On the image set size of differentially uniform functions and related bounds on their nonlinearity and their distance to affine functions

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

E-mail: `claude.carlet@gmail.com`

Abstract

We revisit and take a closer look at a (not so well known) result of a 2017 paper, showing that the differential uniformity of any vectorial function is bounded from below by an expression depending on the size of its image set. We make explicit the resulting tight lower bound on the image set size of differentially δ-uniform functions. This leads to an open problem on APN functions. We also improve an upper bound on the nonlinearity of vectorial functions obtained in the same reference and involving their image set size. We study when the resulting bound improves upon the covering radius bound. We obtain as a by-product a lower bound on the Hamming distance between differentially δ-uniform functions and affine functions, which we improve significantly with a second bound. This leads us to study what can be the maximum Hamming distance between vectorial functions and affine functions. We provide two upper bounds and study the tightness of the second (which is tighter when m is near n); this poses an interesting question on APN functions, to which we answer negatively. We finally improve the bound on the differential uniformity, under an additional condition.

1 Introduction

Differentially uniform functions are those functions $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$ (i.e. those (n,m)-functions) such that the maximal size δ_F of the set $\{x \in \mathbb{F}_2^n; F(x) + F(x+a) = b\}$ when $a \in \mathbb{F}_2^n$ and $b \in \mathbb{F}_2^m$, a being nonzero, is small. Their study is fundamental for the evaluation of the resistance of the block ciphers which use them as substitution boxes against the main attacks (like the differential attack and the linear attack). They have then been much studied since the 1990's. In particular, the important papers of K. Nyberg, like [11, 12], have led to the main block cipher used in civil applications, the AES [8]. But still not enough is known on their properties in general, and since few are known, it is

difficult to make conjectures on them. What is known is characterizations by diverse means (see a survey in [2]), but the properties of general differentially uniform functions are essentially unknown (such as their maximum algebraic degree, their minimum and maximum nonlinearities, their minimum and maximum Hamming distances to affine functions, to permutations and to affine permutations, the structure of their image sets, their maximum and minimum numbers of fixed points, etc.). One of the rare papers giving properties of all differentially uniform functions is [6]. It is not widely known in the community of vectorial functions for cryptography, since this paper was devoted to side channel attacks.
In the present paper, we first make clear what is known on the size of their image set (since the results given in [6] are far from explicit). Then we develop more this study, which leads us to state an open problem, and we also address their nonlinearity and their Hamming distance to affine functions, which leads us to a second open problem.
When $\delta_F = 2$ (which is optimum), differentially δ-uniform (n, n)-functions are called APN (almost perfect nonlinear). We shall of course be particularly interested in these functions, since they contribute in an optimal way to the resistance against the differential attack.

The paper is organized as follows.
After preliminaries in Section 2, we revisit and study more in detail in Section 3 the result from [6] on the differential uniformity of vectorial functions, given the size of their image sets. We study the equivalent lower bound on the image set size of differentially δ-uniform functions and show it is tight. We apply this lower bound to the sums of F and affine functions and this allows us in Section 4 to state an open problem on APN functions. We observe in Section 5 that an upper bound given in [6] on the nonlinearity of (n, m)-functions by means of their image set size is weak and we derive a much better bound, with the same approach which led to the bound on the differential uniformity. We study when this bound improves upon the covering radius bound. In Section 6, we also bound from below the Hamming distance between differentially uniform functions and affine functions, first as a consequence of the bound on the image set size and then by an improved bound. This leads us in Section 7 to study the maximum Hamming distance between vectorial functions and affine functions and to first slightly improve upon the only known explicit upper bound on it and second significantly improve upon it when m is near n. Showing that this bound is not tight leads to an interesting question on APN functions that we solve. In Section 8, we derive an upper bound on the nonlinearity of any (n, m)-function F by an expression depending on the maximum Hamming distance between vectorial functions and affine functions. In Section 9, we improve the bound on the differential uniformity of F, by introducing another parameter of F.

2 Preliminaries

We shall denote by 0_n (resp. 1_n) the zero vector (resp. the all-1 vector) of length n and by e_i the i-th vector of the canonical basis of the vector space $\mathbb{F}_2^n$. We denote by $w_H(x)$ the Hamming weight of an element x of $\mathbb{F}_2^n$, that is, the size of its support $supp(x) = \{i \in \{1,\ldots,n\};\, x_i = 1\}$. We call co-support of x the complement of its support.

The vector space $\mathbb{F}_2^n$ is sometimes endowed with the structure of the field $\mathbb{F}_{2^n}$ (with null element 0); indeed, 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.

Given an n-variable Boolean function f, that is, a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$, the so-called Walsh transform of f is defined as $W_f(u) = \sum_{x\in\mathbb{F}_2^n}(-1)^{f(x)+u\cdot x}$, where "$\cdot$" is some chosen inner product in $\mathbb{F}_2^n$ (such as $u\cdot x = \sum_{i=1}^n u_i x_i$, or, if $\mathbb{F}_2^n$ is endowed with the structure of $\mathbb{F}_{2^n}$, $u\cdot x = tr_n(ux)$, where $tr_n(x) = \sum_{i=0}^{n-1} x^{2^i}$ is the so-called absolute trace function). The Walsh transform satisfies the so-called *inverse Walsh transform relation*:

$$\sum_{u\in\mathbb{F}_2^n} W_f(u)(-1)^{u\cdot v} = 2^n(-1)^{f(v)}, \forall v \in \mathbb{F}_2^n. \tag{1}$$

For a given (n,m)-function F, that is, a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$, the value $W_F(u,v)$ of the Walsh transform of F at $(u,v) \in \mathbb{F}_2^n \times \mathbb{F}_2^m$ equals that of the Walsh transform of the Boolean function $v\cdot F$ at u.
The *nonlinearity* of a Boolean function f equals its minimum Hamming distance to affine Boolean functions $u\cdot x + \epsilon$, $\epsilon \in \mathbb{F}_2$. It equals then:

$$nl(f) = 2^{n-1} - \frac{1}{2}\max_{u\in\mathbb{F}_2^n} |W_f(u)|. \tag{2}$$

It is bounded above by $2^{n-1} - 2^{\frac{n}{2}-1}$, according to the covering radius bound (see e.g. [2]) and f is called *bent* if it achieves this value. The nonlinearity of an (n,m)-function F equals the minimum nonlinearity of its component functions $v\cdot F$, $v \in \mathbb{F}_2^m \setminus \{0_m\}$. It equals then

$$nl(F) = 2^{n-1} - \frac{1}{2}\max_{\substack{u\in\mathbb{F}_2^n\\ v\in\mathbb{F}_2^m, v\neq 0_m}} |W_F(u,v)|. \tag{3}$$

It is bounded above by $2^{n-1} - 2^{\frac{n}{2}-1}$ as well and F is called *bent* if it achieves this value. Bent functions exist for $m \le \frac{n}{2}$, n even, only [11]. For $m = n$, $nl(F)$ is bounded above by $2^{n-1} - 2^{\frac{n-1}{2}}$, according to the Sidelnikov-Chabaud-Vaudenay (SCV) bound (see e.g. [2]) and F is called *almost bent* (AB) if it achieves this value (AB functions exist for every odd n, see e.g. [2]).

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

$$F(x) = \sum_{I\subseteq\{1,\ldots,n\}} a_I \left(\prod_{i\in I} x_i\right) = \sum_{I\subseteq\{1,\ldots,n\}} a_I\, x^I, \tag{4}$$

where a_I belongs to $\mathbb{F}_2^m$. The global degree of the ANF is called the algebraic degree of F. It equals the maximum algebraic degree of the component functions of F. Affine functions are those functions of algebraic degree at most 1. If $\mathbb{F}_2^n$ is endowed with the structure of the field $\mathbb{F}_{2^n}$, then every (n,n)-function (and then, every (n,m)-functions where m divides n) can be uniquely represented by its univariate representation:

$$F(x) = \sum_{i=0}^{2^n-1} \delta_i x^i \in \mathbb{F}_{2^n}[x]/(x^{2^n} + x). \tag{5}$$

The functions whose univariate expression is a monomial are called *power functions*.
We shall denote the image set $\{F(x); x \in \mathbb{F}_2^n\}$ of F by $Im(F)$.

An (n,m)-function F is called differentially δ-uniform, for a given positive integer δ, if for every $a \in \mathbb{F}_2^n \setminus \{0_n\}$ 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 $\delta_F \geq \max(2, 2^{n-m})$. It is shown in [11] that equality can happen if and only if n is even and $m \leq \frac{n}{2}$.

Note that we can have $\delta_F = 2$ only when $n \geq m$. An (n,n)-function F is called *almost perfect nonlinear* (APN) if it is differentially 2-uniform, that is, if for every $a \in \mathbb{F}_2^n \setminus \{0_n\}$ 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 $x + y + z + t = 0_n$ implies $F(x) + F(y) + F(z) + F(t) \neq 0_n$, that is, the restriction of F to any 2-dimensional flat (*i.e.* affine plane) of $\mathbb{F}_2^n$ is non-affine, that is, for every linearly independent $a, b \in \mathbb{F}_2^n$, the function $D_aD_bF(x)$ does not vanish. There are several characterizations of APN functions (see e.g. the survey [2]).

3 The lower bound on the size of the image sets of differentially uniform functions

In [6, Subsection 4.2] is studied (for reasons related to side channel attacks that we shall not develop here) the differential uniformity of those (n,n)-functions F satisfying, for some d, that $d_H(x, F(x)) \leq d$ for every $x \in \mathbb{F}_2^n$. The differential uniformity of such functions is shown to be bad if d is too small. The authors observe that the condition being that all the images of the function $F(x) + x$ have Hamming weight at most d, the size of the image set of this latter function (which has the same differential uniformity as F) is then bounded above by $D = \sum_{i=0}^{d} \binom{n}{i}$. A lower bound is then proved on the differential uniformity of a function by means of the size of its image set.

We now want to study more deeply this aspect of the size of the image set of (n,m)-functions, since it is only approached as a tool in this paper, whereas it deserves more attention because it corresponds to one of the rare known properties of differentially uniform functions, which is moreover not widely known. For our paper to be self-contained, we briefly recall what is this lower bound and the method for proving it. Let F be any (n,m)-function. We have $\sum_{a\in\mathbb{F}_2^n; a\neq 0_n} |(D_aF)^{-1}(0_m)| = |\{(x,y)\in(\mathbb{F}_2^n)^2; F(x)=F(y)\}| - 2^n = \sum_{b\in Im(F)} |F^{-1}(b)|^2 - 2^n \geq \frac{\left(\sum_{b\in Im(F)} |F^{-1}(b)|\right)^2}{|Im(F)|} - 2^n = \frac{2^{2n}}{|Im(F)|} - 2^n$, where the inequality is obtained by the Cauchy-Schwarz inequality. Since, in every numerical sequence, there exists an element larger than or equal to the arithmetic mean of the sequence, we deduce that there exists $a\in\mathbb{F}_2^n$, nonzero, such that $|D_aF^{-1}(0_m)| \geq \frac{\frac{2^{2n}}{|Im(F)|}-2^n}{2^n-1}$. We have then:

Proposition 1. *[6] For every (n,m)-function, the differential uniformity of F satisfies:*

$$\delta_F \geq \left\lceil \frac{\frac{2^{2n}}{|Im(F)|} - 2^n}{2^n-1} \right\rceil .$$

Equivalently, we have $\frac{2^{2n}}{|Im(F)|} \leq (2^n-1)\,\delta_F + 2^n$, that is:

$$|Im(F)| \geq \left\lceil \frac{2^{2n}}{(2^n-1)\,\delta_F + 2^n} \right\rceil \geq \left\lceil \frac{2^n}{\delta_F+1} \right\rceil . \tag{6}$$

For $\delta_F = 2$, we have then that for every APN (n,n)-function:

$$|Im(F)| \geq \left\lceil \frac{2^{2n}}{3\cdot 2^n - 2} \right\rceil . \tag{7}$$

We observe that this latter bound is much stronger than the bound $|Im(F)| \geq \frac{1+\sqrt{2^{n+2}-7}}{2}$ obtained in [13]. Note also that it is tight. Indeed, we know that APN power functions in even dimension n have for image set the set of cubes in $\mathbb{F}_{2^n}$ (see e.g. [2]), whose number equals $1 + \frac{2^n-1}{3} = \left\lceil \frac{2^{2n}}{3\cdot 2^n - 2} \right\rceil$, which achieves then the (tight) bound of (7) with equality. This is natural since, for every power APN function F, the size of $F^{-1}(b)$ is independent of $b \neq 0_n$ in $Im(F)$, and the sequence $|F^{-1}(b)|$, $b \in Im(F)$, is then constant except at 0_n, and the Cauchy-Shwarz inequality is close to an equality.

Remark. This bound generalizes straightforwardly to any characteristic: let p be a prime and let F be a δ-uniform function $F: \mathbb{F}_p^n \mapsto \mathbb{F}_p^m$ (which can be viewed as $F: \mathbb{F}_{p^n} \mapsto \mathbb{F}_{p^m}$). Denoting $D_aF(x) = F(x+a) - F(x)$, we have by definition $|(D_aF)^{-1}(b)| \leq \delta$ for every $a \neq 0_n$ in $\mathbb{F}_p^n$ and every b in $\mathbb{F}_p^m$ and this implies $\sum_{a\in\mathbb{F}_p^n; a\neq 0_n} |(D_aF)^{-1}(0_m)| = |\{(x,y)\in(\mathbb{F}_p^n)^2; F(x)=F(y)\}| - p^n =$

$$\sum_{b\in Im(F)} |F^{-1}(b)|^2 - p^n \geq \frac{\left(\sum_{b\in Im(F)} |F^{-1}(b)|\right)^2}{|Im(F)|} - p^n = \frac{p^{2n}}{|Im(F)|} - p^n, \text{ and}$$

there exists $a \in \mathbb{F}_p^n$, nonzero, such that $|D_aF^{-1}(0_m)| \geq \frac{\frac{p^{2n}}{|Im(F)|}-p^n}{p^n-1}$. We have then: $\delta \geq \left\lceil \frac{\frac{p^{2n}}{|Im(F)|}-p^n}{p^n-1} \right\rceil$. This implies $|Im(F)| \geq \left\lceil \frac{p^{2n}}{(p^n-1)\,\delta+p^n} \right\rceil \geq \left\lceil \frac{p^n}{\delta+1} \right\rceil$. $\diamond$

4 On the sums of differentially uniform functions and affine functions

The bound of Proposition 1 applies to $F + L$, where L is any affine function (or equivalently, any linear function). The next corollary is then straightforward but it gives an interesting property, which may for instance eliminate a large number of potential APN candidates.

Corollary 1. *Let F be any differentially δ-uniform (n, m)-function. Let $\mathcal{A}$ be the set of affine (n, m)-functions*[1]*. Then, for every $A \in \mathcal{A}$, we have:*

$$|Im(F + A)| \geq \left\lceil \frac{2^{2n}}{(2^n-1)\delta + 2^n} \right\rceil .$$

In particular, an (n, n)-function can be APN only if, for every $A \in \mathcal{A}$, we have:

$$|Im(F + A)| \geq \left\lceil \frac{2^{2n}}{3 \cdot 2^n - 2} \right\rceil . \tag{8}$$

Hence, for each A, we can eliminate all the functions F such that $|Im(F + A)| < \left\lceil \frac{2^{2n}}{3\cdot 2^n-2} \right\rceil$ as potential APN candidates. Of course, instead of taking A affine, we can take it linear.

Trying to build vectorial functions satisfying (8) for every affine function A (that is, for every linear function) without using the notion of APN function and Corollary 1 seems hard. Even for the simplest examples that are the Gold function x^{2^k+1} over $\mathbb{F}_{2^n}$ for n co-prime with k, and the inverse function x^{2^n-2} for odd n, it seems difficult to prove directly that for every linear function L, the sizes of the sets $\{x^{2^k+1} + L(x); x \in \mathbb{F}_{2^n}\}$ for n co-prime with k and $\{x^{2^n-2} + L(x); x \in \mathbb{F}_{2^n}\}$ for n odd are larger than or equal to $\frac{2^{2n}}{3\cdot 2^n-2}$.

Remark. Let us try to see if we have the same with x^{2^k+1} for n not co-prime with k. Note that taking for L the zero function, the size of $Im(F)$ is equal to $1 + \frac{2^n-1}{\gcd(2^n-1,2^k+1)} = 1 + \frac{(2^n-1)\gcd(2^n-1,2^k-1)}{\gcd(2^n-1,2^{2k}-1)} = 1 + \frac{(2^n-1)(2^{\gcd(n,k)}-1)}{(2^{\gcd(n,2k)}-1)} =$

$$\begin{cases} 2^n & \text{if } val_2(k) \geq val_2(n) \\ 1 + \frac{2^n-1}{2^{\gcd(n,k)}+1} & \text{if } val_2(k) < val_2(n) \end{cases},$$

[1] We use the same symbol as for affine Boolean functions since there is no ambiguity.

where $val_2(k)$ is the 2-valuation of k. Hence, if $val_2(k) < val_2(n)$, we have $|Im(F)| < \frac{2^{2n}}{3\cdot 2^n-2}$, since $2^{\gcd(n,k)}+1 \geq 5$ (because n is assumed not co-prime with k) and $1+\frac{2^n-1}{5} = \frac{2^n+4}{5} < \frac{2^{2n}}{3\cdot 2^n-2}$, for $n \geq 3$. In the case $val_2(k) \geq val_2(n)$, we would need to consider nonzero L and the case seems more complex.
Similarly, it seems difficult to say if we have the same with x^{2^n-2} for n even. For $L = 0$, the inequality is satisfied since the inverse function is a permutation and for all the other affine functions L we already know that $|Im(F)| \geq \left\lceil \frac{2^{2n}}{5\cdot 2^n-4} \right\rceil$, since $F + L$ is differentially 4-uniform, but it seems difficult to say more. For $L(x) = x^{2^k}$, the equation $x^{2^n-2}+L(x) = b$ for $b \neq 0$ is equivalent to $x^{2^k+1}+bx = 1$ and by the change of variable $x \to b^{2^{-k}}x$, to $x^{2^k+1} + x = b^{-(1+2^{-k})}$. This equation can be handled as shown in [14, 9], but it is already complex, and addressing other linear functions L seems difficult. It is even difficult to know what can be the largest possible value of $\sum_{a\in\mathbb{F}_2^n;a\neq 0} |(D_a(F+L))^{-1}(0)| = \sum_{a\in\mathbb{F}_2^n;a\neq 0} |(D_aF)^{-1}(L(a))|$ (which provides the lower bound on $|Im(F+L)|$). We know that $|(D_aF)^{-1}(b)|$ equals 4 if and only if $ab = 1$ and equals 2 if and only if $ab \notin \mathbb{F}_2$ and $tr_n\left(\frac{1}{ab}\right) = 0$. It is difficult to say if there exist linear functions such that $|(D_aF)^{-1}(L(a))| \geq 2$ for every nonzero a, and how many times $|(D_aF)^{-1}(L(a))|$ can then reach 4.
In the case of the Gold function as well, it seems difficult to determine the largest possible value of $\sum_{a\in\mathbb{F}_2^n;a\neq 0} |(D_aF)^{-1}(L(a))| = \sum_{a\in\mathbb{F}_2^n;a\neq 0} |\{x \in \mathbb{F}_{2^n};\, ax^{2^k} + a^{2^k}x = L(a) + a^{2^k+1}\}| = \sum_{a\in\mathbb{F}_2^n;a\neq 0} |\{x \in \mathbb{F}_{2^n};\, x^{2^k} + x = \frac{L(a)}{a^{2^k+1}} + 1\}|$.

It seems still more difficult than with Gold and inverse APN functions to prove directly (without using Corollary 1) that, over $\mathbb{F}_{2^n}$, Kasami functions $x^{4^i-2^i+1}$, $\gcd(i,n) = 1$, Welch functions $x^{2^{\frac{n-1}{2}}+3}$, n odd, Niho functions $x^{2^{(n-1)/2}+2^{(n-1)/4}-1}$ if $n \equiv 1 \pmod 4$, and $x^{2^{(n-1)/2}+2^{(3n-1)/4}-1}$ if $n \equiv 3 \pmod 4$, and Dobbertin functions $x^{2^{\frac{4n}{5}}+2^{\frac{3n}{5}}+2^{\frac{2n}{5}}+2^{\frac{n}{5}}-1}$, $5|n$, satisfy that, for every affine (n,n)-function L, we have $|Im(F+L)| \geq \left\lceil \frac{2^{2n}}{3\cdot 2^n-2} \right\rceil$. ◇

Corollary 1 leads to the natural question whether its converse is true:

Open problem: Given an (n,n)-function F, if for every affine (or every linear) (n,n)-function L, we have $|Im(F+L)| \geq \left\lceil \frac{2^{2n}}{3\cdot 2^n-2} \right\rceil$, then F is it necessarily APN?

5 An upper bound on the nonlinearity by means of the image set size

In [6] is also proved that the nonlinearity of any differentially δ-uniform (n,m)-function F is bounded from above as follows: $nl(F) \leq 2^{n-1} - \frac{\frac{2^{n+m-1}}{|Im(F)|}-2^{n-1}}{2^m-1}$. This bound is very weak, even if we take for $|Im(F)|$ the value which is the smallest and then the most in its favor, that is, according to Relation (6): $|Im(F)| = \left\lceil \frac{2^{2n}}{(2^n-1)\,\delta_F+2^n} \right\rceil$. Indeed, the bound says then that $nl(F)$ is bounded above by approximately $2^{n-1} - \frac{2^{m-n-1}(2^n-1)\,\delta_F+2^{m-1}-2^{n-1}}{2^m-1} \approx 2^{n-1} - \frac{1+\delta_F}{2} + 2^{n-m-1}$ and the bound is very far above the covering radius bound, for functions having reasonable differential uniformity.

Let us show a much better bound with the same approach as for proving Proposition 1. We have seen that, thanks to the Cauchy-Schwarz inequality, we have $|\{(x,y) \in \mathbb{F}_2^n; F(x) = F(y)\}| = \sum_{b\in Im(F)} |F^{-1}(b)|^2 \geq \frac{2^{2n}}{|Im(F)|}$. We deduce $\sum_{v\in\mathbb{F}_2^m} W_F^2(0_n,v) = \sum_{v\in\mathbb{F}_2^m;x,y\in\mathbb{F}_2^n} (-1)^{v\cdot(F(x)+F(y))} = 2^m\,|\{(x,y) \in \mathbb{F}_2^n; F(x) = F(y)\}| \geq \frac{2^{2n+m}}{|Im(F)|}$. Hence, we have $\sum_{v\in\mathbb{F}_2^m, v\neq 0_m} W_F^2(0_n,v) \geq \frac{2^{2n+m}}{|Im(F)|} - 2^{2n}$ and $\max_{u\in\mathbb{F}_2^n, v\in\mathbb{F}_2^m, v\neq 0_m} W_F^2(u,v) \geq \max_{v\in\mathbb{F}_2^m, v\neq 0_m} W_F^2(0_n,v) \geq \frac{\frac{2^{2n+m}}{|Im(F)|}-2^{2n}}{2^m-1}$. According to Relation (3), we deduce:

Proposition 2. *For every positive integers n, m and every (n,m)-function, we have:*

$$nl(F) \leq 2^{n-1} - \sqrt{\frac{\frac{2^{2n+m-2}}{|Im(F)|} - 2^{2n-2}}{2^m - 1}}. \tag{9}$$

If we take again $|Im(F)| = \left\lceil \frac{2^{2n}}{(2^n-1)\,\delta_F+2^n} \right\rceil$, then $nl(F)$ is bounded above by approximately $2^{n-1} - \sqrt{\frac{2^{m-2}((2^n-1)\,\delta_F+2^n)-2^{2n-2}}{2^m-1}}$, that is, if $m = n$ for instance, by approximately $2^{n-1} - \sqrt{2^{n-2}\,\delta_F}$. This latter inequality is interesting when $\delta_F > 2$ since it improves then upon the SCV bound. Note that, still for $m = n$, if $\delta = 2$, then (9) writes $nl(F) \leq 2^{n-1} - \sqrt{\frac{\frac{2^{3n-2}}{|Im(F)|}-2^{2n-2}}{2^n-1}}$ and gives no information since, according to Relation (7), it is weaker than the SCV bound, except if $|Im(F)| = \frac{2^{2n}}{3\cdot 2^n-2}$ (in which case, the two bounds would coincide, but this number is not an integer). Let us now compare (9) with the covering radius bound. We know from [11] that this latter bound is not tight for $\frac{n}{2} < m$. The bound (9) of Proposition 2 is strictly sharper than the covering radius bound if and only if we have $\frac{\frac{2^{2n+m-2}}{|Im(F)|}-2^{2n-2}}{2^m-1} > 2^{n-2}$. We have then:

Proposition 3. *For every positive integers n, m and every (n,m)-function, the bound (9) of Proposition 2 is sharper than the covering radius bound if and only if $|Im(F)| < \frac{2^{n+m}}{2^n+2^m-1}$.*

Note that when m ranges between $\frac{n}{2}$ and n, this necessary and sufficient condition ranges from $|Im(F)| \lesssim 2^m$ to $|Im(F)| \lesssim 2^{m-1}$ and the bound (9) of Proposition 2 improves upon the covering radius bound on a larger interval when m is larger.

6 On the Hamming distance between differentially uniform functions and affine vectorial functions

Observing that, for every L, we have $d_H(F,L) = |\{x \in \mathbb{F}_2^n; F(x) \neq L(x)\}| \geq |Im(F+L)|-1$, since $(F+L)(x)$ takes at least $|Im(F+L)|-1$ nonzero values, at least once each, we have then that the Hamming distance from any differentially δ-uniform (n,n)-function F to $\mathcal{A}$ satisfies:

$$d_H(F,\mathcal{A}) \geq \left\lceil \frac{2^{2n}}{(2^n-1)\delta + 2^n} \right\rceil - 1. \tag{10}$$

In particular, any APN (n,n)-function lies at Hamming distance at least $\left\lceil \frac{2^{2n}}{3\cdot 2^n - 2} \right\rceil - 1$ from $\mathcal{A}$.

This kind of bound is interesting because of its relation with the important open problem (that we shall revisit in Section 9) of determining whether APN functions have necessarily a good nonlinearity. The minimum Hamming distance $d_H(F,\mathcal{A})$ between a vectorial function F and affine vectorial functions, contrary to the nonlinearity of F, is not directly linked to the linear attack, but as the nonlinearity does, it quantifies to which extent F is different from affine functions. This parameter has been studied in [4, 5, 10] where it was denoted in diverse ways (we shall keep here the notation $d_H(F,\mathcal{A})$). If for some $b,v,x \in \mathbb{F}_2^n$ and some linear function L, we have $v\cdot F(x) \neq v\cdot(L(x)+b)$, that is, denoting by L^* the adjoint operator of L, if we have $v\cdot(F(x)+b) \neq L^*(v)\cdot x$, then we have $F(x) \neq L(x)+b$. Since for $v \neq 0_n$, $L^*(v)$ ranges over $\mathbb{F}_2^n$ when L ranges over the space $\mathcal{L}$ of linear (n,n)-functions and $v\cdot b$ ranges over $\mathbb{F}_2$, this directly proves that $d_H(F,\mathcal{A}) = \min_{b\in\mathbb{F}_2^n, L\in\mathcal{L}} |\{x \in \mathbb{F}_2^n; F(x) \neq L(x)+b\}| \geq \max_{v\in\mathbb{F}_2^n, v\neq 0} \min_{a\in\mathbb{F}_2^n, \epsilon\in\mathbb{F}_2} d_H(v\cdot F, a\cdot x+\epsilon) \geq \max_{v\in\mathbb{F}_2^n, v\neq 0} nl(v\cdot F) \geq \min_{v\in\mathbb{F}_2^n, v\neq 0} nl(v\cdot F) = nl(F)$. The inequality $d_H(F,\mathcal{A}) \geq nl(F)$ was already observed in [10] and lower and upper bounds were given in [5] when F is a bent function. Such lower bound on $d_H(F,\mathcal{A})$ does not then imply a lower bound on the nonlinearity of APN functions, but it gives some insight.

Let us show now that a much stronger bound than (10) is valid:

Proposition 4. *Let F be any δ-uniform (n,n)-function, then we have:*

$$d_H(F,\mathcal{A}) \geq 2^n - \sqrt{2^n + \delta\,(2^n-1)}.$$

In particular, let F be any APN function, then we have:

$$d_H(F,\mathcal{A}) \geq 2^n - \sqrt{3\cdot 2^n - 2}.$$

Proof. We have:

$$\begin{aligned}\max_{b\in\mathbb{F}_2^n}|F^{-1}(b)| &= \max_{b\in\mathbb{F}_2^n}\sqrt{|F^{-1}(b)|^2}\\ &\leq \sqrt{\sum_{b\in\mathbb{F}_2^n}|F^{-1}(b)|^2}\\ &= \sqrt{|\{(x,y)\in(\mathbb{F}_2^n)^2; F(x)=F(y)\}|}\\ &= \sqrt{\sum_{a\in\mathbb{F}_2^n}|(D_aF)^{-1}(0_n)|}\\ &\leq \sqrt{2^n+\delta\,(2^n-1)}.\end{aligned}$$

Applying this for $b=0_n$ to $F+L$ instead of F, we deduce:

$$d_H(F,L)=|\{x\in\mathbb{F}_2^n; (F+L)(x)\neq 0_n\}|\geq 2^n-\sqrt{2^n+\delta\,(2^n-1)}.$$

□

7 On the maximum possible value of $d_H(F,\mathcal{A})$

The number $2^n-\sqrt{3\cdot 2^n-2}$ is rather close to 2^n (which is of course an upper bound for $d_H(F,\mathcal{A})$). This poses the question of determining what is the largest possible value of $d_H(F,\mathcal{A})$ for all (n,n)-functions F and more generally for all (n,m)-functions F (that is, finding equivalents, for this other nonlinearity parameter, of the covering radius bound for $m<n$ and of the SCV bound for $m\geq n$, see e.g. [2]) and still more interestingly, what are the functions which reach it (which would be the equivalent of bent functions and of almost bent functions for this notion of nonlinearity). The following upper bound was given in a paper in Chinese and reproduced in [10]: $d_H(F,\mathcal{A})<(1-2^{-m})(2^n-1)$. The proof deals with character sums. Let us briefly present it (in a slightly simpler and more complete way): for every linear function L, we have $|\{x\in\mathbb{F}_2^n;\ F(x)+L(x)=F(0_n)\}|=2^{-m}\sum_{x\in\mathbb{F}_2^n,v\in\mathbb{F}_2^m}(-1)^{v\cdot(F(x)+L(x)+F(0_n))}$. Denoting by $\mathcal{L}$ the vector space of linear (n,m)-functions, we have that, if $v\neq 0_m$, then $\sum_{L\in\mathcal{L}}(-1)^{v\cdot L(x)}$ equals $|\mathcal{L}|$ if $x=0_n$ and equals 0 otherwise. This implies (distinguishing the cases $v=0_m$ and $v\neq 0_m$) $\sum_{L\in\mathcal{L}}|\{x\in\mathbb{F}_2^n;\ F(x)+L(x)=F(0_n)\}|=(2^{n-m}+2^{-m}(2^m-1))|\mathcal{L}|$ and therefore: $\max_{L\in\mathcal{L}}|\{x\in\mathbb{F}_2^n;\ F(x)+L(x)=F(0_n)\}|\geq 2^{n-m}+1-2^{-m}$ and this gives indeed $d_H(F,\mathcal{A})\leq 2^n-2^{n-m}-1+2^{-m}$. Note that since $d_H(F,\mathcal{A})$ is an integer, this bound is equivalent to $d_H(F,\mathcal{A})\leq 2^n-2^{n-m}-1$ for $m\leq n$ and to $d_H(F,\mathcal{A})\leq 2^n-2$ for $m\geq n$.

In the next proposition, we obtain a bound that is slightly stronger when $m<n$ (and is identical when $m=n$).

Proposition 5. *For every positive integers n, m and every (n,m)-function F, we have:*

$$d_H(F,\mathcal{A}) \leq 2^n - \left\lceil 2^{\frac{n}{2}-m}\sqrt{2^n+2^m-1}\right\rceil,$$

where $\mathcal{A}$ is the vector space of all affine functions over $\mathbb{F}_2^n$ and $d_H(F,\mathcal{A})$ is the minimum Hamming distance between F and affine functions.

Proof. For every linear (n,m)-function L, we have:

$$\max_{b\in\mathbb{F}_2^m} |\{x\in\mathbb{F}_2^n;\, F(x)+L(x)=b\}|^2 \geq \frac{\sum_{b\in\mathbb{F}_2^m} |\{x\in\mathbb{F}_2^n;\, F(x)+L(x)=b\}|^2}{2^m} =$$

$$2^{-m}|\{(x,y)\in(\mathbb{F}_2^n)^2;\, F(x)+L(x)=F(y)+L(y)\}| =$$
$$2^{-2m}\sum_{x,y\in\mathbb{F}_2^n, v\in\mathbb{F}_2^m} (-1)^{v\cdot(F(x)+F(y)+L(x+y))}.$$

We have, for every $x,y\in\mathbb{F}_2^n$ and every nonzero $v\in\mathbb{F}_2^m$ that $\sum_{L\in\mathcal{L}}(-1)^{v\cdot L(x+y)}$ equals $|\mathcal{L}|$ if $x+y=0_n$ and equals 0 otherwise. We deduce (distinguishing the cases $v=0_m$ and $v\neq 0_m$):

$$\sum_{L\in\mathcal{L}} \max_{b\in\mathbb{F}_2^m} |\{x\in\mathbb{F}_2^n;\, F(x)+L(x)=b\}|^2 \geq (2^{2n-2m}+(2^m-1)2^{n-2m})|\mathcal{L}|,$$

and therefore:

$$\begin{aligned}\max_{L\in\mathcal{L},b\in\mathbb{F}_2^m} |\{x\in\mathbb{F}_2^n;\, F(x)+L(x)=b\}|^2 &\geq 2^{2n-2m}+2^{n-m}-2^{n-2m}\\ &= 2^{n-2m}(2^n+2^m-1).\end{aligned}$$

We deduce

$$\begin{aligned}d_H(F,\mathcal{A}) &= 2^n - \max_{L\in\mathcal{L},b\in\mathbb{F}_2^m} |(F(x)+L(x)+b)^{-1}(0_m)|\\ &\leq 2^n - 2^{\frac{n}{2}-m}\sqrt{2^n+2^m-1}.\end{aligned}$$

This completes the proof. □

For $m<n$, we get $d_H(F,\mathcal{A}) \leq 2^n - \left\lceil 2^{n-m}\sqrt{1+2^{m-n}-2^{-n}}\right\rceil$, which is sharper than the bound $d_H(F,\mathcal{A}) \leq 2^n - 2^{n-m} - 1$ of [10].
For $m=n$, we get $d_H(F,\mathcal{A}) \leq 2^n - \left\lceil\sqrt{2-2^{-n}}\right\rceil = 2^n-2$, the same as in [10].
For $m>n$, we get $d_H(F,\mathcal{A}) \leq 2^n - \left\lceil 2^{\frac{n}{2}-\frac{m}{2}}\sqrt{1+2^{n-m}-2^{-m}}\right\rceil$ which may be worse by one unit than $d_H(F,\mathcal{A}) \leq 2^n-2$ proved in [10].

Remark. To avoid the loss of information due to the first inequality in the proof above, a slightly different approach consists in fixing b (taking later the best possible value), as done in [10], but keeping the square of $|\{x\in\mathbb{F}_2^n;\ F(x)+L(x)=b\}|$. We have

$$\sum_{L\in\mathcal{L}}|\{x\in\mathbb{F}_2^n;\ F(x)+L(x)=b\}|^2=$$

$$2^{-2m}\sum_{L\in\mathcal{L}}\sum_{x,y\in\mathbb{F}_2^n,v,w\in\mathbb{F}_2^m}(-1)^{v\cdot(F(x)+L(x)+b)+w\cdot(F(y)+L(y)+b)}=$$

$$2^{-2m}\sum_{x,y\in\mathbb{F}_2^n,v,w\in\mathbb{F}_2^m}(-1)^{v\cdot(F(x)+b)+w\cdot(F(y)+b)}\left(\sum_{L\in\mathcal{L}}(-1)^{v\cdot L(x)+w\cdot L(y)}\right).$$

We have:

$$\sum_{L\in\mathcal{L}}(-1)^{v\cdot L(x)+w\cdot L(y)}=\begin{cases}|\mathcal{L}| & \text{if } x=y=0_n\\ |\mathcal{L}| & \text{if } x=y\neq 0_n \text{ and } v=w\\ 0 & \text{if } x=y\neq 0_n \text{ and } v\neq w\\ |\mathcal{L}| & \text{if } x\neq y \text{ and } v=w=0_m\\ 0 & \text{if } x\neq y \text{ and } v=w\neq 0_m\\ 0 & \text{if } x\neq y \text{ and } v\neq w.\end{cases}$$

This implies:

$$\sum_{L\in\mathcal{L}}|\{x\in\mathbb{F}_2^n;\ F(x)+L(x)=b\}|^2=$$

$$2^{-2m}\sum_{v,w\in\mathbb{F}_2^m}(-1)^{v\cdot(F(0_n)+b)+w\cdot(F(0_n)+b)}|\mathcal{L}|+$$

$$2^{-2m}|\{(x,v)\in\mathbb{F}_2^n\times\mathbb{F}_2^m,x\neq 0_n\}||\mathcal{L}|+$$

$$2^{-2m}|\{(x,y)\in\mathbb{F}_2^n\times\mathbb{F}_2^n,x\neq y\}||\mathcal{L}|.$$

Here again, the value is maximal when $b=F(0_n)$ and then we have

$$\sum_{L\in\mathcal{L}}|\{x\in\mathbb{F}_2^n;\ F(x)+L(x)=F(0_n)\}|^2=|\mathcal{L}|+2^{-m}(2^n-1)|\mathcal{L}|+2^{-2m}2^n(2^n-1)|\mathcal{L}|.$$

and therefore:

$$\max_{L\in\mathcal{L}}|\{x\in\mathbb{F}_2^n;\ F(x)+L(x)=b\}|^2\geq 1+2^{-m}(2^n-1)+2^{-2m}2^n(2^n-1),$$

that is:

$$d_H(F,\mathcal{A})\leq 2^n-\left\lceil\sqrt{2^{2n-2m}+2^{n-m}-2^{n-2m}+1-2^{-m}}\right\rceil.$$

This bound is probably (since we are taking the ceiling) exactly the same as the one in Proposition 5 (with maybe rare exceptions where it would be lower by 1).$\diamond$

We can see with Proposition 5 and with the remark above that the bound of [10] is not easy to significantly improve with an approach by character sums.

And for $m = n$, which is an important practical case, we have no improvement at all with such method. We shall now obtain, by a completely different approach, another bound which is stronger than both bounds for $m \geq n$ (and more generally for $m \geq n - \ln n$ where ln is the natural logarithm).
Let us choose some $a \in \mathbb{F}_2^n$ and n linearly independent elements $a_1, \ldots, a_n$ of $\mathbb{F}_2^n$; there exists a unique affine (n,m)-function A such that $A(a) = F(a)$ and $A(a + a_i) = F(a + a_i)$ for $i = 1, \ldots, n$. Let us briefly recall how this well-known fact can be shown: an (n,m)-function A is affine and such that $A(a) = F(a)$ if and only if the function L defined by $A(x) = F(a) + L(a + x)$, that is, $L(x) = F(a) + A(a + x)$, is linear (since a function is affine and takes zero value at 0_n if and only if it is linear), and thanks to the fact that $(a_1, \ldots, a_n)$ is a basis of the vector space $\mathbb{F}_2^n$, there exists a unique linear function L satisfying $L(a_i) = F(a) + F(a + a_i)$ for $i = 1, \ldots, n$.
Then we have $d_H(F, A) \leq 2^n - (n + 1)$ since A and F coincide at the $n + 1$ distinct points $a, a + a_1, \ldots, a + a_n$.
We have then:

Proposition 6. *For every positive integer and every (n,m)-function F, we have:*

$$d_H(F, \mathcal{A}) \leq 2^n - n - 1.$$

If this bound is tight, then, for every $a \in \mathbb{F}_2^n$ and every linearly independent elements $a_1, \ldots, a_n$ of $\mathbb{F}_2^n$, any (n,m)-function F achieving it with equality must coincide with the affine function A defined above only at $a, a + a_1, \ldots, a + a_n$ (note that this condition is necessary but may not be sufficient). Hence, such F must satisfy $F(a) + F(a + \sum_{i=1}^n \epsilon_i a_i) \neq L(\sum_{i=1}^n \epsilon_i a_i) = \sum_{i=1}^n \epsilon_i L(a_i) = \sum_{i=1}^n \epsilon_i (F(a) + F(a + a_i))$, for every $a \in \mathbb{F}_2^n$, every basis $(a_1, \ldots, a_n)$ of $\mathbb{F}_2^n$, and every $\epsilon \in \mathbb{F}_2^n$ of Hamming weight at least 2. If we look a little more precisely at the cases where $w_H(\epsilon)$ is even, respectively odd, we see that the condition is equivalent to saying that, for every even number $2 \leq r \leq n$ of linearly independent elements $a_1, \ldots, a_r$, the function $F(x) + \sum_{i=1}^r F(x + a_i) + F(x + \sum_{i=1}^r a_i)$ never vanishes.
This is an interesting condition, which includes differential 2-uniformity (indeed, for $r = 2$, it is equivalent to saying that F is differentially 2-uniform - we do not write “APN”, since this term is traditionally used for (n,n)-functions only, and here m may be different from n). For $n \geq 4$, the condition seems much stronger than differential 2-uniformity. If we fix for instance $r = 4$, the resulting condition is equivalent to saying that $D_{a_1} D_{a_2} F(x) + D_{a_3} D_{a_4} F(x) + D_{a_1 + a_2} D_{a_3 + a_4} F(x)$ never vanishes; hence, not only each of these three second-order derivatives do not vanish, their sum would not either. However:

Proposition 7. *For every integer $n \geq 4$ and every positive integer m, any differentially 2-uniform (n,m)-function is such that, for every linearly independent elements $a_1, \ldots, a_4$ of $\mathbb{F}_2^n$, the function $F(x) + \sum_{i=1}^4 F(x + a_i) + F(x + \sum_{i=1}^4 a_i)$ never vanishes if and only if:*

$$\sum_{u,v \in \mathbb{F}_2^n, v \neq 0_n} W_F^6(u,v) = -2^{6n} + 18 \cdot 2^{4n+m} - 39 \cdot 2^{3n+m} + 22 \cdot 2^{2n+m}.$$

No such (n,m)-function F exists.

Proof. For obtaining such characterization, we shall need to address all cases for $a_1,\dots,a_4$, even when they are linearly dependent. Let us then first study the behavior of the function $\phi_{a_1,a_2,a_3,a_4}(x) := F(x)+\sum_{i=1}^4 F(x+a_i)+F(x+\sum_{i=1}^4 a_i)$ when $a_1,\dots,a_4$ are linearly dependent.

If one element among $a_1,\dots,a_4$ is equal to zero (say $a_4 = 0_n$), then the function $\phi_{a_1,a_2,a_3,a_4}(x)$ equals $F(x+a_1)+F(x+a_2)+F(x+a_3)+F(x+a_1+a_2+a_3) = D_{a_1+a_2}D_{a_1+a_3}F(x+a_1)$ and since F is differentially 2-uniform:
- either a_1+a_2 and a_1+a_3 are linearly dependent (that is, $a_1 = a_2$ or $a_1 = a_3$ or $a_2 = a_3$) and $\phi_{a_1,a_2,a_3,0_n}(x)$ is the zero function,
- or they are linearly independent (that is, a_1, a_2, a_3 are distinct) and $\phi_{a_1,a_2,a_3,0_n}(x)$ does not vanish.

If no element is zero among $a_1,\dots,a_4$ and the sum of two elements is zero (say $a_1 + a_2 = 0_n$), then $\phi_{a_1,a_2,a_3,a_4}(x)$ equals $D_{a_3}D_{a_4}F(x)$ and since F is differentially 2-uniform:
- either a_3 and a_4 are linearly dependent (that is, $a_3 = a_4$) and $\phi_{a_1,a_2,a_3,a_4}(x)$ is the zero function,
- or they are linearly independent (that is, distinct) and $\phi_{a_1,a_2,a_3,a_4}(x)$ does not vanish.

If no sum of at least one and at most two elements among $a_1,\dots,a_4$ is zero and the sum of three elements is zero (say $a_2+a_3+a_4 = 0_n$), then $\phi_{a_1,a_2,a_3,a_4}(x)$ equals $D_{a_3}D_{a_4}F(x)$ and we are back to the same situation, but then a_3 and a_4 cannot be linearly dependent since $a_3 = a_4$ would imply $a_2 = 0_n$ and then $\phi_{a_1,a_2,a_3,a_4}(x)$ does not vanish.

If no sum of at least one and at most three elements among $a_1,\dots,a_4$ is zero and the sum of all four elements is zero, then $\phi_{a_1,a_2,a_3,a_4}(x)$ equals $F(x+a_1)+F(x+a_2)+F(x+a_3)+F(x+a_4) = D_{a_1+a_2}D_{a_1+a_3}F(x+a_1)$ and a_1+a_2 and a_1+a_3 cannot be linearly dependent since this would mean that $a_1 = a_2$ or $a_1 = a_3$ or $a_2 = a_3$, which is excluded; then $\phi_{a_1,a_2,a_3,a_4}(x)$ does not vanish since F is differentially 2-uniform.

Summarizing, the condition in Proposition 7 is equivalent to: for every $a_1,\dots,a_4$ such that:
- one element is zero and the others are not distinct, or two elements are equal and the two others are equal too, then $\phi_{a_1,a_2,a_3,a_4}(x)$ is the zero function,
- in all the other cases, $\phi_{a_1,a_2,a_3,a_4}(x)$ does not vanish.

The number N of quadruples (a_1,a_2,a_3,a_4) such that one element is zero and the others are not distinct, or two elements are equal and the two others are equal too, can be evaluated as follows. Counting each case once and once only, by considering successively the subcases where the number of zero elements equals $4,3,2,1,0$, and in each such subcase, considering the two possibilities above gives: $N = 1+4(2^n-1)+6(2^n-1)+4(2^n-1)(3\cdot 2^n-5)+(2^n-1)(6\cdot 2^n-11) = 1+(2^n-1)(18\cdot 2^n-21)$. Hence we have: $N = 18\cdot 2^{2n} - 39\cdot 2^n + 22$.
Since, for every element b of $\mathbb{F}_2^m$, the sum $\sum_{v\in\mathbb{F}_2^m}(-1)^{v\cdot b}$ equals 2^m if $b = 0_m$

and equals zero otherwise, the condition in Proposition 7 is equivalent to:

$$\sum_{\substack{x,a_1,\dots,a_4\in\mathbb{F}_2^n\\ v\in\mathbb{F}_2^m}} (-1)^{v\cdot(F(x)+\sum_{i=1}^4 F(x+a_i)+F(x+\sum_{i=1}^4 a_i))} = 2^{n+m}N.$$

Using the inverse Walsh transform relation, we have:

$$\sum_{\substack{x,a_1,\dots,a_4\in\mathbb{F}_2^n\\ v\in\mathbb{F}_2^m}} (-1)^{v\cdot(F(x)+\sum_{i=1}^4 F(x+a_i)+F(x+\sum_{i=1}^4 a_i))} =$$

$$2^{-6n} \sum_{\substack{x,a_1,\dots,a_4,\\ u_1,\dots,u_6\in\mathbb{F}_2^n, v\in\mathbb{F}_2^m}} \prod_{i=1}^{6} W_F(u_i,v)(-1)^{(u_1+\cdots+u_6)\cdot x+\sum_{i=1}^4 (u_{i+1}+u_6)\cdot a_i}$$

$$= 2^{-n}\sum_{u,v\in\mathbb{F}_2^n} W_F^6(u,v).$$

Hence, the condition is equivalent to:

$$\sum_{u,v\in\mathbb{F}_2^n} W_F^6(u,v) = 18\cdot 2^{4n+m} - 39\cdot 2^{3n+m} + 22\cdot 2^{2n+m},$$

that is to:

$$\sum_{u,v\in\mathbb{F}_2^n, v\neq 0_n} W_F^6(u,v) = -2^{6n} + 18\cdot 2^{4n+m} - 39\cdot 2^{3n+m} + 22\cdot 2^{2n+m}.$$

This is impossible for $m = n \geq 4$ since the number on the right-hand side is then negative, which is all the more true if $n \geq 4$ and $m \leq n$ (the only possibility for F to be differentially 2-uniform). □

Hence, for every $n \geq 4$ and every m, the inequality in Proposition 6 is in fact strict.

Remark. Proposition 7 proves that, for every $n \geq 4$ and every m, and for every (n,m)-function F, there exist a basis $(a_1,\dots,a_n)$ of $\mathbb{F}_2^n$ and two vectors x,ϵ in $\mathbb{F}_2^n$, such that $w_H(\epsilon) \geq 2$ and $F(x)+F(x+\sum_{i=1}^n \epsilon_i a_i)+\sum_{i=1}^n \epsilon_i(F(x)+F(x+a_i)) = 0_m$. It would be interesting to determine whether, for every function F and every basis $(a_1,\dots,a_n)$ of $\mathbb{F}_2^n$, there exist two vectors x,ϵ in $\mathbb{F}_2^n$ having such property, but it seems difficult to do so. Denoting by $(e_1,\dots,e_n)$ the canonical basis of $\mathbb{F}_2^n$ (of those Hamming weight 1 vectors), this is equivalent (by composing F by the linear automorphism mapping $(a_1,\dots,a_n)$ to $(e_1,\dots,e_n)$), to saying that, for every (n,m)-function F, there exist two vectors x,ϵ in $\mathbb{F}_2^n$ such that $w_H(\epsilon) \geq 2$ and $D_\epsilon F(x)+\sum_{i=1}^n \epsilon_i D_{e_i}F(x) = 0_m$. It seems difficult to check if there can exist F such that, for every ϵ in $\mathbb{F}_2^n$ of Hamming weight at least 2 and every x, this latter expression is nonzero. We recall that we have seen that the case where $w_H(\epsilon)$ is odd reduces itself to the case where $w_H(\epsilon)$ is even, so we shall assume that we are in this latter case. If we use the inverse Walsh transform relation again, the number of x such that $D_\epsilon F(x)+\sum_{i=1}^n \epsilon_i D_{e_i}F(x)=0_m$ equals, denoting the support of ϵ by I (whose size is even) and writing the elements of $(\mathbb{F}_2^n)^I$ in the form $U=(u_i)_{i\in I}$:

$$2^{-m}\sum_{x\in\mathbb{F}_2^n,v\in\mathbb{F}_2^m}(-1)^{v\cdot(F(x)+F(x+\epsilon)+\sum_{i\in I}F(x+e_i))}=$$

$$2^{-(|I|+2)n-m}\sum_{x,u_\emptyset,u_\epsilon\in\mathbb{F}_2^n,U\in(\mathbb{F}_2^n)^I,v\in\mathbb{F}_2^m}W_F(u_\emptyset,v)W_F(u_\epsilon,v)\prod_{i\in I}W_F(u_i,v)$$

$$\cdot(-1)^{(u_\emptyset+u_\epsilon+\sum_{i\in I}u_i)\cdot x+\sum_{i\in I}(u_\epsilon+u_i)\cdot e_i}=$$

$$2^{-(|I|+1)n-m}\sum_{u_\epsilon\in\mathbb{F}_2^n,U\in(\mathbb{F}_2^n)^I,v\in\mathbb{F}_2^m}W_F(u_\epsilon+\sum_{i\in I}u_i,v)W_F(u_\epsilon,v)$$

$$\cdot\prod_{i\in I}W_F(u_i,v)(-1)^{\sum_{i\in I}(u_\epsilon+u_i)\cdot e_i}.$$

It seems difficult to go further. ◇

Reference [10] conjectures that $d_H(F,\mathcal{A})\leq(1-2^{-m})(2^n-2^{\frac{n}{2}})$. For n even, since $d_H(F,\mathcal{A})$ is an integer, this gives $d_H(F,\mathcal{A})\leq 2^n-(2^{\frac{n}{2}}+1)$. We see that, according to Proposition 4, APN functions are good candidates for approaching this conjectured bound (if it is true) or for disproving it (if it is false).

8 An upper bound on the nonlinearity by means of the minimum distance to affine functions

We have seen in Section 6 that, for any (n,m)-function F, we have $nl(F)\leq d_H(F,\mathcal{A})$. Proposition 2 implies another upper bound on the nonlinearity by an expression depending on $d_H(F,\mathcal{A})$. Indeed, let us apply this proposition to $F+A$ where A is the best affine approximation of F. Since $F+A$ equals 0 at $2^n-d_H(F,A)$ points, we have $|Im(F+A)|\leq d_H(F,A)+1=d_H(F,\mathcal{A})+1$. Hence:

Corollary 2. *For every positive integers n,m and every (n,m)-function F, we have:*

$$nl(F)\leq 2^{n-1}-\sqrt{\frac{\frac{2^{2n+m-2}}{d_H(F,\mathcal{A})+1}-2^{2n-2}}{2^m-1}}.$$

This bound must be compared with the bound $nl(F)\leq d_H(F,\mathcal{A})$. It improves upon it if and only if $2^{n-1}-\sqrt{\frac{\frac{2^{2n+m-2}}{d_H(F,\mathcal{A})+1}-2^{2n-2}}{2^m-1}}<d_H(F,\mathcal{A})$, that is, $(2^m-1)(d_H(F,\mathcal{A}))^3+(2^m-1)(1-2^n)(d_H(F,\mathcal{A}))^2+(2^{2n-2}+(2^m-1)(2^{2n-2}-2^n))d_H(F,\mathcal{A})+2^{2n-2}+(2^m-1)2^{2n-2}-2^{2n+m-2}<0$. This inequality has the form $A(d_H(F,\mathcal{A}))^3+B(d_H(F,\mathcal{A}))^2+Cd_H(F,\mathcal{A})+D<0$ with $A=2^m-1>0$, $B=(2^m-1)(1-2^n)<0$, $C=2^{2n-2}+(2^m-1)(2^{2n-2}-2^n)=2^{2n+m-2}-(2^m-1)2^n>0$ and $D=2^{2n-2}+(2^m-1)2^{2n-2}-2^{2n+m-2}=0$. we have then $2^{n-1}-\sqrt{\frac{\frac{2^{2n+m-2}}{d_H(F,\mathcal{A})+1}-2^{2n-2}}{2^m-1}}<d_H(F,\mathcal{A})$ when $d_H(F,\mathcal{A})$ is between the two zeros $\frac{(2^m-1)(2^n-1)\pm\sqrt{(2^m-1)^2(2^n-1)^2-4(2^m-1)(2^{2n+m-2}-(2^m-1)2^n)}}{2(2^m-1)}$, that is, after simplification, $\frac{(2^m-1)(2^n-1)\pm\sqrt{(2^m-1)}\sqrt{2^{n+m+1}+2^m-2^{2n}-2^{n+1}-1}}{2(2^m-1)}$, which are located between 0 and 2^n.

9 An improvement of the lower bound of Proposition 1

Let us show that the bound of Proposition 1 can be made stronger for some functions. Let us denote by Δ the set $\{x+y;(x,y)\in(\mathbb{F}_2^n)^2, x\neq y, F(x)=F(y)\}$. For every nonzero $a\notin\Delta$, we have $|(D_aF)^{-1}(0_m)|=0$. Let us assume that F is not injective. Then we have $|\Delta|>0$. We can then refine the calculations that led to Proposition 1 as follows: $\sum_{a\in\Delta}|(D_aF)^{-1}(0_m)|=\sum_{a\in\mathbb{F}_2^n}|(D_aF)^{-1}(0_m)|-2^n\geq\frac{2^{2n}}{|Im(F)|}-2^n$, and we deduce that there exists $a\in\mathbb{F}_2^n$, nonzero, such that $|D_aF^{-1}(0_m)|\geq\frac{\frac{2^{2n}}{|Im(F)|}-2^n}{|\Delta|}$. Hence:

Proposition 8. *For every non-injective (n,m)-function, the differential uniformity of F satisfies:*

$$\delta_F\geq\left\lceil\frac{\frac{2^{2n}}{|Im(F)|}-2^n}{|\Delta|}\right\rceil, \tag{11}$$

where $\Delta=\{x+y;(x,y)\in(\mathbb{F}_2^n)^2, x\neq y, F(x)=F(y)\}$.

Relation (11) improves upon Proposition 1 when $|\Delta|<2^n-1$.

Remark. We have:

$$\begin{aligned}|\Delta| &\leq \frac{1}{2}|\{(x,y)\in(\mathbb{F}_2^n)^2;F(x)=F(y)\}|-2^{n-1}\\ &= 2^{-(m+1)}\sum_{x,y\in\mathbb{F}_2^n,v\in\mathbb{F}_2^m}(-1)^{v\cdot(F(x)+F(y))}-2^{n-1}\\ &= 2^{-(m+1)}\sum_{v\in\mathbb{F}_2^m}W_F^2(0_n,v)-2^{n-1},\end{aligned}$$

and this bound is tight since it is achieved by those functions such that, in the multiset $*\{x+y;(x,y)\in(\mathbb{F}_2^n)^2, x\neq y, F(x)=F(y)\}*$, each value is matched at most twice. ◇

Conclusion

In this paper, we have revisited and clarified a result on the size of the image set of any APN function and we have studied its consequences. We have also shown that differentially uniform functions lie at large Hamming distance from affine functions and preserve then the block ciphers in which they are used as substitution boxes from attacks based on affine approximation, which completes the fact that they preserve them from differential attacks. The fact that the image set size of the sum of any differentially uniform function with any linear function is bounded from below may provide a new theoretical and computational approach of differentially uniform functions, and in particular of APN functions, which is worth future studies, as well as the Hamming distance between vectorial (possibly APN) functions and affine functions.
We have posed an open problem on the characterization of APN functions by the image set size of their sums with linear functions and shown the non-existence of APN functions satisfying $F(x) + \sum_{i=1}^{4} F(x + a_i) + F(x + \sum_{i=1}^{4} a_i) \neq 0$ for every x when $a_1, \ldots, a_4$ are linearly independent.

Acknowledgement

We thank Sihem Mesnager for useful information.

References

[1] C. Carlet. Characterizations of the differential uniformity of vectorial functions by the Walsh transform, *IEEE Transactions on Information Theory* 64 (9), pp. 6443-6453, 2018. (preliminary version available in *IACR Cryptology ePrint Archive* http://eprint.iacr.org/ 2017/516, 2017).

[2] C. Carlet. Boolean Functions for Cryptography and Coding Theory. Monograph in *Cambridge University Press*, 2020.

[3] 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.

[4] C. Carlet and C. Ding. Highly Nonlinear Mappings. *Special Issue "Complexity Issues in Coding and Cryptography", dedicated to Prof. Harald Niederreiter on the occasion of his 60th birthday, Journal of Complexity* 20, pp. 205-244, 2004.

[5] C. Carlet, C. Ding and J. Yuan. Linear codes from perfect nonlinear mappings and their secret sharing schemes. *IEEE Transactions on Information Theory* 51(6), pp. 2089-2102, 2005.

[6] C. Carlet, A. Heuser and S. Picek. Trade-Offs for S-Boxes: Cryptographic Properties and Side-Channel Resilience. *Proceedings of ACNS 2017, Lecture Notes in Computer Science* 10355, pp. 393-414, 2017.

[7] F. Chabaud and S. Vaudenay. Links between Differential and Linear Cryptanalysis. *Proceedings of EUROCRYPT 1994, Lecture Notes in Computer Science* 950, pp. 356-365, 1995.

[8] J. Daemen and V. Rijmen, *The Design of Rijndael: AES - The Advanced Encryption Standard*. Springer, 2002.

[9] K. H. Kim and S. Mesnager. Solving $x^{2^k+1} + x + a = 0$ in $GF(2^n)$ with $\gcd(n, k) = 1$. *Finite Fields and Their Applications (FFA)* 63, 101630, 2020.

[10] J. Liu, S. Mesnager and L. Chen. On the nonlinearity of S-boxes and linear codes. *Cryptography and Communications* 9 (3), pp. 345-361, 2017.

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

[12] K. Nyberg. Differentially uniform mappings for cryptography. *Proceedings of EUROCRYPT 1993, Lecture Notes in Computer Science* 765, pp. 55-64, 1994.

[13] V. A. Vitkup. On the symmetric properties of APN functions. *Journal of Applied and Industrial Mathematics* 10 (1), pp. 126-135, 2016.

[14] K. S. Williams, Note on cubics over $GF(2^n)$ and $GF(3^n)$. *Journal of Number Theory* 7 (4), pp. 361-365, 1975.