x,y), L_2(x,y)\big)$ of $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$, where $L_1, L_2 : \mathbb{F}_{2^n} \times \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, we have

$$\mathcal{L}(G_F) = \{\big(F_1(x), F_2(x)\big) : x \in \mathbb{F}_{2^n}\} \tag{1}$$

where $F_1(x) = L_1(x, F(x))$ and $F_2(x) = L_2(x, F(x))$.

Note that $\mathcal{L}(G_F)$ is the graph of a function if and only if F_1 is a permutation. The function CCZ-equivalent to F whose graph equals $\mathcal{L}(G_F)$ is then $F' = F_2 \circ F_1^{-1}$. The composition by the inverse of F_1 modifies the algebraic degree in general, except, for instance, when $L_1(x,y)$ depends only on x, which corresponds to EA-equivalence of F and F' [8]. It is also proven in [8] that CCZ-equivalence is strictly more general that EA-equivalence combined with taking inverses of permutations.

Proposition 1 ([8])**.** Let F and F' be functions from $\mathbb{F}_2^n$ to itself. The function F' is EA-equivalent to the function F or to the inverse of F (if it exists) if and only if there exists an affine permutation $\mathcal{L} = (L_1, L_2)$ on $\mathbb{F}_2^{2n}$ such that $\mathcal{L}(G_F) = G_{F'}$ and L_1 depends only on one variable, i.e. $L_1(x,y) = L(x)$ or $L_1(x,y) = L(y)$.

It is worth listing some properties that remain invariant under CCZ-equivalence. Let the functions F and F' be CCZ-equivalent. Then

- $\{\Delta_F(a,b) : a, b \in \mathbb{F}_{2^n}, a \neq 0\} = \{\Delta_{F'}(a,b) : a, b \in \mathbb{F}_{2^n}, a \neq 0\}$ [4], [8];
- if F is APN then F' is APN too;
- $\mathcal{NL}_F = \mathcal{NL}_{F'}$ [12];
- if F is AB then F' is AB too.

III. Changing points in general

A construction in which an (n,n)-function G is obtained by changing a single value of a given (n,n)-function F is investigated in [6]. More precisely, given a function F over $\mathbb{F}_{2^n}$, the construction is performed by defining a function G over the same field by

$$G(x) = \begin{cases} F(x) & x \neq u \\ v & x = u \end{cases}$$

for some fixed elements $u, v \in \mathbb{F}_{2^n}$. Since G can be written as $G(x) = F(x) + (F(u) + v)(1 + (x+u)^{2^n - 1})$, it is easy to see that the algebraic degree of at least one of F and G must be equal to n; furthermore, any function G of algebraic degree n can be written in this form for some F of algebraic degree less than n. Indeed, the motivation behind the study of this construction is the unresolved questions of whether APN functions of algebraic degree n can exist over $\mathbb{F}_{2^n}$; the authors investigate the possibility of obtaining an APN function G using the construction, with particular attention being paid to the case when F is itself APN. Two main characterizations of the APN-ness of G are obtained in [6], one involving the Walsh coefficients of F, and one based on the properties of the derivatives $D_a F$. These characterizations are then applied in order to conclude that no function G obtained by such a one-point change from a given F which is a power, plateaued, quadratic or almost bent function can be APN, except possibly for $n \leq 2$ in the case of plateaued functions. For instance, $F(x) = x$ is plateaued and $G(x) = F(x) + x^{2^n - 1} = x^3 + x$ is APN over $\mathbb{F}_{2^2}$; in the case of power, quadratic and almost bent functions, we only have trivial examples over $\mathbb{F}_2$, e.g. when F is the identity function $F(x) = x$ and G is the constant zero function $G(x) = 0$. A number of additional non-existence results are also shown, which support the conjecture that no APN function of algebraic degree n may exist over $\mathbb{F}_{2^n}$; nonetheless, the question in general remains open.

Some properties of the special case when the values of F at two given points are swapped have previously been investigated in [24], and the general case of changing the values of F at two points has been considered in [17]. The authors of the former article have generalized their method to changing points lying on a cycle [18], and have been able to construct involutions over $\mathbb{F}_{2^n}$ using this method [15]. In [17], two main characterizations of the APN-ness of a new function G obtained by modifying two values of a given F are obtained, one in terms of the power moments of the Walsh transform, and one in terms of the differential properties F. We observed that if F and G are at distance two, then at most one of F and G can be AB, and at most one of them can be plateaued; furthermore, if the algebraic degree of F is less than $n-1$, then G can be neither AB nor plateaued for any $n \geq 3$. In the case of swapping the values of a function at 0 and 1, we obtained a sufficient condition for disproving the APN-ness of G by computing a lower bound on the sum $\sum_{y \in \mathbb{F}_{2^n}} \Delta_F(y, F(y)+1) + \Delta_F(y+1, F(y))$. We also showed how to compute a lower bound on this quantity in the case of power functions by finding multiple solutions to the equation $F(x) + F(a+x) + F(a) = 1$ when F is a power function.

The idea of investigating pairs of functions at a small distance to one another is interesting per se, and the aforementioned construction can be naturally extended so that the value of F is changed at more than one point. In the following, we investigate whether, and under what conditions, it is possible to obtain an APN function by changing the values of *multiple* points in a given APN function F. More precisely, given K distinct elements $u_1, u_2, \ldots, u_K$ from $\mathbb{F}_{2^n}$ (referred to as *points*) and K arbitrary elements $v_1, v_2, \ldots, v_K$ from $\mathbb{F}_{2^n}$ (referred to as *shifts*), we are interested in the APN-ness of the function

$$G(x) = F(x) + \sum_{i=1}^{K} 1_{u_i}(x) v_i = F(x) + \sum_{i=1}^{K} (1 + (x+u_i)^{2^n-1}) v_i \tag{2}$$

whose value coincides with the value of F on all points $x \notin \{u_1, u_2, \ldots, u_K\}$ and satisfies $G(u_i) = F(u_i) + v_i$ for $i \in \{1, 2, \ldots, K\}$.

In order to facilitate the following discussion, we introduce some notation related to the construction. We denote by U the set $U = \{u_1, u_2, \ldots, u_K\}$ of points whose value will change. For a given element $a \in \mathbb{F}_{2^n}$, we denote by $a + U$ the set $\{a + u : u \in U\}$. For any given natural number n, we write $[n] = \{1, 2, \ldots, n\}$; in particular, $[K]$ is the set of indices of the points from U. For any given $a \in \mathbb{F}_{2^n}^*$ we define the set $U_a = \{u \in U : a + u \in U\}$, and $\overline{U_a} = U \setminus U_a$. In addition, we define a function p_a on the indices $\{i \in [K] : u_i \in U_a\}$ by the prescription $p_a(i) = j$ where j is such that $u_i + a = u_j$. Since the definition of an APN function is given in terms of differential equations, a natural way to investigate the properties of G is to examine the derivatives $D_a G$ and their relation to the derivatives $D_a F$ of F. From the definition of G in (2) we can immediately see that for any $a \in \mathbb{F}_{2^n}^*$, the derivative $D_a G$ takes the form

$$D_a G(x) = D_a F(x) + \sum_{i=1}^{K} 1_{u_i, a+u_i}(x) v_i. \tag{3}$$

Although all the points u_i are assumed distinct, it is possible that for some $i \neq j$ we have $a + u_i = u_j$ and the sets $\{u_i, a + u_i\}$ and $\{u_j, a + u_j\}$ will coincide. This can be seen more easily if (3) is written in the form

$$D_a G(x) = D_a F(x) + \sum_{i \in U_a : i < p_a(i)} 1_{u_i, u_{p_a(i)}}(x)(v_i + v_{p_a(i)}) + \sum_{i \in \overline{U_a}} 1_{u_i, a+u_i}(x) v_i. \tag{4}$$

A characterization of the conditions under which G is APN can be derived immediately from (3) and the definition of an APN function by examining under what conditions a triple of elements $(a, x, y) \in \mathbb{F}_{2^n}^3$ with $a \neq 0$, $D_a G(x) = D_a G(y)$ and $x + y \notin \{0, a\}$ may exist.

Proposition 2. Let $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, let $u_1, u_2, \ldots, u_K$ be K distinct points from $\mathbb{F}_{2^n}$ and let $v_1, v_2, \ldots, v_K$ be K arbitrary elements from $\mathbb{F}_{2^n}$. Then the function G defined by (2) is APN if and only if all of the following conditions are satisfied for every derivative direction $a \in \mathbb{F}_{2^n}^*$:

(i) $D_a F$ is 2-to-1 on $\mathbb{F}_{2^n} \setminus (U \cup a + U)$;
(ii) $D_a F(u_i) + D_a F(u_j) \neq v_i + v_j + v_{p_a(i)} + v_{p_a(j)}$ for $u_i, u_j \in U_a$ unless $u_i = u_j$ or $u_i + u_j = a$;
(iii) $D_a F(u_i) + D_a F(u_j) \neq v_i + v_j + v_{p_a(i)}$ for $u_i \in U_a, u_j \in \overline{U_a}$;
(iv) $D_a F(u_i) + D_a F(u_j) \neq v_i + v_j$ for $u_i, u_j \in \overline{U_a}$ unless $u_i = u_j$;
(v) $D_a F(u_i) + D_a F(x) \neq v_i + v_{p_a(i)}$ for $u_i \in U_a, x \notin (U \cup a + U)$;
(vi) $D_a F(u_i) + D_a F(x) \neq v_i$ for $u_i \in \overline{U_a}, x \notin (U \cup a + U)$.

Proof. Recall that G is APN if and only if there does not exist a triple $(a, \bar{x}, \bar{y}) \in \mathbb{F}_{2^n}^3$ such that $D_a G(\bar{x}) = D_a G(\bar{y})$ with $a \neq 0$ and $\bar{x} \notin \{\bar{y}, a + \bar{y}\}$. Suppose that such a triple does exist. We will now go through several possible cases, depending on whether $\bar{x}$ and $\bar{y}$ are in $(U \cup a + U)$ or not. In the first case, we will assume that neither $\bar{x}$ nor $\bar{y}$ is in $(U \cup a + U)$; in the second case, we will assume that both $\bar{x}$ and $\bar{y}$ are in $(U \cup a + U)$; and in the third case, we will assume that precisely one of $\bar{x}$ and $\bar{y}$ is in $(U \cup a + U)$:

1) If neither $\bar{x}$ nor $\bar{y}$ belong to $(U \cup a + U)$, then $D_a G(\bar{x}) = D_a F(\bar{x})$ and $D_a G(\bar{y}) = D_a F(\bar{y})$ so that $D_a G(\bar{x}) = D_a G(\bar{y})$ implies $D_a F(\bar{x}) = D_a F(\bar{y})$. Thus $D_a F$ cannot be 2-to-1 over $\mathbb{F}_{2^n} \setminus (U \cup a + U)$. Conversely, if $D_a F$ is 2-to-1 over $\mathbb{F}_{2^n} \setminus (U \cup a + U)$, this guarantees that no such triple can exist with $\bar{x}, \bar{y} \notin (U \cup a + U)$. This leads to the first condition.
2) If both $\bar{x}$ and $\bar{y}$ are points from U or $a + U$, say $\bar{x} = u_i$ and $\bar{y} = u_j$, then we have $D_a G(u_i) = D_a G(u_j)$. We now examine three cases depending on whether one, both or none of u_i and u_j are in U_a:
   a) If $D_a G(u_i) = D_a G(u_j)$ with $u_i, u_j \in U_a$, then we have $D_a F(u_i) + v_i + v_{p_a(i)} = D_a F(u_j) + v_j + v_{p_a(j)}$ from the definition of G (2). If G is APN, this is possible only if $u_i = u_j$ or $u_i = a + u_j$, which leads to the second condition.
   b) If say u_i is in U_a but u_j is not, then $D_a G(u_i) = D_a G(u_j)$ becomes $D_a F(u_i) + D_a F(u_j) = v_i + v_j + v_{p_a(i)}$. Note that we can have neither $u_i = u_j$, nor $u_i + a = u_j$ since u_i is in U_a and u_j is in its complement. This leads to the third condition.
   c) If neither u_i nor u_j is in U_a, then $D_a G(u_i) = D_a G(u_j)$ becomes $D_a F(u_i) + D_a F(u_j) = v_i + v_j$; this can occur if $u_i = u_j$, but $u_i = a + u_j$ is impossible due to $u_j \notin U$. This gives the fourth condition.
3) In the remaining case, we assume that we have $\bar{x} = u_i$ (or $\bar{x} = a + u_i$) but $\bar{y} \notin (U \cup a + U)$, so that we have $D_a G(u_i) = D_a F(\bar{y})$. We examine two sub-cases:
   a) If $D_a G(u_i) = D_a G(\bar{y})$ with $u_i \in U_a$, then $D_a F(u_i) + D_a F(\bar{y}) = v_i + v_{p_a(i)}$. Since both u_i and $u_i + a$ are in U, we cannot have $u_i \in \{y, a + y\}$. This gives the fifth condition.
   b) If, conversely, $D_a G(u_i) = D_a G(\bar{y})$ but $u_i \in \overline{U_a}$, then we have $D_a F(u_i) + D_a F(\bar{y}) = v_i$. As before, we cannot have $u_i \in \{\bar{y}, a + \bar{y}\}$. This gives the sixth and final condition.

The above conditions are clearly necessary for G to be APN, and they are also sufficient since if we have $D_a G(\bar{x}) = D_a G(\bar{y})$ then one of these conditions implies $\bar{x} = \bar{y}$ or $\bar{x} = a + \bar{y}$. □

The following observation shows how condition (vi) of Proposition 2 can be equivalently expressed in terms of the shifted derivatives of F. This is slightly more intuitive in the sense that it allows us to consider the image of a single shifted derivative (instead of the sum of two derivatives as in the original formulation) and is used throughout the next section.

Observation 1. Assume the same notation as in Proposition 2. If G is APN, then for any $a \in \mathbb{F}_{2^n}^*$ for which there exists an $i \in [K]$ such that $D_a^{u_i} F$ maps to $F(u_i) + v_i$ and $a + u_i \notin U$ we must have

$$D_a^{u_i} F(u_j) + F(u_i) = v_i,$$

for some $i \neq j \in [K]$.

Characterizing the APN-ness of G is difficult in the general case due to the large number of choices for the points $u_1, u_2, \dots, u_K$ and shifts $v_1, v_2, \dots, v_K$. For this reason, in the following sections we concentrate on various simplifications of this problem, e.g. by assuming that the points $u_1, u_2, \dots, u_K$ or the number K are fixed.

IV. The case of fixed $u_1, u_2, \dots, u_K$

If we fix the set U of points to change, we can use Observation 1 to dramatically reduce the number of potential candidate values for the shifts $v_1, v_2, \dots, v_K$. Besides filtering out impossible candidates for the shifts v_i, this allows us to obtain a lower bound on the distance between a given APN function F and its closest APN neighbor. This lower bound is given in terms of the number of shifted derivatives of F that map to the elements of $\mathbb{F}_{2^n}$. This quantity can be computed efficiently in practice and can be used to bound from below the number of points K that need to be changed in order to obtain an APN function G. Finally, we observe that this lower bound is invariant under CCZ-equivalence.

A. Filtering out shift candidates

We can immediately apply Observation 1 in practice by fixing some function F over $\mathbb{F}_{2^n}$ along with K points $u_1, u_2, \dots, u_K$ and then, for every $i \in [K]$, making a list of all values $\bar{v} \in \mathbb{F}_{2^n}$ for which setting $v_i = \bar{v}$ violates the necessary condition from Proposition 2. Then only values v_i which are not in this list have to be examined, and their number is typically much smaller than the number 2^n of all possible values. In many cases, no values at all are left for some v_i, which then immediately indicates that no APN functions can be obtained by shifting the points in U.

A more precise description of this procedure is given as Algorithm 1.

Algorithm 1: Reducing the domains of v_i using Observation 1

Data: A function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ and a set of K distinct points $U = \{u_1, u_2, \dots, u_K\} \subseteq \mathbb{F}_{2^n}$.
Result: A domain $D_i \subseteq \mathbb{F}_{2^n}$ for every v_i such that if $G(x)$ is APN, then $v_i \in D_i$ for every $i \in [K]$.
begin
  for *every* $i \in [K]$ **do**
    set $D_i \leftarrow \mathbb{F}_{2^n}$
    compute $A \leftarrow \{ D_a^{u_i} F(x) + F(u_i) : x, a \in \mathbb{F}_{2^n}, a \neq 0, a + u_i \notin U, x \notin (U \cup a + U) \}$
    update $D_i \leftarrow D_i \setminus A$

As already mentioned, the efficiency of this method is particularly prominent in cases when the points $u_1, u_2, \dots, u_K$ cannot be shifted into an APN function (in the sense that G is never APN regardless of the choice of $v_1, v_2, \dots, v_K$). For example, given the function $F(x) = x^3$ over $\mathbb{F}_{2^5}$ and the set of points $U = \{\alpha^i : i \in \{0\} \cup [5]\}$, where α is a primitive element of $\mathbb{F}_{2^5}$, checking every combination of shifts $(v_1, \dots, v_K) \in \mathbb{F}_{2^5}^6$ using an exhaustive search (that is, generating G as defined in (2) and testing whether it is APN for every such combination of shifts) is estimated to take about 75 hours; using the filtering approach described above, however, we can conclude that no APN function G can be obtained by any combination of shifts after only about 0.140 seconds of computation. These experiments were performed on our department server, with the search procedures implemented in the *Magma* programming language.

On the contrary, in some situations (especially when the set of points U can be shifted into an APN function) the filtering procedure may leave rather large domains for the shift candidates, which necessitates long computations. As two contrasting examples, we examine the function x^3 over $\mathbb{F}_{2^5}$ and over $\mathbb{F}_{2^6}$. In the case of $\mathbb{F}_{2^5}$, taking the set U of the eight points generated (in the sense of additive closure) by $\{\alpha^i : i \in \{0\} \cup [2]\}$ leaves the singleton domain $\{\alpha^{25}\}$ for all v_i; indeed, the function G obtained by shifting every point from U by α^{25} is APN and is CCZ-equivalent to x^5. However, when we take $F(x) = x^3$ over $\mathbb{F}_{2^6}$ with U being generated by $\{1, \beta, \beta^4, \beta^{21}\}$ (with β primitive in $\mathbb{F}_{2^n}$), the domains for each v_i after filtering become $D = \{\beta^7, \beta^{14}, \beta^{28}, \beta^{35}, \beta^{49}, \beta^{56}\}$. Taking $v_1 = v_2 = \cdots = v_{16} = v$ for any $v \in D$ then yields an APN function G that is CCZ-equivalent to $x^6 + x^9 + \beta^7 x^{48}$. Conversely, if at least two different values are selected for the shifts, the resulting function is not APN; thus, there are only $|D| = 6$ possible shift combinations that lead to an APN function, but 6^{16} potential combinations that are left after filtering and need to be "manually" checked. Therefore, although our method reduces the size of the domains from $2^6 = 64$ to just 6, the resulting search space is still quite large and requires a significant amount of time in order to be completely explored.

However, additional restrictions may be imposed on the values of v_i by applying conditions (i)-(v) from Proposition 2 which allow the search to be performed more efficiently. More precisely, condition (iv) allows us to remove pairs, condition (iii) allows us to remove triples and condition (ii) allows us to remove quadruples of incompatible elements from the domains. Condition (i) depends entirely on the function F and the set U and can be used to reject a given set U entirely, although it cannot be used for filtering the domains.

These conditions do not allow us to remove any values from the domains of v_i directly, but they do make it possible to restrict some domains after a first few initial choices. For example, having selected a concrete value $\bar{v}_i$ for v_i from its domain, we can for all $j \neq i$, remove values $\bar{v}_j$ from the domain of v_j for which condition (iv) is violated. It is worth noting that this is the most useful of the three conditions given above in the case that the number of points U is relatively small, since it encompasses the greatest number of derivative directions; as K increases, the latter two conditions become more useful. In any case, ensuring that all the conditions from Proposition 2 are satisfied is sufficient to ensure that G is APN.

Coming back to the example of $F(x) = x^3$ over $\mathbb{F}_{2^6}$ discussed above, we can see how much this improves the search efficiency: evaluating all combinations of shifts from the domains (without any filtering) would require approximately 110 years; applying conditions (i)-(iv) from Proposition 2 as described, however, finds all six possibilities in about two seconds.

B. Lower bound on the distance between APN functions

Note that in the statement of Observation 1, we assume that the resulting function G is APN but we do not make any assumptions about F. If, in addition to the hypothesis of the theorem, we assume that F is itself APN, we can obtain the following corollary which gives a lower bound on the Hamming distance between a given APN function and its nearest APN "neighbor".

Corollary 1. Let F and G be as in the statement of Observation 1 with $v_i \neq 0$ for $i \in [K]$, and assume, in addition, that F is APN;

consider some fixed $i \in [K]$. Then no more than $3(K-1)$ derivatives of the form $D_a^{u_i} F$ map to $G(u_i)$.

Proof. First, consider all derivative directions $a \in \mathbb{F}_{2^n}^*$ with $a + u_i \notin U$. By Observation 1 we must have

$$D_a^{u_i} F(u_j) = G(u_i)$$

for some $j \neq i$ if $D_a^{u_i} F$ maps to $G(u_i)$. We now determine for how many $a \in \mathbb{F}_{2^n}$ we may have $D_a^{u_i} F(u_j) = G(u_i)$ for fixed i and j. Suppose that we have both $D_a^{u_i} F(u_j) = G(u_i)$ and $D_{a'}^{u_i} F(u_j) = G(u_i)$ for some $a \neq a'$. Then $D_a^{u_i} F(u_j) = D_{a'}^{u_i} F(u_j)$ can be rewritten as $F(u_j) + F(a+u_j) + F(a+u_i) = F(u_j) + F(a'+u_j) + F(a'+u_i)$ so that we have $D_{u_i+u_j} F(a+u_i) = D_{u_i+u_j} F(a'+u_i)$.

Since i and j (and therefore u_i and u_j) are fixed and since F is APN, this implies either $a = a'$ or $a + a' = u_i + u_j$. In other words, at most two distinct shifted derivatives may map u_j to $G(u_i)$.

Now suppose that i is fixed and j ranges over $[K]$. Since we consider only $j \neq i$ and since there are K indices in total, there are $(K-1)$ choices for j for any fixed i. For each such j, there are at most two shifted derivatives $D_a^{u_i} F$ mapping u_j to $G(u_i)$. Therefore, at most $2(K-1)$ shifted derivatives may take $G(u_i)$ as value when $a + u_i \notin U$.

We now consider the derivative directions $a \in \mathbb{F}_{2^n}$ for which $a + u_i \in U$. There are precisely K such directions a, viz. $u_1 + u_i$, $u_2 + u_i$, $\ldots$, $u_K + u_i$. Furthermore, $D_0^{u_i} F$ cannot map to $G(u_i)$ unless $v_i = 0$, so that there are at most $(K-1)$ derivatives of this type which may map to $G(u_i)$.

Thus, in total, there can be no more than $2(K-1) + (K-1) = 3(K-1)$ derivative directions a for which $D_a^{u_i} F$ maps to $G(u_i)$. □

Note that in the proof above, the number of derivative directions a (with $a + u_i \notin U$) such that $D_a^{u_i} F(u_j) = G(u_i)$ for some fixed i and j is limited to two because F is assumed to be APN. If we take F to be differentially δ-uniform instead, the upper bound on the number of derivatives $D_a^{u_i} F$ mapping to $G(u_i)$ will be $(\delta+1)(K-1)$.

Corollary 1 can now be used to compute a lower bound on the distance between a given F and its nearest APN "neighbor". In order to facilitate the following discussion, we introduce some notation related to the shifted derivatives. In particular, we define $\Pi_F^\beta(b)$ to be the set of derivative directions a for which $D_a^\beta F$ maps to b, i.e.

$$\Pi_F^\beta(b) = \{a \in \mathbb{F}_{2^n} : b \in H_a^\beta F\} = \{a \in \mathbb{F}_{2^n} : (\exists x \in \mathbb{F}_{2^n})(D_a^\beta F(x) = b)\}.$$

By Corollary 1, we need to count the numbers $|\Pi_F^{u_i}(G(u_i))|$ for $i \in [K]$ and ensure that none of them is greater than $3(K-1)$. The minimum value of $|\Pi_F^\beta(b)|$ through all possible values of β and b is certainly a lower bound on $\min_{i \in [K]} |\Pi_F^{u_i}(G(u_i))|$; if this minimum value is greater than $3(K-1)$ for some given K, then no function G within distance K of F can be APN.

Thus, we can apply the lower bound from Corollary 1 by computing the minimum value of $|\Pi_F^\beta(b)|$ through all $\beta, b \in \mathbb{F}_{2^n}$. In certain cases, such as for quadratic functions (see Proposition 5 below), it suffices to consider a fixed value of β and to only go through all $b \in \mathbb{F}_{2^n}$. For this reason, we define the set Π_F^β as the spectrum of the values of $|\Pi_F^\beta(b)|$ for a fixed shift β, i.e.

$$\Pi_F^\beta = \{|\Pi_F^\beta(b)| : b \in \mathbb{F}_{2^n}\}$$

and Π_F as the spectrum of $|\Pi_F^\beta(b)|$ for all shifts β and all values b:

$$\Pi_F = \bigcup_{\beta \in \mathbb{F}_{2^n}} \Pi_F^\beta = \{|\Pi_F^\beta(b)| : \beta, b \in \mathbb{F}_{2^n}\}.$$

For convenience, we also denote by m_F the minimal element of Π_F, i.e. $m_F = \min\{|\Pi_F^\beta(b)| : \beta, b \in \mathbb{F}_{2^n}\}$. The lower bound on the distance between APN functions can now be stated as follows.

Corollary 2. Let F be an APN function over $\mathbb{F}_{2^n}$ and let m_F be the number

$$m_F = \min \Pi_F = \min_{b, \beta \in \mathbb{F}_{2^n}} |\Pi_F^\beta(b)| = \min_{b, \beta \in \mathbb{F}_{2^n}} |\{a \in \mathbb{F}_{2^n} : (\exists x \in \mathbb{F}_{2^n})(D_a^\beta F(x) = b)\}|.$$

Then for any APN function $G \neq F$ over $\mathbb{F}_{2^n}$, the Hamming distance $d(F,G)$ between F and G satisfies

$$d(F,G) \geq \left\lceil \frac{m_F}{3} \right\rceil + 1. \tag{5}$$

Proof. By Corollary 1, if F and G are APN functions at distance K of one another, than no more than $3(K-1)$ shifted derivatives $D_a^{u_i} F$ may map to $G(u_i)$ for any fixed $i \in [K]$. For a fixed i, this quantity can be written as $|\{a \in \mathbb{F}_{2^n} : (\exists x \in \mathbb{F}_{2^n})(D_a^{u_i}(x) = G(u_i))\}|$. If we now go through all possible values of $i \in [K]$, we get that

$$\min_{i \in [K]} |\{a \in \mathbb{F}_{2^n} : (\exists x \in \mathbb{F}_{2^n})(D_a^{u_i}(x) = G(u_i))\}| \leq 3(K-1).$$

Deriving a lower bound on K from this expression, however, would require knowledge of $D_a^{u_i}(x)$ and $G(u_i)$ for each $i \in [K]$. However, since u_i and $G(u_i)$ are elements of the finite field $\mathbb{F}_{2^n}$, going through all possible choices β for u_i and all possible choices b for $G(u_i)$, we clearly have

$$\min_{b,\beta} |\{a \in \mathbb{F}_{2^n} : (\exists x \in \mathbb{F}_{2^n})(D_a^\beta(x) = b)\}| \leq \min_{i \in [K]} |\{a \in \mathbb{F}_{2^n} : (\exists x \in \mathbb{F}_{2^n})(D_a^{u_i}(x) = G(u_i))\}| \leq 3(K-1).$$

If we denote the left-most quantity by m_F, as in the statement of the Corollary, we then have

$$m_F \leq 3(K-1)$$

which immediately implies the lower bound. □

C. Invariance Properties

As discussed above, the lower bound on the Hamming distance between a given APN function F and its closest APN "neighbor" is given in terms of the number m_F which in turn can be expressed via the sets $\Pi_F^\beta(b)$, Π_F^β and Π_F. It is therefore interesting to observe that the set Π_F is invariant under CCZ-equivalence, as shown in the following proposition. This then makes the lower bound obtained via Corollary 2 for some given function F valid for all members of its CCZ-equivalence class.

Proposition 3. Suppose F is APN and is CCZ-equivalent to F' via the affine permutation $\mathcal{L} = (L_1, L_2)$ of $\mathbb{F}_{2^n}^2$. Then $\Pi_F^\beta(t) = \Pi_{F'}^{L_1(\beta,t)}(L_2(\beta,t))$ for any $\beta, t \in \mathbb{F}_{2^n}$. Consequently, the set Π_F is invariant under CCZ-equivalence.

Proof. To show the first part of the statement, define $F_1(x) = L_1(x, F(x))$ and $F_2(x) = L_2(x, F(x))$ as in (1); then F_1 is a permutation and $F' = F_2 \circ F_1^{-1}$.

If we consider the set of all pairs (a,x) such that $D_a^\beta F(x) = t$, we can obtain using the affinity of $\mathcal{L}$:

$$\begin{aligned}
&|\{(a,x) \in \mathbb{F}_{2^n} : F(x) + F(a+x) + F(a+\beta) = t\}| = \\
&|\{(x,y,z) \in \mathbb{F}_{2^n}^3 : (x, F(x)) + (y, F(y)) + (z, F(z)) = (\beta, t)\}| = \\
&|\{(x,y,z) : (F_1(x), F_2(x)) + (F_1(y), F_2(y)) + \\
&\quad (F_1(z), F_2(z)) = \mathcal{L}(\beta, t)\}| = \\
&|\{(x,y,z) : (x, F'(x)) + (y, F'(y)) + \\
&\quad (z, F'(z)) = (L_1(\beta,t), L_2(\beta,t))\}| = \\
&|\{(a,x) : F'(x) + F'(a+x) + F'(a + L_1(\beta,t)) = L_2(\beta,t)\}|.
\end{aligned}$$

In the third step we use the fact that F_1 is a permutation and go through all triples $(F_1^{-1}(x), F_1^{-1}(y), F_1^{-1}(z))$ instead of (x,y,z).

Now, since $|\Pi_F^\beta(t)|$ counts the number of derivative directions a for which $D_a^\beta F$ maps to t, and since all (shifted) derivatives of F and F' are 2-to-1 due to F and F' being APN, we have

$$\begin{aligned}
2|\Pi_F^\beta(t)| = |\{(a,x) \in \mathbb{F}_{2^n} : F(x) + F(a+x) + F(a+\beta) = t\}| = \\
|\{(a,x) : F'(x) + F'(a+x) + F'(a + L_1(\beta,t)) = \\
L_2(\beta,t)\}| = 2|\Pi_{F'}^{L_1(\beta,t)}(L_2(\beta,t))|. \quad (6)
\end{aligned}$$

The invariance of Π_F then follows from the fact that $\mathcal{L} = (L_1, L_2)$ is a permutation and $\Pi_F = \{|\Pi_F^\beta(t)| : \beta, t \in \mathbb{F}_{2^n}\}$, so that when computing Π_F we go through all possible pairs (β, t). □

As EA-equivalence is a special case of CCZ-equivalence, it is evident that EA-equivalence leaves the set Π_F invariant as well. Under EA-equivalence, however, a stronger invariance holds.

Proposition 4. For any fixed $\beta \in \mathbb{F}_{2^n}$, if F' and F are EA-equivalent APN functions via $F' = A_1 \circ F \circ A_2 + A$, where A_1, A_2 and A are affine and A_1, A_2 are bijective, we have

$$(\forall t \in \mathbb{F}_{2^n})(|\Pi_{F'}^\beta(t)| = |\Pi_F^{A_2(\beta)}(A_1^{-1}(t + A(\beta)))|).$$

Consequently, $\Pi_{F'}^\beta = \Pi_F^{A_2(\beta)}$.

Proof. We have, thanks to F and F' being APN and their derivatives being 2-to-1 functions,

$$\begin{aligned}
&2|\Pi_{F'}^\beta(t)| = \\
&|\{(a,x) \in \mathbb{F}_{2^n}^2 : F'(x) + F'(a+x) + F'(a+\beta) = t\}| = \\
&|\{(a,x) : A_1(F(A_2(x))) + A_1(F(A_2(a+x))) + \\
&A_1(F(A_2(a+\beta))) + A(x) + A(a+x) + A(a+\beta) = t\}| = \\
&|\{(a,x) : A_1(F(A_2(x)) + F(A_2(a)) + \\
&\quad F(A_2(a+x+\beta))) = t + A(\beta)\}| = \\
&|\{(a,x) : A_1(F(x) + F(a) + F(a+x+A_2(\beta))) = t + A(\beta)\}| = \\
&|\{(a,x) : F(x) + F(a) + F(a+x+A_2(\beta)) = \\
&A_1^{-1}(t + A(\beta))\}| = 2|\Pi_F^{A_2(\beta)}(A_1^{-1}(t + A(\beta)))|.
\end{aligned} \quad (7)$$

In the second step we use that for any affine function A we have $A(x+y+z) = A(x) + A(y) + A(z)$ for any x, y, z, and also count through $(x, a+x)$ instead of (x, a). In the third step we use the fact that A_2 is a permutation and count through all pairs $(A_2(a), A_2(x))$ instead of (a,x); then $A_2(x)$ becomes x, $A_2(a)$ becomes a and $A_2(x+a+\beta) = A_2(x) + A_2(a) + A_2(\beta)$ becomes $x + a + A_2(\beta)$.

Then clearly

$$\begin{aligned}
\Pi_{F'}^\beta &= \{|\Pi_{F'}^\beta(t)| : t \in \mathbb{F}_{2^n}\} = \\
&\{|\Pi_F^{A_2(\beta)}(A_1^{-1}(t + A(\beta)))| : t \in \mathbb{F}_{2^n}\} = \\
&\{|\Pi_F^{A_2(\beta)}(t)| : t \in \mathbb{F}_{2^n}\} = \Pi_F^{A_2(\beta)},
\end{aligned}$$

thereby concluding the proof. □

D. *The case of quadratic functions*

For a quadratic function F, the set Π_F^β does not depend on the choice of β, which greatly reduces the amount of computation needed to calculate m_F.

Proposition 5. Let F be a quadratic (n,n)-function. Then $\Pi_F^\beta = \Pi_F^{\beta'}$ for any $\beta, \beta' \in \mathbb{F}_{2^n}$.

Proof. Since F is quadratic, its derivatives $D_a F$ for any $a \neq 0$ are affine functions, i.e. they satisfy

$$D_a F(x) + D_a F(y) = D_a F(x+y) + D_a F(0)$$

for any $x, y \in \mathbb{F}_{2^n}$. We thus have

$$\begin{aligned}
D_a^\beta F(x) + D_a^0 F(x+\beta) &= \\
D_a F(x) + D_a F(x+\beta) + F(a+\beta) + F(a) &= \\
D_a F(\beta) + D_a F(0) + F(a+\beta) + F(a) &= \quad (8)\\
F(\beta) + F(a+\beta) + F(0) + F(a) + F(a+\beta) + F(a) &= \\
F(\beta) + F(0)&
\end{aligned}$$

so that we have

$$D_a^\beta F(x) = D_a^0 F(x+\beta) + s$$

for some constant s which depends only on F and β.

We have then

$$|\Pi_F^\beta(t)| = |\Pi_F^0(t+s)|$$

so that, indeed,

$$\Pi_F^\beta = \{|\Pi_F^\beta(t)| : t \in \mathbb{F}_{2^n}\} = \{|\Pi_F^0(t+s)| : t \in \mathbb{F}_{2^n}\} = \Pi_F^0$$

as claimed. □

E. *Examples and computation results*

In some cases, the value m_F can be computed mathematically. As an example, we consider the function $F(x) = x^3$ over the finite field $\mathbb{F}_{2^n}$. We derive an exact formula for the size of $\Pi_F^\beta(b)$, which allows us to express Π_F^β and, consequently, m_F as a function of the dimension n. From this we can then immediately derive a lower bound on the distance between x^3 and the closest APN function. Note that since x^3 is quadratic, by Proposition 5 we have that $m_F = \min \Pi_F^\beta$ for an arbitrary $\beta \in \mathbb{F}_{2^n}$.

Proposition 6. Let $F(x) = x^3$ be over $\mathbb{F}_{2^n}$ and let $b, \beta \in \mathbb{F}_{2^n}$ be arbitrary. Then

$$|\Pi_F^\beta(b)| = \begin{cases}
2^n - 1 & b = \beta^3; \\
2^{n-1} - 1 & b \neq \beta^3, n \text{ odd}; \\
2^{n-1} + 2^{n/2} - 1 & b \neq \beta^3, b + \beta^3 \text{ is a cube}, \\
& \quad n \text{ even}, n/2 \text{ odd}; \\
2^{n-1} - 2^{n/2-1} - 1 & b \neq \beta^3, b + \beta^3 \text{ is not a cube}, \\
& \quad n \text{ even}, n/2 \text{ odd}; \\
2^{n-1} - 2^{n/2} - 1 & b \neq \beta^3, b + \beta^3 \text{ is a cube}, \\
& \quad n \text{ even}, n/2 \text{ even}; \\
2^{n-1} + 2^{n/2-1} - 1 & b \neq \beta^3, b + \beta^3 \text{ is not a cube}, \\
& \quad n \text{ even}, n/2 \text{ even}.
\end{cases} \quad (9)$$

The value $\min_{b \in \mathbb{F}_{2^n}} |\Pi_F^\beta(b)|$ is then equal to

$$m_F = \min \Pi_F^\beta = \begin{cases}
2^{n-1} - 1 & n \text{ is odd}; \\
2^{n-1} - 2^{n/2-1} - 1 & n \text{ is even}, n/2 \text{ is odd}; \\
2^{n-1} - 2^{n/2} - 1 & n \text{ is even}, n/2 \text{ is even};
\end{cases} \quad (10)$$

and the lower bound on the distance to the closest APN function G can be explicitly written as

$$d(F,G) \geq \begin{cases} \frac{2^{n-1}+2}{3} & n \text{ is odd;} \\ \frac{2^{n-1}-2^{n/2-1}+2}{3} & n \text{ is even}, n/2 \text{ is odd;} \\ \frac{2^{n-1}-2^{n/2}+2}{3} & n \text{ is even}, n/2 \text{ is even.} \end{cases} \quad (11)$$

Proof. The shifted derivative $D_a^\beta F$ of the Gold function $F(x) = x^3$ takes the form

$$D_a^\beta F(x) = x^3 + (x+a)^3 + (a+\beta)^3 = a^2(x+\beta) + a(x+\beta)^2 + \beta^3$$

for any $a, \beta \in \mathbb{F}_{2^n}$.

For convenience, we introduce the "equality indicator" $I(A,B)$, where A and B are some arbitrary expressions, defined as

$$I(A,B) = \begin{cases} 1 & A = B \\ 0 & A \neq B. \end{cases}$$

Recall that the value of $|\Pi_F^\beta(b)|$ is the number of derivative directions $a \in \mathbb{F}_{2^n}$ for which $D_a^\beta F$ maps to b. Since F is APN, $|\Pi_F^\beta(b)|$ can be expressed as

$$\begin{aligned} &|\Pi_F^\beta(b)| = \\ &\frac{1}{2}|\{(a,x) \in \mathbb{F}_{2^n}^* \times \mathbb{F}_{2^n} : a^2(x+\beta)+a(x+\beta)^2+\beta^3 = b\}| + I(b,\beta^3) = \\ &\frac{1}{2}|\{(a,x) \in \mathbb{F}_{2^n}^* \times \mathbb{F}_{2^n} : a^2x + ax^2 = b + \beta^3\}| + I(b,\beta^3) \quad (12) \end{aligned}$$

by substituting $x + \beta$ for x.

Note that for $a = 0$, (12) becomes $b = \beta^3$, so that the number of solutions x is $2^n I(\beta, b^3)$; however, all of these solutions correspond to the same derivative direction $a = 0$. For any fixed $a \neq 0$, we can divide both sides of the equation

$$a^2(x+\beta) + a(x+\beta)^2 = b + \beta^3$$

by a^3 and substitute $ax + \beta$ for x in order to obtain

$$x^2 + x = \frac{b+\beta^3}{a^3}. \quad (13)$$

Since x^2+x is linear with roots 0 and 1, it is a 2-to-1 mapping, and its image set over $\mathbb{F}_{2^n}$ is precisely the set of all elements with zero trace. Therefore, for a fixed $a \neq 0$, equation (13) has two solutions if $\mathrm{Tr}_n\left(\frac{b+\beta^3}{a^3}\right) = 0$, and no solutions otherwise. Consequently, if we define the function $h : \mathbb{F}_{2^n} \to \mathbb{F}_2$ as

$$h(a) = \begin{cases} \mathrm{Tr}_n\left(\frac{b+\beta^3}{a^3}\right) + 1 & a \neq 0 \\ 0 & a = 0, \end{cases}$$

we can express $|\Pi_F^\beta(b)|$ as

$$|\Pi_F^\beta(b)| = I(b, \beta^3) + \mathrm{wt}(h) \quad (14)$$

where $\mathrm{wt}(h)$ is the Hamming weight of h, i.e. the number of elements $a \in \mathbb{F}_{2^n}$ for which $h(a)$ is non-zero.

The weight of the Boolean function $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$ defined as $f(a) = \mathrm{Tr}_n(\lambda a^3)$ for some given constant $\lambda \in \mathbb{F}_{2^n}$ is known from [13]. More precisely, $\mathrm{wt}(f)$ takes the following values:

$$\mathrm{wt}(f) = \begin{cases} 0 & \lambda = 0; \\ 2^{n-1} & n \text{ odd }, \lambda \neq 0; \\ 2^{n-1} - 2^{n/2} & n \text{ even}, n/2 \text{ odd}, \\ & \quad \lambda \text{ is a cube}, \lambda \neq 0; \\ 2^{n-1} + 2^{n/2-1} & n \text{ even}, n/2 \text{ odd}, \\ & \quad \lambda \text{ is not a cube}, \lambda \neq 0; \\ 2^{n-1} + 2^{n/2} & n \text{ even}, n/2 \text{ even}, \\ & \quad \lambda \text{ is a cube}, \lambda \neq 0; \\ 2^{n-1} - 2^{n/2-1} & n \text{ even}, n/2 \text{ even}, \\ & \quad \lambda \text{ is not a cube}, \lambda \neq 0. \end{cases} \quad (15)$$

Note that in the case of $a \neq 0$ we can express the weight of h as

$$wt(h) = 2^n - wt(f) - 1 \quad (16)$$

for $f(a) = \mathrm{Tr}_n(\lambda a^3)$ with $\lambda = (b + \beta^3)$. □

From Proposition 6 we can easily see that the distance $d(x^3, G)$ tends to infinity with n. Observe that the value Π_F^β does not actually depend on the shift β; this is true for all quadratic functions as per Proposition 5.

Table I gives the values of m_F (for $F(x) = x^3$) and the lower bound on the distance between x^3 and the nearest APN function for all dimensions n in the range $1 \leq n \leq 20$. Note that for $1 \leq n \leq 4$ the bound is tight as witnessed by:

- $u_1 = 0$, $v_1 = 1$ for $n = 1$;
- $u_1 = 0$, $v_1 = \alpha$ for $n = 2$, where α is a primitive element of $\mathbb{F}_{2^2}$;
- $u_1 = 0$, $u_2 = 1$, $v_1 = 1$, $v_2 = \alpha$ for $n = 3$, where α is a primitive element of $\mathbb{F}_{2^3}$;
- $u_1 = 0$, $u_2 = 1$, $v_1 = 1$, $v_2 = 1$ for $n = 4$.

However, as soon as $n \geq 5$, the bound is no longer tight in general. Indeed, in the case of $n = 5$, we have verified that the smallest distance to an APN function is equal to 8, which shows that the bound is not tight anymore. It is worth noting, furthermore, that in this case all possible APN functions at distance 8 from x^3 were obtained by shifting 8 points from $\mathbb{F}_{2^n}$ by the same value $v \in \mathbb{F}_{2^n}$. Whether the bound is tight for some $n > 5$ remains an open question.

TABLE I
VALUES OF m_F AND LOWER BOUNDS ON $d(F,G)$ FOR ANY G APN FOR $F(x) = x^3$ OVER $\mathbb{F}_{2^n}$

Dimension	m_{x^3}	Lower bound on minimum distance
1	0	1
2	0	1
3	3	2
4	3	2
5	15	6
6	27	10
7	63	22
8	111	38
9	255	86
10	495	166
11	1023	342
12	1983	662
13	4095	1366
14	8127	2710
15	16383	5462
16	32511	10838
17	65535	21846
18	130815	43606
19	262143	87382
20	523263	174422

By Proposition 3, we know that the value m_F for some given APN function F and the lower bound K on the distance to the closest APN function derived from it are valid not only for F itself, but for all functions belonging to its CCZ-equivalence class. Since all APN functions of dimensions four and five have been classified up to CCZ-equivalence [3], Corollary 2 can now be used to obtain a lower bound on the Hamming distance between any two APN functions over $\mathbb{F}_{2^n}$ with $n \in \{4, 5\}$ by examining a single representative from each. For higher dimensions, we can compute the lower bound for the known CCZ-classes.

Table II gives the values of m_F for representatives from all switching classes [16] over $\mathbb{F}_{2^n}$ with $n \in \{4, 5, 6, 7, 8\}$. In the case of $n \in \{4, 5\}$ the selected functions encompass representatives from all CCZ-equivalence classes of the corresponding dimension. In the case of $n \in \{6, 8\}$, the functions are given and indexed according to Table 5 from [16]. Note that for $n = 7$, we obtain the same bound for all functions listed in [16] except for the inverse function. Since APN functions in dimensions $n \leq 5$ have been completely classified up to CCZ-equivalence [3], this means that for $n \leq 5$ we now have a lower bound on the distance to the closest APN function for all APN functions over $\mathbb{F}_{2^n}$.

In addition, we compute the values of Π_F and m_F for new 471, resp. 8157 APN functions over $\mathbb{F}_{2^7}$, resp. $\mathbb{F}_{2^8}$ listed in [23]. In the case of $n = 7$, we obtain $m_F = 63$ for all functions F giving a lower bound of 22 on the minimum distance to the closest APN function. In the case of $n = 8$, m_F takes values $69, 75, 81, 87, 93, 99, 105$, so that the lower bound on the Hamming distance is always at least 24. We thus have a lower bound on the distance to the closest APN function for all known APN functions in dimensions $n = 7$ and $n = 8$. The multiset Π_F takes 6665 distinct values for these 8157 functions. A detailed summary of these computational results can be found online at *https://boolean.h.uib.no/mediawiki/*.

The next-to-last column of the table gives the minimum distance from a given function F to the nearest APN function; this can be computed simply as $\lceil m_F/3 \rceil + 1$ but is explicitly given here for convenience. The last column gives the minimum distance to the closest APN function that can be obtained from F by shifting some number of points by the same shift, as described in Section V. These values can be computed efficiently and effectively provide an upper bound on the minimum distance to the closest APN function.

For the case of $n = 5$, we use the filtering methods described above to compute the exact minimum distance to the closest APN function for a representative from each EA-equivalence class; APN functions have been completely classified in this dimension up to EA-equivalence [3]. This shows, in particular, that the single shift distance can, in general, be larger than the minimum distance to an APN function, and that this minimum distance is not preserved under CCZ-equivalence. The results are given in Table III. In the column labeled "Number of shifts", we given the number of distinct shifts that lead to an APN function; e.g. for BCP-2, either all points from U must be assigned the same shift, or they should be divided into four pairs, with each pair of points shifted by the same value. The last column of Table III gives the CCZ-class to which the function obtained by shifting points from F belongs. The functions labeled "BCP-1" and "BCP-2" are constructed in [8], and constitute the earliest example of an APN function EA-inequivalent to a power function.

V. Single shift

A significantly simplified construction involves shifting all the points $u_1, u_2, \ldots, u_K$ by the same value $v \in \mathbb{F}_{2^n}^*$. In this case,

TABLE II
Values of m_F, lower bounds on $d(F, G)$ and minimum single shift distance for any $G \neq F$ APN for $F(x)$ from [16]

Dimension	F	m_F	Lower bound on minimum distance	Minimum single-shift distance
4	x^3	3	2	2
5	x^3	15	6	8
5	x^5	15	6	8
5	x^{15}	9	4	10
6	1.1	27	10	16
6	1.2	27	10	16
6	2.1	15	6	16
6	2.2	27	10	16
6	2.3	27	10	16
6	2.4	15	6	8
6	2.5	15	6	16
6	2.6	15	6	8
6	2.7	15	6	8
6	2.8	15	6	8
6	2.9	21	8	16
6	2.10	21	8	8
6	2.11	15	6	16
6	2.12	15	6	8
7	7.1	54	19	?
7	all others	63	22	?
8	1.1	111	38	?
8	1.2	111	38	?
8	1.3	111	38	?
8	1.4	111	38	?
8	1.5	111	38	?
8	1.6	111	38	?
8	1.7	111	38	?
8	1.8	111	38	?
8	1.9	111	38	?
8	1.10	111	38	?
8	1.11	111	38	?
8	1.12	111	38	?
8	1.13	111	38	?
8	1.14	99	34	?
8	1.15	111	38	?
8	1.16	111	38	?
8	1.17	111	38	?
8	2.1	111	38	?
8	3.1	111	38	?
8	4.1	99	34	?
8	5.1	105	36	?
8	6.1	105	36	?
8	7.1	111	38	?

TABLE III
Distance between APN EA-representatives from $\mathbb{F}_{2^5}$ and closest APN function

F	Lower bound	Actual distance	Single-shift distance	Number of shifts	CCZ-class
x^3	6	8	8	1	x^5
x^5	6	8	8	1	x^3
BCP-2	6	8	8	1,4	x^3
BCP-1	6	8	8	1,4	x^5
x^7	6	10	12	10	x^7
x^{11}	6	10	12	10	x^{11}
x^{15}	4	10	10	10	x^{15}

characterizing the APN-ness of

$$G(x) = F(x) + v\left(\sum_{i\in[K]} 1_{u_i}(x)\right)$$

becomes easier regardless of whether F is assumed to be APN or not.

For a given triple $(a, x, y) \in \mathbb{F}_{2^n}^3$, let us denote by $N_{a,x,y}$ the parity of the number of elements from $\{x, y, a + x, a + y\}$ that are in U, i.e.

$$N_{a,x,y} = |\{x, y, a + x, a + y\} \cap U| \mod 2.$$

Observe that a differential equation of the form $D_a G(x) = b$ for given $a \in \mathbb{F}_{2^n}^*, b \in \mathbb{F}_{2^n}$ can have more than two solutions if and only if

$$D_a F(x) + D_a F(y) = v N_{a,x,y}$$

for $x, y \in \mathbb{F}_{2^n}$ with $x + y \neq a$.

Given some initial function F over $\mathbb{F}_{2^n}$, the following procedure can then be used to find all APN functions G that can be obtained from F by shifting some set of points U by a given shift $v \in \mathbb{F}_{2^n}^*$:

1) assign a Boolean variable $u_x \in \mathbb{F}_2$ to every field element $x \in \mathbb{F}_{2^n}$; the value of u_x will indicate whether x is in U or not;
2) find all tuples $(x, y, a) \in \mathbb{F}_{2^n}^3$ for which $D_a F(x)+D_a F(y) = v$ with $a \neq 0, x \neq y, a + y$;
3) for every such tuple, consider the equation $u_x + u_y + u_{a+x} + u_{a+y} = 0$;
4) find also all tuples $(x, y, a) \in \mathbb{F}_{2^n}^3$ for which $D_a F(x) + D_a F(y) = 0$ with $a \neq 0, x \neq y, a + y$;
5) for every such tuple, consider the equation $u_x + u_y + u_{a+x} + u_{a+y} = 1$;
6) solve the system of all such equations; this can be done by e.g. constructing an $e \times (2^n)$ matrix over $\mathbb{F}_2$, where e is the number of tuples of both types considered above;
7) the solutions to this system now correspond to precisely those sets $U \subseteq \mathbb{F}_{2^n}$ for which G is APN.

Note that in the case that F is APN, no equations of the type $D_a F(x) + D_a F(y) = 0$ exist for $x + y \neq a$ so that steps four and five above can be skipped.

This method is quite useful in practice, as it can be applied rather efficiently (the main part of the computations consists of finding all tuples (x, y, a) satisfying one of the conditions given above) and since it can be applied to an arbitrary function F (not only APN). Note that the same method can be obtained from Theorem 9 in [16] for the case that F is APN, where it is presented as a special case of the so-called "switching construction". A construction in which a Boolean function is added to an (n, n)-function is also studied in [7].

VI. Conclusion

We examined a construction in which a given vectorial Boolean function F is modified at K different points in order to obtain a new function G. We introduced a new CCZ-invariant for APN functions Π_F which to the best of our knowledge is the first such new invariant for the last ten years. We computed the values of Π_F for all known APN functions over $\mathbb{F}_{2^n}$ for $n \leq 8$. We obtained sufficient and necessary conditions for G to be APN, from which we derived an efficient procedure for searching for APN functions at a given distance from F as well as a lower bound on the distance to the closest APN function in terms of Π_F and m_F. Based on this, we computed a lower bound on the Hamming distance to the closest APN function for all APN functions over $\mathbb{F}_{2^n}$ for $n \leq 5$, and for all known APN functions over $\mathbb{F}_{2^n}$ for $n \leq 8$. We also gave a formula expressing this lower bound for the Gold function x^3 over $\mathbb{F}_{2^n}$ for any dimension n. An additional method for characterizing the APN-ness of G was given for the special case when all the shifts $v_1, v_2, \ldots, v_K$ are identical.

There is a lot of room for future work, and a number of questions and research directions remain open. The methods used here for the characterizations of APN functions may be applied to other classes such as differentially 4-uniform functions. A theoretical lower bound on the value m_F would be valuable, as well as additional results related to its computation. Finding relations between m_F and other properties of F may be very important, and applying the filtering procedure in practice may lead to new examples of APN functions.

Acknowledgements

The research presented in this paper was supported by the Trond Mohn Foundation, and by the Research Council of Norway under contract 247742/O70.

References

[1] Thierry P. Berger, Anne Canteaut, Pascale Charpin, and Yann Laigle-Chapuy. On Almost Perfect Nonlinear Functions Over $\mathbb{F}_2^n$. *IEEE Transactions on Information Theory*, 52(9):4160–4170, 2006.

[2] Eli Biham and Adi Shamir. Differential Cryptanalysis of DES-like Cryptosystems. *Journal of Cryptology*, 4(1):3–72, Jan 1991.

[3] Marcus Brinkmann and Gregor Leander. On the Classification of APN Functions up to Dimension Five. *Designs, Codes and Cryptography*, 49:273–288, 2008.

[4] Lilya Budaghyan. *The Equivalence of Almost Bent and Almost Perfect Nonlinear Functions and Their Generalizations*. PhD thesis, Otto-von-Guericke-University Magdeburg, 2005.

[5] Lilya Budaghyan, Claude Carlet, Tor Helleseth, and Nikolay Kaleyski. Changing Points in APN Functions. The 3rd International Workshop on Boolean Functions and their Applications (BFA), June 17-22, Loen, Norway.

[6] Lilya Budaghyan, Claude Carlet, Tor Helleseth, Nian Li, and Bo Sun. On Upper Bounds for Algebraic Degrees of APN Functions. *IEEE Transactions on Information Theory*, 64(6):4399–4411, 2018.

[7] Lilya Budaghyan, Claude Carlet, and Gregor Leander. Constructing New APN Functions from Known Ones. *Finite Fields and Their Applications*, 15(2):150–159, 2009.

[8] Lilya Budaghyan, Claude Carlet, and Alexander Pott. New Classes of Almost Bent and Almost Perfect Nonlinear Polynomials. *IEEE Transactions on Information Theory*, 52(3):1141–1152, 2006.

[9] Anne Canteaut, Pascale Charpin, and Hans Dobbertin. Binary m-sequences with three-valued crosscorrelation: a proof of welch's conjecture. *IEEE Transactions on Information Theory*, 46(1):4–8, 2000.

[10] Claude Carlet. Vectorial Boolean Functions for Cryptography. *Boolean models and methods in mathematics, computer science, and engineering*, 134:398–469, 2010.

[11] Claude Carlet. Boolean and Vectorial Plateaued Functions and APN Functions. *IEEE Transactions on Information Theory*, 61(11):6272–6289, 2015.

[12] Claude Carlet, Pascale Charpin, and Victor A. Zinoviev. Codes, Bent Functions and Permutations Suitable for DES-like Cryptosystems. *Designs, Codes and Cryptography*, 15(2):125–156, 1998.

[13] Leonard Carlitz. Explicit Evaluation of Certain Exponential Sums. *Mathematica Scandinavica*, 44:5–16, 1979.

[14] Florent Chabaud and Serge Vaudenay. Links between Differential and Linear Cryptanalysis. In *Workshop on the Theory and Application of Cryptographic Techniques, EUROCRYPT 94*, volume 950, pages 356–365, 1994.

[15] Pascale Charpin, Sihem Mesnager, and Sumanta Sarkar. Involutions over the Galois field $\mathbb{F}_{2^n}$. *IEEE Transactions on Information Theory*, 62(4):2266–2276, 2016.

[16] Yves Edel and Alexander Pott. A New Almost Perfect Nonlinear Function which is not Quadratic. *Advances in Mathematics of Communications*, 3(1):59–81, 2009.

[17] Nikolay S. Kaleyski. Changing APN Functions at Two Points. *Cryptography and Communications*, Apr 2019.

[18] Yongqiang Li, Mingsheng Wang, and Yuyin Yu. Constructing Differentially 4-uniform Permutations over $GF(2^{2k})$ from the Inverse Function Revisited. *IACR Cryptology ePrint Archive*, 2013:731, 2013.

[19] Kaisa Nyberg. Perfect nonlinear s-boxes. In Donald W. Davies, editor, *Advances in Cryptology — EUROCRYPT '91*, pages 378–386, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg.

[20] Kaisa Nyberg. Differentially Uniform Mappings for Cryptography. *Lecture Notes in Computer Science*, 765:55–64, 1994.

[21] Kaisa Nyberg. S-boxes and round functions with controllable linearity and differential uniformity. In *International Workshop on Fast Software Encryption*, pages 111–130, 1994.

[22] Lawrence C. Washington and Wade Trappe. *Introduction to Cryptography with Coding Theory*. Prentice Hall, 2002.

[23] Yuyin Yu, Mingsheng Wang, and Yongqiang Li. A Matrix Approach for Constructing Quadratic APN Functions.

[24] Yuyin Yu, Mingsheng Wang, and Yongqiang Li. Constructing Differentially 4 Uniform Permutations from Known Ones. *Chinese Journal of Electronics*, 22(3):495–499, 2013.