On APN functions whose graphs are maximal Sidon sets

Claude Carlet[1,2]

[1]Department of informatics, University of Bergen, Norway
[2]LAGA, Department of Mathematics, University of Paris 8, (and Paris 13 and CNRS), France

Abstract

The graphs $\mathcal{G}_F = \{(x, F(x)); x \in \mathbb{F}_2^n\}$ of those (n,n)-functions $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^n$ that are almost perfect nonlinear (in brief, APN; an important notion in symmetric cryptography) are, equivalently to their original definition by K. Nyberg, those Sidon sets (an important notion in combinatorics) S in $(\mathbb{F}_2^n \times \mathbb{F}_2^n, +)$ such that, for every $x \in \mathbb{F}_2^n$, there exists a unique $y \in \mathbb{F}_2^n$ such that $(x, y) \in S$. Any subset of a Sidon set being a Sidon set, an important question is to determine which Sidon sets are maximal relatively to the order of inclusion. In this paper, we study whether the graphs of APN functions are maximal (that is, optimal) Sidon sets. We show that this question is related to the problem of the existence / non-existence of pairs of APN functions lying at distance 1 from each others, and to the related problem of the existence of APN functions of algebraic degree n. We revisit the conjectures that have been made on these latter problems.

Keywords: Almost perfect nonlinear function; Sidon set in an Abelian group; symmetric cryptography.

1 Introduction

Almost perfect nonlinear (APN) functions, that is, vectorial functions $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^n$ whose derivatives $D_aF(x) = F(x) + F(x + a)$; $a \neq 0$, are 2-to-1, play an important role in symmetric cryptography (see for instance the book [9]), since they allow an optimal resistance against the differential cryptanalysis of the block ciphers that use them as substitution boxes. Their mathematical study is an important domain of research, whose results (and in particular those by K. Nyberg in the early nineties) made possible the invention of the Advanced Encryption Standard (AES), chosen as a standard by the U.S. National Institute of Standards and Technology (NIST) in 2001, and today used worldwide as a cryptosystem dedicated to civilian uses. APN functions also play an important role in coding theory (see [11]).

Sidon sets, which are subsets S in Abelian groups such that all pairwise sums $x + y$ (with $\{x, y\} \subset S$, $x \neq y$), are different, are an important notion in

combinatorics [1], whose name refers to the Hungarian mathematician Simon Sidon, who introduced the concept in relation to Fourier series.

These two notions are related: by definition, a vectorial function $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^n$ is APN if and only if its graph $\mathcal{G}_F = \{(x, F(x)); x \in \mathbb{F}_2^n\}$ is a Sidon set in $(\mathbb{F}_2^n \times \mathbb{F}_2^n, +)$. Since, given a Sidon set S, every subset of S is also a Sidon set, it is useful to study optimal Sidon sets (that is, Sidon sets that are maximal with respect to inclusion). In the present paper, we study the optimality of the graphs of APN functions as Sidon sets. We characterize such optimality in different ways (by the set $\mathcal{G}_F + \mathcal{G}_F + \mathcal{G}_F$ and by the Walsh transform of F) and we relate it to the two problems of the existence / non-existence of pairs of APN functions at Hamming distance 1 from each others, and of APN functions of algebraic degree n. We revisit the conjectures that have been made on these two problems. We address the case of the so-called plateaued APN functions by exploiting further a trick that Dillon used for showing that, for every APN function and every $c \neq 0$, there exist x, y, z such that $F(x) + F(y) + F(z) + F(x + y + z) = c$. The situation is more demanding in our case, but thanks to previous results on plateaued functions, we find a way to reduce the difficulty and this provides a much simpler proof that a plateaued APN function modified at one point cannot be APN, implying that its graph is an optimal Sidon set. We leave open the case of non-plateaued functions and list the known APN functions whose graphs could possibly be non-optimal Sidon sets (for values of n out of reach by computers).

2 Preliminaries

We call (n, m)-function any function F from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$ (we shall sometimes write that F is "in n variables"). It can be represented uniquely by its algebraic normal form (ANF) $F(x) = \sum_{I \subseteq \{1,\ldots,n\}} a_I \prod_{i \in I} x_i$, where $a_I \in \mathbb{F}_2^m$. The algebraic degree of an (n, m)-function equals the global degree of its ANF. Function F is affine if and only if its algebraic degree is at most 1; it is called quadratic if its algebraic degree is at most 2; and it has algebraic degree n if and only if $\sum_{x \in \mathbb{F}_2^n} F(x) \neq 0$. In particular, if F is Boolean (that is, valued in $\mathbb{F}_2$, with $m = 1$) then its algebraic degree is n if and only if it has odd Hamming weight $w_H(F) = |\{x \in \mathbb{F}_2^n;\ F(x) \neq 0\}$.
The vector space $\mathbb{F}_2^n$ can be identified with the field $\mathbb{F}_{2^n}$, since this field is an n-dimensional vector space over $\mathbb{F}_2$. If F is an (n, n)-function viewed over $\mathbb{F}_{2^n}$, then it can be represented by its (also unique) univariate representation $F(x) = \sum_{i=0}^{2^n-1} a_i x^i$, $a_i \in \mathbb{F}_{2^n}$. Its algebraic degree equals then the maximum Hamming weight of (the binary expansion of) those exponents i in its univariate representation whose coefficients a_i are nonzero.

An (n, n)-function is called *almost perfect nonlinear* (APN) [17, 2, 16] if, for every nonzero $a \in \mathbb{F}_2^n$ and every $b \in \mathbb{F}_2^n$, the equation $D_a F(x) := F(x) + F(x + a) = b$ has at most two solutions. Equivalently, the system of equations

$$\begin{cases} x + y + z + t = 0 \\ F(x) + F(y) + F(z) + F(t) = 0 \end{cases}$$

has for only solutions quadruples (x, y, z, t) whose elements are not all distinct (i.e. are pairwise equal). The notion is preserved by extended affine (EA) equivalence (in other words, if F is APN then any function obtained by composing it

Table 1: Known APN exponents on $\mathbb{F}_{2^n}$ up to equivalence and to inversion.

Functions	Exponents d	Conditions
Gold	2^i+1	$\gcd(i,n)=1$
Kasami	$2^{2i}-2^i+1$	$\gcd(i,n)=1$
Welch	2^t+3	$n=2t+1$
Niho	$2^t+2^{\frac{t}{2}}-1$, t even $2^t+2^{\frac{3t+1}{2}}-1$, t odd	$n=2t+1$
Inverse	$2^{2t}-1$ or 2^n-2	$n=2t+1$
Dobbertin	$2^{4t}+2^{3t}+2^{2t}+2^t-1$	$n=5t$

on the left and on the right by affine permutations $x \to x \times M + u$, where M is a nonsingular $n \times n$ matrix over $\mathbb{F}_2$ and $u \in \mathbb{F}_2^n$, and adding an affine function to the resulting function is APN). It is also preserved by the more general CCZ-equivalence (two functions F and G are called CCZ-equivalent if their graphs $\mathcal{G}_F = \{(x, F(x)); x \in \mathbb{F}_2^n\}$ and $\mathcal{G}_G = \{(x, G(x)); x \in \mathbb{F}_2^n\}$ are the image of each other by an affine permutation of $((\mathbb{F}_2^n)^2, +)$, see more in [9]).
APN functions have been characterized by their Walsh transform [14]. Let us recall that the value at $u \in \mathbb{F}_2^n$ of the *Fourier-Hadamard transform* of a real-valued function φ over $\mathbb{F}_2^n$ is defined as $\widehat{\varphi}(u) = \sum_{x \in \mathbb{F}_2^n} \varphi(x)(-1)^{u \cdot x}$, (where "$\cdot$" denotes an inner product in $\mathbb{F}_2^n$). The Fourier-Hadamard transform is bijective. The value of the *Walsh transform* of F at $(u, v) \in \mathbb{F}_2^n \times \mathbb{F}_2^n$ equals the value at u of the Fourier-Hadamard transform of the function $(-1)^{v \cdot F(x)}$, that is, $W_F(u, v) = \sum_{x \in \mathbb{F}_2^n} (-1)^{v \cdot F(x) + u \cdot x}$. In other words, the Walsh transform of F equals the Fourier-Hadamard transform of the indicator function of its graph (which takes value 1 at the input (x, y) if and only if $y = F(x)$). Then F is APN if and only if $\sum_{u,v \in \mathbb{F}_2^n} W_F^4(u, v) = 3 \cdot 2^{4n} - 2^{3n+1}$. This is a direct consequence of the easily shown equality: $\sum_{u,v \in \mathbb{F}_2^n} W_F^4(u, v) = 2^{2n} |\{(x, y, z); F(x) + F(y) + F(z) + F(x + y + z) = 0\}|$.

The nonlinearity[1] of F equals the minimum Hamming distance between the *component functions* $v \cdot F$, $v \neq 0$, and the affine Boolean functions $u \cdot x + \begin{cases} 0 \\ 1 \end{cases}$. It equals $nl(F) = 2^{n-1} - \frac{1}{2} \max_{\substack{u,v \in \mathbb{F}_2^n \\ v \neq 0}} |W_F(u, v)|$.

A large part of known APN functions is made of functions EA-equivalent to power functions, that is, to functions of the form $F(x) = x^d$, after identification of $\mathbb{F}_2^n$ with the field $\mathbb{F}_{2^n}$ (which is possible since this field is an n-dimensional vector space over $\mathbb{F}_2$). The known APN power functions are all those whose exponents d are the conjugates $2^i d \pmod{2^n - 1}$ of those d given in Table 1 below, or of their inverses when they are invertible in $\mathbb{Z}/(2^n - 1)\mathbb{Z}$.

A subset of an elementary 2-group is called a *Sidon set* if it does not contain four distinct elements x, y, z, t such that $x+y+z+t = 0$. The notion is preserved by affine equivalence: if S is a Sidon set and A is an affine permutation, then $A(S)$ is a Sidon set.

By definition, an (n, n)-function F is then APN if and only if its graph $\mathcal{G}_F$

[1]The relationship between nonlinearity and almost perfect nonlinearity is not clear. The question whether all APN functions have a rather large nonlinearity is open.

is a Sidon set in the elementary 2-group $((\mathbb{F}_2^n)^2, +)$.

Any set included in a Sidon set being a Sidon set, the most important for the study of Sidon sets in a given group is to determine those which are maximal (that is, which are not contained in larger Sidon sets); the knowledge of all maximal Sidon sets allows knowing all Sidon sets. A particular case of maximal set is when the set has maximal size, but the maximal size of Sidon sets is unknown. As far as we know, only an upper bound is known: the size $|S|$ of any Sidon set of $((\mathbb{F}_2^n)^2, +)$ satisfies $\binom{|S|}{2} = \frac{|S|\,(|S|-1)}{2} \leq 2^{2n} - 1$, that is (see e.g. [13]), $|S| \leq \left\lfloor \frac{1+\sqrt{2^{2n+3}-7}}{2} \right\rfloor \approx 2^{n+\frac{1}{2}}$. And an obvious lower bound on the maximal size of Sidon sets in $((\mathbb{F}_2^n)^2, +)$ is of course $|S| \geq 2^n$ since there exist APN functions whatever is the parity of n.

We shall see that there are many cases of APN functions whose graphs are maximal Sidon set. The size 2^n of the graph is roughly $\sqrt{2}$ times smaller than what gives the upper bound on the size of Sidon sets, and there seems to be room for the existence of APN functions whose graphs are non-maximal Sidon sets. However, there is no known case where the graph is non-maximal. We relate the question of such existence to a known conjecture on APN functions, and this may lead to conjecturing that no APN function exists whose graph is non-maximal as a Sidon set (however, many conjectures made in the past on APN functions have subsequently been disproved; it may then be risky to state explicitly such conjecture).

3 Characterizations

Note that the property that $\mathcal{G}_F$ is an optimal Sidon set is preserved by CCZ equivalence.

The graph of an APN function F is a non-optimal Sidon set if and only if there exists an ordered pair (a, b) such that $b \neq F(a)$ and such that $\mathcal{G}_F \cup \{a, b\}$ is a Sidon set. It is easily seen that $\mathcal{G}_F \cup \{a, b\}$ is a Sidon set if and only if the system of equations

$$\begin{cases} x + y + z + a & = & 0 \\ F(x) + F(y) + F(z) + b & = & 0 \end{cases} \tag{1}$$

has no solution. Indeed, if this system has a solution (x, y, z) then x, y, z are necessarily distinct, because $b \neq F(a)$, and then, $\mathcal{G}_F \cup \{a, b\}$ is not a Sidon set, since the four points $(x, F(x)), (y, F(y)), (z, F(z)$, and (a, b) are pairwise distinct (because (a, b) by hypothesis cannot equal one of the other points) and sum to $(0, 0)$. Conversely, if the system (1) has no solution, then $\mathcal{G}_F \cup \{a, b\}$ is a Sidon set because, F being APN, four distinct points in $\mathcal{G}_F$ cannot sum to 0 and three points in $\mathcal{G}_F$ cannot sum to (a, b) either. Hence, the graph of an APN function is an optimal Sidon set if and only if, for every ordered pair (a, b) such that $b \neq F(a)$, the system (1) has a solution, that is, since $(x+y+z, F(x)+F(y)+F(z))$ lives outside $\mathcal{G}_F$ when x, y, z are distinct because F is APN[2], $\{(x + y + z, F(x) + F(y) + F(z)); x, y, z \in \mathbb{F}_2^n\}$ covers the whole set $(\mathbb{F}_2^n)^2 \setminus \mathcal{G}_F$. And since, for every (n, n)-function F, $(x+y+z, F(x)+F(y)+F(z))$ covers $\mathcal{G}_F$ when x, y, z are not distinct in $\mathbb{F}_2^n$, we have:

[2]This is a necessary and sufficient condition for APNness.

Proposition 3.1 *The graph of an APN (n,n)-function F is an optimal Sidon set in $((\mathbb{F}_2^n)^2,+)$ if and only if the set*

$$\mathcal{G}_F+\mathcal{G}_F+\mathcal{G}_F=\{(x+y+z,F(x)+F(y)+F(z));x,y,z\in\mathbb{F}_2^n\}$$

covers the whole space $(\mathbb{F}_2^n)^2$.

Remark. For a vectorial function, APNness implies a behavior as different as possible from that of affine functions from the viewpoint of derivatives, since for F APN, $D_aF(x)=F(x)+F(x+a)$ covers a set of (maximal) size 2^{n-1} for every nonzero a, while for an affine function, this set has (minimal) size 1. Having a graph that is an optimal Sidon set also implies a behavior as different as possible from affine functions, from the viewpoint of $\mathcal{G}_F+\mathcal{G}_F+\mathcal{G}_F$, since if F is affine, then $(x+y+z,F(x)+F(y)+F(z))=(x+y+z,F(x+y+z))$ covers a set of size 2^n, which is minimal. $\diamond$

Remark. J. Dillon (private communication) observed that, for every nonzero $c\in\mathbb{F}_{2^n}$, the equation $F(x)+F(y)+F(z)+F(x+y+z)=c$ must have a solution. In other words, there exists a in $\mathbb{F}_2^n$ such that the system in (1) with $b=F(a)+c$ has a solution.
Dillon's proof is given in [9] (after Proposition 161). Let us revisit this proof and say more: let v and c be nonzero elements of $\mathbb{F}_2^n$ and let $G(x)=F(x)+(v\cdot F(x))\,c$. Then we have $G(x)+G(y)+G(z)+G(x+y+z)=F(x)+F(y)+F(z)+F(x+y+z)+(v\cdot(F(x)+F(y)+F(z)+F(x+y+z)))\,c$ and $G(x)+G(y)+G(z)+G(x+y+z)=0$ if and only if $t:=F(x)+F(y)+F(z)+F(x+y+z)$ satisfies $t=(v\cdot t)\,c$. If $v\cdot c=1$, then this is equivalent to $t\in\{0,c\}$. Hence, we have $|\{(x,y,z);G(x)+G(y)+G(z)+G(x+y+z)=0\}|=|\{(x,y,z);F(x)+F(y)+F(z)+F(x+y+z)\in\{0,c\}\}|$. The common size of these two sets is strictly larger than the number of triples (x,y,z) such that x,y,z are not distinct (that is, $3\cdot 2^{4n}-2^{3n+1}$) since G having zero nonlinearity because $v\cdot G=0$ (still assuming that $v\cdot c=1$), it cannot be APN, as proved in [7]. This proves that $|\{(x,y,z);F(x)+F(y)+F(z)+F(x+y+z)=c\}|>0$, since F being APN, we have $|\{(x,y,z);F(x)+F(y)+F(z)+F(x+y+z)=0\}|$ if and only if x,y,z are not distinct.
Dillon's result shows (as we already observed) that for every nonzero c, there exists a in $\mathbb{F}_2^n$ such that the system in (1) with $b=F(a)+c$ has a solution, while in Proposition 3.1, we want that this same system has a solution for every a and every nonzero c in $\mathbb{F}_2^n$.
In the case of a quadratic function F, since the derivative $D_aF(x)=F(x)+F(x+a)$ is affine, its image set $Im(D_aF)=\{D_aF(x);x\in\mathbb{F}_2^n\}$ when $a\neq 0$ is an affine hyperplane, say equals u_a+H_a where u_a is an element of $\mathbb{F}_2^n$ and H_a is a linear hyperplane of $\mathbb{F}_2^n$, say $H_a=\{0,v_a\}^\perp$, where $v_a\neq 0$. Since $F(x)+F(y)+F(z)+F(x+y+z)$ equals $D_aF(x)+D_aF(z)$ with $a=x+y$, and since $Im(D_aF)+Im(D_aF)=H_a+H_a=H_a$, Dillon's result means then in this particular case that $\bigcup_{\substack{a\in\mathbb{F}_2^n\\ a\neq 0}}H_a$ equals $\mathbb{F}_2^n$. $\diamond$

Let us now translate Proposition 3.1 in terms of the Walsh transform (by a routine method):

Corollary 3.2 *The graph of an APN (n,n)-function F is an optimal Sidon set*

in $((\mathbb{F}_2^n)^2, +)$ *if and only if:*

$$\forall (a,b) \in (\mathbb{F}_2^n)^2, \sum_{(u,v)\in(\mathbb{F}_2^n)^2} (-1)^{v\cdot b+u\cdot a}\, W_F^3(u,v) \neq 0. \tag{2}$$

Indeed we have:

$$\sum_{(u,v)\in(\mathbb{F}_2^n)^2} (-1)^{v\cdot b+u\cdot a}\, W_F^3(u,v) =$$
$$\sum_{x,y,z\in\mathbb{F}_2^n} \sum_{(u,v)\in(\mathbb{F}_2^n)^2} (-1)^{v\cdot(F(x)+F(y)+F(z)+b)+u\cdot(x+y+z+a)} =$$
$$2^{2n}|\{(x,y,z) \in (\mathbb{F}_2^n)^3; (x+y+z, F(x)+F(y)+F(z)) = (a,b)\}.$$

Remark. An APN function F has then a graph that is non-maximal as a Sidon set if and only if, making the product of all the expressions in Corollary 3.2 for (a,b) ranging over $(\mathbb{F}_2^n)^2$, we obtain 0:

$$\sum_{\substack{\mathcal{U}=(u_{a,b},v_{a,b})_{(a,b)\in(\mathbb{F}_2^n)^2}\\ \in((\mathbb{F}_2^n)^2)^{((\mathbb{F}_2^n)^2)}}} (-1)^{\sum_{(a,b)\in(\mathbb{F}_2^n)^2}(v_{a,b}\cdot b+u_{a,b}\cdot a)} \prod_{(a,b)\in(\mathbb{F}_2^n)^2} W_F^3(u_{a,b},v_{a,b}) = 0.$$

$\diamond$

Remark. Without loss of generality (by changing $F(x)$ into $F(x)+F(0)$), let $F(0)=0$. Then, since F is APN, we know that $\sum_{(u,v)\in(\mathbb{F}_2^n)^2} W_F^3(u,v) = 3\cdot 2^{3n} - 2^{2n+1}$ (this can be easily calculated since $\sum_{(u,v)\in(\mathbb{F}_2^n)^2} W_F^3(u,v) = 2^{2n}|\{(x,y,z)\in(\mathbb{F}_2^n)^3;\ x+y+z=F(x)+F(y)+F(z)=0\}| = 2^{2n}|\{(x,y,z)\in(\mathbb{F}_2^n)^3;\ x=0 \text{ and } y=z \text{ or } y=0 \text{ and } x=z \text{ or } z=0 \text{ and } x=y\}|$). Hence, Inequality (2) is, under the condition $F(0)=0$, equivalent to:

$$\forall (a,b) \in (\mathbb{F}_2^n)^2, \sum_{\substack{(u,v)\in(\mathbb{F}_2^n)^2\\ v\cdot b+u\cdot a=0}} W_F^3(u,v) \neq 3\cdot 2^{3n-1} - 2^{2n}.$$

Searching for APN (n,n)-functions whose graphs are non-maximal Sidon sets corresponds then to searching for APN (n,n)-functions F and linear hyperplanes H of $(\mathbb{F}_2^n)^2$ such that $\sum_{(u,v)\in H} W_F^3(u,v) = 3\cdot 2^{3n-1} - 2^{2n}$. $\diamond$

4 Relation with the problem of the (non)existence of pairs of APN functions at distance 1 from each others

The question of the existence of pairs of APN functions lying at Hamming distance 1 from each others, and the related question of the existence of APN functions of algebraic degree n have been studied in [4]. The question of the possible distance between APN functions has been studied further in [3]. The following proposition will show the close relationship between these two questions and the maximality of the graphs of APN functions as Sidon sets.

Proposition 4.1 *Let n be any positive integer and F any APN (n,n)-function. The graph of F is non-maximal as a Sidon set if and only if there exists an APN (n,n)-function G which can be obtained from F by changing its value at one single point (i.e. such that G lies at Hamming distance 1 from F).*

Proof. Assume first that the graph of F is non-maximal as a Sidon set. Then there exists $(a,b) \in (\mathbb{F}_2^n)^2$ such that $b \neq F(a)$ and $\mathcal{G}_F \cup \{(a,b)\}$ is a Sidon set. Then the set $(\mathcal{G}_F \cup \{(a,b)\}) \setminus \{(a,F(a))\}$ being automatically a Sidon set, the function G such that $G(x) = \begin{cases} F(x) \text{ if } x \neq a \\ b \text{ if } x = a \end{cases}$ is also APN.
Conversely, if a pair (F,G) of APN functions at distance 1 from each other exists, then there exists a unique $a \in \mathbb{F}_2^n$ such that $F(a) \neq G(a)$ (and $F(x) = G(x)$ for any $x \neq a$). Let us show that the set equal to the union of the graphs of F and G is then a Sidon set (and the graphs of F and G are then non-optimal as Sidon sets): otherwise, let X, Y, Z, T be distinct ordered pairs in this union and such that $X + Y + Z + T = 0$. Since F and G are APN, two elements among X, Y, Z, T have necessarily a for left term. Without loss of generality, we can assume that $Z = (a, F(a))$ and $T = (a, G(a))$. But then we have $X = (x, F(x))$ and $Y = (y, F(y))$ for some x, y and since $X + Y + Z + T = 0$ we must then have $x = y$ and therefore $X = Y$, a contradiction. □

Note that if F and G are defined as in Proposition 4.1 and F has algebraic degree smaller than n, then G has algebraic degree n, since $\sum_{x \in \mathbb{F}_2^n} G(x) = \sum_{x \in \mathbb{F}_2^n} F(x) + b + F(a) = b + F(a) \neq 0$. Hence one function at least among F and G has algebraic degree n.

The following conjecture was stated in [4] (we number it as in this paper):

Conjecture 2: any function obtained from an APN function F by changing one value is not APN.

In other words, there do not exist two APN functions at Hamming distance 1 from each other. Another conjecture was even stated as follows (we number it as in [4] as well):

Conjecture 1: there does not exist any APN function of algebraic degree n for $n \geq 3$.

This conjecture is stronger than Conjecture 2 since if it is true then a pair (F,G) of APN functions at distance 1 from each other would need to be made with functions of degrees less than n, and this is impossible according to Proposition 4.1 and to the observation below it.
According to Proposition 4.1, Conjecture 2 is equivalent to:

Conjecture 3: the graphs of all APN functions are maximal Sidon sets.

Conjectures 1 and 2-3 are still completely open. Ref. [3] has studied further the Hamming distance between APN functions, but no progress was made on Conjectures 1 and 2.

5 The case of plateaued APN functions

Let C be a class of (n,n)-functions that is globally preserved by any translation applied to the input of the functions or to their output. For proving that the graphs of all the APN functions in C are optimal Sidon sets by using Proposition 3.1, it is enough, thanks to a translation of the input by a and of the output by $F(a)$, to prove that, for any APN function F in this class, the system (1) with $a = F(0) = 0$ (and $b = c$) has a solution[3]. Moreover, according to what we have seen in the remark recalling Dillon's observation, if we define $G(x) = F(x) + (v \cdot F(x))\, c$, where $v \cdot c = 1$, if G also belongs to C for every F in C, it is enough to show that, for every function $G \in C$ such that $G(0) = 0$ and having zero nonlinearity, the equation $G(x) + G(y) + G(x+y) = 0$ has solutions (x, y) where x and y are linearly independent over $\mathbb{F}_2$. Indeed, we have $|\{(x,y); G(x)+G(y)+G(x+y) = 0\}| = |\{(x,y); F(x)+F(y)+F(x+y) \in \{0,c\}\}|$, and since F is APN such that $F(0) = 0$, the equality $F(x)+F(y)+F(x+y) = 0$ requires that x and y are linearly dependent. Hence, the equation $F(x)+F(y)+F(x+y) = c$ has solutions if and only if the equation $G(x)+G(y)+G(x+y) = 0$ has solutions (x, y) where x and y are linearly independent.

Recall that an (n,n)-function is called *plateaued* (see e.g. [9]) if, for every $v \in \mathbb{F}_2^n$, there exists a number $\lambda_v \geq 0$ (which is necessarily a power of 2) such that $W_F(u,v) \in \{0, \pm\lambda_v\}$ for every $u \in \mathbb{F}_2^n$. All quadratic APN functions (and more generally all generalized crooked functions, that is, all functions F such that for every $a \neq 0$, the image set H_a of D_aF is an affine hyperplane[4] are plateaued and some other non-quadratic functions are plateaued as well (e.g. all Kasami APN functions, see [19], and all AB functions).

The class of plateaued functions is preserved by translations of the input and by translations of the output; moreover, if F is plateaued then $G(x) = F(x) + (v \cdot F(x))\, c$, where $v \cdot c = 1$, is plateaued (since the component functions of G are also component functions of F) and is non-APN since it has zero nonlinearity. We know from [8, Proposition 7] that, when G is plateaued, the condition "the equation $G(x) + G(y) + G(x+y) = 0$ has linearly independent solutions x, y" is equivalent to non-APNness. This provides a much simpler proof of the next proposition, which has been initially proved in [4, Theorem 3], but the proof was long, globally, and technical for n even.

Corollary 5.1 *Given any plateaued APN (n,n)-function F, changing F at one input gives a function which is not APN. Hence, the graphs of plateaued APN (n,n)-functions are all optimal Sidon sets.*

The proof is straightforward thanks to the observations above and to Proposition 4.1.

According to Proposition 3.1, we have then that, for every plateaued APN function, $\mathcal{G}_F + \mathcal{G}_F + \mathcal{G}_F$ covers the whole space $(\mathbb{F}_2^n)^2$, and according to Corollary 3.2, that $\forall (a,b) \in (\mathbb{F}_2^n)^2$, $\sum_{(u,v)\in(\mathbb{F}_2^n)^2} (-1)^{v\cdot b + u\cdot a}\, W_F^3(u,v) \neq 0$.

Among plateaued APN functions are almost bent functions. A vectorial Boolean function $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is called *almost bent* (AB) [14] if its nonlinearity achieves the best possible value $2^{n-1} - 2^{\frac{n-1}{2}}$ (with n necessarily odd), that is, if

[3] Note that we could also reduce ourselves to $a = b = 0$ but we could not reduce ourselves to $a = F(0) = b = 0$ without loss of generality. This is why we consider G in the sequel.

[4] See more in [9], where is recalled that no non-quadratic crooked function is known.

all of the component functions $v \cdot F$, $v \neq 0$, satisfy $W_F(u,v) \in \{0, \pm 2^{\frac{n+1}{2}}\}$. All AB functions are APN. The converse is not true in general, even when n is odd, but it is true for n odd in the case of plateaued functions (and more generally in the case of functions whose Walsh transform values are all divisible by $2^{\frac{n+1}{2}}$). In Table 1, the AB functions are all Gold and Kasami functions for n odd and Welch and Niho functions.

Remark. The fact that the graphs of AB functions are optimal Sidon sets can also be directly shown by using the van Dam and Fon-Der-Flaass characterization of AB functions [18]: any (n,n)-function is AB if and only if the system $\begin{cases} x+y+z & = & a \\ F(x)+F(y)+F(z) & = & b \end{cases}$ admits $3 \cdot 2^n - 2$ solutions if $b = F(a)$ (i.e. F is APN) and $2^n - 2$ solutions otherwise. It can also be deduced from Corollary 3.2; F being AB, we have:

$$\sum_{(u,v) \in (\mathbb{F}_2^n)^2} (-1)^{v \cdot b + u \cdot a} \, W_F^3(u,v) =$$

$$2^{3n} + 2^{n+1} \left(\sum_{(u,v) \in (\mathbb{F}_2^n)^2} (-1)^{v \cdot b + u \cdot a} \, W_F(u,v) - 2^n \right)$$

and this equals $2^{3n} + 2^{3n+1} - 2^{2n+1} \neq 0$ if $b = F(a)$ and $2^{3n} - 2^{2n+1} \neq 0$ otherwise. $\diamond$

For n even there also exist plateaued APN functions: all Gold and all Kasami functions.

Quadratic functions (i.e. functions of algebraic degree at most 2) are plateaued, as well as the APN function in 6 variables that is now commonly called the Brinkmann-Leander-Edel-Pott function [15]:

$$x^3 + \alpha^{17}(x^{17} + x^{18} + x^{20} + x^{24}) + \alpha^{14}[\alpha^{18}x^9 + \alpha^{36}x^{18} + \alpha^9 x^{36} + x^{21} + x^{42}] + \alpha^{14}\, Tr_1^6(\alpha^{52}x^3 + \alpha^6 x^5 + \alpha^{19}x^7 + \alpha^{28}x^{11} + \alpha^2 x^{13}),$$

where α is primitive (see [9] for the history of this function).

5.1 An interesting particular case

Some APN functions have all their component functions $v \cdot F$ unbalanced (i.e. of Hamming weight different from 2^{n-1}, that is, such that $W_F(0,v) \neq 0$); this is the case for instance of all APN power functions in even number n of variables. A simpler characterization than by Proposition 3.1 (and Corollary 3.2) is possible in such case, providing an interesting property of such functions:

Corollary 5.2 *For every APN plateaued (n,n)-function whose component functions are all unbalanced, the set $ImF + ImF = \{F(x) + F(y); (x,y) \in (\mathbb{F}_2^n)^2\}$ (where ImF is the image set of F) covers the whole space $\mathbb{F}_2^n$.*

Proof. Since we have $W_F(0,v) \neq 0$ for every v, we have then $W_F^3(u,v) = W_F^2(0,v)W_F(u,v)$, for every u, v and therefore:

$$\begin{aligned}
\sum_{(u,v)\in(\mathbb{F}_2^n)^2} (-1)^{v\cdot b+u\cdot a} W_F^3(u,v) &= \sum_{(u,v)\in(\mathbb{F}_2^n)^2} (-1)^{v\cdot b+u\cdot a} W_F(u,v)W_F^2(0,v) \\
&= \sum_{v\in\mathbb{F}_2^n} (-1)^{v\cdot b} W_F^2(0,v)\Big(\sum_{u\in\mathbb{F}_2^n} (-1)^{u\cdot a} W_F(u,v)\Big) \\
&= 2^n \sum_{v\in\mathbb{F}_2^n} (-1)^{v\cdot b} W_F^2(0,v)(-1)^{v\cdot F(a)} \\
&= 2^n \sum_{v,x,y\in\mathbb{F}_2^n} (-1)^{v\cdot(b+F(x)+F(y)+F(a))} \\
&= 2^{2n}|\{(x,y)\in(\mathbb{F}_2^n)^2; F(x)+F(y)+F(a)=b\}|.
\end{aligned}$$

Hence, since the graph of F is an optimal Sidon set, for every (a,b), the set $\{(x,y) \in (\mathbb{F}_2^n)^2; F(x)+F(y)+F(a) = b\}$ is not empty, that is, we have $ImF + ImF = \mathbb{F}_2^n$. □.

This can also be deduced from [9, Theorem 19] and Dillon's result recalled above.

We have then $ImF + ImF = \mathbb{F}_2^n$ in particular for every APN power function in even dimension n. Of course, this is also true for n odd, since APN power functions are in this case bijective.

Note the difference between the condition in Proposition 3.1, "$(x+y+z, F(x)+F(y)+F(z))$ covers the whole space $(\mathbb{F}_2^n)^2$" which lives in $(\mathbb{F}_2^n)^2$ and deals with three elements x, y, z, and that in Corollary 5.2, "$F(x)+F(y)$ covers the whole space $\mathbb{F}_2^n$", which lives in $\mathbb{F}_2^n$, involves two elements x, y and is simpler.

Remark. We know from [12, 10] that the size of the image set of any APN (n,n)-function is at least $\frac{2^n+1}{3}$ when n is odd and $\frac{2^n+2}{3}$ when n is even. Since both numbers are considerably larger than $2^{\frac{n}{2}}$, the size of ImF is plenty sufficient for allowing the condition of Corollary 5.2 to be satisfied. Of course, the fact that ImF has size much larger than $2^{\frac{n}{2}}$ is not sufficient and the question whether some APN functions may have graphs that are not optimal as Sidon sets remains open. ◇

6 Candidate APN functions for having non-optimal graphs as Sidon sets

Plateaued functions are a large part of all known APN functions but they are most probably a tiny part of all APN functions. The only known APN functions that are not plateaued are power functions (the inverse and Dobbertin functions in Table 1) and some APN functions found in [6] and in [5].

In [4] is proved that if F is an APN power function, then given $u \neq 0$, $ux^{2^n-1}+F$ is not APN (or equivalently $u\delta_0 + F$ is not APN, where δ_0 is the indicator of $\{0\}$) when either $u = 1$ or F is a permutation. But, for covering all cases of change at one point of an APN power function, we would need to address

$ux^{2^n-1} + F$ for n even and F not plateaued, and $u(x+1)^{2^n-1} + F$ for every n and F not plateaued. This was done with the multiplicative inverse function x^{2^n-2} for n odd (which is APN): changing it at one point (any one) gives a function that is not APN. According to Proposition 4.1, the graph of the APN multiplicative inverse function is then a maximal Sidon set. But there is some uncertainty about general APN power functions (however, it was checked with a computer that for $n \leq 15$, changing any APN power function at one point makes it non-APN).

Given the APN functions covered in [4], the only possibility of finding known APN functions with a graph that is not maximal as a Sidon set is with:

- functions EA equivalent to Dobbertin functions in a number of variables divisible by 5, at least 20,
- the functions obtained in [6] as CCZ equivalent to Gold functions in even numbers of variables (because in odd numbers of variables, they are AB, since ABness is preserved by CCZ equivalence), that is: $x^{2^i+1} + (x^{2^i} + x + 1)tr(x^{2^i+1})$, $n \geq 4$ even, $\gcd(i,n) = 1$, and $[x + Tr_3^n(x^{2(2^i+1)} + x^{4(2^i+1))}) + tr(x)Tr_3^n(x^{2^i+1} + x^{2^{2i}(2^i+1))}]^{2^i+1}$, where $6|n$ and $\gcd(i,n) = 1$, and $Tr_3^n(x) = x + x^8 + x^{8^2} + \cdots + x^{8^{\frac{n}{3}-1}}$,
- and the following functions found in [5]: $x^3 + tr(x^9) + (x^2+x+1)tr((x^3)$, where $n \geq 4$ is even and $\gcd(i,n) = 1$, and $\Big(x + Tr_3^n(x^6 + x^{12}) + tr(x)Tr_3^n(x^3 + x^{12})\Big)^3 + tr\Big((x + Tr_3^n(x^6 + x^{12}) + tr(x)Tr_3^n(x^3 + x^{12}))^9\Big)$, where $6|n$ and $\gcd(i,n) = 1$.

But these two last cases cannot provide graphs that are non-maximal Sidon sets. Indeed, the functions to which these functions are CCZ-equivalent are quadratic - and then plateaued - and their graphs are then maximal Sidon sets. Plateauedness is not CCZ-invariant, but the fact that the graph of a function is a maximal Sidon set is CCZ-invariant.
With no surprise, the investigation made in [4] did not find any example of a known APN function with a graph that is not maximal as a Sidon set (which means that the graphs of the Dobbertin functions in 5, 10 and 15 variables are maximal Sidon sets). It seems difficult to push it to larger values of n.

Acknowledgement. We thank Lilya Budaghyan and Nian Li for useful information.

References

[1] László Babai and Vera T Sós. Sidon sets in groups and induced subgraphs of Cayley graphs. *European Journal of Combinatorics*, 6(2):101–114, 1985.

[2] Thomas Beth and Cunsheng Ding. On almost perfect nonlinear permutations. In *Proceedings of Eurocrypt' 93, Lecture Notes in Computer Science 765*, pages 65–76. Springer, 1994.

[3] Lilya Budaghyan, Claude Carlet, Tor Helleseth, and Nikolay Kaleyski. On the distance between APN functions. *IEEE Transactions on Information Theory*, 66(9):5742–5753, 2020.

[4] 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, 2017.

[5] Lilya Budaghyan, Claude Carlet, and Gregor Leander. Constructing new APN functions from known ones. *Finite Fields and Their Applications*, 15(2):150–159, 2009.

[6] 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.

[7] Claude Carlet. *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, chapter Vectorial Boolean Functions for Cryptography. Cambridge University Press, 2010.

[8] Claude Carlet. Boolean and vectorial plateaued functions and APN functions. *IEEE Transactions on Information Theory*, 61(11):6272–6289, 2015.

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

[10] Claude Carlet. Bounds on the nonlinearity of differentially uniform functions by means of their image set size, and on their distance to affine functions. *IEEE Transactions on Information Theory*, 67(12):8325–8334, 2021.

[11] Claude Carlet, Pascale Charpin, and Victor A. Zinoviev. Codes, bent functions and permutations suitable for DES-like cryptosystems. *Des. Codes Cryptogr.*, 15(2):125–156, 1998.

[12] Claude Carlet, Annelie Heuser, and Stjepan Picek. Trade-offs for S-boxes: Cryptographic properties and side-channel resilience. In *International conference on applied cryptography and network security*, pages 393–414. Springer, 2017.

[13] Claude Carlet and Sihem Mesnager. On those multiplicative subgroups of $\mathbb{F}_{2^n}^*$ which are Sidon sets and/or sum-free sets. *Journal of Algebraic Combinatorics*, 55(1):43–59, 2022.

[14] Florent Chabaud and Serge Vaudenay. Links between differential and linear cryptanalysis. In *Proceedings of EUROCRYPT'94, Lecture Notes in Computer Science 950*, pages 356–365. Springer, 1994.

[15] 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.

[16] Kaisa Nyberg. Differentially uniform mappings for cryptography. In *Proceedings of EUROCRYPT' 93, Lecture Notes in Computer Science 765*, pages 55–64. Springer, 1993.

[17] Kaisa Nyberg and Lars Ramkilde Knudsen. Provable security against differential cryptanalysis. In *Proceedings of CRYPT0' 92, Lecture Notes in Computer Science 740*, pages 566–574. Springer, 1992.

[18] Edwin R van Dam and Dmitry Fon-Der-Flaass. Codes, graphs, and schemes from nonlinear functions. *European Journal of Combinatorics*, 24(1):85–98, 2003.

[19] Satoshi Yoshiara. Plateaudness of Kasami APN functions. *Finite Fields and Their Applications*, 47:11–32, 2017.