Two generalizations of almost perfect nonlinearity

Claude Carlet,

E-mail: claude.carlet@gmail.com

Universities of Bergen, Norway and Paris 8, France.

Abstract. Almost perfect nonlinear (in brief, APN) functions are vectorial functions $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ playing roles in several domains of information protection, at the intersection of computer science and mathematics. Their definition comes from cryptography and is also related to coding theory. When they are used as substitution boxes (S-boxes, which are the only nonlinear components in block ciphers), APN functions contribute optimally to the resistance against differential attacks. This makes of course a strong cryptographic motivation for their study, which has been very active since the 90's, and has posed interesting and difficult mathematical questions, some of which are still unanswered.

Since the introduction of differential attacks, more recent types of cryptanalyses have been designed, such as integral attacks. No notion about S-boxes has been identified which would play a similar role with respect to integral attacks. In this paper, we study two generalizations of APNness that are natural from a mathematical point of view, since they directly extend classical characterizations of APN functions. We call these two notions strong non-normality and sum-freedom. The former existed already for Boolean functions and the latter is new. We study how they are related to cryptanalyses (the relation is stronger for sum-freedom). The two notions behave differently from each other while they have similar definitions. They behave differently from differential uniformity, which is a well-known generalization of APNness. We study the different ways to define them, and on the example of Kasami functions, how difficult they are. We prove their satisfiability, their monotonicity, their invariance under classical equivalence relations and we characterize them by the Walsh transform.

We begin a study of the multiplicative inverse function (used as a substitution box in the Advanced Encryption Standard and other block ciphers) from the viewpoint of these two notions. In particular, we find a simple expression of the sum of the values taken by this function over affine subspaces of $\mathbb{F}_{2^n}$ that are not vector subspaces. This formula shows that, in such case, the sum never vanishes (which is a remarkable property of the inverse function). We also give a formula for the case of a vector space defined by one of its bases.

Keywords: Vectorial function, Substitution box, Almost perfect nonlinearity

1 Introduction

One of the main attacks on block ciphers, in symmetric cryptography, is the differential attack [6]. *Almost perfect nonlinear* (APN) (n,n)-functions $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$, introduced in [51, 50], are those (vectorial Boolean) functions which, at the local level of their role as a substitution box (S-box) in a round, optimally contribute to the resistance against this attack; see also [48] and [7, 18, 20]. Such S-boxes are essential for contributing to what C. Shannon called confusion (note however that, even if S-boxes are the only nonlinear components of block ciphers, they do not provide confusion on their own, independently of the linear layers). APNness is not strictly required for resistance to differential attacks and in fact most block ciphers are not based on APN functions (they use S-boxes of a low differential uniformity, see below) because they must satisfy also other constraints such as an efficient implementation (which needs n to be most often even, and very often a power of two) and bijectivity. But if an efficiently implementable and bijective S-box could be found, it would most likely be preferred. APN functions are a very interesting mathematical research topic in relation to cryptography.
Almost perfect nonlinearity can be characterized in at least three equivalent ways (the first of which is the original definition):
(i) for every nonzero $a \in \mathbb{F}_2^n$, the derivative[1] $D_aF(x) = F(x) + F(x+a)$ is 2-to-1 (that is, every element of the co-domain has either two pre-images or none by D_aF);
(ii) the restriction of F to any affine plane $\{x, y, z, x+y+z\}$ of $\mathbb{F}_2^n$ (with $x, y, z, x+y+z$ distinct, that is, with x, y, z distinct) is not an affine function;
(iii) the sum of the values taken by $F(x)$ when x ranges over any affine plane is nonzero (that is, $F(x)+F(y)+F(z)+F(x+y+z)$ is nonzero for every distinct x, y, z).
There is also a characterization in terms of coding theory [20] that we shall not use in this paper.

The notion of APN function is mathematically interesting since its definition is very simple and it poses difficult questions, that have remained open for more than thirty years now, despite an active related research activity in several domains of discrete mathematics. It is also important cryptographically, of course. For instance, the choice of the substitution boxes in the most important block cipher for civil use, the Advanced Encryption Standard (AES) [26], is directly related to the work of Kaisa Nyberg in [50] about APN functions (see also [27], which extends and corrects this analysis). Much still needs to be understood on the structure and the properties of APN functions. For instance, finding an APN permutation in an even number of variables larger than 6 would be an important theoretical and practical advance, as well as determining whether APN functions necessarily have a non-weak nonlinearity[2] for every n (that is, whether

[1] To distinguish this derivative from the classical derivative of a polynomial, we could specify "discrete derivative".

[2] The nonlinearity is a parameter of vectorial functions related to the resistance against linear attacks, another very important class of attacks.

the nonzero linear combinations of their coordinate functions are always at a reasonably large Hamming distance from all affine Boolean functions $x \mapsto a \cdot x + \epsilon$, $a \in \mathbb{F}_2^n$, $\epsilon \in \mathbb{F}_2$). This latter question is clarified for quadratic APN functions for $n \leq 8$ (see [4] and [3]) but it remains open for $n \geq 6$ in the case of general APN functions. A lower bound is known for a subclass of APN functions including all known APN functions, see [19].

A well-known generalization of APNness, also related to the differential attack, is *differential uniformity* [48, 49], which extends the first of the three definitions of APNness above: given three positive integers n, m and δ, an (n, m)-function $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is differentially δ-uniform if, for every $a \neq 0$ in $\mathbb{F}_2^n$ and every b in $\mathbb{F}_2^m$, the equation $D_aF(x) = b$ has at most δ solutions. The value $\delta = 2$ is the smallest possible, since for every function F, the number of these solutions is even, because $D_aF(x) = D_aF(x+a)$ (the situation is different in odd characteristic), and cannot be always zero. APNness is equivalent to differential 2-uniformity and the term is reserved for (n, n)-functions.
A notion completing the information given by differential uniformity is that of *vanishing flats* [43], that is, of affine planes over which the function does not sum to 0, and whose number can be seen as a measure of the distance between an (n, n)-function and the set of almost perfect nonlinear functions.

Other generalizations of APNness have been introduced in the literature. An (n, m)-function F is called weakly APN in [1] if its nonzero derivatives all have image set size larger than 2^{n-2}, and it is called partially APN in [12] if, for some $c \in \mathbb{F}_2^n$, the sum of the values $F(x)$ when x ranges over any affine plane containing c is nonzero. One more generalization, called almost perfect c-nonlinearity (APcN), was introduced recently in [30]; its definition is similar to APNness and is related to the c-differential uniformity of vectorial functions, defined for a function F as the maximal number of solutions $(a, b) \in \mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ of the equation $F(x) + cF(x + a) = b$ (with $a \neq 0$ if $c = 1$). These ad hoc generalizations are not directly related to efficient attacks and will not play a role in the present paper.
A generalization of APN functions is given in [38] for odd characteristic p, in relation with the following modified version of the derivative $\tilde{D}_aF(x) = \sum_{i \in \mathbb{F}_p} F(x + ia)$.

More attacks have been designed by the cryptographic community after differential attacks (and linear attacks). *Higher order differential attacks* [39, 35], instead of studying, as differential attacks do, the propagation of differences between two texts, study the propagation of differences from a larger set of texts. They work in particular when the nonlinear functions $F(x)$ implemented in the algorithm have some of their component functions $\langle \beta, F \rangle$ (where $\langle , \rangle$ is an inner product[3]) that have low algebraic degree, that is, are equal to the sums of monomials $\prod_{i \in I} x_i$ of low degree. The attack can work also when the restrictions of these functions obtained by fixing some of their input bits x_i (in guess and determine attacks), have low algebraic degree. Denoting the algebraic degree by d, this makes that the $(d+1)$-th order derivatives of such component functions are equal to the constant zero function. Calculating such $(d+1)$-th order derivatives

[3] We shall use also the notation $\beta \cdot F$.

corresponds (as we shall detail more later) to making sums of outputs corresponding to inputs belonging to affine spaces of dimension $d+1$. Note that the additions of round keys and the linear layers being affine transformations, they propagate these sums.

The principle of higher order differential attacks can be extended to the case where only certain $(d+1)$-th derivatives vanish. Still more generally, the principle of *integral attacks* [36, 55] is based on key-independent sums of cipher-texts for a well chosen subset of plaintexts: given a set (or a multiset) X of vectors, an integral over X is defined as the sum (with a possible ponderation by field elements, see [52]) of all vectors in X and one tries to predict the values in the integrals after a certain number of rounds of encryption. This yields in some cases a distinguisher, which can be turned into a key recovery attack, as for the differential attack. Integral cryptanalysis applies to some ciphers which are not vulnerable to the differential and linear cryptanalyses, and the 128-bit AES limited to six rounds (instead of ten), while it resists the latter, is vulnerable to the former. Cube attacks [28] and Square attacks [25, 2] can be seen as variants or particular cases of integral attacks.

Providing arguments that a given cipher is resistant against integral attacks is difficult, if the sums can a priori be made over any set. This does not distinguish these attacks for instance from (generalized) linear cryptanalysis when any non-linear function can be used, but the difficulty of ensuring resistance seems still harder in the case of integral attacks, despite the fact that, in practice, the approach in integral cryptanalysis is based on very specific input sets. A difference between integral attacks and previous attacks is that studying each S-box independently of the rest of the algorithm, as done by Nyberg and Knudsen for the differential attack [50, 51], has less sense, and there is no equivalent to APNness for such attacks. Of course, we know that, even for the regular APN concept, there is no 100% correspondence between strong S-boxes and strong ciphers, but the situation with integral attacks is worse, and reducing the question of confusion to the role of S-boxes would be more unsuitable about integral attacks than about differential and linear attacks, because linear functions generally introduce a large number of trails in their context.

The functions made of the composition of rounds in block ciphers are complex (these rounds are rather simple for implementation reasons but they have many variables and composition results then in complex functions). Tools are then necessary for allowing to find integral distinguishers. In [54, 8, 34], a distinguishing property on block ciphers called the *division property*, generalizing integral and higher-order differential distinguishers, and introduced by Todo, led to the first cryptanalysis of the full block cipher Misty. For $u \in \mathbb{F}_2^n$, the division property considers the sum $\sum_{x \in X} x^u$ where $x^u := \prod_{i=1}^n x_i^{u_i}$ for a given set X. This sum equals the parity modulo 2 of the size of the intersection between X and the affine space $\{x \in \mathbb{F}_2^n; supp(u) \subseteq supp(x)\}$ where $supp$ denotes the support. The set X is said to have the division property at the order l if this sum equals 0 for every u of Hamming weight less than l. Taking the terminology of [8], X has

the division property if the *parity set* of X, defined as

$$\mathcal{U}(X) = \{u \in \mathbb{F}_2^n; \sum_{x \in X} x^u = 1\},$$

is included in $\{u \in \mathbb{F}_2^n; w_H(u) \geq l\}$. It is well-known since the 70's that this is equivalent to the fact that the indicator function f of X (taking value 1 on X and 0 elsewhere) has algebraic degree at most $n - l$. Note that $\sum_{x \in X} x^u$ equaling $\sum_{supp(u) \subseteq supp(x)} f(x)$, the indicator function of the parity set of X is different from the well-known Möbius transform of f, defined by $g(u) = \sum_{supp(x) \subseteq supp(u)} f(x)$ (see e.g. [18]), but it is closely related since it equals $\sum_{supp(x+1_n) \subseteq supp(u+1_n)} f(x)$, where 1_n is the all-1 vector. For this reason, it is also involutive, and this makes it simple to express that an input division property u propagates to an output division property v through a function F: $v \in \mathcal{U}(F(\mathcal{U}(\{u\}))$.

In [5] is initiated a theory to describe integral and divisional cryptanalyses in a way similar to linear cryptanalysis and (quasi) differential cryptanalysis, where the Linear Approximation Table (correlation matrix) and the Difference Distribution Table are replaced by a quasidifferential transition matrix, which has the nice property that the transition matrix of a composition of functions is the product of their corresponding transition matrices, and there is a simple similar result for concatenation. In addition to the theoretical advance that this notion represents and the computational improvements that it allows (through algorithms computing division properties and efficiently searching for [extended] integral properties), it induces progress in the direction which interests us in the present paper: highlighting the features of vectorial functions allowing them to contribute to the resistance of block ciphers using them as an S-box against integral attacks. But it does not give yet a specific and simple criterion on S-boxes for their contribution to the resistance against these attacks.

Further improvements could lead to such criteria in the future, but it seems useful already to try helping the designers to make choices between S-boxes, in order to improve the resistance of block ciphers against integral attacks. Defining such features seems easier if we restrict ourselves to those attacks where the set over which are considered the integrals is taken as an affine subspace (and actually, this is the case in most attacks, see e.g. [33]; often, but not always, see e.g. [40], this affine space corresponds to fixing some bits in the plaintext). In a similar way as the existence probability of differentials for a block cipher depends on the existence of sufficiently non-uniform derivatives for the involved S-boxes, it seems natural that the condition of the unpredictability of the propagation of integrals is more difficult to achieve if, for some S-boxes used in the cipher, there exist affine spaces A over which they sum to zero. Actually, [54] considers explicitly the possibility that the sum of the values taken by an S-box over an affine space of inputs is zero (this scenario introduced in [36] is denoted by $\mathcal{B}$ in [54, End of Section 2 and Section 3]). Of course, even the non-existence of such affine spaces A of any dimension in the domain of the S-boxes in a block cipher (property that we shall call informally *sum-freedom*), assuming it is possible,

which is not clear, would not ensure that no attack can be found, but we shall see in Subsection 3.2 that it would oblige the cryptanalyst to take X different from an affine space. Note that the condition of not summing to zero over affine spaces of dimension one corresponds for (n, n)-functions to bijectivity, and for dimension two, it corresponds to APNness. We wish to consider it for larger dimensions.

In the present paper, we study then the notions which generalize in a natural way the two characterizations (ii) and (iii) above of APNness (replacing "affine plane" by "k-dimensional affine space", with $k \geq 2$). We call them k-strong non-normality (see below why we choose this term) and kth-order sum-freedom, respectively. We shall see that there is a rather strong relation between sum-freedom and integral attacks (more precisely, the division property), which partially contradicts that the S-box alone cannot be an obstacle to analysis. There does not seem to exist such a relation in the case of strong non-normality. Both notions are related to other attacks such as guess-and-determine attacks involving higher-order differential attacks, but strong non-normality seems (in the case of vectorial functions) less interesting for its own sake than for a comparison with sum-freedom. We shall see that each of the two notions is significantly different from differential uniformity, and that there are big differences between them too as well.

The notion of k-strong non-normality is a generalization to all vectorial functions of the contrary of a notion on Boolean functions (see [18, Definition 28]), generalizing (e.g. in [46]) the normality notion proposed by Dobbertin in [29]: given $k \leq n$, an n-variable Boolean function f is called k-normal (resp. k-weakly normal) if there exists a k-dimensional affine space (a k-flat) on which f is constant (resp. affine). For n even, $\frac{n}{2}$-normal functions are simply called normal and $\frac{n}{2}$-weakly normal functions are called weakly normal. Such Boolean functions are considered peculiar when k is large enough (and indeed, almost all[4] n-variable Boolean functions are k_n-strongly non-normal when the sequence k_n satisfies $k_n \geq c \log_2 n$ for some $c > 1$, but almost all[5] known bent Boolean functions are $\frac{n}{2}$-normal, see *e.g.*[18]). The generalization of k-weak normality to vectorial functions has been considered in [9] (in which paper are mainly studied the densities of the sets of k-normal and k-weakly normal (n, m)-functions and of a few other families, and algorithms for checking these properties) but not the generalization of k-strong non-normality. The notion of kth-order sum-freedom corresponds (for $k \geq 2$) to a strengthening (a considerable one if k is large enough) of the notion of k-strong non-normality: the restriction of F to any k-dimensional flat has (optimal) algebraic degree k.

There is some relation between sum-freedom and invariant subspace attacks, that have been studied in a larger generality in [42]. An invariant subspace is an affine subspace A whose image (by some permutation F, which can be an S-box, or more interestingly, the part of a round that is preceding the addition of the round key) is a coset of A, so that there exist round keys that are such that the

[4] In the sense of probability, n ranging over $\mathbb{N}^*$.
[5] In the common sense.

image after the addition of the key equals A. If this happens, then F sums to 0 over A (indeed, the sum of the elements of any affine space of a dimension at least 2 equals 0). Hence, sum-freedom protects against the existence of invariant subspaces (and is much more demanding than avoiding invariant subspaces).

The paper is organized as follows. After preliminaries in Section 2, we define the two notions (strong non-normality and sum-freedom) in Section 3, we study the different ways of expressing them, we show the relation of sum-freedom with higher-order derivatives, we study the relation of each notion with cryptanalyses (which provides a rather strong motivation for studying sum-freedom, and some motivation for studying the k-strong non-normality of vectorial functions), we show the difficulty of studying sum-freedom with the example of Kasami functions, and we verify the existence of functions satisfying each notion. We study in Section 4 the properties of the two notions, which show important differences between them and in some cases with APNness. After studying in Subsection 4.1 the constraints on the algebraic degree implied by these notions and in Subsection 4.2 their (non-)monotonicity, we generalize in Subsection 4.3 the Chabaud-Vaudenay characterization of APNness by the Walsh transform to k-strong non-normality and to kth-order sum-freedom (both characterizations happen to be more difficult to obtain than for APNness, and their expressions are more complex, but they give more insight, even on APNness). We study in Subsection 4.4 the invariance under the classical equivalences of both notions. In Section 5, we begin a study, with respect to these two notions, of the multiplicative inverse function $x \in \mathbb{F}_{2^n} \mapsto x^{2^n-2}$ (which is clearly, since Nyberg's works and the invention of the AES, one of the most important infinite classes of vectorial functions to be studied from a cryptographic point of view). We show in particular that this function has the strong property of summing to nonzero values over all affine subspaces of $\mathbb{F}_{2^n}$ that are not linear subspaces, whatever is their dimension, and we give a well-structured expression of the sum of inverses over linear spaces given by a basis.

2 Preliminaries

2.1 Boolean and vectorial functions

Given two positive integers n and m, the functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$ are called (n,m)-functions. When n and/or m are not specified, these functions are called vectorial functions. In the particular case of $m = 1$, they are called n-variable Boolean functions, or Boolean functions in dimension n. The vector space of n-variable Boolean functions is denoted by $\mathcal{B}_n$. Every (n,m)-function F admits a unique *algebraic normal form* (ANF), that is, a representation as a multivariate polynomial of the form $F(x) = \sum_{I \subseteq \{1,\ldots,n\}} a_I \prod_{i \in I} x_i = \sum_{I \subseteq \{1,\ldots,n\}} a_I\, x^I$; $x = (x_1, \ldots, x_n) \in \mathbb{F}_2^n, a_I \in \mathbb{F}_2^m$. The degree $\max\{|I|; a_I \neq 0\}$ of this multivariate polynomial is called the *algebraic degree* of F and denoted by $d_{alg}(F)$. Note that the algebraic degree of a vectorial function F is then the maximum algebraic degree of its component functions $v \cdot F$; $v \in \mathbb{F}_2^m \setminus \{0\}$, where "$\cdot$" is an inner

product in $\mathbb{F}_2^m$ (for instance $v \cdot y = \sum_{i=1}^{m} v_i y_i \in \mathbb{F}_2$, or, if $\mathbb{F}_2^m$ is endowed with the structure of the field $\mathbb{F}_{2^m}$, $v \cdot y = tr_m(vy)$, where $tr_m(y) = y + y^2 + y^{2^2} + \cdots + y^{2^{m-1}}$ is the trace function from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$).
Any function F is affine, that is, satisfies $F(x) + F(y) + F(z) + F(x+y+z) = 0$ for every $x, y, z \in \mathbb{F}_2^n$ if, and only if, its algebraic degree is at most 1. We shall say that a function is *quadratic* if it has algebraic degree at most 2 (hence, affine functions are particular quadratic functions in this terminology, which is nowadays widely accepted since defining quadratic functions as having algebraic degree exactly 2 would make many statements more complex). Function F has algebraic degree at most $r < n$ if, and only if, it sums to zero over every affine space of dimension $r+1$ (and then over every affine space of dimension $k > r$). An (n, m)-function has algebraic degree n (the maximum) if, and only if, it sums to a nonzero value over $\mathbb{F}_2^n$. An n-variable Boolean function f has then algebraic degree n if and only if it has an odd Hamming weight. If f has an even Hamming weight then it has algebraic degree $n-1$ exactly if, and only if, $\sum_{x \in \mathbb{F}_2^n} (xf(x)) \neq 0$ (see e.g. [18]). This latter result comes from the fact that we have $\sum_{x \in \mathbb{F}_2^n} (xf(x)) \neq 0$ if and only if there exists v such that $v \cdot \left(\sum_{x \in \mathbb{F}_2^n} (xf(x)) \right) = \sum_{x \in \mathbb{F}_2^n} \left((v \cdot x) f(x) \right) \neq 0$, that is, function f is not orthogonal to the Boolean function $v \cdot x$ of algebraic degree 1, and we know (see [44, 18]) that the orthogonal[6] of the $\mathbb{F}_2$-vector space of functions of algebraic degree at most 1 (the so-called Reed-Muller code of order 1) is the vector space of Boolean functions of algebraic degree at most $n-2$ (the Reed-Muller code of order $n-2$).

If $\mathbb{F}_2^n$ is endowed with the structure of the field $\mathbb{F}_{2^n}$ (which is always possible since we know that $\mathbb{F}_{2^n}$ is an n-dimensional vector space over $\mathbb{F}_2$), then every (n, n)-function (and thus, every (n, m)-function where m divides n) can be uniquely represented by its univariate representation:

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

(we call power functions the functions of univariate representation $F(x) = x^i$). The algebraic degree of function F in (1) equals the largest Hamming weight of the binary expansion of those exponents i whose coefficients δ_i are nonzero. The Hamming weight of the binary expansion of an integer i is called its 2-weight and is denoted by $w_2(i)$. Note that any Boolean function f over $\mathbb{F}_{2^n}$ is also an (n, n)-function because its co-domain $\mathbb{F}_2$ is a subfield of $\mathbb{F}_{2^n}$. For such a function, we have $\delta_0, \delta_{2^n-1} \in \mathbb{F}_2$ and $\delta_{2i} = \delta_i^2$ for every $i \in \{1, \ldots, 2^n - 2\}$ (where the index $2i$ is taken modulo $2^n - 1$). Denoting by tr_n the absolute trace function over $\mathbb{F}_{2^n}$: $tr_n(x) = \sum_{i=0}^{n-1} x^{2^i}$ (which satisfies $tr_n(x^2) = tr_n(x)$ and is valued in $\mathbb{F}_2$), we can then write the univariate representation of f in the form $\delta_0 + tr_n(\sum_{i=0}^{2^n-1} b_i x^i)$ (but there is no more uniqueness of the b_i; the representation with uniqueness is more complex, see e.g. [18])).

[6] In coding theory, we say dual.

2.2 Walsh transform

The *Walsh transform* of a Boolean function f is the function from $\mathbb{F}_2^n$ to $\mathbb{Z}$ defined as follows:

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

where "$\cdot$" is some inner product in $\mathbb{F}_2^n$. The Walsh transform satisfies the so-called *inverse Walsh transform relation*:

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

The Walsh transform of an (n, m)-function F takes value $W_{v \cdot F}(u)$ at input $(u, v) \in \mathbb{F}_2^n \times \mathbb{F}_2^m$:

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

where "$\cdot$" denotes, by abuse of notation, two inner products, one in $\mathbb{F}_2^n$ and one in $\mathbb{F}_2^m$.

The Walsh transform allows to define or to characterize almost all the important cryptographic criteria on Boolean and vectorial functions. For instance, the nonlinearity $nl(f)$ of any Boolean function f (i.e. its Hamming distance to affine Boolean functions) is nicely expressed by means of the Walsh transform: $nl(f) = 2^{n-1} - \frac{1}{2} |\max_{u \in \mathbb{F}_2^n} W_f(u)|$. This allows to prove that $nl(f)$ cannot be larger than $2^{n-1} - 2^{\frac{n}{2}-1}$; the Boolean functions achieving this maximum (with n necessarily even) are called *bent* functions.

2.3 Equivalence notions

Two (n, m)-functions F and G are called affine equivalent if there exist two affine permutations L over $\mathbb{F}_2^m$ and L' over $\mathbb{F}_2^n$ such that $G = L \circ F \circ L'$. In the case of Boolean functions, L is taken equal to identity (which makes the affine equivalence of Boolean functions slightly different from what gives the definition of the affine equivalence of vectorial functions when $m = 1$, since we should normally also consider the case of the identity plus constant 1, but this would not preserve the Hamming weight). More generally, F and G are called extended-affine (EA) equivalent if there exists an affine function L from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$ such that F and $G + L$ are affine equivalent. Still more generally, they 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 affine equivalent (that is, one is the image of the other by an affine permutation over $\mathbb{F}_2^{n+m}$). Writing the affine automorphism mapping $\mathcal{G}_F$ to $\mathcal{G}_G$ as $(x, y) \mapsto (L_1(x, y), L_2(x, y))$ where $L_1 : \mathbb{F}_2^{n+m} \to \mathbb{F}_2^n$ and $L_2 : \mathbb{F}_2^{n+m} \to \mathbb{F}_2^m$ are affine functions, then defining $F_1(x) = L_1(x, F(x))$ and $F_2(x) = L_2(x, F(x))$, we have that F_1 is a permutation of $\mathbb{F}_2^n$ and $G = F_2 \circ F_1^{-1}$, see e.g. [18]. A particular case of CCZ equivalence is between any (n, n)-permutation and its inverse, since the two graphs are the swaps of each other. In the case of Boolean functions, CCZ equivalence reduces to EA equivalence (see e.g. [18]).

We shall say that a notion is affine invariant (respectively, EA invariant, CCZ invariant) if it is preserved by affine equivalence (respectively, EA equivalence, CCZ equivalence). For theoretical and practical reasons, it is important to determine the most general equivalence, among the above notions of equivalence, preserving each notion introduced.

2.4 Differential uniformity, almost perfect nonlinearity

We have seen in the introduction that an (n, m)-function is called differentially δ-uniform if $|\{x \in \mathbb{F}_2^n; F(x) + F(x + a) = b\}| \leq \delta$ for every nonzero $a \in \mathbb{F}_2^n$ and every $b \in \mathbb{F}_2^m$. Differential uniformity is a CCZ-invariant. As observed initially by Nyberg, we have $\delta \geq 2^{n-m}$, with equality if, and only if, F is bent, that is, all the nonzero linear combinations of the coordinate functions of F are bent; such functions exist if, and only if, n is even and $m \leq \frac{n}{2}$, as proved in [48]. We shall speak of almost perfect nonlinear function when $\delta = 2$ and $m = n$. When $m = n - 1$ such functions do not exist since they would be bent and we know that this is not possible unless $n = 2$. When $m \geq n + 1$ we keep the term of differential 2-uniformity. Chabaud and Vaudenay have characterized in [23] the APNness of (n, n)-functions by the Walsh transform: F is APN if, and only if, $\sum_{u \in \mathbb{F}_2^n} \sum_{v \in \mathbb{F}_2^n} W_F^4(u, v) = 3 \cdot 2^{4n} - 2^{3n+1}$ (and this characterization has been generalized in diverse ways to the characterization of differentially uniform functions in [17], including more characterizations of APNness).

3 Two new generalizations of APNness

In this section, we introduce the two extensions of the notion of APN functions and we detail the equivalent ways to define them; we study an example to see how difficult they are to satisfy (and to check), and we study their satisfiability.

Definition 1. *Let $2 \leq k \leq n$ and m be positive integers. An (n, m)-function F is called k-strongly non-normal (resp. kth-order sum-free) if, for every k-dimensional affine subspace (i.e. k-flat) A of $\mathbb{F}_2^n$, the restriction of F to A is not an affine function (resp. the sum $\sum_{x \in A} F(x)$ is nonzero).*

Remark. When F is not kth-order sum-free, it is good (but difficult) to determine over how many k-dimensional flats it sums to 0 (as this is done for $k = 2$ in [43], in relation with APNness). From a practical viewpoint, it is more interesting (but still more difficult) to consider specific affine spaces over which the sum is supposed to be nonzero. In the case where the function is the multiplicative inverse function, we shall study in Subsection 5.2 the sum taken over affine spaces being not linear spaces and in Subsection 5.3 the sum taken over a linear space given by a basis. ◇

Clearly, kth-order sum-freedom implies k-strong non-normality, since the sum of the values taken by an affine function over an affine space of dimension at least 2 equals 0. We shall see that kth-order sum-freedom is a strong property

and k-strong non-normality is a much weaker one.
By the definition of affineness, F is k-strongly non-normal if, and only if, every k-dimensional affine space in $\mathbb{F}_2^n$ contains an affine plane (a 2-dimensional affine space) on which F does not sum to 0. Consequently, if a function is k-strongly non-normal, then it is l-strongly non-normal for every $l \geq k$ (see more in Subsection 4.2). In particular, every APN function is k-strongly non-normal for every $k \geq 2$. Of course, for every $k \leq l \leq n$, F is k-strongly non-normal (resp. kth-order sum-free) if, and only if, its restriction to any l-dimensional affine space of $\mathbb{F}_2^n$ is k-strongly non-normal (resp. kth-order sum-free).

Note that F is kth-order sum-free if, and only if, for every $(k-1)$-dimensional affine subspace A of $\mathbb{F}_2^n$ and every coset $a + A \neq A$, we have $\sum_{x\in A} F(x) \neq \sum_{x\in A} F(x+a)$, that is, $\sum_{x\in A} D_aF(x) \neq 0$ (note that this does not mean that D_aF is $(k-1)$th-order sum-free, despite the similarity, since a must not belong to the underlying linear space of A, for ensuring $a+A \neq A$). Equivalently, for every $(k-1)$-dimensional vector subspace E of $\mathbb{F}_2^n$, the mapping $\phi_E : a \in \mathbb{F}_2^n/E \to \sum_{x\in a+A} F(x)$ is injective. There is then a connection between 3rd-order sum-freedom and the so-called D-property (saying that the union of the image sets of all such mappings ϕ_E, when E ranges over the set of all affine planes, covers $\mathbb{F}_2^m \setminus \{0\}$). D-property is so named in [53] because Dillon was the first to consider it, by showing that it is satisfied by every APN (n,n)-function (see his result reported in [18] after Proposition 161).
An (n,m)-function F is third-order sum-free if and only if, for every $a \neq 0$, the system of equations

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

has no solution (x,y,z,t) with x,y,z,t distinct such that $a \notin \langle x+y, x+z\rangle = \{0, x+y, x+z, x+t\}$ (where we denote by $< S >$ the vector space spanned by a set S in a vector space). Equivalently, for every nonzero $u \in \mathbb{F}_2^n$ and every $v \in \mathbb{F}_2^m$, the system

$$\begin{cases} x+y=u \\ D_aF(x)+D_aF(y)=v \end{cases}$$

has at most one solution as an unordered pair $\{x,y\}$ in a linear hyperplane H such that $u \in H$ and $a \notin H$.
This is a convenient characterization when the derivative of F is simple enough. But when D_aF is complex (we shall see the example of Kasami functions below), it may be better to state the condition by means of F rather than its derivative: for every $v \in \mathbb{F}_2^m$, the system

$$\begin{cases} x+y+z+t=0 \\ F(x)+F(y)+F(z)+F(t)=v \end{cases} \tag{3}$$

does not have two solutions $\{x,y,z,t\}$ and $\{x',y',z',t'\}$ with x,y,z,t distinct in $\mathbb{F}_2^n$ and such that the common value of $x+x'=y+y'=z+z'=t+t'$ does not belong to the direction of the affine plane $\{x,y,z,t\}$ (that is, to the linear

plane $\{0, x+y, x+z, x+t\}$). Note that if F is APN, then we can without loss of generality assume that $v \neq 0$. Note also that the APNness of F is not necessarily implied, since for APNness we need that (3) is never satisfied with $v = 0$ and x, y, z, t distinct while here we can accept one solution.

3.1 Relation of sum-freedom with higher-order derivatives

Any k-dimensional affine subspace A of $\mathbb{F}_2^n$ has the form $a+ < a_1, \ldots, a_k >$ where $a \in \mathbb{F}_2^n$ and $a_1, \ldots, a_k$ are linearly independent in $\mathbb{F}_2^n$ over $\mathbb{F}_2$, and $\sum_{x \in A} F(x)$ equals then the value $D_{a_1} D_{a_2} \ldots D_{a_k} F(a)$ of the so-called kth-order (discrete) derivative $D_{a_1} D_{a_2} \ldots D_{a_k} F$, which is the iteration of the first-order derivative $D_a F(x) = F(x) + F(x+a)$. This is well-known (see [39]). Hence, F is kth-order sum-free if, and only if, every kth-order derivative $D_{a_1} \ldots D_{a_k} F$ with $a_1, \ldots, a_k$ $\mathbb{F}_2$-linearly independent never takes the zero value. This illustrates again the difficulty of proving that a given (n, n)-function is kth-order sum-free: if we for instance represent it as a polynomial over $\mathbb{F}_{2^n}$, we have to prove that some polynomial functions (the derivatives $D_{a_1} \ldots D_{a_k} F$) do not vanish, which is in general quite hard. And indeed, for $k = 2$ already, the proofs by Dobbertin and his co-authors of the APNness of the known APN monomial functions are quite difficult, and we shall see below with the Kasami functions that, even when these proofs could be simplified, checking 3rd-order sum-freedom may still be quite tough.
Another related method consists of showing that the restriction of F to any k-dimensional affine space A, viewed as a (k, m)-function through the choice of a basis of the vector space equal to the direction of A (all such (k, m)-functions are affine equivalent), has algebraic degree k, exactly, but this method seems hard to implement, except when k is close to n.

Remark. The work made in [32] about the Kasami Boolean bent functions $f_\lambda(x) = tr_n(\lambda K_i(x))$ (where λ is not a cube in $\mathbb{F}_{2^n}$) has some similarity with the kth-order sum-freedom of the Kasami functions K_i that we shall tacle below for $k = 3$ (without going to the end of the proof); but it is in fact much simpler: it proves that the derivatives of orders $i-1$ and $i-2$ of $f_\lambda(x)$ do not completely vanish under some conditions on n. To prove this, the author had to calculate $D_{a_1} D_{a_2} \ldots D_{a_k} f_\lambda$ and to prove that for any such λ, there exists x in $\mathbb{F}_{2^n}$ such that $D_{a_1} D_{a_2} \ldots D_{a_k} f_\lambda(x) \neq 0$. For this, it is enough to show that at least one monomial (that the author could choose) in the univariate representation of this latter Boolean function has a nonzero coefficient, while showing kth-order sum-freedom by calculating the kth-order derivative leads to showing that the value of $D_{a_1} D_{a_2} \ldots D_{a_k} F(x)$ is nonzero *for every* $x \in \mathbb{F}_{2^n}$, which needs to take into account all the monomials with their coefficients, and to do a job about them which seems very hard. $\diamond$

3.2 Relationship between the two notions and cryptanalyses

Before addressing the case of modern integral attacks, let us make some preliminary observations.

There is a relation between the sum-freedom of vectorial functions and the resistance of the block ciphers using them as S-boxes to guess-and-determine attacks combined with higher-order differential (HOD) attacks. HOD attacks work when the S-boxes in an algorithm have low algebraic degrees. If a guess allows to know that the input to an S-box lives in some affine space A of dimension k, then:

- The fact that the restriction of F to A is affine, that is, has lowest possible algebraic degree, is the worst case for the resistance against this guess-and-determine attack; k-strong non-normality avoids such worst case to happen.
- The fact that $\sum_{x\in A} F(x) = 0$ is equivalent to the fact that the restriction of F to A has algebraic degree less than k and this may allow the attack to be successful if k is not too large; kth-order sum-freedom avoids such weakness.

There is also some relation between sum-freedom and higher order differential attacks themselves. A vectorial function F over $\mathbb{F}_2^n$ has algebraic degree d if and only if (see e.g. [18]) the sums $\sum_{x\in A} F(x)$ of its values over all $(d+1)$-dimensional affine spaces A are equal to zero (and we know that we can reduce ourselves to A being a vector space of the form $\{x \in \mathbb{F}_2^n; supp(x) \subseteq I\}$ where $I \subset \{1, \ldots, n\}$ has size $d+1$). If F has a larger degree than d, but however sums to zero over many $(d+1)$-dimensional affine spaces A, then this would be an undesirable property, that may allow a distinguisher in some cases, in relation with the early integral attacks. Of course, the weakness implied by the fact that the number of such A is large will be more in favor of the attacker when such A can be detected faster than by exhaustive search. If F is kth-order sum-free, then no A of dimension k exists, which may play at least a role in the probability that an attack is possible.

Let us now address the case of modern integral attacks. The relation could seem much weaker, because the sums are taken over the composition of a multivariate monomial with the function, but it is not so. We have seen in Introduction that a set X has the division property at an order l if and only if the indicator function 1_X of X has algebraic degree at most $n - l$. In particular, it has the division property at the order 1 if and only if X has an even size. Note that having the division property is monotonic in the sense that having the property at the order l implies having the property at any order $l' \leq l$; hence in practice X will have an even size. The following lemma will show that if $\sum_{x\in X} F(x) \neq 0$ then this division property is completely lost (except at the order 1) when the (n, m)-function F is applied i.e. the propagation of the division property is a complete failure.

The image of X by F we need to consider is not the classic one, that is, $F(X) = \{y \in \mathbb{F}_2^m; F^{-1}(y) \neq \emptyset\}$ but the subset:

$$F((X)) = \{y \in \mathbb{F}_2^m; |F^{-1}(y)| \text{ is odd}\},$$

where $|\ldots|$ denotes the cardinality. We prefer using the notation $F((X))$ rather than using $F(X)$ (as in [34]). Note that we have $\sum_{x\in X} F(x) = \sum_{y\in F((X))} y$

(while in general, we have $\sum_{x\in X} F(x) \neq \sum_{y\in F(X)} y$). We also have that if X has an even size then $F((X))$ has also an even size (which is in general not the case for $F(X)$). Let us recall the following lemma (that is known since the last century and that we prove again for self-completeness):

Lemma 1. *Let Y be any subset of $\mathbb{F}_2^m$. If $\sum_{y\in Y} y \neq 0$, then the indicator function 1_Y of Y in $\mathbb{F}_2^m$ has algebraic degree at least $m-1$ (exactly $m-1$ if Y has an even size).*

This is a direct consequence of the properties recalled in Section 2: if Y has an odd size then its indicator function $f = 1_Y$ has algebraic degree m and if it has an even size and satisfies $\sum_{y\in Y} y \neq 0$, then f has algebraic degree exactly $m-1$, because $\sum_{y\in\mathbb{F}_2^m}(yf(y)) \neq 0$. We shall exclude in the following corollary that X (and then $F((X))$) has an odd size, since it will simplify the statement and X with an even size will be the practical situation.

Corollary 1. *Let X be any even size subset of $\mathbb{F}_2^n$. We have $\sum_{x\in X} F(x) \neq 0$ if and only if $F((X))$ does not have the division property at the order 2.*

Integral attacks often (but not always) focus on X being an affine space. We see that kth-order sum-freedom forces the attacker, when searching for sets satisfying the division property, to take X not being a k-dimensional affine subspace.

Remark. The relation between strong non-normality and integral attacks seems much weaker, since the fact that the restriction of F to an affine space X is not affine does not tell much about the algebraic degree of $1_{F((X))}$. ⋄

3.3 The example of the Kasami almost bent functions illustrating the difficulty of studying sum-freedom

Note that the simplest infinite class of APN functions, that of Gold functions, defined over $\mathbb{F}_{2^n}$ by $G_i(x) = x^{2^i+1}$ with $i < n/2$ and $\gcd(i,n) = 1$, are not kth-order sum-free for $k \geq 3$ since they have algebraic degree 2 and sum then to 0 over every k-dimensional affine space with $k \geq 3$. Being APN, they are k-strongly non-normal for every $k \geq 2$ and second-order sum-free, of course.
The Kasami functions are the power functions over $\mathbb{F}_{2^n}$ defined by $K_i(x) = x^{2^{2i}-2^i+1}$, with $i < n/2$ and $\gcd(i,n) = 1$. For any n, K_i is APN (and is then k-strongly non-normal for every $k \geq 2$). If additionally, n is odd, K_i also contributes to an optimal resistance against the linear attack (it is what we call an almost bent function). See more details in [18] and the references therein. Kasami functions are used as S-boxes (with n odd) in the Misty and Kasumi block ciphers [45, 31].
For $i = 1$, the Kasami function equals the cube function x^3 (the simplest Gold function). We know then that it is not 3rd-order sum-free. By curiosity let us check that the system (3) $\begin{cases} x+y+z+t=0 \\ x^3+y^3+z^3+t^3=v \end{cases}$ has two solutions $\{x,y,z,t\}$ and $\{x',y',z',t'\}$ with x,y,z,t distinct in $\mathbb{F}_2^n$ and such that $x+x' = y+y' =$

$z + z' = t + t' \neq 0$. This will show that studying sum-freedom is in some particular cases simple. Given a solution (x, y, z, t) let us check the existence of $a \neq 0$ such that $(x + a, y + a, z + a, t + a)$ is also a solution. The second equation in the system becomes $a(x^2 + y^2 + z^2 + t^2) + a^2(x + y + z + t) = 0$ and it is in fact true for every a since $x + y + z + t = 0$.
For general value of i, since K_i has algebraic degree $i+1$, then for every $k \geq i+2$, it is not kth-order sum-free. And if k divides n, then $\mathbb{F}_{2^k}$ is a subfield of $\mathbb{F}_{2^n}$ and if k is additionally odd, then the restriction of K_i to $\mathbb{F}_{2^k}$ being (according to a result on APN functions by Dobbertin) a permutation of $\mathbb{F}_{2^k}$ and summing then to 0 over $\mathbb{F}_{2^k}$, K_i is not kth-order sum-free[7]. The problem is to determine whether it can be kth-order sum-free for $k \leq i + 1$ (whatever is the parity of n), and not dividing n. Let us consider $k = 3$ and n odd (being not a multiple of 3).

Remark. Since $2^{2i} - 2^i + 1 = \frac{2^{3i}+1}{2^i+1}$, we have then $K_i = G_{3i} \circ G_i^{-1}$ where G_i^{-1} is the compositional inverse of G_i (which is a permutation). Denoting $x = G_i(\mathtt{x})$, $y = G_i(\mathtt{y})$, $z = G_i(\mathtt{z})$, and $t = G_i(\mathtt{t})$, System (3) above becomes:

$$\begin{cases} G_i(\mathtt{x}) + G_i(\mathtt{y}) + G_i(\mathtt{z}) + G_i(\mathtt{t}) = 0 \\ G_{3i}(\mathtt{x}) + G_{3i}(\mathtt{y}) + G_{3i}(\mathtt{z}) + G_{3i}(\mathtt{t}) = v \end{cases}, \tag{4}$$

with $\mathtt{x}, \mathtt{y}, \mathtt{z}, \mathtt{t}$ distinct. Function F is 3rd-order sum-free if and only if this system does not admit two different solutions $(\mathtt{x}, \mathtt{y}, \mathtt{z}, \mathtt{t})$ and $(\mathtt{x}', \mathtt{y}', \mathtt{z}', \mathtt{t}')$ in $\mathbb{F}_{2^n}^4$ with $\mathtt{x}, \mathtt{y}, \mathtt{z}, \mathtt{t}$ distinct and such that

$$G_i(\mathtt{x}) + G_i(\mathtt{x}') = G_i(\mathtt{y}) + G_i(\mathtt{y}') = G_i(\mathtt{z}) + G_i(\mathtt{z}') = G_i(\mathtt{t}) + G_i(\mathtt{t}') \tag{5}$$

does not belong to the direction of the affine plane $\{G_i(\mathtt{x}), G_i(\mathtt{y}), G_i(\mathtt{z}), G_i(\mathtt{t})\}$. Note that since $\mathtt{x}, \mathtt{y}, \mathtt{z}, \mathtt{t}$ are distinct, G_i is a permutation and K_i is APN, we have that v must be nonzero.
It seems difficult to go further. An alternative approach consists in using that for n odd, we have $K_i(x) = G_i \circ L \circ G_i^{-1} + L'$ where $L(x) = x^{2^{2i}} + x$ and $L'(x) = x^{2^{2i}} + x^{2^i} + x$, as proved in [10]. Since L' being linear, it sums to 0 over every affine space of dimension at least 2, we can consider the function $G_i \circ L \circ G_i^{-1}$ instead of K_i. Then, still denoting $x = G_i(\mathtt{x})$, $y = G_i(\mathtt{y})$, $z = G_i(\mathtt{z})$, and $t = G_i(\mathtt{t})$, we have instead of System (4):

$$\begin{cases} G_i(\mathtt{x}) + G_i(\mathtt{y}) + G_i(\mathtt{z}) + G_i(\mathtt{t}) = 0 \\ G_i \circ L(\mathtt{x}) + G_i \circ L(\mathtt{y}) + G_i \circ L(\mathtt{z}) + G_i \circ L(\mathtt{t}) = v, \end{cases} \tag{6}$$

with the same condition (5). Here also, since $\mathtt{x}, \mathtt{y}, \mathtt{z}, \mathtt{t}$ are distinct, G_i is a permutation and $G_i \circ L \circ G_i^{-1}$ is APN, we have that v must be nonzero. And it seems difficult to go further. ◇

We checked by a computer investigation that $K_2(x) = x^{13}$ is 3rd-order sum-free over $\mathbb{F}_{2^5}$, but is not over $\mathbb{F}_{2^7}$ nor over $\mathbb{F}_{2^9}$. Similarly, $K_3(x) = x^{57}$ is not

[7] This observation is true more generally for any APN power function when k is an odd divisor of n.

3rd-order sum-free over $\mathbb{F}_{2^7}$ nor over $\mathbb{F}_{2^9}$, and $K_4(x) = x^{241}$ is not 3rd-order sum-free over $\mathbb{F}_{2^9}$.

3.4 Existence of k-strongly non-normal functions and of kth-order sum-free functions

The existence of k-strongly non-normal (n, n)-functions for every $k \geq 2$ is clear for every $n \geq k$, since all APN functions (which exist for every n) are k-strongly non-normal for every $2 \leq k \leq n$. Moreover, given a k-strongly non-normal (n, m)-function F, any $(n, m+1)$-function obtained by adding any coordinate function to F is k-strongly non-normal, and any $(n-1, m)$-function obtained by restricting F to a hyperplane is k-strongly non-normal if $k \leq n-1$; we deduce then the existence of k-strongly non-normal (n, m)-functions for every $m \geq n \geq k \geq 2$.

We know that differentially 2-uniform (n, m)-functions (that is, 2-strongly non-normal (n, m)-functions) do not exist for $m < n$ when $n > 2$. But for $k > 2$, the set of those triples (n, m, k) for which k-strongly non-normal (n, m)-functions do not exist is not clear in general. For $m = 1$, we know (see [18, Subsection 6.1.10]) that the restriction of any bent Boolean function f (n even) to any affine space of dimension at least $\frac{n}{2} + 1$ cannot be affine, so k-strongly non-normal Boolean functions exist for any $k \geq \frac{n}{2} + 1$. This is then also true for (n, m)-functions for $m \geq 1$. The question of the existence of non-normal and strongly non-normal bent functions has been an open question[8] until [13] provides a non-normal bent function for any even $n \geq 10$ and a strongly non-normal bent function for any even $n \geq 14$. For n odd, we can take f as the restriction of an $(n+1)$-variable bent function and so k-strongly non-normal Boolean functions exist for any $k \geq \frac{n+1}{2} + 1$ (which is also true for any $m \geq 1$) and even for $k \geq \frac{n+1}{2}$ if n is large enough. It seems difficult to determine, for every (n, k), the maximal value of m such that there does not exist any k-strongly non-normal function. Actually, for $m = 1$, determining the region of all (n, k) such that k-strongly non-normal Boolean functions exist seems open. What is known is that every Boolean function on $\mathbb{F}_2^n$ with $n \leq 7$ is $\lfloor \frac{n}{2} \rfloor$-normal, and (as we already recalled) almost all (in the sense of probability) n-variable Boolean functions are k_n-strongly non-normal when the sequence k_n satisfies $k_n \geq c \log_2 n$ for some $c > 1$ (see [18, Proposition 34]), and this of course is also true for $m \geq 1$. There exists then a positive integer N such that, for every $n \geq N$, k_n-nonnormal (n, m)-functions exist. For $k_n = \lfloor \frac{n}{2} \rfloor$ and $m = 1$, we can take $N = 12$. Also, since it is proved (cf. [18, Proposition 33]) that k-weakly normal Boolean functions on $\mathbb{F}_2^n$ have a nonlinearity smaller than or equal to $2^{n-1} - 2^{k-1}$, when Boolean functions of a larger nonlinearity exist, these functions are k-strongly non-normal, but the complete determination of the region of those triples (n, m, k) such that k-strongly non-normal (n, m)-functions exist is open, even if, according to the results recalled above, the largest part of it is

[8] The contrary had even been conjectured.

known since for k as low as $c\log_2(n)$ such functions exist asymptotically.

Remark. In an approach by the ANF, we can try to consider the ANF of the restriction of a general (n,m)-function to an affine space, and look whether this restriction has algebraic degree larger than 1. Having the choice of the ANF of the (n,m)-function may help ensuring that the restriction of the function to affine spaces of equations "$x_i = a_i; i \in I$" has full degree. For the other affine spaces, by Jordan's reduction, an affine space has equations $x_{i_j} = L_j(x_{i_{j+1}}, x_{i_{j+2}}, \dots) + a_j$ and this seems harder. $\diamond$

We shall see that kth-order sum-freedom has a more complex behavior than k-strong non-normality, and even the question of the existence of functions satisfying it is not straighforward. We need then to address it first, for avoiding studying an empty class for $k > 2$. Let us show that the cube function (which is APN and then second-order sum-free) is the first element of an infinite sequence of kth-order sum-free (n,n)-functions.

Proposition 1. *Let $2 \leq k \leq n$ be integers. Let $P_k(x)$ be the power function x^{2^k-1} over $\mathbb{F}_{2^n}$. Denoting by G_k the set of bijections from $\{1,\dots,k\}$ to $\{0,\dots,k-1\}$, we have:*

$$D_{a_1}\dots D_{a_k}P_k(x) = \sum_{\sigma\in G_k}\prod_{i=1}^{k} a_i^{2^{\sigma(i)}} = \begin{vmatrix} a_1 & a_2 & \dots & a_k \\ a_1^2 & a_2^2 & \dots & a_k^2 \\ \vdots & \vdots & \dots & \vdots \\ a_1^k & a_2^k & \dots\dots & a_k^k \end{vmatrix} = \prod_{l\in\mathbb{F}_2^k, l\neq 0}\left(\sum_{i=1}^{k} l_i a_i\right) \quad (7)$$

and P_k is kth-order sum-free.

Proof. We have $2^k - 1 = \sum_{i=1}^{k} 2^{i-1}$, hence $P_k(x) = \prod_{i=1}^{k} L_i(x)$, where $L_i(x) = x^{2^{i-1}}$. It is well-known (as first proved in [14]) that, if $L_0,\dots,L_{k-1}$ are linear, then for every $a_1,\dots,a_k$ in $\mathbb{F}_{2^n}$:

$$D_{a_1}\dots D_{a_k}\Big(\prod_{i=0}^{k-1} L_i\Big)(x) = \sum_{\sigma\in G_k}\prod_{i=1}^{k} L_{\sigma(i)}(a_i). \quad (8)$$

Let us recall the proof of Relation (8), since the thesis [14] is not easily available and is in French.
For every function F, we have the classic formula (which is a direct consequence of the definition of derivatives): $D_{a_1}\dots D_{a_k}F(x) = \sum_{\epsilon\in\mathbb{F}_2^k} F(\sum_{i=1}^{k}\epsilon_i a_i)$. Assuming that the formula $D_{a_1}\dots D_{a_k}(\prod_{i=0}^{k-1} L_i)(x) = \sum_{\sigma\in G_k}\prod_{i=1}^{k} L_{\sigma(i)}(a_i)$ (which is obviously true for $k=1$) is true for some value of $k \geq 1$ and any $a_1,\dots,a_k$, let us prove it for the value $k+1$ and any $a_1,\dots,a_{k+1}$. We have, denoting

$F(x) = \prod_{i=0}^{k-1} L_i(x)$:

$$D_{a_1} \dots D_{a_{k+1}} \Big(\prod_{i=0}^{k} L_i \Big)(x) = \sum_{\epsilon \in \mathbb{F}_2^{k+1}} \Big[F\big(\sum_{i=1}^{k+1} \epsilon_i a_i \big) L_k\big(\sum_{j=1}^{k+1} \epsilon_j a_j \big) \Big]$$
$$= \sum_{\epsilon \in \mathbb{F}_2^{k+1}} \Big[F\big(\sum_{i=1}^{k+1} \epsilon_i a_i \big) \Big(\sum_{j=1}^{k+1} \epsilon_j L_k\big(a_j\big) \Big) \Big]$$
$$= \sum_{j=1}^{k+1} \Big(L_k\big(a_j\big) \Big) \sum_{\substack{\epsilon \in \mathbb{F}_2^{k+1} \\ \epsilon_j = 1}} \Big[F\big(\sum_{i=1}^{k+1} \epsilon_i a_i \big) \Big].$$

Hence we have:

$$D_{a_1} \dots D_{a_{k+1}} \Big(\prod_{i=0}^{k} L_i \Big)(x) =$$

$$\sum_{j=1}^{k+1} \Big(L_k\big(a_j\big) \Big) \Big(D_{a_1} \dots D_{a_{j-1}} D_{a_{j+1}} \dots D_{a_{k+1}} \Big(\prod_{i=0}^{k-1} L_i \Big)(x + a_j) \Big) =$$

$$\sum_{j=1}^{k+1} \Big(L_k\big(a_j\big) \Big) \Big(D_{a_1} \dots D_{a_{j-1}} D_{a_{j+1}} \dots D_{a_{k+1}} \Big(\prod_{i=0}^{k-1} L_i \Big)(x) \Big), \tag{9}$$

this last equality coming from the fact that $D_{a_1} \dots D_{a_{k+1}} \Big(\prod_{i=0}^{k-1} L_i \Big)(x) = 0$, thanks to the induction hypothesis (or to the fact that any $(k+1)$th-order derivative of any vectorial function of algebraic degree at most k equals 0). The induction hypothesis now applied to (9) completes the proof of (8).
We have then:

$$D_{a_1} \dots D_{a_k} P_k(x) = \sum_{\sigma \in G_k} \prod_{i=1}^{k} a_i^{2^{\sigma(i)}}.$$

The two other equalities in (7) are well-known since the involved determinant is a Moore determinant [47]. □

As in the case of k-strong non-normality, given a kth-order sum-free (n, m)-function F, any $(n, m+1)$-function obtained by adding any coordinate function to F is kth-order sum-free, and any $(n-1, m)$-function obtained by restricting F to a hyperplane is kth-order sum-free if $k \leq n-1$. We deduce then from Proposition 1 the existence of kth-order sum-free (n, m)-functions for every $m \geq n \geq k \geq 2$. Here also, for $m < n$, no second-order sum-free function exists for $n > 2$, but for $k > 2$, the condition on n, m, k such that no (n, m)-function can be kth-order sum-free is not clear, except when $k = n$, since the existence is then obvious whatever is m.
Note that, for $m = 1$, no kth-order sum-free Boolean function f can exist, whatever is $1 \leq k \leq n-1$, since there are of course only two possible values 0

and 1 for the sums $\sum_{x\in A} f(x)$, and it is impossible that all sums such that A has dimension k take value 1. Indeed, even for $k = n-1$ (for which the number of k-dimensional affine subspaces is the smallest), it is impossible because a Boolean function f having odd Hamming weight restrictions to the four hyperplanes of equations $x_n = 0, x_n = 1, x_{n-1} = 0$ and $x_{n-1} = 1$ has necessarily even Hamming weight restrictions to the hyperplanes of equations $x_n + x_{n-1} = 0$, and $x_n + x_{n-1} = 1$ (whatever is the parity ϵ of the weight of the restriction to $x_n = x_{n-1} = 0$, the weights of the restrictions to the affine spaces of equations "$x_n = 1$ and $x_{n-1} = 0$" and "$x_n = 0$ and $x_{n-1} = 1$" must have the same parity $\epsilon + 1$).
The question is to determine, for each k, from what value of m do kth-order sum-free (n, m)-functions exist. Note that, according to Proposition 1, if we start from the (n, n)-function $P_k(x) = x^{2^k-1}$ over $\mathbb{F}_{2^n}$, we do not obtain a kth-order sum-free (n, m)-function with $m < n$ by linearly projecting the image x^{2^k-1} over an m-dimensional subspace, when $\gcd(n, k) = 1$. Indeed, since multiplying each a_i by a same nonzero coefficient λ multiplies $\prod_{l\in\mathbb{F}_2^k, l\neq 0}\left(\sum_{i=1}^k l_i a_i\right)$ by λ^{2^k-1}, all nonzero values in $\mathbb{F}_{2^n}$ are reached by $D_{a_1}\dots D_{a_k}P_k(x)$ and some are necessarily mapped to 0 by the projection.
It seems difficult to determine, for every (n, k), the maximal value of m such that there does not exist any kth-order sum-free function.
Considering the problem with the viewpoint of the ANF, it is easy to build, for some n, m, k, functions of algebraic degree at least k which are not kth-order sum-free. For instance, if we take for F a function of degree k, if its degree k part is not homogeneous, then considering a degree k monomial missing in its ANF and fixing to zero the variables that are not present in it, we get a restriction of degree less than k. But this does not tell whether functions satisfying the property can exist when $m < n$.

4 Properties of the two notions

Let us go into more details with the properties of the two notions that we briefly saw after introducing their definition.
We first state explicitly what we observed immediately after Definition 1 and study the converse:

Proposition 2. *For every $2 \le k \le n$ and m, if an (n, m)-function is kth-order sum-free, then it is k-strongly non-normal.*

About the converse of this implication:

- for $k = 2$, it is of course valid, since the two notions coincide (with APNness),
- for $k \ge 3$, the converse of Proposition 2 is not true; there exist indeed, for every $n \ge k$ and every $m \ge 1$, k-strongly non-normal (n, m)-functions which are not kth-order sum-free, because there are non-affine Boolean functions, even quadratic ones, which sum to zero, that is, which have an even Hamming weight.

An interesting particular case in this regard is when $k \geq 2$ is a divisor of n and F is a polynomial function over $\mathbb{F}_{2^n}$ whose coefficients all belong to $\mathbb{F}_{2^k}$ (in other words, $F(x^{2^k})$ equals $(F(x))^{2^k}$ for every $x \in \mathbb{F}_{2^n}$; for $k = 1$, such F is called an idempotent). Then F maps the subfield $\mathbb{F}_{2^k}$ (which is a k-dimensional vector space) into itself, and if it maps $\mathbb{F}_{2^k}$ onto itself, that is, if it is a permutation of $\mathbb{F}_{2^k}$, then $\sum_{x\in\mathbb{F}_{2^k}} F(x) = 0$ and F is then not kth-order sum-free while F can be k-strongly non-normal. For instance, an APN power (n, n)-function F cannot be kth-order sum-free for k an odd divisor of n (we know from Dobbertin, as reported in [18], that F is then a permutation of $\mathbb{F}_{2^k}$). Of course, if k is even and such that F is a permutation of $\mathbb{F}_{2^k}$ (for instance, when F itself is a permutation), we have the same situation.

4.1 Algebraic degree

Recall that any (n, m)-function has an algebraic degree bounded above by some integer $d \leq n$ if, and only if, it sums to zero over every affine space whose dimension is strictly larger than d (this is well-known for Boolean functions, see e.g. [18], and it directly generalizes to (n, m)-functions). We have then (as we already observed above about Gold and Kasami functions):

Proposition 3. *For every $2 \leq k \leq n$ and every m, all kth-order sum-free (n, m)-functions have necessarily algebraic degree at least k (and this latter necessary condition is also sufficient if $k = n$).*

This makes a difference with k-strong non-normality, since all APN functions, among which are quadratic ones, are k-strongly non-normal for every $k \geq 2$.
In fact, an (n, m)-function F is kth-order sum-free if, and only if, the restriction of F to any k-dimensional affine space, viewed as a k-variable function through the choice of a basis of the vector space equal to the direction of this affine space (i.e. such that the affine space is a coset - a translate - of the linear space), has algebraic degree k.

Remark. Since the algebraic degree of the indicator 1_A of any k-dimensional affine subspace A of $\mathbb{F}_2^n$ equals $n-k$ and the algebraic degree of the product of a Boolean function and a vectorial function is bounded above by the sum of their algebraic degrees, F cannot be kth-order sum-free when $n - k + d_{alg}(F) < n$, since the algebraic degree of $1_A F$ is then smaller than n and $\sum_{x\in A} F(x) = \sum_{x\in\mathbb{F}_2^n} 1_A(x)F(x)$ equals then 0. This gives again that if F is kth-order sum-free, then $d_{alg}(F) \geq k$, as in Proposition 3. It provides additionally that if $d_{alg}(F) = k$, then F is kth-order sum-free if, and only if, $d_{alg}(1_A F) = d_{alg}(1_A) + d_{alg}(F)$, for every k-dimensional affine space A. We say then that F has no degree-drop k-dimensional affine space (see [21] where the case of Boolean functions is studied). $\diamond$

4.2 Monotonicity/non-monotonicity

Monotonicity of strong non-normality We have seen in Section 3 that if a function is k-strongly non-normal then it is l-strongly non-normal for every $l \geq k$ (a slightly different way of seeing this is by observing that the restriction of every affine function to every affine subspace of its domain is affine). The notion is then monotonic. In particular, k-strong non-normality for $k \geq 3$ is a generalization (and a weakening) of APNness, as is differential uniformity, but differently.

The monotonicity of the notion is strict. For instance, there are 3-strongly non-normal functions which are not APN: given an APN (n, n)-function F and a point $a \in \mathbb{F}_2^n$, let $G(x) = \begin{cases} F(x) \text{ if } x \neq a \\ b \text{ if } x = a \end{cases}$, where b is chosen so that G is not APN (it is easy to find b; it is not even clear whether any APN function F and any points a and $b \neq F(a)$ can exist such that G is APN, see [11]). Function G is 3-strongly non-normal because, for every 3-dimensional affine space A, there exists an affine plane included in $A \setminus \{a\}$ and since G coincides with F on this affine plane, it is not affine on it; hence G is not affine on A.

We shall see in Subsection 5.1 that the multiplicative inverse function is also an example of a 3-strongly non-normal function that is not APN, when n is even.

Non-monotonicity of sum-freedom Propositions 1 and 3 imply the existence of functions that are kth-order sum-free and not lth-order sum-free for some $l \geq k$. Note also that kth-order sum-freedom is not decreasing monotonic either (that is, preserved when we decrease k). For instance, take a non-APN (n, n)-function of algebraic degree n; then F is nth-order sum-free and it is not second-order sum-free.

We see that the behavior of kth-order sum-freedom is pretty complex, while it is more related to integral attacks (and in particular, higher order differential attacks) than k-strong non-normality.

4.3 Characterization by the Walsh transform

It is usual, when a notion is studied, to try to characterize it by the Walsh transform. Many important cryptographic properties of Boolean and vectorial functions can be translated in terms of the Walsh transform. Having characterizations of the notions by the Walsh transform gives then a chance of relating them by bounds. When Chabaud and Vaudenay studied in [23] the notions of almost perfect nonlinearity and almost bentness, they characterized them by the Walsh transform (and after that, it took 24 years before a characterization could be found for differentially uniform functions in [17]). Unfortunately, their characterization did not really allow results on the nonlinearity of APN functions, so far. Partial results were found in [19] thanks to another characterization by the Walsh transform.

We give now characterizations of the two notions by the Walsh transform. They

are rather complex; this was expected since both notions are more complex than APNness, and even a slight increase in the complexity of the definition of a notion implies a much greater increase in the difficulty of finding a characterization by the Walsh transform, and in the complexity of the characterization that we can obtain.

***k*-strong non-normality** Given a k-dimensional affine subspace A (over $\mathbb{F}_2$) of $\mathbb{F}_2^n$, the restriction of $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ to A is affine if and only if, for every $v \in \mathbb{F}_2^m$, the Boolean function $v \cdot F$, where "$\cdot$" is some inner product in $\mathbb{F}_2^m$, is affine on A. According to the Parseval relation (which, for a k-variable Boolean function f, writes $\sum_{u\in\mathbb{F}_2^k} W_f^2(u) = 2^{2k}$, see e.g. [18]) and to the inverse Walsh transform relation, this is equivalent to $\sum_{x\in A}(-1)^{v\cdot F(x)+u\cdot x} \in \{0, \pm 2^k\}$ for all $u \in \mathbb{F}_2^n$, where (by an abuse of notation), we use the same notation "$\cdot$" for inner products in $\mathbb{F}_2^n$ and $\mathbb{F}_2^m$. Hence, the restriction of F to A is affine if and only if, for every $u \in \mathbb{F}_2^n$ and $v \in \mathbb{F}_2^m$, we have:

$$\Big(\sum_{x\in A}(-1)^{v\cdot F(x)+u\cdot x}\Big)^2\Big(2^{2k} - \Big(\sum_{x\in A}(-1)^{v\cdot F(x)+u\cdot x}\Big)^2\Big) = 0.$$

Note that (the left-hand side of) this latter expression is always non-negative for every function F. Therefore, the restriction of F to A is affine if and only if:

$$\sum_{u\in\mathbb{F}_2^n, v\in\mathbb{F}_2^m}\Big(\sum_{x\in A}(-1)^{v\cdot F(x)+u\cdot x}\Big)^2\Big(2^{2k} - \Big(\sum_{x\in A}(-1)^{v\cdot F(x)+u\cdot x}\Big)^2\Big) = 0.$$

For every $v \in \mathbb{F}_2^m$, we have:

$$\sum_{u\in\mathbb{F}_2^n}\Big(\sum_{x\in A}(-1)^{v\cdot F(x)+u\cdot x}\Big)^2 = \sum_{x,y\in A}(-1)^{v\cdot(F(x)+F(y))}\sum_{u\in\mathbb{F}_2^n}(-1)^{u\cdot(x+y)} = 2^{n+k}.$$

Writing $A = a+E$, where $a \in \mathbb{F}_2^n$ and E is a k-dimensional vector space, the Poisson summation formula (see *e.g.* [18, Relation (2.41)]) writes $\sum_{x\in A}(-1)^{v\cdot F(x)+u\cdot x} = \pm 2^{k-n}\sum_{w\in u+E^\perp}(-1)^{a\cdot w}W_F(w,v)$, where $E^\perp = \{w \in \mathbb{F}_2^n; \forall x \in E, w \cdot x = 0\}$, and therefore, we have:

$$\sum_{u\in\mathbb{F}_2^n, v\in\mathbb{F}_2^m}\Big(\sum_{x\in A}(-1)^{v\cdot F(x)+u\cdot x}\Big)^4 =$$

$$2^{4k-4n}\sum_{u\in\mathbb{F}_2^n, v\in\mathbb{F}_2^m}\Big(\sum_{w\in u+E^\perp}(-1)^{a\cdot w}W_F(w,v)\Big)^4 =$$

$$2^{4k-4n}\sum_{\substack{u\in\mathbb{F}_2^n, v\in\mathbb{F}_2^m\\(T_1,T_2,T_3,T_4)\in(E^\perp)^4}}(-1)^{a\cdot\sum_{i=1}^4 T_i}\prod_{i=1}^4 W_F(u+T_i,v).$$

We deduce:

Proposition 4. *For every $2 \leq k \leq n$ and m, any (n,m)-function is k-strongly non-normal if and only if, for every $a \in \mathbb{F}_2^n$ and every k-dimensional vector subspace E of $\mathbb{F}_2^n$, we have:*

$$2^{n+m+3k} - 2^{4k-4n} \sum_{\substack{u\in\mathbb{F}_2^n, v\in\mathbb{F}_2^m \\ (T_1,T_2,T_3,T_4)\in(E^\perp)^4}} (-1)^{a\cdot\sum_{i=1}^4 T_i} \prod_{i=1}^4 W_F(u+T_i,v) > 0.$$

Remark. Paradoxically (since the properties of k-strong non-normality are in general rather easy to show contrary to those of kth-order sum-freedom), it seems difficult to characterize k-strong non-normality in a simpler way by means of the Walsh transform, while we shall be able to characterize below kth-order sum-freedom by a single formula. ◇

***k*th-order sum-freedom** Still taking $A = a + E$, where E is a k-dimensional $\mathbb{F}_2$-vector subspace of $\mathbb{F}_2^n$, we have $\sum_{x\in A} F(x) \neq 0$ if and only if we have: $\sum_{v\in\mathbb{F}_2^m} (-1)^{v\cdot\left(\sum_{x\in A} F(x)\right)} = 0$. Hence, fixing E and letting a range over $\mathbb{F}_2^n$, we have $\sum_{x\in a+E} F(x) \neq 0$ for every $a \in \mathbb{F}_2^n$ if and only if:

$$\begin{aligned}\sum_{a\in\mathbb{F}_2^n} \left(\sum_{v\in\mathbb{F}_2^m} (-1)^{v\cdot(\sum_{x\in a+E} F(x))} \right)^2 &= \sum_{\substack{a\in\mathbb{F}_2^n \\ v,v'\in\mathbb{F}_2^m}} (-1)^{(v+v')\cdot(\sum_{x\in E} F(a+x))} \\ &= 2^m \sum_{\substack{a\in\mathbb{F}_2^n \\ v\in\mathbb{F}_2^m}} (-1)^{v\cdot(\sum_{x\in E} F(a+x))} \qquad (10)\end{aligned}$$

equals 0, that is, using the inverse Walsh transform formula as follows: $(-1)^{v\cdot F(a+x)} = 2^{-n} \sum_{u\in\mathbb{F}_2^n} W_F(u,v)(-1)^{(a+x)\cdot u}$, and denoting by $U = (u_x)_{x\in E}$ the elements of $(\mathbb{F}_2^n)^E$:

$$\sum_{\substack{a\in\mathbb{F}_2^n \\ v\in\mathbb{F}_2^m}} \sum_{U\in(\mathbb{F}_2^n)^E} \left(\prod_{x\in E} W_F(u_x,v)(-1)^{\sum_{x\in E}(a+x)\cdot u_x} \right) = \qquad (11)$$

$$\sum_{U\in(\mathbb{F}_2^n)^E} \sum_{v\in\mathbb{F}_2^m} \left(\prod_{x\in E} W_F(u_x,v) \right) \left(\sum_{a\in\mathbb{F}_2^n} (-1)^{a\cdot(\sum_{x\in E} u_x)} \right) (-1)^{\sum_{x\in E} x\cdot u_x} = 0,$$

that is:

$$\sum_{\substack{U\in(\mathbb{F}_2^n)^E \\ \sum_{x\in E} u_x = 0}} \sum_{v\in\mathbb{F}_2^m} \left(\prod_{x\in E} W_F(u_x,v) \right) (-1)^{\sum_{x\in E} x\cdot u_x} = 0, \qquad (12)$$

since $\sum_{a\in\mathbb{F}_2^n} (-1)^{a\cdot(\sum_{x\in E} u_x)}$ equals 0 if $\sum_{x\in E} u_x \neq 0$.
Let us now write $E =< a_1, \ldots, a_k >$, where $(a_1, \ldots, a_k)$ is a basis of E over

$\mathbb{F}_2$. Then writing $\sum_{i=1}^{k} x_i a_i$ (where $x = (x_1, \ldots, x_k) \in \mathbb{F}_2^k$) instead of $x \in E$, Relation (12) becomes:

$$\sum_{\substack{U \in (\mathbb{F}_2^n)^{\mathbb{F}_2^k} \\ \sum_{x \in \mathbb{F}_2^k} u_x = 0}} \sum_{v \in \mathbb{F}_2^m} \left(\prod_{x \in \mathbb{F}_2^k} W_F(u_x, v) \right) (-1)^{\sum_{x \in \mathbb{F}_2^k} (\sum_{i=1}^{k} x_i a_i) \cdot u_x} = 0. \qquad (13)$$

When $a_1, \ldots, a_k$ are not $\mathbb{F}_2$-linearly independent, going back from (13) to the corresponding versions of (11) and (10), we see that, the expression on the left-hand side of (13) has value $2^{-n} \sum_{\substack{a \in \mathbb{F}_2^n \\ v \in \mathbb{F}_2^m}} \sum_{U \in (\mathbb{F}_2^n)^{\mathbb{F}_2^k}} \left(\prod_{x \in E} W_F(u_x, v) (-1)^{\sum_{x \in E} (a+x) \cdot u_x} \right) =$ $2^{-n+2^k n} \sum_{\substack{a \in \mathbb{F}_2^n \\ v \in \mathbb{F}_2^m}} (-1)^{v \cdot (\sum_{x \in \mathbb{F}_2^k} F(a + \sum_{i=1}^{k} x_i a_i))}$, and since $\sum_{x \in \mathbb{F}_2^k} F(a + \sum_{i=1}^{k} x_i a_i) =$ $D_{a_1} \ldots D_{a_k} F(a)$ equals 0 for every a (because $a_1, \ldots, a_k$ are not $\mathbb{F}_2$-linearly independent), the expression on the left-hand side of (13) equals $2^{2^k n + m}$, for each k-tuple $(a_1, \ldots, a_k)$ with $a_1, \ldots, a_k$ not $\mathbb{F}_2$-linearly independent (since the value of $\sum_{\substack{a \in \mathbb{F}_2^n \\ v \in \mathbb{F}_2^m}} (-1)^{v \cdot (\sum_{x \in \mathbb{F}_2^k} F(a + \sum_{i=1}^{k} x_i a_i))}$ when $\sum_{x \in \mathbb{F}_2^k} F(a + \sum_{i=1}^{k} x_i a_i) = 0$ is 2^{n+m}). The number of k-tuples $(a_1, \ldots, a_k)$ of linearly dependent elements equals $2^{kn} - (2^n - 1)(2^n - 2) \cdots (2^n - 2^{k-1})$. Hence, F is kth-order sum-free if, and only if, the sum for $(a_1, \ldots, a_k)$ ranging over $(\mathbb{F}_2^n)^k$ of the left-hand side of (13) is equal to $2^{2^k n + m} \left(2^{kn} - (2^n - 1)(2^n - 2) \cdots (2^n - 2^{k-1}) \right)$. This sum equals:

$$\sum_{\substack{U \in (\mathbb{F}_2^n)^{\mathbb{F}_2^k} \\ \sum_{x \in \mathbb{F}_2^k} u_x = 0}} \sum_{v \in \mathbb{F}_2^m} \left(\prod_{x \in \mathbb{F}_2^k} W_F(u_x, v) \right) \sum_{(a_1, \ldots, a_k) \in (\mathbb{F}_2^n)^k} (-1)^{\sum_{x \in \mathbb{F}_2^k} (\sum_{i=1}^{k} x_i a_i) \cdot u_x} =$$

$$\sum_{\substack{U \in (\mathbb{F}_2^n)^{\mathbb{F}_2^k} \\ \sum_{x \in \mathbb{F}_2^k} u_x = 0}} \sum_{v \in \mathbb{F}_2^m} \left(\prod_{x \in \mathbb{F}_2^k} W_F(u_x, v) \right) \sum_{(a_1, \ldots, a_k) \in (\mathbb{F}_2^n)^k} \prod_{i=1}^{k} (-1)^{a_i \cdot (\sum_{x \in \mathbb{F}_2^k} x_i u_x)} =$$

$$\sum_{\substack{U \in (\mathbb{F}_2^n)^{\mathbb{F}_2^k} \\ \sum_{x \in \mathbb{F}_2^k} u_x = 0}} \sum_{v \in \mathbb{F}_2^m} \left(\prod_{x \in \mathbb{F}_2^k} W_F(u_x, v) \right) \prod_{i=1}^{k} \left(\sum_{a \in \mathbb{F}_2^n} (-1)^{a \cdot (\sum_{x \in \mathbb{F}_2^k} x_i u_x)} \right) =$$

$$2^{nk} \sum_{\substack{U \in (\mathbb{F}_2^n)^{\mathbb{F}_2^k}; \forall i = 1, \ldots, k, \\ \sum_{x \in \mathbb{F}_2^k} x_i u_x = \sum_{x \in \mathbb{F}_2^k} u_x = 0}} \sum_{v \in \mathbb{F}_2^m} \left(\prod_{x \in \mathbb{F}_2^k} W_F(u_x, v) \right).$$

Note that U can be identified with the (k,n)-function $x \mapsto u_x$ and the condition $\forall i = 1,\ldots,k; \sum_{x\in\mathbb{F}_2^k} x_i u_x = \sum_{x\in\mathbb{F}_2^k} u_x = 0$, is that each of its n coordinate functions $x \mapsto (u_x)_j$, $j = 1,\ldots,n$; is orthogonal to each of the k-variable coordinate Boolean functions $x \mapsto x_i$, $i = 1,\ldots,k$; and is also orthogonal to the k-variable constant function 1, that is, each coordinate function of U belongs to the dual of the vector space over $\mathbb{F}_2$ of affine Boolean functions, called the first-order Reed-Muller code $RM(1,k)$. We know (see e.g. [44]) that the dual of $RM(1,k)$ equals $RM(k-2,k)$, the vector space of Boolean functions of algebraic degree at most $k-2$, generated by all monomials $\prod_{i\in I} x_i$ where $0 \le |I| \le k-2$. The characterization of kth-order sum-freedom writes then:

Proposition 5. *For every $2 \le k \le n$ and m, any (n,m)-function is kth-order sum-free if and only if:*

$$\sum_{U\in[RM(k-2,k)]^n} \sum_{v\in\mathbb{F}_2^m} \left(\prod_{x\in\mathbb{F}_2^k} W_F(u_x, v) \right) =$$

$$2^{n(2^k-k)+m} \left(2^{kn} - (2^n-1)(2^n-2)\cdots(2^n-2^{k-1}) \right).$$

Note that for $m = n$ and $k = 2$, Proposition 5 gives a characterization of APN (n,n)-functions, and since $RM(k-2,k)$ equals $\{(0,0,0,0),(1,1,1,1)\}$, the condition in this characterization writes:

$$\sum_{u\in\mathbb{F}_2^n} \sum_{v\in\mathbb{F}_2^n} W_F^4(u,v) = 2^{3n}\left(2^{2n} - (2^n-1)(2^n-2)\right) = 3\cdot 2^{4n} - 2^{3n+1};$$

this is exactly the Chabaud-Vaudenay characterization [23].

Remark. Relation (12) provides a characterization by the Walsh transform of the fact that an (n,m)-function sums to nonzero values over all (parallel) affine spaces of a given direction E. The method for proving Proposition 5, when we consider it in the particular case of $k = 2$, is then different from the known proof of the Chabaud-Vaudenay characterization of APN functions, which addresses all affine planes globally. Relation (12) for $k = 2$ writes

$$\sum_{\substack{(u_1,u_2,u_3,u_4)\in(\mathbb{F}_2^n)^4 \\ \sum_{i=1}^4 u_i = 0}} \sum_{v\in\mathbb{F}_2^m} \left(\prod_{i=1}^4 W_F(u_i,v) \right) (-1)^{\sum_{i=1}^4 x_i\cdot u_i} = 0,$$

where $E = \{x_1, x_2, x_3, x_4)$ (with, then, $x_1,\ldots,x_4$ distinct and $x_4 = x_1+x_2+x_3$), or equivalently:

$$\sum_{\substack{(u_1,u_2,u_3) \\ \in(\mathbb{F}_2^n)^3; v\in\mathbb{F}_2^m}} \Big(\prod_{i=1}^3 W_F(u_i,v) \Big) W_F(u_1+u_2+u_3, v)(-1)^{x_1\cdot u_1 + x_2\cdot u_2 + x_3\cdot u_3 + (x_1+x_2+x_3)\cdot(u_1+u_2+u_3)}$$

$$= 0.$$

It is a new result in itself. ◇

4.4 Invariance under equivalence

In Boolean function theory, when we study a property of (n, m)-functions, an important point is to determine the groups of permutations σ of $\mathbb{F}_2^n$ and τ of $\mathbb{F}_2^m$ such that, if F satisfies the property, then $\tau \circ F \circ \sigma$ does too, and more generally the groups of permutations Σ of $\mathbb{F}_2^n \times \mathbb{F}_2^m$ such that, if F satisfies the property and if Σ maps the graph of F to the graph of an (n, m)-function G, then G satisfies the property. We say that the composition by σ, τ (resp. Σ) preserves the property and this leads to a notion of equivalence between (n, m)-functions preserving the property. Let us determine the equivalences preserving k-strong non-normality and kth-order sum-freedom (and being a priori the most general as such); this will show one more difference between the two introduced notions. We assume that σ, τ are affine functions since otherwise the affineness of an affine space A is not preserved when applying σ to A and the weak normality/strong non-normality of the restriction of F (resp. the fact that its sum of values equals zero or is nonzero) is not preserved when composing with τ. Also, we assume that Σ is an affine function. We are then led to determining whether each notion is preserved by EA equivalence, respectively, by CCZ equivalence.

Proposition 6. *For every $k \geq 2$, every $n \geq k$ and every m, the property of being k-strongly non-normal, for an (n, m)-function, is CCZ invariant.*

Proof. Let L be an affine automorphism of $\mathbb{F}_2^n \times \mathbb{F}_2^m$ and let F and G be two (n, m)-functions such that the graph $\{(x, G(x)); x \in \mathbb{F}_2^n\}$ of G equals the image by L of the graph $\{(x, F(x)); x \in \mathbb{F}_2^n\}$ of F. Let F_1 and F_2 be defined as recalled in Section 2: $F_1(x) = L_1(x, F(x))$ and $F_2(x) = L_2(x, F(x))$, where $L = (L_1, L_2)$. Recall that we have $G = F_2 \circ F_1^{-1}$. If F is not k-strongly non-normal, then let A be a k-dimensional affine subspace of $\mathbb{F}_2^n$ over which F is affine. Then F_1 is affine over A. Moreover, the image A' of A by F_1 is a k-dimensional affine subspace of $\mathbb{F}_2^n$, and F_1^{-1} is affine over A'. Besides, F_2 is affine over A. Then G is affine over A'. Hence CCZ equivalence preserves the fact of not being k-strongly non-normal. This completes the proof, by contraposition. □

Proposition 7. *For every $k \geq 3$, the property of being kth-order sum-free for an (n, m)-function is EA-invariant, but not CCZ-invariant in general, and for $k = 2$, it is CCZ invariant.*

Proof. The notion is clearly EA invariant for every $k \geq 2$, $n \geq k$ and m, by contraposition again, since the fact that a function sums to zero over at least one k-dimensional affine space is preserved by affine equivalence and by the addition of affine functions (it is even preserved by the addition of functions of algebraic degree at most $k - 1$). The notion is CCZ invariant for $k = 2$ since APNness is is CCZ invariant, see for instance [18, Subsection 3.4.1]). For $k \geq 3$, it is easy to find examples of kth-order sum-free (n, n)-permutations whose compositional inverses are not kth-order sum-free. For instance, in $\mathbb{F}_{2^5}$, the power function x^7 is third-order sum-free as we saw with Proposition 1, while its inverse equals x^9 (indeed $9 \times 7 = 63 \equiv 1 \pmod{31}$) and is then not third-order sum-free, since it is quadratic. □

5 The case of the multiplicative inverse function

The multiplicative inverse function is the function from the field $\mathbb{F}_{2^n}$ to itself whose univariate representation (see Section 2) equals x^{2^n-2} (which will also be denoted by x^{-1} since the exponents live in $\mathbb{Z}/(2^n-1)\mathbb{Z}$). This power function coincides with the inverse function $x \mapsto \frac{1}{x}$ over $\mathbb{F}_{2^n}^*$ and maps 0 to 0. Its algebraic degree equals $n-1$. It is used (with n even for computational reasons) in the S-boxes of many of the most important block ciphers such as AES. It contributes (in principle) optimally[9] to the resistance of ciphers using it as an S-box when n is odd (it is then APN) and sub-optimally when n is even (it is then differentially 4-uniform).

5.1 The k-strong non-normality of multiplicative inverse function

For n odd, since the inverse function is APN [50], it is k-strongly non-normal for every $k \geq 2$.
For n even, it is only differentially 4-uniform [50] and we need to study whether it is 3-strongly non-normal. Let a, b be any $\mathbb{F}_2$-linearly independent elements of $\mathbb{F}_{2^n}$. If x is $\mathbb{F}_2$-linearly independent of a, b, then $x^{-1}+(x+a)^{-1}+(x+b)^{-1}+(x+a+b)^{-1} = \frac{ab(a+b)}{x(x+a)(x+b)(x+a+b)}$ does not vanish. If $x \in \mathbb{F}_{2^n}$ is $\mathbb{F}_2$-linearly dependent of a, b, we have that $x^{-1}+(x+a)^{-1}+(x+b)^{-1}+(x+a+b)^{-1} = \frac{1}{a}+\frac{1}{b}+\frac{1}{a+b} = \frac{a^2+b^2+ab}{ab(a+b)} = \frac{b}{a(a+b)}\left(\left(\frac{a}{b}\right)^2+\frac{a}{b}+1\right)$ equals 0 if, and only if, $a \in \{wb, w^2b\}$ where w is a primitive element of $\mathbb{F}_4$. Let A be any 3-dimensional affine space, then there exist $\mathbb{F}_2$-linearly independent elements a, b in the direction of A which are not such that $a \in \{wb, w^2b\}$ (indeed, b being chosen, the set $\{0, b, wb, w^2b\}$ is a vector space of dimension 2 only) and then the restriction of F to A is not affine. We deduce that F is 3-strongly non-normal, that is, k-strongly non-normal for every $k \geq 3$.

5.2 Sums of the values taken by the multiplicative inverse function over affine spaces not containing 0

In this subsection, we obtain an explicit expression of the sum of the values of the multiplicative inverse function taken over affine subspaces of $\mathbb{F}_{2^n}$ that are not vector subspaces. This allows us to prove that such sum is always nonzero.
Let E_k be any k-dimensional vector subspace of $\mathbb{F}_{2^n}$. It is well-known that the polynomial $L_{E_k}(x) = \prod_{u \in E_k}(x+u)$ is a linearized polynomial. This can be proved by induction: given a basis $(a_1, \ldots, a_k)$ of $\mathbb{F}_2^k$, let E_{k-1} be generated by

[9] In a local sense, since it may happen that a suboptimal S-box leads to a better resistance to differential cryptanalysis if its interaction with the linear layer of the cipher is more favourable to the designer.

$a_1, \ldots, a_{k-1}$, then if $L_{E_{k-1}}$ is linear, we have:

$$\begin{aligned} L_{E_k}(x) &= L_{E_{k-1}}(x) L_{E_{k-1}}(x + a_k) \\ &= L_{E_{k-1}}(x)(L_{E_{k-1}}(x) + L_{E_{k-1}}(a_k)) \\ &= \left(L_{E_{k-1}}(x)\right)^2 + L_{E_{k-1}}(a_k) L_{E_{k-1}}(x) \end{aligned} \tag{14}$$

and L_{E_k} is linear. Let us write then:

$$L_{E_k}(x) = \sum_{i=0}^{k} b_{k,i} x^{2^i}, \tag{15}$$

where $b_{k,k} = 1$ and $b_{k,0} = \prod_{u \in E_k, u \neq 0} u \neq 0$.
Since we are in characteristic 2, the (polynomial) derivative of $L_{E_k}(x)$ equals $L'_{E_k}(x) = b_{k,0}$, while according to the classical formula on the derivative of a product, we have: $L'_{E_k}(x) = \sum_{u \in E_k} \prod_{v \in E_k, v \neq u} (x + v)$. For $x \in E_k$, this does not give any information (indeed, it gives $b_{k,0} = \prod_{v \in E_k, v \neq x} (x + v)$), but for $x \notin E_k$, this gives $b_{k,0} = \left(\sum_{u \in E_k} \frac{1}{x+u}\right) L_{E_k}(x)$. We have then:

Theorem 1. *For every $0 \leq k \leq n$, let E_k be any k-dimensional $\mathbb{F}_2$-subspace of $\mathbb{F}_{2^n}$ and let $F(x) = x^{2^n - 2} = x^{-1}$ be the multiplicative inverse function over $\mathbb{F}_{2^n}$. We have:*

$$\forall x \notin E_k, \sum_{u \in E_k} F(x + u) = \sum_{u \in E_k} \frac{1}{x + u} = \frac{\prod_{u \in E_k, u \neq 0} u}{\prod_{u \in E_k} (x + u)} = \frac{b_{k,0}}{L_{E_k}(x)} \neq 0, \tag{16}$$

where $L_{E_k}(x) = \prod_{u \in E_k} (x + u)$ and $b_{k,0}$ is its coefficient of x. Hence, F sums to a nonzero value over the affine space $x + E_k$.

Remark. For every $0 \leq k \leq n$, the restriction of the multiplicative inverse function to any k-dimensional affine subspace of $\mathbb{F}_{2^n}$ that is not a vector space has then maximal algebraic degree k, when viewed as a (k, n)-function. This property seems rare among all permutations over $\mathbb{F}_2^n$. Summing the values taken over affine spaces is probably a good way of distinguishing the multiplicative inverse function from random (n, n)-functions or permutations. It is not clear whether this may allow to guess that a secret S-box used in a block cipher is equivalent to the inverse function or, when we know that an S-box used in a cipher is the multiplicative inverse function, if this can be exploited in cryptanalyses. ◇

Remark. In [37] is shown that the only affine spaces that are mapped by the inverse function to affine spaces are the multiplicative cosets of the subfields of $\mathbb{F}_{2^n}$ (0 included). The result of [37, Theorem 1], stating that the affine spaces that are not vector spaces cannot be mapped by the inverse function to affine spaces, is a direct consequence of Theorem 1 in the present paper. Indeed, the sum of the values in an affine space of dimension at least 2 equals 0. ◇

5.3 Sums of the values taken by the multiplicative inverse function over linear subspaces

The case of subspaces containing 0 (that is, linear subspaces) is much more complex (except, of course, when the linear space is a subfield, since any permutation of this subfield sums to zero over it). Computer investigations made for $6 \leq n \leq 12$ show that the inverse function is not kth-order sum-free, whatever is $k \in \{3, \ldots, n-3\}$ (but we could see that for $n = 5$, it is 3rd-order sum-free). Proving this for every n will probably need much work (we could not find a general result allowing to prove it).
Anyway, in practice, cryptanalyses may not consider all affine spaces, and it may then be more important to develop tools for evaluating the sum taken by the inverse function over specific linear spaces (whether one has been able to show that it is not kth-order sum-free, or one could not determine it). The simplest way of describing a linear space is of course by a basis. Let us give then an expression of the sum of inverses by means of such a basis.

Let $\phi_k(x) = \prod_{u \in E_k, u \neq 0} (x+u) = \frac{L_{E_k}(x)}{x}$. According to Relation (15), we have $\phi_k(x) = \sum_{i=0}^{k} b_{k,i} x^{2^i - 1}$. Then $\phi_k(0) = \prod_{u \in E_k, u \neq 0} u = b_{k,0}$ and $\phi_k'(0) = b_{k,1}$. The derivative of a product formula gives $\phi_k'(x) = \sum_{u \in E_k, u \neq 0} \prod_{v \neq 0, v \neq u} (x+v)$ and then

$$\sum_{u \in E_k, u \neq 0} \frac{1}{u} = \frac{\phi_k'(0)}{\phi_k(0)} = \frac{b_{k,1}}{b_{k,0}}. \tag{17}$$

We have seen in Relation (7) that, for every $a_1, \ldots, a_k$, we have:

$$\sum_{\sigma \in G_k} \prod_{i=1}^{k} a_i^{2^{\sigma(i)}} = \prod_{l \in \mathbb{F}_2^k, l \neq 0} \left(\sum_{i=1}^{k} l_i a_i \right) = b_{k,0}.$$

Changing k into $k+1$ and denoting a_{k+1} by x, we obtain:

$$\begin{aligned} \sum_{\sigma \in G_{k+1}} x^{2^{\sigma(k+1)}} \prod_{i=1}^{k} a_i^{2^{\sigma(i)}} &= \prod_{l \in \mathbb{F}_2^{k+1}, l \neq 0} \left(l_{k+1} x + \sum_{i=1}^{k} l_i a_i \right) \\ &= \left(\prod_{l \in \mathbb{F}_2^k, l \neq 0} \left(\sum_{i=1}^{k} l_i a_i \right) \right) \left(\prod_{l \in \mathbb{F}_2^k,} \left(x + \sum_{i=1}^{k} l_i a_i \right) \right). \end{aligned}$$

Let $(a_1, a_2, \ldots, a_k)$ be a basis of E_k, we obtain then:

$$\sum_{\sigma \in G_{k+1}} x^{2^{\sigma(k+1)}} \prod_{i=1}^{k} a_i^{2^{\sigma(i)}} = \phi_k(0) L_{E_k}(x).$$

Hence, $b_{k,1}$ (the coefficient of x^2 in $L_{E_k}(x)$) equals the coefficient of x^2 in $\sum_{\sigma \in G_{k+1}} x^{2^{\sigma(k+1)}} \prod_{i=1}^{k} a_i^{2^{\sigma(i)}}$, divided by $\phi_k(0) = b_{k,0}$, and we deduce:

Proposition 8. *Let $2 \leq k \leq n$. Let E_k be any k-dimensional $\mathbb{F}_2$-subspace of $\mathbb{F}_2^n$ and $(a_1, \ldots, a_k)$ a basis of E_k. Let G_k be the set of bijective functions from $\{1, \ldots, k\}$ to $\{0, \ldots, k-1\}$ and G'_k the set of bijective functions from $\{1, \ldots, k\}$ to $\{0, 2, \ldots, k\}$. We have:*

$$\sum_{u \in E_k, u \neq 0} \frac{1}{u} = \frac{\sum_{\sigma \in G'_k} \prod_{i=1}^{k} a_i^{2^{\sigma(i)}}}{\left(\sum_{\sigma \in G_k} \prod_{i=1}^{k} a_i^{2^{\sigma(i)}}\right)^2}.$$

This formula may not allow to reduce the complexity of the computation of the sum of inverses, but it shows a nice mathematical structure.

Corollary 2. *Let $2 \leq k \leq n$. The multiplicative inverse function over $\mathbb{F}_{2^n}$ is kth-order sum-free if and only if the function:*

$$\sum_{\sigma \in G'_k} \prod_{i=1}^{k} a_i^{2^{\sigma(i)}}, \tag{18}$$

where G'_k is the set of bijective functions from $\{1, \ldots, k\}$ to $\{0, 2, \ldots, k\}$, vanishes only at the $2^{kn} - (2^n - 1)(2^n - 2)\ldots(2^n - 2^{k-1})$ k-tuples $(a_1, \ldots, a_k) \in (\mathbb{F}_{2^n})^k$ whose terms are $\mathbb{F}_2$-linearly dependent elements of $\mathbb{F}_{2^n}$.

Indeed, we know according to Relation (7) that the expression $\sum_{\sigma \in G_k} \prod_{i=1}^{k} a_i^{2^{\sigma(i)}} = \prod_{l \in \mathbb{F}_2^k, l \neq 0} \left(\sum_{i=1}^{k} l_i a_i\right)$ vanishes if and only if $a_1, \ldots, a_k$ are $\mathbb{F}_2$-linearly dependent.

Conclusion

We have introduced and studied two natural generalizations of almost perfect nonlinearity (APN), called k-strong non-normality (notion which already existed for Boolean functions) and kth-order sum-freedom. The latter is mathematically and practically more interesting but we briefly studied k-strong non-normality to compare it with the other notion. While, at the smallest possible order 2, these two generalizations both coincide with APNness, they behave for larger orders quite differently from each other (in particular, the latter is much stronger than the former) and from APNness. We have seen that their study poses interesting questions. We have stated the following open problems:

- Completely determine for $k > 2$, the set of those triples (n, m, k) for which k-strongly non-normal (n, m)-functions exist *and determine these functions.*
- Completely determine for $k > 2$, the set of those triples (n, m, k) for which kth-order sum-free (n, m)-functions exist (for $m = n$, we answered this question) *and determine these functions.*
- Find a simpler characterization of the k-strong non-normality of vectorial functions by means of the Walsh transform.

- Prove that the Kasami APN functions are not 3rd-order sum-free for $n \geq 6$.
- Study the k-strong non-normality and kth-order sum-freedom of the other known infinite classes of APN functions.
- Prove in particular that the multiplicative inverse function is not kth-order sum-free for every $k \in \{3, \ldots, n-3\}$ $(n \geq 6)$.

The (partial) study of the behavior of the multiplicative inverse function over $\mathbb{F}_{2^n}$ with respect to kth-order sum-freedom, led to an interesting property of this particular but cryptographically important infinite class of functions: it sums to non-zero values over all affine subspaces of their domain that are not linear subspaces. This property is rare among random functions.
Since the k-strong non-normality notion seems too weak when applied to vectorial functions and the kth-order sum-freedom seems too strong, we could consider the following intermediate notion: an (n, m)-function is called (k, l)-degree-constrained, for $2 \leq l \leq k \leq n$ if, for every k-dimensional affine space A, the restriction of F to A has algebraic degree at least l (so that k-strong non-normality corresponds to $l = 2$ and kth-order sum-freedom corresponds to $l = k$).

Acknowledgement. The author is grateful to Vincent Rijmen for his useful indications on the cryptographic relevance of the two notions studied in this paper, to Gregor Leander, Willi Meier and Jorge Nakahara for their kind information about cryptanalyses, and to Stjepan Picek for his kind help with computations. He also thanks the anonymous reviewers whose comments improved the quality of the paper and simplified the proof of Proposition 1.

References

1. R. Aragona, M. Calderini, D. Maccauro and M. Sala. On some differential properties of Boolean functions. *Applicable Algebra in Engineering, Communication and Computing* 27 (5), pp. 359-372, 2016.
2. P. Barreto, V. Rijmen, J. Nakahara, B. Preneel, J. Vandewalle, H.-Y. Kim. Improved SQUARE Attacks against Reduced-Round HIEROCRYPT. Fast Software Encryption (FSE '01), *Lecture Notes in Computer Science* 2355, pp. 165-173.
3. C. Beierle and C. Carlet. Gold functions and switched cube functions are not 0-extendable in dimension $n > 5$. *Designs, Codes and Cryptography* 91(2), pp. 433-449, 2023.
4. C. Beierle, G. Leander, and L. Perrin. Trims and extensions of quadratic APN functions. *Designs, Codes and Cryptography*, 90(4):10091036, 2022.
5. T. Beyne and M. Verbauwhede. Integral Cryptanalysis Using Algebraic Transition Matrices. *IACR Transactions on Symmetric Cryptology*, 2023 (4), pp. 244-269. See also *Cryptology ePrint Archive*, 2023.
6. E. Biham and A. Shamir. Differential Cryptanalysis of DES-like Cryptosystems. *Journal of Cryptology*, Vol 4, no.1, pp. 3-72, 1991.
7. C. Blondeau and K. Nyberg. Perfect nonlinear functions and cryptography. *Finite Fields and Their Applications* 32, pp. 120-147, 2015.
8. C. Boura and A. Canteaut. Another view of the division property. Proceedings of CRYPTO 2016, Part I. *Lecture Notes in Computer Science* 9814 pp. 654-682, 2016.

9. A. Braeken, C. Wolf, and B. Preneel. Normality of vectorial functions. *Proceedings of 10th IMA International Conference, Lecture Notes in Computer Science* 3796, pp. 186-200, 2005.
10. L. Budaghyan, M. Calderini, C. Carlet, D. Davidova and N. S. Kaleyski. On Two Fundamental Problems on APN Power Functions. *IEEE Trans. Inf. Theory* 68(5), pp. 3389-3403, 2022.
11. L. Budaghyan, C. Carlet, T. Helleseth and N. Kaleyski. On the Distance Between APN Functions. *IEEE Transactions on Information Theory* 66 (9), pp. 5742-5753, 2020. See also: Changing Points in APN Functions. *IACR Cryptology ePrint Archive* (http://eprint.iacr.org/) 2018/1217.
12. L. Budaghyan, N. Kaleyski, S. Kwon, C. Riera and P. Stănică. Partially APN Boolean functions and classes of functions that are not APN infinitely often. *Special Issue on Boolean Functions and Their Applications 2018, Cryptography and Communications* 12 (3), pp. 527-545.
13. A. Canteaut, M. Daum, H. Dobbertin and G. Leander. Finding nonnormal bent functions. *Discrete Applied Mathematics* 154, pp. 202 - 218, 2006. See also "Normal and Non-Normal Bent Functions". *Proceedings of the Workshop on Coding and Cryptography 2003*, pp. 91-100, 2003".
14. C. Carlet. *Codes de Reed-Muller, codes de Kerdock et de Preparata.* PhD thesis. Publication of LITP, Institut Blaise Pascal, Université Paris 6, 90.59, 1990.
15. C. Carlet. Two new classes of bent functions. *Proceedings of EUROCRYPT 1993, Lecture Notes in Computer Science* 765, pp. 77-101, 1994.
16. C. Carlet, On Cryptographic Complexity of Boolean Functions. *Finite Fields with Applications to Coding Theory, Cryptography and Related Areas* (Proceedings of the Conference Fq6), Springer-Verlag, Berlin, pp. 53-69, 2002.
17. 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).
18. C. Carlet. Boolean Functions for Cryptography and Coding Theory. Monograph in *Cambridge University Press*, 562 pages, 2021.
19. C. Carlet. On the properties of the Boolean functions associated to the differential spectrum of general APN functions and their consequences. *IEEE Transactions on Information Theory* 67(10), pp.6926-6939, 2021.
20. 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.
21. C. Carlet, S. Feukoua, and A. Sălăgean. On the algebraic degree stability of Boolean functions when restricted to affine spaces. To appear in the proceedings of the *Thirteenth International Workshop on Coding and Cryptography* WCC 2024.
22. C. Carlet, K. H. Kim, and S. Mesnager. A direct proof of APN-ness of the Kasami functions. *Designs, Codes and Cryptography*, vol. 89, p. 441-446, 2021.
23. 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.
24. W.E. Clark, X.D. Hou and A. Mihailovs. The affinity of a permutation of a finite vector space. *Finite Fields and Their Applications* 13(1), pp. 80-112, 2007.
25. J. Daemen, L. Knudsen and V. Rijmen. The block cipher square. *Proceedings of Fast Software Encryption FSE 1997, Lecture Notes in Computer Science*, vol. 1267, pp. 149165, 1997.

26. J. Daemen and V. Rijmen, *The Design of Rijndael: AES - The Advanced Encryption Standard*. Springer, 2002.
27. J. Daemen and V. Rijmen. Understanding two-round differentials in AES. *Proceedings of International Conference on Security and Cryptography for Networks, Lecture Notes in Computer Science* 4116, pp. 78-94, 2006.
28. I. Dinur and A. Shamir. Cube attacks on tweakable black box polynomials. *Proceedings of EUROCRYPT 2009, Lecture Notes in Computer Science* 5479, pp. 278299. Springer (2009).
29. H. Dobbertin. Construction of bent functions and balanced Boolean functions with high nonlinearity. *Proceedings of Fast Software Encryption FSE 1994, Lecture Notes in Computer Science* 1008, pp. 61-74, 1995.
30. P. Ellingsen, P. Felke, C. Riera, P. Stănică and A. Tkachenko. C-differentials, multiplicative uniformity, and (almost) perfect c-nonlinearity. *IEEE Transactions on Information Theory* 66(9), pp.5781-5789, 2020.
31. European Telecommunications Standards Institute, Technical Specification 135 202 V9.0.0: Universal mobile telecommunications system (UMTS); LTE; specification of the 3GPP confidentiality and integrity algorithms; Document 2: KASUMI specification (3GPP TS 35.202 V9.0.0 Release 9).
32. A.A. Frolova. The essential dependence of Kasami bent functions on the products of variables. *Journal of Applied and Industrial Mathematics* 7, pp.166-176, 2013.
33. P. Hebborn, B. Lambin, G. Leander and Y. Todo. Lower bounds on the degree of block ciphers. *Proceedings of ASIACRYPT 2020, Part I, Lecture Notes in Computer Science*, vol. 12491, pp. 537566, 2020.
34. P. Hebborn, G. Leander, and A. Udovenko. Mathematical aspects of division property. *Cryptography and Communications* 15, no. 4, pp. 731-774, 2023.
35. L. Knudsen. Truncated and higher order differentials. *Proceedings of Fast Software Encryption FSE 1995, Lecture Notes in Computer Science* 1008, pp. 196-211, 1995.
36. L. Knudsen and D. Wagner. Integral cryptanalysis. *Proceedings of Fast Software Encryption FSE 2002, Lecture Notes in Computer Science* vol. 2365, pp. 112127, 2002.
37. N. Kolomeec and D. Bykov. On the image of an affine subspace under the inverse function within a finite field. To appear in *Designs, Codes and Cryptograhy*.
38. M. Kuroda and S. Tsujie. A generalization of APN functions for odd characteristic. *Finite fields and their applications* 47, pp. 64-84, 2017.
39. X. Lai. Higher order derivatives and differential cryptanalysis. *Proceedings of the "Symposium on Communication, Coding and Cryptography", in honor of J. L. Massey on the occasion of his 60'th birthday*, pp. 227-233, 1994.
40. B. Lambin, P. Derbez and P. Fouque. Linearly equivalent s-boxes and the division property. *Designs, Codes and Cryptography* 88 (10), pp. 22072231, 2020.
41. S. Lang. Algebra, Graduate Texts in Mathematics, 211 (Revised third ed.), New York: Springer-Verlag, 2002.
42. G. Leander, B. Minaud and S. Rønjom. A generic approach to invariant subspace attacks: Cryptanalysis of Robin, iSCREAM and Zorro. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 254-283, 2015. Berlin, Heidelberg: Springer Berlin Heidelberg.
43. S. Li, W. Meidl, A. Polujan, A. Pott, C. Riera, and P. Stănică. Vanishing flats: A combinatorial viewpoint on the planarity of functions and their application. *IEEE Transactions on Information Theory* 66 (11), pp.7101-7112, 2020.
44. F. J. MacWilliams and N. J. Sloane. *The theory of error-correcting codes*, North Holland. 1977.

45. M. Matsui. Block encryption algorithm MISTY. *Proceedings of Fast Software Encryption FSE 1997, Lecture Notes in Computer Science* 1267, pp. 54-68, 1997.
46. M. J. Mihaljevic, S. Gangopadhyay, G. Paul, H. Imai. Generic cryptographic weakness of k-normal Boolean functions in certain stream ciphers and cryptanalysis of grain-128. *Period. Math. Hung.* 65(2): 205-227, 2012.
47. E. H. Moore. A two-fold generalization of Fermat's theorem. *Bulletin of the American Mathematical Society* 2 (7), pp. 189199, 1896.
48. K. Nyberg. Perfect non-linear S-boxes. *Proceedings of EUROCRYPT' 91, Lecture Notes in Computer Science* 547, pp. 378-386, 1992.
49. K. Nyberg. On the construction of highly nonlinear permutations. *Proceedings of EUROCRYPT' 92, Lecture Notes in Computer Science* 658, pp. 92-98, 1993.
50. K. Nyberg. Differentially uniform mappings for cryptography. *Proceedings of EUROCRYPT' 93, Lecture Notes in Computer Science* 765, pp. 55-64, 1994.
51. K. Nyberg and L. R. Knudsen. Provable security against differential cryptanalysis. *Proceedings of CRYPT0' 92, Lecture Notes in Computer Science* 740, pp. 566-574, 1993.
52. B. Sun, L. Qu and C. Li. New cryptanalysis of block ciphers with low algebraic degree. *Proceedings of FSE 2009, Lecture Notes in Computer Science* 5665, pp. 180-192, 2009.
53. H. Taniguchi. D-property for APN functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2^{n+1}$. *Cryptography and Communications* 15, no. 3 , pp. 121, 2023.
54. Y. Todo. Structural evaluation by generalized integral property. *Proceedings of EUROCRYPT 2015, Lecture Notes in Computer Science* 9056, pp. 287-314, 2015.
55. M.R. Zaba, H. Raddum, M. Henricksen and E. Dawson. Bit-Pattern Based Integral Attack. *Proceedings of FSE 2008, Lecture Notes in Computer Science* 5086, pp. 363-381, 2008.