On the APNness and differential uniformity of some classes of (n, n)-functions over $\mathbb{F}_2^n$

Claude Carlet,*
Universities of Paris 8, France and Bergen, Norway.

E-mail: claude.carlet@gmail.com

We call (n, m)-functions the functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$. We call n-variable Boolean functions the $(n, 1)$-functions.
An (n, m)-function F is called differentially δ-uniform, for a given positive integer δ, if for every $a \in \mathbb{F}_2^n \backslash \{0\}$ and every $b \in \mathbb{F}_2^m$, the equation $F(x)+F(x+a) = b$ has at most δ solutions. We denote the minimum of these integers δ by δ_F and call it the differential uniformity of F. For every (n, m)-function F, we have $\delta_F \geq \max(2, 2^{n-m})$. Equality can happen for $m < n$ if and only if n is even and $m \leq \frac{n}{2}$ [8].
We can have $\delta_F = 2$ only when $m \geq n$. We shall consider such case only when $m = n$. An (n, n)-function F is called *almost perfect nonlinear* (APN) if it is differentially 2-uniform, that is, if for every $a \in \mathbb{F}_2^n \setminus \{0\}$ and every $b \in \mathbb{F}_2^n$, the equation $F(x) + F(x + a) = b$ has 0 or 2 solutions (i.e. the derivative $D_aF(x) = F(x)+F(x+a)$ is 2-to-1). Equivalently, $|\{D_aF(x), x \in \mathbb{F}_2^n\}| = 2^{n-1}$. Still equivalently, for distinct elements x, y, z, t of $\mathbb{F}_2^n$, the equality $x+y+z+t = 0$ implies $F(x) + F(y) + F(z) + F(t) \neq 0$, that is, the restriction of F to any 2-dimensional flat (*i.e.* affine plane) of $\mathbb{F}_2^n$ is non-affine. There are several other characterizations of APN functions, see e.g. [4].
A subclass of APN functions is that of AB functions, which for n odd, have (optimal) nonlinearity $2^{n-1} - 2^{\frac{n-1}{2}}$, the nonlinearity being the maximum Hamming distance between the component functions of F (i.e. the nonzero linear combinations of its coordinate functions) and the affine Boolean functions.

APN functions have been much studied since the end of the last century. But still little is known on them; in particular, very few infinite classes of APN functions are known, all the more if we avoid quadratic functions (vectorial functions, when they are used as substitution boxes (S-boxes) in block ciphers, have better an algebraic degree larger than 2, that is, are preferred non-quadratic, see e.g. [4]). Moreover, no APN permutation is known in even dimension n larger than 6, while for implementation reasons, n is preferred to be even (and if possible to be a power of 2 - concretly, $n = 8$ for robust block ciphers and

*The research of the author is partly supported by the Norwegian Research Council.

$n = 4$ for lightweight ones). The designers of substitution-permutation networks (one of the two main models of block ciphers, which needs to implement permutations) are then driven, for ensuring the resistance against differential attacks, to choose S-boxes that are differentially 4-uniform, such as the inverse function x^{2^n-2}.

Any (n, m)-function can be uniquely represented by its algebraic normal form (ANF):

$$F(x) = \sum_{I \subseteq \{1,\ldots,n\}} a_I \left(\prod_{i \in I} x_i \right) = \sum_{I \subseteq \{1,\ldots,n\}} a_I \, x^I, \tag{1}$$

where a_I belongs to $\mathbb{F}_2^m$. The global degree of the ANF is called the algebraic degree of F.

The vector space $\mathbb{F}_2^n$ is sometimes endowed with the structure of the field $\mathbb{F}_{2^n}$: this field being an n-dimensional vector space over $\mathbb{F}_2$, each of its elements can be identified with the binary vector of length n of its coordinates relative to a fixed basis. Then any (n, n)-function can be uniquely represented by its univariate representation:

$$F(x) = \sum_{i=0}^{2^n-1} \delta_i x^i \in \mathbb{F}_{2^n}[x]/(x^{2^n} + x). \tag{2}$$

When a class of APN functions is found, it is important to determine whether it is new, up some equivalence notions preserving APNness. Let us recall what are these notions. By increasing order of generality:

1. Two (n, m)-functions F and $L' \circ F \circ L$ where

   $$L : (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n \mapsto (x_1, x_2, \ldots, x_n) \times M \in \mathbb{F}_2^n$$

   is an $\mathbb{F}_2$-linear automorphism of $\mathbb{F}_2^n$, M being a nonsingular $n \times n$ matrix over $\mathbb{F}_2$, and L' being an $\mathbb{F}_2$-linear automorphism of $\mathbb{F}_2^m$, are called *linearly equivalent*. If $\mathbb{F}_2^n$ is identified with $\mathbb{F}_{2^n}$, then "$L : (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n \mapsto (x_1, x_2, \ldots, x_n) \times M \in \mathbb{F}_2^n$, M nonsingular" must be replaced by "$L : x \in \mathbb{F}_{2^n} \mapsto \sum_{i=0}^{n-1} a_i x^{2^i}$, where $a_i \in \mathbb{F}_{2^n}$ and $\{x \in \mathbb{F}_{2^n};\ \sum_{i=0}^{n-1} a_i x^{2^i} = 0\} = \{0\}$".

2. Two (n, m)-functions F and $L' \circ F \circ L$, where

   $$L : (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n \mapsto (x_1, x_2, \ldots, x_n) \times M + (a_1, a_2, \ldots, a_n)$$

   is an affine automorphism of $\mathbb{F}_2^n$ and L' is an affine automorphism of $\mathbb{F}_2^m$, are called *affinely equivalent* or affine equivalent.

3. Two (n, m)-functions F and $L' \circ F \circ L + L''$, where L is an affine automorphism of $\mathbb{F}_2^n$, L' is an affine automorphism of $\mathbb{F}_2^m$ and $L'' : (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n \mapsto (x_1, x_2, \ldots, x_n) \times M'' + (a_1, a_2, \ldots, a_m) \in \mathbb{F}_2^m$ is an affine (n, m)-function, M'' being an $n \times m$ binary matrix, are called (extended affine) *EA equivalent*.

4. Two (n,m)-functions F and G whose graphs $\mathcal{G}_F = \{(x,y) \in \mathbb{F}_2^n \times \mathbb{F}_2^m;\ y = F(x)\}$ and $\mathcal{G}_G = \{(x,y) \in \mathbb{F}_2^n \times \mathbb{F}_2^m;\ y = G(x)\}$ are affinely equivalent (*i.e.* such that $L(\mathcal{G}_F) = \mathcal{G}_G$ for some affine automorphism L on $\mathbb{F}_2^n \times \mathbb{F}_2^m$) are called *CCZ equivalent* (the notion is from [5] and the name from [1]).

When an APN function is found, it can be considered new only if it is EA inequivalent to all known APN functions. If it is also CCZ inequivalent, this is still better. But finding functions CCZ equivalent and EA inequivalent to a given function is a hard task, which has been successfully made, in [1], only when the initial function was a Gold function x^{2^i+1}, where $x \in \mathbb{F}_{2^n}$. So, EA inequivalence to all known APN functions may be considered as sufficient.

No infinite class of non-quadratic APN functions that are new up to the standard EA equivalence has been found since 2005 (see [1])[1]. The last class found that is new up to the more general CCZ equivalence dates back to 2000 (see [6]). Moreover, all the known infinite classes of non-quadratic APN functions are CCZ equivalent to power functions and live then in a very narrow family of functions (in [7] has been found an APN (6,6)-function CCZ inequivalent to power functions and to quadratic functions but no infinite class has been built starting from it). New approaches for studying APN functions seem then necessary.

All the known studies of APN (and more general differentially uniform) S-boxes for block ciphers consider functions over the finite field $\mathbb{F}_{2^n}$. No known infinite class of APN functions, or even of differentially uniform functions of a reasonably low order, is given by the algebraic normal form (ANF) of the functions, viewed over the vector space $\mathbb{F}_2^n$ (all known APN functions are indeed given in univariate representation over the field $\mathbb{F}_{2^n}$). Of course, every univariate representation of a specific (n,n)-function leads to an ANF after choosing a basis of the vector space $\mathbb{F}_{2^n}$ over $\mathbb{F}_2$, but this does not work for infinite families since this transformation should be done for all values of n, while in fact the different values of n lead to different expressions. General expressions of functions over $\mathbb{F}_2^n$, leading to infinite families of APN functions given by their ANF, seem useful for several reasons:
- S-boxes are naturally defined over $\mathbb{F}_2^n$ and APN functions with simple ANFs would be good candidates for concrete block ciphers; moreover, a theoretical approach of such S-boxes may give additional insight on them;
- it becomes very difficult to find infinite classes of APN or differentially 4-uniform functions which are new, at least up to EA equivalence and if possible up to CCZ equivalence; new approaches are then useful since they potentially can lead to such new classes;
- a puzzling question is the existence of an infinite class of APN functions with a bad nonlinearity (this low nonlinearity would be nonzero, as proved in [2], but we would like it small compared to the bound $2^{n-1} - 2^{\frac{n-1}{2}}$). Such functions would not present interest for their applications in cryptography, but showing their existence (or their non-existence) would clearly be quite interesting theoretically.

[1] Infinite classes of quadratic APN functions have been regularly found.

If some can exist then the approach over $\mathbb{F}_2^n$ may allow to find a part of them, because some natural classes of functions over $\mathbb{F}_2^n$, such as monotone functions, have all their elements having bad nonlinearity, as shown in [3]. From this viewpoint, it is interesting to know, as we shall prove in the present paper, that monotone APN functions can exist only in small dimensions. Is this an indicator that APN functions cannot have low nonlinearity?
After recalling some basic facts and making a few general observations on the search for differentially uniform functions over $\mathbb{F}_2^n$, we shall study symmetric (n,n)-functions:

Definition 1. *We call* symmetric *any (n,n)- function $F(x) = (f_1(x), \ldots, f_n(x))$ such that $F(\sigma(x)) = \sigma(F(x))$ for every $x \in \mathbb{F}_2^n$ and every permutation σ of $\{1, \ldots, n\}$.*

We shall prove the following negative result:

Proposition 1. *No symmetric (n,n)-function can be APN for $n \geq 4$.*

We shall propose then superclasses that are large enough for our negative result not to apply, and well fitted enough for maybe easing the study of APN-ness in them. We shall study more in detail the class of rotation symmetric functions (that commute with the cyclic shift) and provide for APN rotation symmetric functions a generalization of Dobbertin's result that any APN power function for n odd is a permutation. We shall also give a lower bound on the differential uniformity of monotone (n,n)-functions:

Proposition 2. *For every positive integers n and δ such that $2 \leq \delta \leq n$ and every monotone differentially δ-uniform (n,n)-function, we have:*

$$\delta\left(1 + n + \binom{n}{2} + \cdots + \binom{n}{\lfloor 1 + 2\log_2(\delta)\rfloor}\right) \geq 2^n.$$

In particular, no monotone APN (n,n)-function exists for $n \geq 8$.

We shall propose as well superclasses in which we could hope finding APN functions.

References

[1] L. Budaghyan, C. Carlet and A. Pott. New Classes of Almost Bent and Almost Perfect Nonlinear Polynomials. *Proceedings of the Workshop on Coding and Cryptography 2005*, pp. 306-315, 2005.

[2] C. Carlet. Vectorial Boolean Functions for Cryptography. Chapter of the monograph *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, Y. Crama and P. Hammer eds, Cambridge University Press, pp. 398-469, 2010.

[3] C. Carlet. On the nonlinearity of monotone Boolean functions. Special Issue SETA 2016 of *Cryptography and Communications* 10 (6), pp. 1051-1061, 2018.

[4] C. Carlet. *Boolean Functions for Cryptography and Coding Theory.* Cambridge University Press, 2021.

[5] C. Carlet, P. Charpin, and V. Zinoviev. Codes, bent functions and permutations suitable for DES-like cryptosystems. *Designs, Codes and Cryptography*, 15 (2), pp. 125-156, 1998.

[6] H. Dobbertin. Almost perfect nonlinear power functions on GF(2^n): a new case for n divisible by 5. *Proceedings of Finite Fields and Applications* Fq5, Augsburg, Germany, Springer, pp. 113-121, 2000.

[7] Y. Edel and A. Pott. A new almost perfect nonlinear function which is not quadratic. *Advances in Mathematics of Communications* 3 (1), pp. 59-81, 2009. *Proceedings of EUROCRYPT 1998, Lecture Notes in Computer Science* 1403, pp. 475-488, 1998.

[8] K. Nyberg. Perfect non-linear S-boxes. *Proceedings of EUROCRYPT 1991, Lecture Notes in Computer Science* 547, pp. 378-386, 1992.